Investigation of the Expression of the O:54 Antigen from *Salmonella enterica*

By

Brittany Alexandra Hunt

A Thesis Presented to
The Faculties of Graduate Studies
of
The University of Guelph

In partial fulfilment of requirements for the degree of
Master of Science
in
Molecular and Cellular Biology

Guelph, Ontario, Canada

© Brittany Alexandra Hunt, 2013
ABSTRACT

INVESTIGATION OF THE EXPRESSION OF THE O:54 ANTIGEN FROM
SALMONELLA ENTERICA

Brittany Alexandra Hunt
University of Guelph, 2013

Advisor: Professor Chris Whitfield

WbbF is a proposed member of the synthase family of dual glycosyltransferase-exporter proteins. WbbF assembles the O:54 antigen from Salmonella enterica. The wb*O:54 operon (encoding wbbE, wbbF and mnaA) is found on a naturally-occurring plasmid (pWQ799). Previous research suggested only WbbE and WbbF are necessary for O:54 antigen assembly due to the presence of a chromosomal gene encoding WecB, a functional homolog of MnaA. WecB and MnaA synthesize the essential UDP-ManNAc donor. However, this study determined optimal production of O:54 in E. coli was dependent on the presence of MnaA. A growth defect was observed in cells producing O:54 antigen, but unable to ligate it to the lipopolysaccharide acceptor (due to ΔwaaL). Cells recovered growth over time via mutations that turned off (or down) O:54 antigen biosynthesis. During this study, a cysteine-free version of WbbF that retained function was established, providing a key resource for future topology mapping experiments.
# Table of Contents

**List of Figures** ........................................................................................................................................ v

**List of Tables** .......................................................................................................................................... vii

**List of Abbreviations** ............................................................................................................................... ix

## 1.0 INTRODUCTION ................................................................................................................................... 1

1.1 Glycosyltransferases .............................................................................................................................. 1

1.1.1 Properties of glycosyltransferases .................................................................................................... 1

1.1.2 Processive glycosyltransferases ......................................................................................................... 5

1.1.3 Conserved and unique domains and motifs found in the HasA sub-family of processive GTs ............... 5

1.1.4 Hyaluronan synthases ......................................................................................................................... 6

1.1.5 Classification of HASs ......................................................................................................................... 9

1.1.6 Cellulose synthases ........................................................................................................................... 12

1.1.7 Alginate synthases ............................................................................................................................. 14

1.1.8 Type III capsule synthase in *Streptococcus pneumoniae* ................................................................. 16

1.1.9 O:54 biosynthesis and the role of the putative synthase ................................................................. 17

1.2 Lipopolysaccharide ............................................................................................................................... 18

1.2.1 Gram-negative cell envelope .......................................................................................................... 18

1.2.2 Structure and function of lipopolysaccharide ................................................................................... 20

1.2.3 Overview of LPS biosynthesis .......................................................................................................... 23

1.2.4 O-antigen synthesis .......................................................................................................................... 25

1.2.5 The origin of serovar O:54 .............................................................................................................. 27

1.2.6 Proposed synthase pathway for O:54 biosynthesis .......................................................................... 31

1.2.7 WbbF, the O:54 synthase from *S. enterica* sv. Borreze ................................................................. 33

1.3 Aims of this study and objectives ......................................................................................................... 34
2.0 MATERIALS AND METHODS .................................................................................................................. 38

2.1 Bacterial strains, plasmids and growth conditions .............................................................................. 38
2.2 Bacterial growth analysis ...................................................................................................................... 38
2.3 General DNA methods ......................................................................................................................... 42
2.4 Construction of ΔwecB and ΔwaaL in E. coli Top10 ............................................................................. 48
2.5 Site-directed mutagenesis .................................................................................................................... 49
2.6 Lipopolysaccharide analysis methods ................................................................................................... 49
2.7 Agglutination reactions ........................................................................................................................ 52
2.8 Protein analysis techniques .................................................................................................................. 52
2.10 In silico analysis methods ................................................................................................................... 54

3.0 RESULTS .............................................................................................................................................. 56

3.1 Optimization of O-antigen detection techniques .................................................................................. 56
3.2 Effects of expressing WbbE and WbbF from separate plasmids on O:54 LPS profiles......................... 68
3.3 Upstream sequence of wbbE does not improve O:54 production ....................................................... 74
3.4 Construction of a wecB deletion mutant in E. coli Top10 ................................................................. 77
3.5 Presence of MnaA required for optimal O:54 production in E. coli Top10 ....................................... 80
3.6 Effect of MnaA, WbbE and WbbF protein levels on O:54 LPS production ........................................... 82
3.7 Deletion of waaL from E. coli Top10 .................................................................................................... 89
3.8 Production of O:54 in ΔwaaL deletion mutant causes a growth defect that is recovered via endogenous inhibition of O:54 production .......................................................... 90
3.9 Development of a functional cysteine-free WbbF ............................................................................. 99
3.10 Preliminary protein experiments ....................................................................................................... 100

4.0 DISCUSSION ....................................................................................................................................... 108

REFERENCES .............................................................................................................................................. 128
List of Figures

Figure 1.1: Two major structural folds of GTs ................................................................. 2
Figure 1.2: Inverting and retaining catalytic mechanisms of GTs ........................................ 4
Figure 1.3: Topology map of hyaluronan synthase from Streptococcus pyogenes (SpHAS) .... 11
Figure 1.4: Crystal structure of the BcsA-BcsB complex with a translocating polysaccharide at 3.25 Å resolution .................................................................................................................. 15
Figure 1.5: Depiction of the cell envelope of Gram-negative bacteria .................................. 19
Figure 1.6: Structure of a typical lipopolysaccharide (LPS) molecule, and the lipid A and core structures in S. enterica ................................................................................................................. 21
Figure 1.7: Cartoon models of the three assembly pathways for O-Ag biosynthesis .............. 26
Figure 1.8: Structure of O:54 antigen and Organization of pWQ799 .................................. 28
Figure 1.9: Proposed synthase-dependent pathway for O:54 antigen biosynthesis ............... 32
Figure 1.10: Clustal Omega amino acid alignment .................................................................. 35
Figure 3.1: Confirmation of O:54 production in E. coli Top10 strain in the presence of WbbE and WbbF, but not WbbF alone ........................................................................................................... 57
Figure 3.2: Resolution quality of LPS samples on SDS-PAGE gels with different compositions. ........................................................................................................................................... 60
Figure 3.3: Comparison of silver staining methods for optimization of O:54 LPS visibility..... 61
Figure 3.4: Comparison of O:54 LPS separation using a glycine-based and tricine-based SDS-PAGE gel systems .......................................................................................................................... 63
Figure 3.5: Effect of sample buffer composition on LPS mobility and resolution ................. 66
Figure 3.6: Effect of sample buffer SDS concentration on LPS mobility and resolution. ....... 67
Figure 3.7: Growth curve of E. coli Top10 and E. coli Top10 containing pWQ800. ............... 69
Figure 3.8: LPS profiles of E. coli Top10 cells in mid-logarithmic and stationary growth phases. ........................................................................................................................................... 70
Figure 3.9: Cell counts (colony-forming units) for S. enterica sv. Borreze, E. coli Top10 and CWG1167 (ΔwecB) corresponding to an OD_{600} of 1.0 ................................................................. 71
Figure 3.10: O:54 production in *E. coli* Top10 by WbbE + WbbF expressed from separate constructs and expressed from pWQ800 (wb*O:54ΔmnaA*) compared to *S. enterica* sv. Borreze.

Figure 3.11: Nucleotide sequence of *wbbE* including 108 base pairs upstream.

Figure 3.12: Inclusion of DNA upstream of the potential start codon of WbbE does not improve O:54 production.

Figure 3.13: Generation of CWG1167, an *E. coli* Top10 ΔwecB mutant.

Figure 3.14: Expression of O:54 LPS in CWG1167 (ΔwecB).

Figure 3.15: Examination of effect of different sources of UDP-GlcNAc 2-epimerase on O:54 antigen production.

Figure 3.16: Titration of MnaA protein expression level affects quantity of O:54 LPS.

Figure 3.17: WbbF protein is not detected in boiled samples.

Figure 3.18: Titration of WbbE and WbbF protein expression levels affects O:54 production.

Figure 3.19: Production of CWG1168, an *E. coli* Top10 ΔwaaL deletion mutant.

Figure 3.20: Production of O:54 in CWG1168 (ΔwaaL).

Figure 3.21: Growth curve of CWG1168 (ΔwaaL mutant) containing pWQ203 with, and without induction of O:54 biosynthesis.

Figure 3.22: Brightfield microscopy using a differential interference contrast filter of CWG1168 (ΔwaaL) cells before and during O:54 production.

Figure 3.23: Production of O:54 over time in CWG1168 (ΔwaaL).

Figure 3.24: Amino acid sequence of WbbF with cysteine residues highlighted in green and underlined.

Figure 3.25: Analysis of WbbF Cys → Ala mutants.

Figure 3.26: Analysis of function of WbbF Cys → Ala mutant after the sequential substitution of each cysteine residue to alanine.

Figure 3.27: Examination of WbbF expression from different plasmids.

Figure 3.28: Subcellular localization of highly expressed WbbF.

Figure 3.29: Subcellular localization WbbF under low-level expression conditions.
Figure 4.1: Clustal Omega amino acid sequence alignments of WecB from *E. coli* and MnaA from pWQ799 .......................................................... 116

Figure 4.2: *In silico* predictions of the membrane topology of WbbF. ........................................ 125
List of Tables

Table 1.1: *Salmonella enterica* serovars expressing O:54 antigen .................................................. 29

Table 2.1: Bacterial strains used in this study .................................................................................. 39

Table 2.2: Plasmids used in this study ............................................................................................. 40

Table 2.3: Primer sequences used in this study .............................................................................. 43
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>AhT</td>
<td>Anhydrotetracycline</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulfate</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CB</td>
<td>Cascade Blue</td>
</tr>
<tr>
<td>CFU</td>
<td>colony forming units</td>
</tr>
<tr>
<td>cm</td>
<td>chloramphenicol</td>
</tr>
<tr>
<td>ECA</td>
<td>enterobacterial common antigen</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FRT</td>
<td>FLP recognition target</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylglucosamine</td>
</tr>
<tr>
<td>GlcUA</td>
<td>glucuronic acid</td>
</tr>
<tr>
<td>GT</td>
<td>glycosyltransferase</td>
</tr>
<tr>
<td>HA</td>
<td>Hyaluronan</td>
</tr>
<tr>
<td>HAS</td>
<td>hyaluronan synthase</td>
</tr>
<tr>
<td>His&lt;sub&gt;6&lt;/sub&gt;</td>
<td>hexa-histidine</td>
</tr>
<tr>
<td>HRP</td>
<td>horseradish peroxidase</td>
</tr>
<tr>
<td>IM</td>
<td>inner membrane</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>Kdo</td>
<td>3-deoxy-D-&lt;i&gt;manno&lt;/i&gt;-oct-2-ulopyranosonic acid</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>L,D-Hep</td>
<td>L-glycero-D-&lt;i&gt;manno&lt;/i&gt;-heptopyranose</td>
</tr>
</tbody>
</table>
LPS      lipopolysaccharide
ManNAc-6-P  $N$-acetylmannosamine-6-phosphate
ManNAc  $N$-acetylmannosamine
NDP      nucleotide diphosphate
O-Ag     O-antigen
OD$_{600}$ optical density (600 nm)
OM       outer membrane
PAGE     polyacrylamide gel electrophoresis
PBS      phosphate buffered saline
PCR      polymerase chain reaction
PhoA     alkaline phosphatase
PHPT     polyisoprenyl-phosphate hexose-1-phosphate transferases
PNPT     polyisoprenyl-phosphate $N$-acyethylhexosamine-1-phosphate transferases
RBS      ribosome binding site
R-LPS    rough lipopolysaccharide
SDS      sodium dodecyl sulfate
S-LPS    smooth lipopolysaccharide
TBE      Tris-borate-EDTA
TBST     tris-buffered saline tween
TEMED    $N,N,N',N'$-tetramethylethylenediamine
TM       transmembrane
UDP      uridine diphosphate
Und-P    undecaprenyl phosphate
Wt       wildtype
1.0 INTRODUCTION
1.1 Glycosyltransferases

1.1.1 Properties of glycosyltransferases

Glycosyltransferases (GTs) are a large group of enzymes that are ubiquitous in nature. GTs catalyze the transfer of an activated donor sugar to a specific acceptor molecule, forming a glycosidic bond (Lairson et al., 2008). The most common donors are nucleotide diphospho-sugars, like UDP-glucose, but polyprenol-phospho/pyrophospho sugars, and sugar-1-phosphates/pyrophosphates can also be used (Breton et al., 2006). The donor is generally transferred to a growing carbohydrate chain. However, it can also be transferred to any other compound that can be modified by glycosylation (Breton et al., 2012; Lairson et al., 2008). GTs are currently classified into families based on amino acid sequence similarities (Coutinho et al., 2003). Ninety-four families have been identified and they are listed in the CAZy database (http://www.cazy.org/).

The first X-ray crystal structure of a GT was reported by Vrielink et al. (1994) for the bacteriophage T4 β-glucosyltransferase (BGT). As more structures were reported, two conserved folds emerged, termed GT-A and GT-B (Figure 1.1). These folds were initially observed in SpsA from Bacillus subtilis and the BGT structure, respectively (Charnock and Davies, 1999; Vrielink et al., 1994). The GT-A fold consists of two tightly associated β/α/β domains with a central β-sheet and is similar to Rossman-like folds, which are typical of nucleotide-binding proteins (Figure 1.1A) (Lairson et al., 2008). The N-terminal domain is a structurally conserved Rossman-like nucleotide binding domain (Breton et al., 1998; Breton and Imberty, 1999). Topologically, there is another structurally conserved region at the C-terminus. However, it is more variable among the GTs because this region must incorporate acceptor specificity and
**Figure 1.1:** Two major structural folds of GTs (modified from Chang et al., 2011). **A)** GT-A fold of the Human Blood Group GT with the identified signature DXD motif labeled. The required Mn\(^{2+}\) ion and UDP are shown (PDB ID: 3IOJ). Flexible loops (L3’ and L7’) are also labeled **B)** GT-B fold of then oleandomycin GT with its E/D-Q signature region labeled, and UDP and erythromycin identified (PDB ID: 2IYF).
regiospecificity to be able to recognize the substrate (Ünligil and Rini, 2000). The GT-B fold consists of two β/α/β Rossmann-like domains with a connecting linker, and the active site cleft is located between the two domains (Figure 1.1B). The acceptor substrate-binding site is found at the N-terminus and a conserved nucleotide binding domain is found at the C-terminus (Breton et al., 2006; Lairson et al., 2008). A third, less common structural fold, GT-C, has also been described. The C-terminal domain consists of a novel α-helical fold surrounded by three β-sheet domains and an N-terminal transmembrane domain that is vital for activity (Igura et al., 2008). The first extracellular loop in the N-terminal domain is believed to be required for binding the lipid region of the lipid-linked donor. Only two crystal structures of the C-terminal globular domain of GTs with the GT-C fold have been reported thus far (Igura et al., 2008; Lairson et al., 2008; Maita et al., 2010).

GTs are classified as either inverting or retaining enzymes. Inverting GTs catalyze the transfer of a sugar molecule with an inverse stereochemistry compared to the anomeric reaction center of the donor substrate, whereas retaining GTs maintain the same stereochemical configuration as the donor substrate (Figure 1.2) (Lairson et al., 2008; Saxena et al., 1995). The type of GT fold does not have any implications on the stereochemical outcome of the reaction it catalyzes (Bourne and Henrissat, 2001). The mechanism of inverting GTs is a direct displacement S_N2-like reaction (Figure 1.2A) (Murray et al., 1996). The side chain of an acidic amino acid acts as a base and deprotonates the hydroxyl group of the acceptor substrate, facilitating a nucleophilic attack on the anomeric carbon of the donor substrate (Lairson et al., 2008). This causes displacement of the activated phosphate leaving group. The reaction mechanism for retaining GTs is controversial as there is support for two reaction mechanisms: a double displacement reaction involving the formation of a covalently bound glycosyl-enzyme
Figure 1.2: Inverting and retaining catalytic mechanisms of GTs (Breton et al., 2012). A) Inverting GTs follow a direct displacement $S_{N2}$-like reaction, resulting in the addition of a sugar with an inverted configuration at the anomeric carbon. This occurs via a single oxocarbenium ion-like transition state. B) Proposed double-displacement reaction mechanism for retaining GTs involving a covalently bound glycosyl-enzyme intermediate formed. C) Proposed $S_{Ni}$-like retaining reaction mechanism, involving the formation of an oxocarbenium phosphate ion pair intermediate.
intermediate and an \( S_N \)-like mechanism where the acceptor hydroxyl group executes a nucleophilic attack on the anomeric carbon of the donor on the same side as the leaving group departs (Figure 1.2B, C) (Lairson et al., 2008; Monegal and Planas, 2006; Soya et al., 2011). Currently, there is more support for the latter reaction mechanism.

1.1.2 Processive glycosyltransferases

GTs can be further distinguished by their processivity: nonprocessive GTs catalyze the addition of only one sugar residue to an acceptor in a reaction step, whereas processive GTs transfer more than one sugar residue to the acceptor (Saxena et al., 2001). More specifically, processive GTs use one or two independent catalytic active sites to transfer one or more donor sugars with different linkages (Weigel and DeAngelis, 2007). Processive GTs are collectively referred to as synthases. Examples of known or predicted synthases include GTs that synthesize cellulose, chitin, hyaluronic acid (HA), alginate and the O:54 antigen region of the lipopolysaccharide of some \textit{Salmonella enterica} serovars (Keenleyside and Whitfield, 1996; Merighi et al., 2007; Saxena et al., 2001). Synthases that have been verified or predicted to possess a single active site for the processive GT function have been grouped into the HasA subfamily, which is named after one of the most studied enzymes, HasA found in several \textit{Streptococcus} species (DeAngelis et al., 1993; Raetz and Whitfield, 2002). The HasA sub-family is part of the large GT2 family of \( \beta \)-GTs, which contains both processive and non-processive enzymes in the CAZy database. Our current understanding about the structure and function of synthases is discussed below.

1.1.3 Conserved and unique domains and motifs found in the HasA sub-family of processive GTs

Two conserved functional domains were identified by hydrophobic cluster analysis, and are known as Domain A and Domain B (Saxena et al., 1995). Domain A is located near the N-
terminus, and is conserved in most GTs. Conversely, Domain B is located at the C-terminus and is conserved in only processive GTs. It was concluded by the authors of the initial study that Domain B may be a marker for processive GTs. Most processive GTs contain both domains.

Within Domain A, two conserved aspartic acid residues were identified. These residues make up the common DXD catalytic motif found in most GTs across all GT families (Breton and Imberty, 1999; Wiggins and Munro, 1998). The carboxylate groups in this motif interact primarily with the phosphate groups of the nucleotide donor, through the coordination of a divalent cation cofactor (typically Mn$^{2+}$ or Mg$^{2+}$) (Breton et al., 1998; Wiggins and Munro, 1998). Domain B contains at least one other conserved Asp and a QXXRW motif (Saxena et al., 1995). In cellulose synthases, these features are part of a larger motif defined as the D, D, D$_{35}$Q(Q/R)XRW motif (Saxena and Brown, 1997). The Asp residues are believed to be involved in both binding the UDP-sugar donors and in the glycosyl transfer reactions. The QXXRW motif is believed to be involved in holding the glycan within the active site as more sugar residues are added (Saxena et al., 2001). Thus far, only the QXXRW motif present in a bacterial cellulose synthase has been shown to bind the growing glycan at the terminal acceptor disaccharide (Morgan et al., 2013). A conserved FFCGS sequence (only identified in cellulose synthases) was also bound to the growing glycan. The Pseudomonas aeruginosa alginate synthase (Alg8) possesses a LXXRW motif rather than a QXXRW motif (Oglesby et al., 2008). The reason for the change from Q to L is unknown, but it was suggested the residue found in the (Q/L)XXRW motif may be dependent on the identity of the nucleotide diphosphate of the donor.

1.1.4 Hyaluronan synthases

Hyaluronan (HA) is a large, unbranched polysaccharide composed of disaccharide repeat units of alternating N-acetylglucosamine (GlcNAc) and glucuronic acid (GlcUA) connected by
β-1,3 and β-1,4 glycoside linkages (Meyer, 1958). This polymer is expressed ubiquitously in vertebrates, mainly residing on the cell surface or in the extracellular space, but it is also produced by several bacterial species including some *Streptococcus* species, *Bacillus cereus*, *Pasturella multocida*, as well as in some chloroviral species that infect freshwater green algae such as *Paramecium bursaria* chlorella virus (DeAngelis, 1999). HA is synthesized by the hyaluronan synthase (HAS), and was the first processive GT to be discovered (Markovitz et al., 1959).

Although HA is a relatively simple structure, it is considered an important signalling molecule that affects numerous physiological processes, particularly during vertebrate development. Examples include cell migration, cell adhesion, and inflammatory processes (Bollyky et al., 2012; Jiang et al., 2007; Lee and Spicer, 2000; Toole, 2000). The size and concentration of HA are major factors in determining how HA affects these processes. Thus, HASs generally produce a broad range of high molecular weight HA to regulate these physiological processes (Itano and Kimata, 2002). Mammalian HASs are also capable of producing low molecular weight HA that acts as a potent signalling molecule, but affects cellular processes differently than the higher molecular weight form (Lee and Spicer, 2000; Weigel and DeAngelis, 2007).

Three isoforms of HAS were identified in mammals and named HAS1, -2 and -3, in order of discovery (Itano and Kimata, 1996; Shyjan et al., 1996; Spicer et al., 1997; Watanabe and Yamaguchi, 1996). Despite the high degree of homology among the three isoforms (Itano and Kimata, 2002), each have distinctive enzymatic properties. Varying features include the $K_m$ values for the two substrates (UDP-GlcNAc and UDP-GlcUA), enzyme stabilities, and HA elongation rate (Itano et al., 1999). HAS2 and HAS3 were determined to produce HA with the
highest and lowest average molecular masses, respectively. Remarkably, incubation of HAS1 with only UDP-GlcNAc was shown to produce a chito-oligosaccharide, leading other investigators to propose that higher eukaryotic HASs likely evolved from chitin synthases (Lee and Spicer, 2000; Yoshida et al., 2000). Lee and Spicer (2000) also proposed mammalian HASs may have evolved from cellulose synthases.

The viral HAS shares approximately 50% amino acid sequence identity with vertebrate HASs suggesting the has gene was acquired by horizontal gene transfer from a eukaryotic host (DeAngelis et al., 1997). The viral HAS produces HA in the cell wall of the infected host Chlorella cells, making Chlorella cells the only known plant cells to express HA (Graves et al., 1999; Weigel and DeAngelis, 2007). Unfortunately, the biological function of chlorovirus HA remains unclear. The Chlorella virus HAS is being developed as a novel, industrial HA production system (Yamada and Kawasaki, 2005).

Bacterial HA is located on the cell surface as an extracellular capsule and plays a critical role in colonization and infection (Wessels and Bronze, 1994). HA is considered an important virulence factor for group A Streptococcus that protects cells from phagocytosis and complement during infection (Wessels et al., 1994; Wessels et al., 1991; Whitnack et al., 1981). Like HA produced by eukaryotes, the length of bacterial HA is significant, as mutants expressing shorter HA are less virulent due to an increase in susceptibility to host defenses (Chung et al., 2001; Wessels et al., 1994).

The hasA gene, encoding the bacterial HAS enzyme was first discovered in Streptococcus pyogenes, SpHAS (DeAngelis et al., 1993; Dougherty and van de Rijn, 1994). It is located in the has operon, comprised of hasA and one or more genes, depending on the bacterial species, which encode enzymes that are involved in HA precursor sugar biosynthesis (Blank et
al., 2008). The most conserved genes in the has operons are hasA and hasB. HasB is a UDP-glucose dehydrogenase enzyme required to produce the UDP-GlcUA precursor. The enzymes HAS and HasB are considered essential for HA biosynthesis in bacteria. Recently, however, a paralog of HasB encoded in a different area of the chromosome was discovered in *S. pyogenes* (Cole et al., 2012). Thus, HasB is no longer considered essential in this species due to the presence of the functionally redundant paralog.

### 1.1.5 Classification of HASs

HAS enzymes are currently divided into two classes based upon amino acid sequence similarities. Thus far, HAS from *P. multocida* (PmHAS) is the only representative of Class II HASs, and there are large differences between PmHAS and other HASs, as discussed below. All other vertebrate, viral and bacterial HASs have been placed in Class I (DeAngelis, 1999; DeAngelis et al., 1998). The Class I HASs, along with related synthases, form the HasA sub-family of GT-2 enzymes. Members of this group are all membrane-bound, possess a GT-A fold, and are either known or predicted to be processive β-GTs. However, they also possess predicted exporter or translocater function to export the newly synthesized oligosaccharide/polysaccharide across the membrane.

Despite the division of HASs into two classes, all of these enzymes are inverting GTs that only require activated donor UDP-sugars and the presence of Mg$^{2+}$ or Mn$^{2+}$ to assemble the HA product. Thus, these enzymes do not require a lipid or protein carrier, or a primer, to initiate HA glycan assembly (Stoolmiller and Dorfman, 1969; Sugahara et al., 1979; Yoshida et al., 2000). Some HASs add sugars to the reducing end of the growing glycan while other HASs add sugars to the nonreducing terminal (Asplund et al., 1998; Bodevin-Autelet et al., 2005; Prehm, 2006; Tlapak-Simmons et al., 2005).
Class I HASs are predicted to contain four to six transmembrane domains (TMs) and one or two membrane-associated domains (Weigel et al., 1997). Researchers have suggested that Class I HASs are lipid-dependent due to the fact that maximum activity is achieved when membrane-bound or purified HASs are incubated with cardiolipin, although this is not essential for activity (Weigel and DeAngelis, 2007). In contrast to Class I integral membrane proteins, the class II PmHAS is a peripheral protein proposed to have its C-terminus docked to an integral membrane transport protein, commonly found in encapsulated Gram-negative bacteria (Jing and DeAngelis, 2000). Currently, the PmHAS is the only known HAS to synthesize using two independent GT active sites (Jing and DeAngelis, 2000). PmHAS contains two domain A modules, rather than one domain A and one domain B found in all other HAS enzymes (Saxena et al., 1995; Weigel and DeAngelis, 2007).

A topology map of SpHAS has been generated using a fusion reporter system (Figure 1.3) (Heldermon et al., 2001). The N- and C- termini were both located in the cytoplasm, along with a large central loop predicted to contain the catalytic activity. The large loop contains the highest identity to other HASs, as well as chitin and cellulose synthases (Nagahashi et al., 1995; Yoshida et al., 2000). Other members of the HasA sub-family, are predicted to have a similar topology based on high sequence similarities and topology predictions, but this has not been confirmed by direct experimentation.

The proposed export function of the Class I HASs was recently examined by two different research groups. In one study, purified *Streptococcus equisimilis* HAS (SeHAS) was reconstituted in liposomes containing bovine cardiolipin and preloaded with a fluorophore (Cascade Blue, CB). Translocation of CB across the liposomes with reconstituted SeHAS was observed (Medina et al., 2012). The efflux of HA was also eliminated or greatly reduced when
**Figure 1.3:** Topology map of hyaluronan synthase from *Streptococcus pyogenes* (SpHAS) (adapted from Heldermon et al., 2001). The topology was characterized by fusion reporter analysis, using enzymes alkaline phosphotase (PhoA) and β-galatosidase (LacZ), as well as chemical labeling, protease accessibility, and substrate accessibility. SpHAS is proposed to contain four transmembrane domains (labeled 1, 2, 4, and 5) and two membrane-associated domains (labeled 3 and 6). The predicted amino acid residues located at each junction between membrane and cytoplasmic or periplasmic areas are labeled. Outside denotes the periplasmic space and Inside denotes the cytoplasmic space.
residues required for in vivo enzyme activity were substituted, or when SeHAS was heat-
inactivated, oxidized or Cys-modified to inhibit enzyme activity. The GT and export activities of
SeHAS were both recovered after the chemical DTT was slowly reintroduced to the oxidized
enzyme. The second study took a more direct approach, by demonstrating in vitro HA synthesis
and translocation into the lumen of proteoliposomes reconstituted with SeHAS (Hubbard et al.,
2012). Interestingly, this group demonstrated that HA synthesis and translocation are spatially
coupled events, providing support for a proposed intraprotein structural pore required for HA
export. Further investigation of this novel membrane translocation mechanism is required.

1.1.6 Cellulose syntheses

The most abundant biopolymer in the world is cellulose. It is found in plants, some algae,
tunicates, amoeba, and in a wide range of bacterial species including Gluconacetobacter
(formerly known as Acetobacter), Rhizobium, Pseudomonas, Agrobacterium, Escherichia coli
and Salmonella enterica (Ross et al., 1991; Saxena et al., 2001; Zogaj et al., 2001). Cellulose
consists of parallel, unbranched β-1,4-linked D-glucan chains, which form microfibrils through
intra- and interchain hydrogen bonding and Van der Waals forces (Bayley et al., 1957; Green,
1962; Setterfield and Bayley, 1958). The final highly ordered, insoluble crystalline structure is
the main component of plant cell walls, which provides structural rigidity and strength, as well as
protection against environmental stresses (Bacic et al., 1988). Cellulose produced in bacteria is
displayed on the extracellular surface of the cell and plays a role in aggregation to form biofilms
and adherence to other cell types, like plant root hair cells (Jahn et al., 2011; Römling, 2002).

The nascent cellulose glucan chains are assembled from activated UDP-glucose donor
sugars by the GT, cellulose synthase. Like the higher eukaryotic HASs, there are several
isoforms of plant cellulose synthases (CesA) (Crowell et al., 2010; Vergara and Carpita, 2001).
Plant CesA isoforms are strikingly different than HAS and other synthases in that multiple CesA proteins come together at the plasma membrane to form a large complex hypothesized to be composed of a hexamer of hexamers (36 subunits), in order to synthesize the glucan chains (Brown and Montezinos, 1976; Doblin et al., 2002). Similarly, a protein complex of at least 3 subunits is required for the production and secretion of cellulose in bacterial species, and these are encoded by the *bcs* (bacterial cellulose synthesis) operon (Römling, 2002). Two genes are conserved in all *bcs* operons, including the cellulose synthase itself, known as *bcsA*.

An interesting facet of BcsA enzymes is the utilization of a PilZ receptor domain for activation of cellulose via bis-(3',5')-cyclic dimeric GMP (c-di-GMP) binding (Remminghorst and Rehm, 2006a, c; Römling et al., 2005; Ross et al., 1987). This domain is located in the C-terminal domain of BcsA. Another protein known as BcsB (expressed from the second most conserved gene within the *bcs* operon) is associated with BcsA and is also required for cellulose biosynthesis (Römling, 2002; Saxena et al., 1994; Standal et al., 1994). BcsB is a periplasmic protein anchored to the inner membrane by a single transmembrane helix and forms a 1:1 stoichiometric complex with BcsA (Morgan et al., 2013).

Eukaryotic and prokaryotic cellulose synthases are predicted to have a similar topologies, with 8 TMs and one large cytoplasmic loop that is similar to the loop in other synthases (Morgan et al., 2013). The large loop from all predicted and confirmed BcsAs exhibit a high level of sequence similarity (~70% similarity) (Römling, 2002). Very little homology is observed among either the N- or C- terminal domains of plant and bacterial CesAs (Pear et al., 1996; Vergara and Carpita, 2001).

A major breakthrough in the study of cellulose synthases was made recently with the report of a crystal structure of BcsA in complex with BcsB from *Rhodobacter sphaeroides*.
(Figure 1.4) (Morgan et al., 2013). The crystal structure verified structural regions, including the 8 TMs predicted in BcsA, and the one anchor TM for BcsB. The TMs of BcsA are organized in pairs surrounding a narrow channel with an overall arrangement of an elongated cuboid, where the TM of BcsB is packed tightly on one side of the cuboid. The PilZ domain was localized right next to the GT catalytic region within the C-terminus of BcsA. The GT domain was confirmed to resemble the GT-A fold, predicted for the cellulose synthase and all other members in the GT2 family (Lairson et al., 2008). The QXXRW motif was identified on interface helix 2 cooperating with a conserved FFCGS sequence to form a binding site for the terminal acceptor disaccharide of the glucan.

Mechanistic insight into the translocation channel was provided by the presence of a glucan within the channel in the solved structure. The channel used to translocate an assembled chain of cellulose is formed by TM3-8 in BcsA. Morgan et al. (2013) determined the channel is able to accommodate 10 glucose units of the glycan. Interestingly, the glucan extended from the GT domain at a 45° angle towards the membrane, kinked near the putative water-lipid interface to run through the center of the TM region, then sharply bent at the BcsA-BcsB interface, so it exited the channel at a 145° angle.

1.1.7 Alginate synthases

Like HA and the microfibrils of cellulose, alginate is a high molecular weight, linear exopolysaccharide. It consists of β-1,4 linked β-D-mannuronic acid and it’s C5’ epimer, α-L-guluronic acid, in a nonrepeating fashion. Alginate is produced by Brown algae and by two bacterial genera, Pseudomonas and Azotobacter (Remminghorst and Rehm, 2006b). In Azotobacter vinelandii, alginate is only found in capsule-like layers that surround mature cysts formed during a differentiation process (Brivonese and Sutherland, 1989). In Pseudomonas
Figure 1.4: Crystal structure of the BcsA-BcsB complex with a translocating polysaccharide at 3.25 Å resolution (obtained from Morgan et al., 2013). Transmembrane helices of BcsA are coloured green, the C-terminal domain is shown in red and the beige region represents the glycosyltransferase domain. The membrane-associated domain for BcsB is shown in dark blue and the periplasmic domain in light blue. The translocating glycan is shown in cyan and violet. The black horizontal bars represent the membrane junction that divides the cytoplasmic and periplasmic space.
*Pseudomonas aeruginosa*, alginate is an important virulence factor secreted into the surrounding medium in the lungs of patients with cystic fibrosis, to increase the microbe’s resistance to phagocytosis. It also allows for the formation of differentiated biofilms (Nivens et al., 2001; Pedersen et al., 1992; Pier et al., 2001).

The first step in alginate biosynthesis is the assembly of a linear β-1,4 linked homopolymer of D-mannuronic acid residues. The synthase, Alg8, is predicted to be responsible for both the assembly and translocation of the nascent poly-β-D-mannuronic acid across the inner membrane. Thus far, the LXXRW motif and four aspartates in Alg8 have been shown to be required for the alginate polymerization *in vivo* (Oglesby et al., 2008). Another inner membrane protein, Alg44 is believed to act as a copolymerase to the main GT Alg8 (Oglesby et al., 2008; Remminghorst and Rehm, 2006a). Alg44 is also responsible for regulation of Alg8, and ultimately alginate biosynthesis, via the PilZ domain located at the N-terminus of Alg44. Once the nascent homopolymer is translocated across the inner membrane, regions of the glycan are epimerised and further modified to form the mature polymer. The final product is then escorted and translocated across the periplasm and outer membrane by several other Alg proteins (Whitney and Howell, 2013).

**1.1.8 Type III capsule synthase in *Streptococcus pneumoniae***

*Streptococcus pneumoniae* type 3 capsule is a high molecular weight, linear exopolysaccharide that is comprised of a repeating subunit of β-1,3 and β-1,4 linked glucose (Glc) and GlcUA, which are added to the nonreducing end of the growing polymer chain (Cartee et al., 2000; Forsee et al., 2000; Reeves and Goebel, 1941). Assembly of the type 3 capsule by the type 3 capsule synthase (Cps3S) is initiated by the transfer of Glc from UDP-Glc to the lipid phosphatidylglycerol, generating a novel lipid carrier upon which the type 3 capsule
is synthesized (Cartee et al., 2001; Cartee et al., 2005). After initiation, Cps3S assembles oligosaccharides before synthesis of the longer type 3 capsule (Cartee et al., 2001; Forsee et al., 2006). A critical oligosaccharide length of 8 sugar residues must be reached to act as a precursor for the assembly of the longer type 3 capsule polysaccharide (Forsee et al., 2006). The concentration of the UDP-sugars affects the polysaccharide length and the total amount of capsule produced (Ventura et al., 2006). If the UDP-GlcUA concentration is low, the oligosaccharide precursor is unable to attain the critical length, causing chain elongation to terminate and the oligosaccharide is ejected from the synthase (Forsee et al., 2006).

1.1.9 O:54 biosynthesis and the role of the putative synthase

As illustrated above, there are common themes, as well as interesting functional, structural and regulatory disparities amongst the synthase-dependent biosynthesis pathways for the assembly of such large and important polysaccharides. Important features of the synthases are only now beginning to be established. It is imperative to study these enzymes to provide extensive and comprehensive insight in order to enhance our understanding of this unique and important group of enzymes.

Only one synthase, known as WbbF, has been identified to be involved in the biosynthesis of lipopolysaccharide (LPS) O-antigens (O-Ags) for Gram-negative bacteria. WbbF is responsible for producing the O:54 antigen homopolymer in Salmonella serovars. Currently, little is known about WbbF, or about the novel O-Ag assembly pathway in which WbbF participates. The main focus of this thesis is the synthase WbbF, and the components of the synthase-dependent pathway. The goal was to provide a better understanding of the synthase-dependent pathway for O-Ag synthesis, as well as a broader understanding of the HasA subgroup of GTs. The following text puts the O:54 antigen into context of the bacterial cell envelope.
1.2 Lipopolysaccharide

1.2.1 Gram-negative cell envelope

All Gram-negative bacteria contain a cell envelope that consists of a cytoplasmic or inner membrane (IM) and an outer membrane (OM) (Figure 1.5). Located between these membranes is the periplasm, which contains a rigid layer of peptidoglycan and a variety of soluble proteins including degradative enzymes, redox carriers and chaperone molecules that are important for envelope biogenesis (Hobot et al., 1984; Silhavy et al., 2010; Van Wielink and Duine, 1990). The IM is a phospholipid bilayer composed of glycerophospholipids, and contains many critical integral and membrane-associated proteins (Silhavy et al., 2010). The peptidoglycan layer is an essential structural component that helps determine bacterial cell shape and preserves cell integrity by withstanding turgor pressure (Vollmer et al., 2008). In many bacteria, the peptidoglycan layer is linked to the OM via a linker molecule that (in *E. coli*) is called murein lipoprotein or Braun’s lipoprotein (Braun, 1975).

The OM is an essential part of the Gram-negative cell envelope and provides a selectively-permeable barrier. The asymmetrical bilayer contains phospholipids in the inner leaflet and mainly LPS in the outer leaflet (Nikaido, 2003). The OM contains mainly channel-forming, integral membrane proteins known as porins that allow the passive intake of vital small, water soluble nutrients, as well as, the excretion of toxic waste products. The OM protects the bacterium from harmful, often hydrophobic, external compounds such as antibiotics and detergents via the size-exclusivity of these OM proteins (Nikaido, 2003). However, the main factor in barrier properties is LPS.
Figure 1.5: Depiction of the cell envelope of Gram-negative bacteria. The inner membrane is a phospholipid bilayer containing integral membrane proteins (IMP). The outer membrane is an asymmetrical bilayer predominantly composed of phospholipids within the inner leaflet and lipopolysaccharide (LPS) within the outer leaflet. Integrated into the outer membrane are outer membrane proteins (OMP). Lipoproteins are also associated with the outer membrane and anchor the peptidoglycan layer to the outer membrane.
1.2.2 Structure and function of lipopolysaccharide

LPS consists of three structural regions: Lipid A, core oligosaccharide and O-Ag polysaccharide (Figure 1.6A). LPS is anchored into the OM via the hydrophobic lipid A moiety. The structure of Lipid A is highly conserved, and Lipid A is a required membrane component for maintaining the integrity of the OM barrier (Raetz and Whitfield, 2002). One exception to this is *Neisseria meningitidis*, which is viable in the absence of LPS (Bos et al., 2004). The lipid A backbone of the prototypical *E. coli* and *S. enterica* LPSs contains a disaccharide of glucosamine with a β-(1→6)-linkage (Figure 1.6B). Positions 1 and 4’ on the disaccharide are phosphorylated, and the 2, 3, 2’ and 3’ positions are acylated with R-3-hydroxymyristic acid, via either ester- or amine-linkages (Qureshi et al., 1982; Rietschel et al., 1983). Further substitution of the fatty acid chains occurs at the hydroxyl groups at positions 2’ and 3’ with the addition of ester-linked laurate (C12) and myristate (C14), respectively (Rietschel et al., 1987). Various enzymes, such as PagP, PagL and LpxO that modify the *S. enterica* lipid A structure to respond to the local environment, are shown in Figure 1.6B (Trent, 2004).

The core oligosaccharide portion of LPS is linked to lipid A via the sugar residue 3-deoxy-D-manno-oct-2-ulopyranosonic acid (Kdo) (Unger, 1981). The addition of two Kdo residues to lipid A creates the Kdo₂-lipid A moiety, which is well conserved throughout the family Enterobacteriaceae. The Kdo₂-lipid A moiety was initially believed to be the minimal LPS structure required for survival. However, an *E. coli* K-12 mutant that was only able to produce a precursor of lipid A, known as Lipid IVₐ, was viable in a laboratory setting, when combined with the expression of the LPS transporter MsbA (Meredith et al., 2006; Raetz et al., 2007). Nonetheless, this mutant was significantly compromised since it was hypersensitive to hydrophobic compounds, released periplasmic enzymes and was susceptible to phagocytosis.
**Figure 1.6:** Structure of a typical lipopolysaccharide (LPS) molecule, and the lipid A and core structures in *S. enterica*. A) Cartoon structure of LPS in *S. enterica*. There are three distinct regions of LPS: lipid A, core oligosaccharide (subdivided into inner and outer core) and the distal immunodominant O-antigen polysaccharide (general representation). The black circles represent phosphate. The “n” shown outside the brackets represents the number of repeat units. B) Chemical structure of Kdo2-lipid A moiety of *S. enterica* sv. Typhimurium shown with enzymes that modify the structure. The types of modifications that occur are highlighted in red, blue, and pink (modified from Trent, 2004). The number of carbons in each acyl chain is indicated by a number in an enclosed circle. C) Composition of the two outer core oligosaccharides found in *Salmonella* (modified from Whitfield et al., 2003).
The core oligosaccharide is divided into two sections (the inner core and outer core) and can be up to 15 sugars long (Holst, 2002). The inner core is proximal to the lipid A moiety and its conservation is correlated with its role in OM integrity. In S. enterica, the inner core structure is L-D-Hep-(1→7)-L-D-Hep-(1→3)-L-α-D-Hep(1→5)-Kdo with a glucopyranose substituted at the O-3 position of the second Hep, and Hep I and II are phosphorylated (Holst, 2007). Further sugar and phosphate substitutions can be made to this general structure (Frirdich and Whitfield, 2005; Holst, 2007). There are two core structures for Salmonella (sv Arizonae IIIa and sv Typhimurium), which are based mainly on the variability of the outer core (Figure 1.6C) (Kaniuk et al., 2002). The outer core is linked to the hyper variable O-Ag domain, but, in the Enterobacteriaceae, it can also serve as an attachment site for other glycans such as one form of enterobacterial common antigen (ECA), short oligosaccharides of E. coli “group 1” capsular polysaccharides and, in the case of Salmonella, the poorly understood T1 and T2 antigens (Kuhn et al., 1988; Whitfield et al., 1997).

O-Ag is an immunodominant polysaccharide, comprised of repeat subunits. An individual repeat unit contains 1-8 sugar residue(s) with different stereochemistries of the O-glycosidic linkages. Non-carbohydrate substituents such as O- and N-acetyl groups, phosphate, phosphorylethanolamine groups, formyl groups, amino acids and glyceric acid may also be attached to residues. The individual repeat units may be linked in a linear or branched manner, which further contributes to the diversity within the O-Ag structure (Caroff and Karibian, 2003; Raetz and Whitfield, 2002). Ultimately, a combination of each of these factors defines the serological specificity of the organism, providing a useful tool for serotyping, clinical diagnosis, epidemiology and taxonomy. The repeat units are generally polymerized to form a chain with a strain-specific range of chain lengths known as the ‘modal size distribution’. The modal
The distribution of O-Ag in *Salmonella* spp. is heterogeneous, but generally only two different chain lengths are observed in one isolate. They can be as small as 16 repeat units or greater than 100 repeat units in length (Goldman and Leive, 1980; Munford et al., 1980; Peterson and McGroarty, 1985). Alterations to the chain length of an O-Ag has been shown to affect virulence properties of *S. enterica* sv Typhimurium (May and Groisman, 2013; Murray et al., 2003, 2006).

The position of O-Ag places it at the boundary between the bacterium and the external environment. Because of this, O-Ag plays a protective role for the bacterium. O-Ag enables bacteria to evade host immune responses such as phagocytic uptake and the alternative serum complement pathway, thereby allowing persistence and survival within the body tissue or fluids (Alexander and Rietschel, 2001; Fierer and Guiney, 2001). *Salmonella* mutants with only the lipid A-core domain or less, known as rough LPS (R-LPS), are often more sensitive to common antibiotics compared to bacteria with the O-Ag domain, which is known as smooth-LPS (S-LPS) (Nikaido, 1976). Therefore, while the core and O-Ag segments of LPS are not required for bacterial growth within a laboratory setting, they are nonetheless important for protection and barrier permeability.

### 1.2.3 Overview of LPS biosynthesis

The assembly of LPS is fulfilled by two separate pathways. One pathway assembles the lipid A + core oligosaccharide via a series of sequential reactions. The O-Ag domain is synthesized independently by one of three different pathways, depending on the isolate. The two components are subsequently ligated together to form the complete LPS molecule via the ligase WaaL. The assembly of the lipid A-Kdo₂ moiety is conserved among Gram-negative species. This process is called the Raetz pathway (after Dr. Christian Raetz who identified most of the critical steps) and occurs in the cytoplasm and inner leaflet of the IM (Raetz et al., 2007). The
pathway involves nine enzymatic steps that have been well characterized in *E. coli*. The lipid A-core moiety is subsequently transported from the inner leaflet of the IM to the periplasmic leaflet by MsbA (an ABC transporter), where it can be ligated to O-Ag (Raetz and Whitfield, 2002; Zhou et al., 1998).

Completion of LPS occurs at the periplasmic face of the IM where all O-Ag biosynthesis pathways converge with the lipid A-core synthesis pathway. The ligase enzyme, WaaL, is responsible for transferring the O-Ag from the lipid carrier to the lipid A-core moiety (Han et al., 2012). The newly synthesized LPS molecule is then transported and inserted into the OM by the seven essential proteins LptA-G (Ruiz et al., 2009). LptB, -F, and –G form an ABC transporter in the IM that releases the newly formed LPS from the outer leaflet of the IM (Narita and Tokuda, 2009; Ruiz et al., 2008; Sperandeo et al., 2007). Associated with the ABC transporter is the protein LptC, which has recently been shown to interact with the periplasmic protein LptA (Sperandeo et al., 2011). Both LptA and LptC can bind to LPS, though only LptA facilitates transport of LPS across the periplasm (Sperandeo et al., 2007; Tran et al., 2010; Tran et al., 2008) via formation of a trans-periplasm scaffold (Chng et al., 2010; Freinkman et al., 2012; Sperandeo et al., 2011; Tefsen et al., 2005). LptA is capable of forming oligomers in an end-to-end arrangement with itself and has been shown to interact with the N-terminal domain of the OM β-barrel protein LptD (Freinkman et al., 2012; Merten et al., 2012). LptD is embedded within the OM and forms a complex with the lipoprotein LptE (Chng et al., 2010). The LptD-LptE complex is responsible for exporting LPS molecules from the periplasm through the OM. Currently, evidence supports the proposal that LPS transport occurs via a trans-envelope complex model.
1.2.4 O-antigen synthesis

Synthesis of O-Ag occurs via one of three biosynthetic pathways: the ABC-transporter dependent, the Wzy-dependent or the synthase-dependent pathway (Figure 1.7). The first two pathways are widespread among Gram-negative bacteria, whereas only the O:54 factor of S. enterica is synthesized by the synthase-dependent pathway (Keenleyside and Whitfield, 1996). These pathways differ in the cellular location of polymerization steps and the type of transporter involved in shuttling completed O-Ag, or individual O-Ag repeat units, to the periplasmic face of the IM. The assembly of any O-Ag repeating unit begins on the cytoplasmic side of the IM on the membrane-bound carrier undecaprenyl phosphate (und-P), a C₅₅ polyisoprenoid lipid intermediate also required for the biosynthesis of peptidoglycan and capsular polysaccharides (Greenfield and Whitfield, 2012). A sugar-1-P residue is transferred from an NDP-activated donor to und-P, to form und-PP-linked glycosyl. This “primes” the carrier for the assembly of O-Ag (Raetz and Whitfield, 2002). Depending on the microorganism, this is catalyzed by an enzyme belonging either to the PNPT (polyisoprenyl-phosphate N-acetylhexosamine-1-phosphate transferases) or PHPT (polyisoprenyl-phosphate hexose-1-phosphate transferases) family. In most S. enterica serovars, the PHPT common prototype, WbaP transfers a galactose-1-P from UDP-Gal to und-P (Saldías et al., 2008). However, initiation for the production of the Salmonella O:54 factor occurs by the well-distributed PNPT enzyme, WecA, by transferring a GlcNAc-1-P from UDP-GlcNAc to und-P (Alexander and Valvano, 1994; Keenleyside et al., 1994). Chain extension continues on the “primed” und-PP-glycosyl via a process that depends on the synthesis pathway. Only the synthase pathway will be discussed in detail (see below).
**Figure 1.7:** Cartoon models of the three assembly pathways for O-Ag biosynthesis (obtained from Greenfield and Whitfield, 2012). **A)** In the ABC-transporter pathway, the O-Ag chain is sequentially elongated by GTs on a lipid carrier associated with the cytoplasmic face by the addition of monomers to the non-reducing terminus of the growing chain. Once complete, the full length O-Ag is exported to the periplasmic face of the IM by the ATP-Binding Cassette (ABC) transporter known as Wzm-Wzt. **B)** In the Wzx/Wzy-dependent pathway, individual repeat units are assembled on und-P at the cytoplasmic face of the IM by GTs. The flippase, Wzx translocates the und-PP-linked repeat units across the IM to the periplasmic face where the polymerase, Wzy assembles the nascent O-Ag polysaccharide via the addition of repeat units to the reducing terminus of the glycan. **C)** In the synthase-dependent pathway, the putative synthase, WbbF is proposed to assemble the O-Ag on the cytoplasmic face of the IM while simultaneously exporting the growing glycan to the periplasmic face of the IM.
1.2.5 The origin of serovar O:54

Members of the genus *Salmonella* are divided by the Kauffmann-White scheme in two species: *Salmonella bongori* and *Salmonella enterica* (formerly known as *Salmonella choleraesuis*), with the latter divided into six subspecies known as *S. enterica* subsp. *enterica, salamae, arizonae, diarizonae, houtenae*, and *indica* (Tindall et al., 2005). Among these, there are over 2500 distinct serovars based on serological analysis of the antigenic properties of the H-antigen (flagellar filament) and the O-Ag structures expressed on the surface of *Salmonella* (Grimont and Weill, 2007). Contributing to the wide diversity of serovars in *Salmonella* is the interesting discovery of two O-Ag structures co-expressed by an isolate: one O-Ag is chromosomally-encoded and a separate O-Ag (O:54) is encoded on a naturally occurring mobilizable plasmid. The antigenic factor O:54 is a unique homopolymer of N-acetylmannosamine (ManNAc) sugar residues (Figure 1.8A), and it is the only known O-Ag where synthesis is encoded by a plasmid (Keenleyside et al., 1994; Popoff and Le Minor, 1985). The O:54 factor was first discovered by Kaufmann et al. in 1961. Since then, 15 isolates that express the O:54 antigenic factor have been reported (Table 1.1) (Le Minor et al., 1971; Le Minor et al., 1973; Popoff and Le Minor, 1985). Some of the isolates express additional O-factors simultaneously with the O:54 polysaccharide, while other O-factors will only be expressed in the absence of the O:54 factor. Moreover, the expression of O:54 antigen can also be lost while expression of the other O-factor(s) is maintained (Popoff and Le Minor, 1985). These observations led to questions concerning the origin of the biosynthetic gene cluster responsible for expression of the O:54 antigen.
Figure 1.8: Structure of O:54 antigen and Organization of pWQ799. A) Chemical structure of the O:54 antigen repeat unit with primer and adaptor regions. The O:54 factor is a homopolymer comprised of a disaccharide repeat of N-acetylmannosamine (ManNAc) with altering β1,3 and β1,4 linkages. Assembly of this O antigen occurs on an undecaprenyl phosphate lipid carrier primed with a single N-acetylglucosamine residue. Attached to the primer is a single ManNAc residue (linkage unknown) that is an adaptor for which the O-polysaccharide is assembled upon. B) Organization of pWQ799 plasmid (modified from Keenleyside and Whitfield, 1995). The arrows underneath represent size, location, and direction of transcription of the open reading frames (ORFs) for the O:54 biosynthetic gene cluster. The replicon region of pWQ799 consists of the ori, bom, and mob genes, which are the origin of replication, the origin of transfer and the mobilization region, respectively.
Table 1.1: *Salmonella enterica* serovars expressing O:54 antigen.

<table>
<thead>
<tr>
<th>Serovar</th>
<th>O antigen</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonev</td>
<td>21, 54</td>
<td>Le Minor (1968)</td>
</tr>
<tr>
<td>Winnipeg</td>
<td>54</td>
<td>Popoff and Le Minor (1985)</td>
</tr>
<tr>
<td>Rossleben</td>
<td>3, 54</td>
<td>Grimont and Weill (2007)</td>
</tr>
<tr>
<td>Borreze</td>
<td>54</td>
<td>Popoff and Le Minor (1985)</td>
</tr>
<tr>
<td>Uccle</td>
<td>3, 54</td>
<td>Le Minor et al. (1970)</td>
</tr>
<tr>
<td>Newholland</td>
<td>4, 12, 54</td>
<td>Popoff and Le Minor (1985)</td>
</tr>
<tr>
<td>Poeseldorf</td>
<td>8, 20, 54</td>
<td>Le Minor (1988)</td>
</tr>
<tr>
<td>Ochsenwerder</td>
<td>6, 7, 54</td>
<td>Le Minor (1988)</td>
</tr>
<tr>
<td>Montevideo(^1)</td>
<td>(6, 7, 14) (54)</td>
<td>Grimont and Weill (2007)</td>
</tr>
<tr>
<td>Czernyring</td>
<td>54</td>
<td>Grimont and Weill (2007)</td>
</tr>
<tr>
<td>Steinwerder</td>
<td>3, 15, 54</td>
<td>Le Minor (1988)</td>
</tr>
<tr>
<td>Yerba</td>
<td>54</td>
<td>Grimont and Weill (2007)</td>
</tr>
<tr>
<td>Canton</td>
<td>54</td>
<td>Popoff and Le Minor (1985)</td>
</tr>
<tr>
<td>Barry</td>
<td>54</td>
<td>Popoff and Le Minor (1985)</td>
</tr>
<tr>
<td>Mundubbera</td>
<td>54</td>
<td>Grimont and Weill (2007)</td>
</tr>
</tbody>
</table>

\(^1\) Factors O:6, 7, 14 are expressed only in the absence of O:54
A 7-8 Kb plasmid was found to be associated with the expression of O:54 factor (Popoff and Le Minor, 1985). Although the exact role of this plasmid was not established in the initial study, it was proposed that the plasmid contained biosynthetic genes, gene(s) encoding a modifying enzyme, or a regulator gene(s) for O:54 antigen synthesis. This hypothesis was tested by curing a 6.9 Kb plasmid (later named pWQ799, Figure 1.8B; Keenleyside and Whitfield, 1995) from a field isolate, S. enterica sv Borreze, which resulted in the loss of the O:54 factor (Keenleyside et al., 1994). To further examine the role of the 6.9 Kb plasmid, it was cloned into the vector pGEM-7Zf(+), to produce a hybrid recombinant plasmid called pWQ800. E. coli DH5α and E. coli SØ874 (lacking the K-12 O-Ag biosynthesis gene cluster) transformed with pWQ800 expressed the O:54 antigen. It was therefore postulated from these results that the pWQ799 plasmid contained essential O:54 antigen biosynthesis genes (Keenleyside et al., 1994). However, this did not rule out the possible involvement of other chromosomal gene products as participants in O:54 antigen biosynthesis in the transformed E. coli cells.

Based on its nucleotide sequence, pWQ799 is a member of the ColE1-related plasmid group (Keenleyside and Whitfield, 1995). ColE1 plasmids are unable to self-transmit but contain mobilization functions (mob) within the genome, which facilitate transmission in the presence (in trans) of conjugative transfer functions, often provided by another plasmid (Boyd et al., 1989; Varsaki et al., 2003). The oriT (origin of transfer) and bom (basis of mobility) loci sites were both detected in pWQ799, as expected (Figure 1.8B). Therefore, pWQ799 is a mobilizable plasmid that requires supplementation of the proper transfer functions in trans (Keenleyside and Whitfield, 1995). This characteristic of pWQ799 explains why the O:54 factor is found in numerous serovars and can be co-expressed with chromosomally encoded O-factors.
The remaining sequence of pWQ799 contained no detectable homology with any known ColE1-related sequences and contained a lower average G + C content of 39% compared to the average 50-53% G + C content seen in the replicon sequence and (as an overall average) within the genomes of *Salmonella* spp. (Keenleyside and Whitfield, 1995). This is common in gene clusters involved in polysaccharide synthesis (Whitfield and Valvano, 1993). The non-replicon sequence contains three open reading frames, designated *wbbE*, *wbbF* and *mnaA* (formerly known as *rfbA*, *rfbB*, and *rfbC*, respectively) (Keenleyside and Whitfield, 1996; Reeves et al., 1996). These genes were proposed to encode a monofunctional *N*-acetylmannosaminyl transferase, a processive β-glycosyltransferase (the “synthase” in this assembly system), and a sugar nucleotide epimerase, respectively. All the enzymes were shown to be involved in O:54 biosynthesis (Keenleyside and Whitfield, 1996).

**1.2.6 Proposed synthase pathway for O:54 biosynthesis**

The synthesis of the O:54 antigen begins with the formation of und-PP-GlcNAc by WecA (as discussed previously) in order to provide a primer for assembly of the poly-*N*-acetylmannosamine (poly-ManNAc) O:54 polysaccharide (Figure 1.9) (Keenleyside et al., 1994). The donor for chain extension, UDP-ManNAc, is produced from UDP-GlcNAc by the epimerase, MnaA (Keenleyside et al., 1994; Keenleyside and Whitfield, 1996). A single ManNAc residue is transferred to the und-PP-GlcNAc primer by the monofunctional ManNAc transferase, WbbE (Keenleyside et al., 2001; Keenleyside and Whitfield, 1996). The addition of ManNAc commits the und-PP-GlcNAc primer to the synthase pathway and allows for chain extension. WbbF is believed to function as the polymerizing synthase which catalyzes formation of the →4)-β-D-ManpNAc-(1→3)-β-D-ManpNAc-(1→ repeat units to form the nascent polymer
Figure 1.9: Proposed synthase-dependent pathway for O:54 antigen biosynthesis. A) WecA primes undecaprenyl phosphate (Und-P) with GlcNAc-1-phosphate B) MnaA (or the functional homolog WecB) provides the donor sugar by converting UDP-GlcNAc to UDP-ManNAc. C) WbbE adds a single ManNAc residue to Und-PP-GlcNAc. D) WbbF extends the O:54 homopolymer while E) exporting the O:54 antigen to the outer leaflet of the inner membrane.
(Keenleyside and Whitfield, 1996). It is not known what mechanism determines or controls the chain extension reaction. Keenleyside and Whitfield (1996) discovered that smooth O:54 LPS was expressed in an *E. coli* K-12 host strain devoid of its O-Ag biosynthesis genetic locus and containing only WbbE and WbbF. This led to the proposal that these two proteins were sufficient for both polymerization and export, and that WbbF simultaneously exports the nascent O-Ag chain across the IM as it extends the polymer. Furthermore, the plasmid-encoded *mnaA* gene was proposed to be expendable for O:54 biosynthesis in *E. coli* and *Salmonella* due to the presence of chromosomally-encoded WecB (formerly known as RffE), a functional homologue of MnaA. It was proposed that O:54 synthesis represents a third, novel pathway for O-Ag biosynthesis and was subsequently named the synthase pathway due to the predicted activities of WbbF (Keenleyside and Whitfield, 1996; Raetz and Whitfield, 2002).

### 1.2.7 WbbF, the O:54 synthase from *S. enterica* sv. Borreze

WbbF is predicted to be a 53.3 kDa, integral membrane protein composed of 459 amino acids (Keenleyside and Whitfield, 1996). Four transmembrane helices between residues 11–40, 325–340, 385–406, and 416–438, with a large central hydrophilic domain were predicted by hydropathic analysis. A periplasmic loop between residues 340-384 was confirmed by an in-frame PhoA fusion at residue 368. The predicted topology of WbbF resembles the predicted topology of other synthases. Moreover, hydrophobic cluster analysis showed that WbbF contains structural regions of homology with various GTs involved in the synthesis of bacterial cell surface polysaccharides (Keenleyside and Whitfield, 1996). An amino acid sequence alignment with well-known synthases performed by Clustal Omega also shows large regions of similarity (Figure 1.10).
WbbF contains the conserved domains, A and B, that characterize processive enzymes (Keenleyside and Whitfield, 1996; Saxena et al., 1995). Not surprisingly, a conserved Asp and the QXXRW motif were identified in domain B, as well as two conserved Asp residues within domain A. Catalytic residues speculated for WbbF are Asp-92 or Asp-94, Asp-151, Asp-244, and either Glu-268 or Asp-269 based on similarities of related enzymes (Keenleyside and Whitfield, 1996), but a requirement for any or all of these residues have not been confirmed. Despite these similarities, the involvement of WecA in O:54 biosynthesis makes WbbF the only known synthase to operate with polyisoprenoid lipid carrier.

1.3 Aims of this study and objectives

Progress has been made towards understanding the HasA family of GTs, but information on the structure, and mechanisms for both the glycosyl transfer and export functions are still limited, despite the biological importance of several members of this group of GTs. Outstanding questions include how these enzymes bind two distinct UDP-sugars, catalyze the transfer of these two distinct sugars in two different linkages, and transfer the polymer across the membrane. Polymers produced by synthases have application within the medical, biofuel and food industries and further research is needed to assist in designing large-scale production operations to mass produce these polymers (Yamada and Kawasaki, 2005). Thus, studying the putative synthase, WbbF may reveal important details of the assembly of O:54 antigen via the synthase-dependent pathway, as well as provide a broader comprehension of this important enzyme family.

Previous research on WbbF and the synthase pathway has been limited by inadequate methods of detection of O:54. Moreover, the lack of understanding of the O-Ag synthase-dependent pathway itself has prevented the characterization of WbbF. Sufficient detection of the
Figure 1.10: Clustal Omega amino acid alignment (Goujon et al., 2010; Sievers et al., 2011) of synthase proteins 1. WbbF; 2. Cellulose synthase from *Gluconacetobacter xylinus*; 3. Cellulose synthase from *Rhodobacter sphaeroides*; 4. Type 3 capsule synthase from *Streptococcus pneumoniae*; 5. Hyaluronan synthase from *Streptococcus equi*. Black boxes surround the conserved DXD and QXXRW motifs. The webserver program marks a conserved amino acid residue in all sequences with an asterisk (*). Conservation of residues with strongly similar properties is marked with a colon (:); weakly similar properties conserved are marked with a period (.) underneath. Colour of residue indicates the physiochemical property of that amino acid. Red designates small and hydrophobic residues. Blue, magenta, and grey designate acidic, basic, unusual amino acid residues, respectively. Green designates hydroxyl, and sulphydryl.
O:54 polysaccharide is imperative for the study of WbbF in order to verify enzymatic function. Therefore, the first objective of the research for this thesis was to optimize the methods of detection of the O:54 antigen to be able to confirm function (or non-function) of the synthase-dependent pathway after experimental manipulation. A logical way to begin a detailed investigation of WbbF is to resolve the topological structure to understand the organization of possible functional domains. Therefore, the last goal was to develop a functional, cysteine-free WbbF to use for the construction of WbbF derivatives with single cysteine residues at precise reporter sites for cysteine-scanning topological studies.
2.0 MATERIALS AND METHODS

2.1 Bacterial strains, plasmids and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 2.1 and Table 2.2, respectively. Bacterial strains were grown in Miller’s Luria-Bertani (LB) (Bertani, 1951) broth base (Invitrogen Life Technologies) at 37°C in a shaking incubator at 200 rpm, unless otherwise stated. When appropriate, media were supplemented with the antibiotic(s) chloramphenicol (34 μg/mL) and ampicillin (100 μg/mL) for selective growth. Antibiotics were purchased from Sigma-Aldrich. LB media were supplemented with D-glucose (0.4% w/v), Anhydrotetracycline hydrochloride (AhT) (2.5 ng/mL), L-arabinose (0.02% w/v) and/or isopropyl β-D-thiogalactopyranoside (IPTG) (0.5 mM) when necessary. Some experiments required supplementation of a chemical(s) at different concentrations than listed here and these instances are described in the results.

2.2 Bacterial growth analysis

Growth curves were examined for *E. coli* Top10 and CWG1168 (described below) during O:54 antigen production, as well as when these strains were not producing O:54 antigen. Overnight cultures were sub-cultured 1:100 in either 25 mL LB or 50 mL LB. After sub-cultured, the optical density at 600 nm (OD$_{600}$) was measured for certain cultures every 0.5 hr for a span of time (dilution of strain considered time point 0 hr). Other cultures were grown until an OD$_{600}$ of 0.2 was obtained (considered time point 0 hr). After an OD$_{600}$ of 0.2 was reached, these cultures were divided equally into two warmed sterile, flasks and O:54 antigen production was induced in one of the two flasks, by the addition of 2.5 ng/mL AhT. After induction, OD$_{600}$ was measured every 0.5 hr for a span of time.

The relationship between culture OD$_{600}$ and colony-forming units was examined for *E. coli* Top10, CWG1167 (described below), CWG1168 and *S. enterica* sv. Borreze, to properly
Table 2.1: Bacterial strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype/Serotype/Description</th>
<th>Ref. or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top10</td>
<td>F, mcrA, Δ(mrr-hsdRMS-mcrBC), φ80, lacZΔM15, ΔlacX74, deoR, nupG, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(Str5), endA1</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>CWG1167</td>
<td>Top10; ΔwecB</td>
<td>This study</td>
</tr>
<tr>
<td>CWG1168</td>
<td>Top10; Δwaal</td>
<td>This study</td>
</tr>
<tr>
<td><strong>S. enterica serovars</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borreze</td>
<td>54:f,g,s:-1; contains plasmids of 96, 4.5, and 2.3 MDa</td>
<td>(Popoff and Le Minor, 1985)</td>
</tr>
</tbody>
</table>

1O-Ag: H antigen phase 1: H antigen phase 2
Table 2.2: Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBAD24</td>
<td>L-arabinose-inducible expression vector, Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>(Guzman et al., 1995)</td>
</tr>
<tr>
<td>pKD3</td>
<td>Source of Cm&lt;sup&gt;r&lt;/sup&gt; resistance cassette</td>
<td>(Datserenko and Wanner, 2000)</td>
</tr>
<tr>
<td>pKD46</td>
<td>Helper plasmid containing λ Red recombinase genes γ, β, and exo for λ Red mutagenesis</td>
<td>(Datserenko and Wanner, 2000)</td>
</tr>
<tr>
<td>pCP20</td>
<td>Helper plasmid encoding FLP recombinase for λ Red mutagenesis</td>
<td>(Datserenko and Wanner, 2000)</td>
</tr>
<tr>
<td>pWQ572</td>
<td>pBAD24-derivative containing a P&lt;sub&gt;lac&lt;/sub&gt; promoter, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>J. D. King</td>
</tr>
<tr>
<td>pWQ799</td>
<td>Naturally occurring plasmid containing wb&lt;sup&gt;*&lt;/sup&gt;O54 operon</td>
<td>(Keenleyside et al., 1994)</td>
</tr>
<tr>
<td>pWQ800</td>
<td>pGEM-7Zf(+)/pWQ799 hybrid with the pGEM vector inserted at the single EcoRI restriction site of pWQ799, Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>(Keenleyside et al., 1994)</td>
</tr>
<tr>
<td>pKB7bc</td>
<td>pLC69 derivative containing WbbF-His&lt;sub&gt;6&lt;/sub&gt; inserted using EcoRI and PstI restriction sites, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>K. Byatte</td>
</tr>
<tr>
<td>pKB101</td>
<td>pLC69 derivative containing synthetic WbbF-His&lt;sub&gt;6&lt;/sub&gt; codon-optimized for expression in E. coli (Geneart), inserted using EcoRI and PstI restriction sites, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>K. Byatte</td>
</tr>
<tr>
<td>pWQ203</td>
<td>pWQ572 derivative containing the wb&lt;sup&gt;*&lt;/sup&gt;O54 operon from pWQ799 inserted using BamHI and PstI restriction sites, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pBH6</td>
<td>pWQ572 derivative containing 100bp WbbE-WbbF-His&lt;sub&gt;6&lt;/sub&gt; including three potential upstream start codons found in frame of wbbE in 108 bp upstream of WbbE, inserted using EcoRI and PstI restriction sites, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pBH9</td>
<td>pBAD24 derivative containing MnaA-His&lt;sub&gt;6&lt;/sub&gt; inserted using KpnI and HindIII restriction sites, Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pBH13</td>
<td>pWQ572 derivative containing RBS WbbE-WbbF-His&lt;sub&gt;6&lt;/sub&gt; including the predicted RBS site upstream of WbbE and the presumed start codon, inserted using EcoRI and PstI restriction sites, Cm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pBH68</td>
<td>pWQ572 derivative containing wb&lt;sup&gt;*&lt;/sup&gt;O54 operon from pWQ799 with C49A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH69</td>
<td>pWQ572 derivative containing wb&lt;sup&gt;*&lt;/sup&gt;O54 operon from pWQ799 with C102A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH70</td>
<td>pWQ572 derivative containing wb&lt;sup&gt;*&lt;/sup&gt;O54 operon from pWQ799 with C183A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH71</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C195A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pBH72</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C237A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH73</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C328A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH74</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C394A mutation in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH75</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A mutations in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH79</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A, C195A mutations in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH80</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A, C195A, C237A mutations in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH81</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A, C195A, C237A, C328A mutations in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH82</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A, C195A, C237A, C328A, C394A mutations in WbbF</td>
<td>This study</td>
</tr>
<tr>
<td>pBH83</td>
<td>pWQ572 derivative containing ( w^*_{\text{OS4}} ) operon from pWQ799 with C49A, C102A, C183A, C195A, C237A, C328A, C394A</td>
<td>This study</td>
</tr>
</tbody>
</table>
evaluate the amount of O:54 production in each strain. Overnight cultures of each strain were sub-cultured 1:100 into 25 mL LB and grown until an OD$_{600}$ of 1.0 was reached. Cultures were placed on ice immediately to inhibit growth, diluted to $10^{-6}$ via serial dilution and then a 100 μL aliquot of the $10^{-6}$ dilution was plated onto LB agar. Colonies on each plate were counted to calculate the colony forming units (CFU)/mL for each sample. These experiments were performed in triplicate and an average CFU/mL was calculated. Using the average CFU/mL, the variance for each number was calculated and used to determine the standard deviation. From this, the standard error was calculated.

2.3 General DNA methods

Custom oligonucleotide primers used in this study were all purchased from Sigma-Aldrich and are listed in Table 2.3. Polymerase chain reaction (PCR) amplification of DNA fragments was carried out using Pwo DNA polymerase (Roche Applied Science), PfuUltra High Fidelity DNA polymerase (Stratagene), or KOD Hot Start DNA polymerase (Novagen). PCR-cycle parameters were determined according to the manufacturer’s specifications for the polymerase used in the reaction and adjusted if necessary if difficulties arose. Invitrogen PureLink PCR Purification kit was used for purification of DNA fragments generated from PCR reactions and restriction endonuclease digestions. When necessary, DNA fragments were isolated from agarose gels. Separation of DNA fragments was achieved by electrophoresis of DNA through a 0.9% or 1.2% agarose gels at 80 V for 40 minutes. The region of gel containing the target fragment was isolated and DNA purified using the Invitrogen Quick Gel Extraction kit. Plasmid DNA, including native plasmid DNA from S. enterica sv. Borreze, was purified using PureLink Quick Plasmid Miniprep kit (Invitrogen), or Qiagen HiSpeed Midi Kit (Qiagen).
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′→3′)</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBSWbbEF_Fwd</td>
<td>gatcgaattcGCTGTATAGGAAATTAGGG</td>
<td>Forward primer used for the amplification of \textit{wbbE} and \textit{wbbF}, including the RBS and original start site of \textit{wbbE} in the creation of pBH13; \textit{EcoRI} restriction site</td>
</tr>
<tr>
<td>TTGupWbbEF_Fwd</td>
<td>gatcgaattcAGGAGGTTAACAATTTTATTTATTATG</td>
<td>Forward primer used for the amplification of \textit{wbbE} and \textit{wbbF}, including the 108 bp upstream of \textit{wbbE} in the creation of pBH6; \textit{EcoRI} restriction site</td>
</tr>
<tr>
<td>WbbF_Rev</td>
<td>gatcttaccAGGAGGTTAACAATTTTATTTATTATG</td>
<td>Reverse primer used for the amplification of \textit{wbbF} and \textit{wbbE} in the creation of pBH6 and pBH13; \textit{PstI} restriction site</td>
</tr>
<tr>
<td>MnaA_Fwd</td>
<td>gatcggtaccATCGGAGCAAGGTTTATTCTTTTTG</td>
<td>Forward primer used for the amplification of \textit{mmaA} in the creation of pBH9; \textit{KpnI} restriction site</td>
</tr>
<tr>
<td>MnaA_Rev</td>
<td>gatcggtaccATCGGAGCAAGGTTTATTCTTTTTG</td>
<td>Reverse primer used for the amplification of \textit{mmaA} in the creation of pBH9; \textit{HindIII} restriction site</td>
</tr>
<tr>
<td>O54Fwd</td>
<td>gatcggtaccGCTGTATAGGAAATTAGGG</td>
<td>Forward primer used for the amplification of \textit{wbb}_*_O:54 operon in the creation of pWQ203; \textit{BamHI} restriction site</td>
</tr>
<tr>
<td>O54Rev</td>
<td>gatcctgcagTCAGTATATATAGTTCTTGAAGATAAC</td>
<td>Reverse primer used for the amplification of \textit{wbb}_*_O:54 operon in the creation of pWQ203; \textit{PstI} restriction site</td>
</tr>
<tr>
<td>WecB_delFwd</td>
<td>CTGCCGCGGTGAGCGGCGAAGCCGCAGTCGCGGCTTATCGAAGGCGAATCGA GTgttagctggagctggctgcg</td>
<td>Forward primer used for the amplification of \textit{cat} (Cm-resistance cassette) from pKD3 including 50 bp upstream of chromosomally-encoded \textit{wecB} in the creation of CWG1167 (Δ\textit{wecB})</td>
</tr>
<tr>
<td>WecB_delRev</td>
<td>CGTTTGCGACGCCGATATACCGAATCGCAGTCGCGGCTTATCGAAGGCGAATCGA GTgttagctggagctggctgcg</td>
<td>Reverse primer used for the amplification of \textit{cat} (Cm-resistance cassette) from pKD3 including 50 bp downstream of chromosomally-encoded \textit{wecB} in the creation of CWG1167 (Δ\textit{wecB})</td>
</tr>
<tr>
<td>WaaL_delFwd</td>
<td>TCAACAGTCAGCAGCAGTTTGTGGAAA GTTTATCTATTCGGATAGTAAAGC Gttcgctggagctggctgcg</td>
<td>Forward primer used for the amplification of \textit{cat} (Cm-resistance cassette) from pKD3 including 50 bp upstream of chromosomally-encoded \textit{waaL} in the creation of CWG1168 (Δ\textit{waaL})</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>WaaL_delRev</td>
<td>TTGTATAGATAAGAAGTGAGTTTTAACTCACTTCTTTAAACTTTGTTTATCg aatatcctccttagtcc</td>
<td>Reverse primer used for the amplification of cat from pKD3 including 50 bp downstream of chromosomally-encoded waaL in the creation of CWG1168 (ΔwaaL)</td>
</tr>
<tr>
<td>WecBFwdseq</td>
<td>CGTTGCTCGAAATAGCAACAC</td>
<td>Forward primer used for colony PCR to amplify cat insert and 5' junction between wecC and wecB</td>
</tr>
<tr>
<td>WecBRevseq</td>
<td>CGATTGATGGTATCAACCGC</td>
<td>Reverse primer used for colony PCR to amplify cat insert and 3' junction between wzzE and wecB</td>
</tr>
<tr>
<td>WaaLFwdseq</td>
<td>CCTCAACAGTCAAGCAGTTTTTG</td>
<td>Forward primer used for colony PCR to amplify cat insert and 5' junction between waaC and waaL</td>
</tr>
<tr>
<td>WaaLRevseqRev</td>
<td>CGTACTCCCGAAATATCCCTC</td>
<td>Reverse primer used for colony PCR to amplify cat insert and 3' junction between waaK and waaL</td>
</tr>
<tr>
<td>Catint5'Rev</td>
<td>CTTTGTGCCTTGCGTATAA</td>
<td>Reverse primer used for colony PCR to amplify the 5' junction of the cat insert and a portion of the flanking upstream sequence</td>
</tr>
<tr>
<td>Catint3'Fwd</td>
<td>ATCCAGCTGAACGGTCTGGTTTATAGG</td>
<td>Forward primer used for colony PCR to amplify the 3' junction of the cat insert and a portion of the flanking upstream sequence</td>
</tr>
<tr>
<td>pWQ799F1</td>
<td>GTTATTACCTTGCTGGCTTTTTAAATG</td>
<td>Forward primer used to amplify wbbF from pWQ799 to obtain the correct sequence for wbbF</td>
</tr>
<tr>
<td>pWQ799F2</td>
<td>GATGCGAATATCCTTACTGAGC</td>
<td>Forward primer used to amplify wbbF from pWQ799 to obtain the correct sequence for wbbF</td>
</tr>
<tr>
<td>pWQ799R1</td>
<td>CAGTAAGATCTTTTATCGCGTAG</td>
<td>Reverse primer used to amplify wbbF from pWQ799 to obtain the correct sequence for wbbF</td>
</tr>
<tr>
<td>pWQ799R2</td>
<td>CTATGACAAGTGGAGCCATC</td>
<td>Reverse primer used to amplify wbbF from pWQ799 to obtain the correct sequence for wbbF</td>
</tr>
<tr>
<td>KB1</td>
<td>GCTAAAAATAAAAAAGACTATCCTGACGTCCTCTCTGAAAGGCCCG</td>
<td>Forward primer used to generate the C49A mutation of WbbF in the creation of pBH68</td>
</tr>
<tr>
<td>KB2</td>
<td>CAAGAATTAGAAAGCGGGCTTTCAGGAGGAGCGTCAGGATAGTC</td>
<td>Reverse primer used to generate the C49A mutation of WbbF in the creation of pBH68</td>
</tr>
<tr>
<td>KB3</td>
<td>CCACTGATCGGAAGAGAGCTTTCAGGAGGAGCGTCAGGATAGTC</td>
<td>Forward primer used to generate the C102A mutation of WbbF in the creation of pBH69 and pBH75</td>
</tr>
<tr>
<td>KB4</td>
<td>CCACATGCTTTTACTTCTGACTATCAGCGATAAGTCTGGTCC</td>
<td>Reverse primer used to generate the C102A mutation of WbbF in the creation of pBH69 and pBH75</td>
</tr>
<tr>
<td>KB5</td>
<td>GCCCGGAAGCTATACAGGCGTGATCAGGAGCGTCCTGGAAGTCTC</td>
<td>Forward primer used to generate the C183A mutation of WbbF in the creation of pBH69 and pBH75</td>
</tr>
<tr>
<td></td>
<td>Primer Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>KB6</td>
<td>GCCAAAAGAGAGAGATGTTG\nAGTTTTTAGCCTACAGATACG</td>
<td>Reverse primer used to generate the C183A mutation of WbbF in the creation of pBH70 and pBH83</td>
</tr>
<tr>
<td>KB7</td>
<td>CTCAACATCTCTCTCTCTCTGTGC\nTACGCTACATCATACTGGAT</td>
<td>Forward primer used to generate the C195A mutation of WbbF in the creation of pBH71 and pBH79</td>
</tr>
<tr>
<td>KB8</td>
<td>GGAAAAATCGATT CATCAGATACG\nATGATGTCG TAGAGCAAAAG</td>
<td>Reverse primer used to generate the C195A mutation of WbbF in the creation of pBH71 and pBH79</td>
</tr>
<tr>
<td>KB9</td>
<td>CTGATAAATACTGGAGATTGCT\nTTAAAATCTCTGA CCAGAGGA</td>
<td>Forward primer used to generate the C237A mutation of WbbF in the creation of pBH72 and pBH80</td>
</tr>
<tr>
<td>KB10</td>
<td>CCAGTTCAATGT CCTCGTCAGAG\nATTTAAAGCA AATCCTCCCA</td>
<td>Reverse primer used to generate the C237A mutation of WbbF in the creation of pBH72 and pBH80</td>
</tr>
<tr>
<td>c328a_F</td>
<td>GGGCCGTGC TTGCAGGTTGCTAT\nTTTTTCATCAATATCTTTTC</td>
<td>Forward primer used to generate the C328A mutation of WbbF in the creation of pBH73 and pBH81</td>
</tr>
<tr>
<td>c328a_R</td>
<td>GAAAGATATTG TGA AAAATAATA\nGCAAACCTGC AAAGCA CGGCCC</td>
<td>Reverse primer used to generate the C328A mutation of WbbF in the creation of pBH73 and pBH81</td>
</tr>
<tr>
<td>c394a_F</td>
<td>CGTCACATTA ATATCCCATGCTTTA\nTGTTATGCTGATT TTTACC</td>
<td>Forward primer used to generate the C395A mutation of WbbF in the creation of pBH74 and pBH82</td>
</tr>
<tr>
<td>c394a_R</td>
<td>GGTAAAATCA GCATACCATAAGC\nGATGGATATTTA ATGTGACG</td>
<td>Reverse primer used to generate the C395A mutation of WbbF in the creation of pBH74 and pBH82</td>
</tr>
</tbody>
</table>

^ Restriction sites are underlined
\(^b\) Non-chromosomal sequence is shown in lower-case letters
\(^c\) Point mutations are bolded
PCR products were digested using restriction endonucleases (Invitrogen and New England Biolabs) in single and double enzyme reactions at 37°C for 1 hour, depending on compatibility of enzymes used. After each reaction, DNA fragments were purified using a kit, as described above. Ligations of digested DNA products and vectors were carried out using T4 DNA ligase (New England Biolabs) at room temperature overnight. Ligation reactions were transformed into chemically-prepared competent cells by heat-shock at 42°C (in waterbath or heatblock) for 1 min (Hanahan, 1983), or transformed into electrocompetent E. coli cells by electroporation using a MicroPulserTM electroporator (Bio-Rad). Immediately after transformation, fresh LB growth medium was added and the transformed cells were allowed to recover for 1 hr before plating on selective media.

Chemically-competent cells were prepared by a modified protocol based on the method developed by Hanahan (1983). E. coli Top10 cells were incubated overnight in 5 mL LB containing 10 mM MgSO₄ and subsequently diluted (1:100) into 100 mL of the same medium. Cells were grown until an OD₆₀₀ of 0.4-0.6 was reached. The culture was then transferred to pre-chilled 50 mL centrifugation bottles and chilled on ice for 10 min. Cells were collected by centrifugation at 3,000 x g for 10 min. Supernatant was discarded and the cell pellets were resuspended in 17 mL RF1 solution (100 mM RbCl, 50 mM MnCl₂·4H₂O, 30 mM CH₃CO₂K, 10 mM CaCl₂·2H₂O, 15% (v/v) glycerol; adjusted to pH 5.8 using glacial acetic acid). Resuspended cells were incubated on ice for 1 hour then harvested by centrifugation at 2000 x g for 15 min. Again the supernatant was discarded and cell pellet was resuspended in 4 mL RF2 solution (10 mM MOPS, 10 mM RbCl, 75 mM CaCl₂·2H₂O, 15% (v/v) glycerol; adjusted to pH 6.8 with 1M NaOH). Resuspended cells were incubated on ice for 15 min, and then 100 µL volumes were
aliquoted into prepared 1.5 mL microcentrifuge tubes and flash frozen using a bath of 95% ethanol and dry ice. Aliquots were stored at -80°C for later use.

Electrocompetent cells were also prepared by a procedure based on the method described by Dower et al. (1988). Bacteria were grown overnight in 5 mL LB, then diluted (1:100) into fresh medium (5 mL LB) and grown until an OD_{600} of 0.4-0.6 was reached. Cultures were cooled on ice for 10-20 min and then harvested by centrifugation at 10 000 x g for 1 min at 4°C. The cells were then washed twice using ice-cold sterile water and once using 10% ice-cold sterile glycerol. The cells were collected by centrifugation and then resuspended in 50 μL 10% glycerol and used immediately.

Transformants were screened for potential positive clones using an in-well lysis technique (Eckhardt, 1978). Individual colonies were patched to fresh plates to retain potential clones. A sample of each patched out colony was resuspended individually in TE buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0). 20 μL of SRL solution [25% w/v sucrose in Tris-borate-EDTA buffer (TBE; 90 mM Tris, 90 mM Boric acid, 2.5 mM EDTA); 2 U/mL RNase; 1 mg/mL lysozyme] was added to each cell suspension. The mixture was incubated for 5 min and then loaded into individual wells of a 1% agarose gel prepared in TBE containing 0.2% SDS. Electrophoresis was carried out in TBE buffer for 20 min at 20 V to allow samples to slowly enter the gel. The voltage was then increased to 80V for 40 min. Potential recombinant plasmids migrated slower through the gel compared to vector DNA. Analytical restriction endonuclease digestion tests were performed as additional verification of positive constructs identified in the initial in-well lysis screen. Successful constructs were confirmed by DNA sequencing performed at the Advanced Analysis Centre (University of Guelph).
2.4 Construction of ΔwecB and ΔwaaL in E. coli Top10

The λ-red recombinase system (Datsenko and Wanner, 2000) was used to generate individual ΔwecB and ΔwaaL chromosomal mutants in E. coli Top10. This procedure exploits a homologous recombination system derived from bacteriophage λ. A chloramphenicol-resistance (cm-resistance) cassette flanked with FLP recognition target (FRT) sites replaced the target gene. PCR was performed to amplify the cm-resistance cassette from plasmid pKD3. The oligonucleotide primers (Table 2.3) used for this were designed to include 50 nucleotide extensions homologous to the upstream and downstream region of the target gene. PCR reaction products were digested by DpnI (New England Biolabs) for 1 hour at 37°C to remove any remaining template plasmid DNA. Digested PCR products were then purified using the PCR kit as described previously. The DNA was concentrated by adding 5 μL of 5M NaCl and 110 μL 100% ethanol and then the mixture was left at -20°C for at least an hour. The DNA mixture was then left on ice to bring temperature up to 4°C and the DNA was harvested by centrifugation at 16 000 x g for 15 min. The DNA pellet was washed with 75% cold ethanol, air-dried and resuspended in 10 μL H2O.

Before transformation with the mutagenesis fragment, the parental strain E. coli Top10 was transformed with the Red helper plasmid pKD46, encoding the genes γ, β and exo behind a P_{araB} promoter that is inducible with arabinose. The Red helper plasmid helps facilitate recombination of the resistance cassette into the target location on the chromosome. Replication of plasmid pKD46 is temperature sensitive, and so transformed cells were grown at 30°C for maintenance of the plasmid. Electrocompetent cells expressing the Red helper genes were transformed with the PCR-amplified linear mutagenesis fragment. Transformed cells were allowed to recover at 30°C for 3 hours. Once recovered, cells were plated on LB-agar containing
chloramphenicol and grown at 37°C (non-permissive for replication) to eliminate helper plasmid. Colony PCR was performed using primers specific to the resistance gene and the upstream/downstream chromosome regions flanking the mutant (Table 2.3). The antibiotic cassette was then removed by the FLP recombinase encoded on the plasmid, pCP20 using the FRT sites included in the mutagenesis fragment. Successful removal of the cassette was confirmed determined by colony PCR using primers specific to the upstream/ downstream chromosome regions flanking the region of deletion.

2.5 Site-directed mutagenesis

Seven cysteine residues present in WbbF were substituted to alanine residues using a protocol based on the QuikChange site-directed mutagenesis protocol (Stratagene). Mutations were introduced in the \( \text{wb}^{*}\Omega:54 \) operon via PCR using oligonucleotide primers designed to incorporate the appropriate nucleotide base changes (Table 2.3). PCR cycle parameters were set up as follows. Samples were initially heated up to 95°C for 2 min. The cycling steps, repeated 25 times, included 20 sec at 95°C for denaturation of DNA, 20 sec at 55°C for primer annealing, and 2 min at 70°C for elongation. After the PCR reaction, the parental, non-mutated DNA was removed by digestion with \( \text{DpnI} \) for 2 hr at 37°C. The digested PCR products were transformed into \( \text{E. coli} \) Top10 cells. The desired DNA mutation(s) was confirmed by DNA sequencing.

2.6 Lipopolysaccharide analysis methods

LPS was examined in whole-cell lysates using an adaptation of the method developed by Hitchcock and Brown (Hitchcock and Brown, 1983). Overnight cultures were diluted 1:100 into 25 mL LB. Cultures were grown at 37°C until an approximate OD\(_{600}\) of 0.5 -1.0 was reached, unless otherwise stated. Cells were harvested by collecting a volume equivalent of 1.0 mL culture with an OD\(_{600}\) of 1.0 in 1.5 mL microcentrifuge tubes and then centrifuging at 13 000 x g
for 1 min. The supernatant was discarded and the cell pellets were resuspended in 100 μL 1× SDS-PAGE loading buffer containing 1 M Tris-HCl pH 6.8, 10% (v/v) glycerol, 2% SDS, 4% β-mercaptoethanol and 0.0025% (w/v) bromophenol blue (Laemmli, 1970). Samples were boiled for 10 min and then treated with proteinase K (2.5 μL of a 20 mg/mL stock) for at least 2 hr while incubated in a 55°C waterbath. 8 μL of each sample was loaded onto SDS-PAGE gels with a 4% acrylamide stacking layer and a 12.5% acrylamide resolving gel. Electrophoresis of samples was carried out in a Tris-glycine running buffer (25 mM Tris, 200 mM glycine, 0.1% (w/v) SDS) (Laemmli, 1970) at 40 V for 60 min, or until samples ran through the stacking layer. Electrophoresis was then continued at 80 V until the dye front reached approximately 0.5 cm from the bottom of the gel. Some samples were also analyzed by the tricine-SDS-PAGE gel system, developed by Lesse and colleagues for analyzing LPS and lipooligosaccharides (Lesse et al., 1990). 8 μL of each sample was loaded onto tricine SDS-PAGE gels with 4% acrylamide stacking layer and either a 12% or a 14% acrylamide resolving layer. Electrophoresis of samples was carried out with cathode buffer (0.1 M Tris, 0.1 M tricine, 0.1% SDS) in the top chamber and anode buffer (0.2 M Tris-HCl, pH 8.9) in the lower chamber. Samples were electrophoresed at 30 V until samples ran through the stacking layer, then electrophoresis was continued at 105 V until the dye front reached approximately 0.5 cm from the bottom of the gel.

LPS was visualized using the Fomsgaard and colleagues (1990) adaptation of the silver stain method developed by Tsai and Frasch (1982). The entire silver staining protocol was performed at room temperature with constant shaking. Once electrophoresis was completed, the SDS-PAGE gel was incubated in 100 mL of fixing solution (40% ethanol, 5% acetic acid) for a minimum of 1 hr (maximum overnight). Next, the gel was incubated for 15 min in oxidizing solution containing 0.7% (w/v) H₃IO₆ (periodic acid) in fixing solution and subsequently washed
three times for 10 min using Millipore water (from Millipore SuperQ system). After these washes, the gels were stained for 10 min in silver stain containing 0.67% (w/v) AgNO$_3$, 1.3% (v/v) NH$_4$OH hydroxide and 19 mM NaOH. Stained gels were again washed 3 times for 5 min using Millipore water. The stained gels were then developed using developer solution containing 0.025% (w/v) citric acid and 0.0925% (v/v) formaldehyde. Development was stopped by washing repeatedly in water.

O:54 antigen was detected by Western immunoblotting. Samples were transferred from SDS-PAGE gels onto PROTRAN nitrocellulose membranes (PerkinElmer Life Science) by electrophoresis at 200 mA for 60 min in transfer buffer (25 mM Tris, 200 mM glycine, 20% (v/v) methanol). Membranes were incubated at room temperature with shaking for all steps after electrophoresis. Membranes were washed twice in phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10.1 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$), then blocked in 2.5% (w/v) skim milk in PBS for at least 1 hr. Primary anti-O54 serum (Statens Serum Institut, Denmark) diluted 1:1000 in 1% (w/v) bovine serum albumin (BSA) in PBS was applied to membranes overnight. Membranes were washed with PBS solution three times for 10 min and then incubated for 1 hr with alkaline phosphatase-conjugated goat anti-rabbit secondary antibody (Cedar Lane Laboratories) at a dilution of 1:3000 in PBS. The membranes were again washed three times for 5 min each with PBS. O-Ag was detected using a solution of 122 mM nitro-blue tetrazolium chloride (Roche) and 115 mM 5-bromo-4-chloro-3-indolyl phosphate (Roche) in alkaline phosphatase buffer (100 mM Tris-HCl, 100 mM NaCl, 5 mM MgCl$_2$; pH 9.5). Images of the developed membranes were captured digitally using a Bio-Rad GS-800 Calibrated Densitometer.
2.7 Agglutination reactions

The presence of O:54 polysaccharide on the cell surface was detected by an agglutination test. Cells from a liquid culture were harvested by centrifugation at 13 000 x g and resuspended in PBS buffer. Conversely, a single colony was picked from a fresh overnight plate and resuspended in PBS buffer solution. A 10 μL aliquot of the cell suspension was placed on a glass slide and 1 μL of anti-O54 serum was added. Samples were incubated in the anti-O54 serum with gentle rocking at room temperature. Positive reaction for O:54 polysaccharide presence on cell surface was reflected in visible clumping of cells.

2.8 Protein analysis techniques

All protein samples were prepared as follows with the exception of certain steps in the preparation of WbbF protein samples, which will be described in the Results section (altered steps will be indicated when necessary). Overnight cultures were diluted 1:100 in 25 mL LB broth supplemented with the appropriate antibiotic(s). Cultures were grown until an OD_{600} of 0.6 was reached, at which time protein expression was induced using the appropriate inducing chemical (L-arabinose, AhT, or IPTG at the appropriate concentrations), unless otherwise stated. Cultures were incubated for another 1.5 hr. The OD_{600} of the culture was measured to determine the volume of cells necessary to correspond to 1 mL of a culture of cells with an OD_{600} of 1.0. The cells were then collected by centrifugation at 13 000 x g for 1 min and resuspended in 100 μL of 1× SDS-PAGE loading buffer (Laemmli, 1970). Routine samples were boiled for 10 min. Samples containing WbbF were heated at 45°C for 30 min and these samples were physically sheared using a 26 G1/2 needle (Becton Dickinson) before loading on gels.

Sub-cellular location of WbbF was examined using an ultracentrifugation method. Overnight cultures were diluted 1:100 in 250 mL LB broth and grown until an OD_{600} of 0.6 was
reached, at which time protein expression was induced using the appropriate inducing chemical (L-arabinose, AhT, or IPTG at the appropriate concentrations), unless otherwise stated. Cultures were incubated for another 1.5 hr. The cells were then collected by centrifugation at 5 000 x g for 10 min. The cell pellet was resuspended in 25 mL of buffer containing 50 mM Tris-HCl and 50 mM NaCl, pH 7.5, prior to lysis by sonication. A cell-free lysate was prepared by sequential centrifugation steps at 4 000 x g for 10 min and 12 000 x g for 20 min. The supernatant of the 12 000 x g centrifugation spin was then centrifuged at 100 000 x g for 60 min to separate the membrane fraction (pellet) from the soluble fraction. The membrane fraction pellet was resuspended in 750 μL of Tris-HCl, NaCl buffer. Protein samples were prepared at each step in the procedure.

Protein samples were analyzed by SDS-PAGE and Western immunoblotting. 10 μL of each protein sample was loaded onto an SDS-PAGE gel with 4% stacking layer and a 12.5% or 18% resolving gel and subjected to electrophoresis at 150 V in Tris-glycine buffer until the dye front reached approximately 0.5 cm from the bottom of the gel. Proteins were visualized using SimplyBlue SafeStain (Invitrogen) using the manufacturer’s fast stain microwave protocol.

Hexa-histidine (His6-) tagged proteins were detected by Western immunoblotting (Towbin et al., 1992). All steps in the procedure were routinely performed at room temperature with shaking. In some cases, a step was performed overnight in which case membranes were held at 4°C instead. Proteins were transferred from SDS-PAGE gels to PROTRAN nitrocellulose membranes by electrophoresis at 200 mA for 60 min in immunoblot transfer buffer. The membranes were washed twice in Tris-buffered saline tween (TBST) solution (10 mM Tris-HCl, pH 7.5; 150 mM NaCl, 0.05% (w/v) Tween 20), then blocked for at least 1 h in 3% (w/v) BSA in TBST. The membranes were then washed 3 times for 5 min each using TBST solution. The
primary antibody mouse anti-His$_5$ (Qiagen) was diluted 1:2000 in TBST solution containing 1% BSA and was applied to membranes and incubated for minimum 1 hr (maximum overnight at 4°C). The membranes were washed again 3 times for 5 min using TBST solution and then incubated for up to 1 hr in horseradish peroxidase (HRP)-conjugated goat anti-mouse (Jackson ImmunoResearch Laboratories). This secondary antibody was diluted 1:3000 in TBST solution containing 3% skim milk. Chemiluminescence detection was carried out using Luminata Classico or Crescendo (for higher sensitivity) Western HRP substrate (Millipore Corporation) according to the manufacturer’s instructions.

2.9 Microscopy methods

Cells were viewed under a 63× objective lens using a Zeiss Axiovert 200M microscope. Images were captured with a C10600-10B camera (Hamamatsu Photonics K.K., Japan) using the Volocity version 6.2.1 software program (PerkinElmer Life Science).

2.10 In silico analysis methods

Multiple sequence alignments were generated using the Clustal Omega (formally known as ClustalW) server available at the European Bioinformatics Institute website (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Goujon et al., 2010; Sievers et al., 2011). All amino acid sequences used for alignment were obtained from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/). Two nucleotides (nucleotide numbers 975 and 976, CG) listed in the nucleotide sequence for WbbF were found to be incorrect in the NCBI Accession AAC98402. This resulted in two incorrect amino acids (F325 and E326) in the translated sequence for WbbF. This was determined by sequencing \textit{wbbF} using primers specific for different sections of \textit{wbbF} (Table 2.3). The correct nucleotide sequence is GC rather than CG,
and so the proper amino acid sequence is L325 and Q326. This was amended for all cloning and amino acid sequence alignments of *wbbF* presented in this thesis.

Protein topology predictions were created using three online webserver programs, SVMtop (http://bio-cluster.iis.sinica.edu.tw/~bioapp/SVMtop/tmp/index.php) (Lo et al., 2007), HMMTOP (http://www.enzim.hu/hmmtop/) (Tusnády and Simon, 1998, 2001), and TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001; Sonnhammer et al., 1998).
3.0 RESULTS

3.1 Optimization of O-antigen detection techniques

Our laboratory determined that the assembly of the O:54 homopolymer by WbbF cannot occur until WbbE creates the appropriate acceptor (ManNAc-GlcNAc-PP-Und) for polymerization (Keenleyside and Whitfield, 1996). Moreover, that study proposed that only WbbE and WbbF are essential components encoded by the plasmid-encoded operon (wb*O:54) for the biosynthesis of the O:54 antigen. The mnaA gene from wb*O:54 was understood to be expendable due to the presence of chromosomal wecB gene, which encodes a functional homolog of MnaA (Meier-Dieter et al., 1990).

To confirm these observations, whole-cell lysates were prepared from E. coli Top10 cultures containing either pWQ800 (wb*O:54; mnaA); the gene mnaA is disrupted by the vector, pGEM-7Zf (+) which was inserted into an EcoRI site located in mnaA) or pKB7bc (encoding wbbF). For positive and negative controls, whole-cell lysates were prepared for S. enterica sv. Borreze, containing the native pWQ799 plasmid, and E. coli Top10. O:54 LPS was detected in the positive control S. enterica sv. Borreze sample and E. coli Top10 sample expressing WbbE and WbbF on the silver stained SDS-PAGE gel (Figure 3.1). Presence of O:54 antigen was verified by Western immunoblot with anti-O:54 serum. O:54 antigen was not detected in the negative control E. coli Top10, or in the same strain containing only WbbF, thus confirming previous results. Although O:54 LPS was observed for the positive control and the sample containing WbbE and WbbF on the silver stained SDS-PAGE gel, the staining was weak and lacked the typically well-resolved ladder of S-LPS (Figure 3.1A) (Hitchcock and Brown, 1983; Palva and Mäkelä, 1980; Tsai and Frasch, 1982). Instead, the stained O-Ag appeared as a smear. The staining intensity of the lipid A-core molecules visible on the SDS-PAGE gel was greater
Figure 3.1: Confirmation of O:54 production in *E. coli* Top10 strain in the presence of WbbE and WbbF, but not WbbF alone. LPS samples prepared from proteinase K digested whole-cell lysates was resolved by SDS-PAGE and visualized by a (A) silver stained SDS-PAGE gel and (B) corresponding α-O54 Western immunoblot.
compared to other smooth LPS species, suggesting the amount of O:54 antigen being ligated to the lipid A-core (R-LPS) was low. The O:54 antigen detected by Western immunoblotting also appeared as a smear (Figure 3.1B).

These observations were concerning since sensitive detection of the O:54 polysaccharide was critical for the objectives of this thesis. Therefore, it was imperative to address potential issues with the LPS detection methods to increase confidence in the reliability and sensitivity of O:54 antigen detection. There are numerous factors that may contribute to poor detection of LPS containing the O:54 antigen, and so these factors were addressed in a sequential manner.

The properties of the SDS-PAGE gels may contribute to the poor resolution characteristics of the O:54 LPS (Hitchcock, 1983; Peterson and McGroarty, 1985). Therefore, two different SDS-PAGE gel recipes were compared to determine if resolution of the O:54 LPS could be increased by altering the properties of the gel. The original SDS-PAGE gels used to produce Figure 3.1 consisted of 12.5% (w/v) acrylamide, 0.375 M Tris-HCl (pH 8.8), 0.1% SDS, 0.05% (w/v) ammonium persulfate (APS) and 0.0025% (v/v) TEMED (N,N,N′,N′-tetramethylethylenediamine). This will be referred to as the “original SDS-PAGE” gel. An adapted 12.5% acrylamide (w/v) SDS-PAGE gel recipe was tested, consisting of 0.57 M Tris-HCl (pH 8.8), 0.15% (w/v) SDS, 0.15% (w/v) APS and 0.0005% (v/v) TEMED. This system has previously been used successfully in the Whitfield laboratory to examine LPS for numerous bacterial strains including *E. coli* O8 and *E. coli* O9a (Greenfield et al., 2012) and will be referred to as the “new SDS-PAGE gel” recipe. Standard samples of the well-documented *E. coli* serotypes O8 and O9a, along with lysate of cells expressing O:54 LPS, were examined by both gel recipes. The resolution of the LPS samples *E. coli* O8, *E. coli* O9a, *S. enterica* sv. Borreze and *E. coli* Top10 containing pWQ800 was better on the new SDS-PAGE gel, compared to the
original SDS-PAGE method (Figure 3.2). The lipid A-core molecules also appeared more compact on the new SDS-PAGE gel compared to the original, signifying improved resolution; thus the “new SDS-PAGE” gel conditions were used for subsequent experiments. The visibility and the ladder-like appearance of LPS species containing the O:54 polysaccharide were, nonetheless, still relatively poor on the new SDS-PAGE gel.

The staining of LPS containing the O:54 antigen in Figure 3.2 was considerably fainter than the E. coli O8 and O9a antigens. This could be due to lower amounts of O:54 LPS, and so a staining procedure with greater sensitivity may be required for effective O:54 LPS detection. Alternative staining methods were examined in an attempt to increase the sensitivity. Two SDS-PAGE gels were loaded and run with the same whole-cell lysate preparations of E. coli Top10 containing pWQ800 and S. enterica sv. Borreze in a range of increasing sample volumes including the 8 μL loading volume used previously. These gels were stained with one of two silver staining methods: a modified Tsai and Frasch protocol developed by Fomsgaard and colleagues, which was used in Figure 3.1A and Figure 3.2 (Fomsgaard et al., 1990) and the commercial, ultrasensitive silver staining kit Pierce© SilverSNAP Stain Kit II (Thermo Scientific). Development of the SDS-PAGE gel stained using the SilverSNAP kit occurred immediately after the developing solution was applied, causing the stain to appear very dark and severely smeared, consequently the lipid A-core region and the O:54-containing LPS species were virtually indistinguishable (Figure 3.3A). Therefore, this method was determined to not be suitable for O:54 LPS detection. In comparison, O:54 LPS from S. enterica sv. Borreze stained using the Fomsgaard ultrafast silver stain protocol was better resolved with a distinguishable banding pattern with higher loading volumes. A partial banding pattern was observed for E. coli Top10 (pWQ800) LPS with 10 μL loading volume, although this was fainter than the LPS from
**Figure 3.2:** Resolution quality of LPS samples on SDS-PAGE gels with different compositions. LPS preparations made from whole-cell lysates and proteinase K digested were loaded onto A) new SDS-PAGE gel and B) original SDS-PAGE gel. LPS samples were visualized by the silver staining gels.
**Figure 3.3:** Comparison of silver staining methods for optimization of O:54 LPS visibility. **A)** Pierce® SilverSNAP Kit and **B)** Fomsgaard ultrafast, sensitive silver stain methods were used to separately stain LPS whole-cell lysate preparations of *S. enterica* sv. Borreze and *E. coli* Top10 containing pWQ800. The range of loading volumes is indicated above each lane.
S. enterica sv. Borreze. Because visibility of O:54-containing LPS was improved using the Fomsgaard silver staining technique, this technique was used in all subsequent experiments.

Although the study of LPS structure has greatly advanced since the development of the SDS-PAGE gel electrophoresis method for LPS observation, resolution and detection of some of the vastly different O-Ag structures by SDS-PAGE gel electrophoresis has proven difficult (Darveau and Hancock, 1983; Jann et al., 1975; Kittelberger and Hilbink, 1993). One strategy exploited to improve the separation of LPS, as well as lipooligosaccharides, was the tricine SDS-PAGE gel electrophoresis system (Lesse et al., 1990). This system has been used to visualize O:54-containing LPS in previous studies (Keenleyside and Whitfield, 1995, 1996), so a tricine-based gel system was compared to the glycine-based SDS-PAGE gel system. The effect of the amount of acrylamide present in the resolving gel (10% or 12.5% acrylamide for glycine-SDS-PAGE gels, and 12% or 14% for the tricine-SDS-PAGE gels) was also examined. In general, 12.5% acrylamide glycine-SDS-PAGE gels yielded better resolution than the 10% acrylamide glycine-SDS-PAGE gel (Figure 3.4A), whereas, the 12% and 14% acrylamide tricine-SDS-PAGE gels were comparable (Figure 3.4B). As expected, resolution of the lower molecular weight R-LPS species was markedly improved in a tricine gel. The results shown in Figure 3.4 establish that resolution of O:54 LPS by the tricine-SDS-PAGE gel system generates a modest improvement in the O:54 LPS profile detected on silver stained gels. However, the profile of the O:54 antigen on the corresponding Western immunoblot is not as distinguishable as the glycine-based system. Since the silver staining method of detection for O:54-containing LPS yields low sensitivity, potentially resulting in loss of detection of certain O:54 LPS bands, Western immunoblotting was chosen as the appropriate detection method to unequivocally confirm the
Figure 3.4: Comparison of O:54 LPS separation using a glycine-based and tricine-based SDS-PAGE gel systems. Whole-cell lysates of *E. coli* Top10 expressing pWQ800 were separated and detected using the A) glycine SDS-PAGE gel system and B) tricine SDS-PAGE gel system. The upper panel shows the silver-stained gel and the corresponding Western immunoblot is shown below. *E. coli* O9a and *E. coli* Top10 cell LPS preparations were used as positive and negative controls, respectively.
presence of the O:54 antigen. Thus, the glycine-based SDS-PAGE system was used for further analyses.

The role of the composition of the sample buffer was also examined. Most researchers examining LPS profiles by SDS-PAGE gel electrophoresis use a standard sample loading buffer developed by Hitchcock and Brown, although other examples of sample loading buffers can be found (Peterson and McGroarty, 1985; Kittelberger and Hilbrink, 1993; Hitchcock and Brown, 1983). The standard sample loading buffer from Hitchcock and Brown (1983) was used as a reference buffer and is denoted 1× sample buffer. 2× and 4× sample buffers were prepared, where concentrations of all components were increased accordingly (Table 3.1). Previous research established that the concentration of SDS can influence the migration and banding pattern of LPS, and so 1× sample buffers with 2%, 4% and 8% w/v SDS concentrations were also made (Peterson and McGroarty, 1985). Whole-cell lysates of E. coli Top10 with pWQ800 were prepared in each sample loading buffer listed above. The mobility and resolution of O:54-containing LPS were not significantly different with increasing loading buffer concentration (Figure 3.5). Similarly, increasing the percent of SDS in 1× sample buffer did not affect the resolution of the O:54-containing LPS (Figure 3.6). Therefore, 1× sample buffer was used for subsequent sample preparations.

Another element of LPS preparation that might influence O:54 antigen detection is the quantity of O:54 antigen recovered in each LPS preparation. These are standardized by optical density of the culture. Physiological processes do change within an individual bacterial cell over time and growth phases, depending on environmental factors. Variation in O-Ag quantity during different growth phases has been observed in Salmonella serovars Typhi, Typhimurium and Enteritidis (Bittner et al., 2002; Bravo et al., 2008; Rojas et al., 2001). Therefore, it is plausible
Table 3.1: Composition of 1×, 2× and 4× sample loading buffer solutions.

<table>
<thead>
<tr>
<th>Sample loading buffer</th>
<th>Chemical</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1×</td>
<td>Glycerol</td>
<td>10% (v/v)</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>2.5% (v/v)</td>
</tr>
<tr>
<td></td>
<td>SDS</td>
<td>2% (w/v)</td>
</tr>
<tr>
<td></td>
<td>Tris-HCl (pH 6.8)</td>
<td>1.0 M</td>
</tr>
<tr>
<td>2×</td>
<td>Glycerol</td>
<td>20% (v/v)</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>5% (v/v)</td>
</tr>
<tr>
<td></td>
<td>SDS</td>
<td>4% (w/v)</td>
</tr>
<tr>
<td></td>
<td>Tris-HCl (pH 6.8)</td>
<td>1.0 M</td>
</tr>
<tr>
<td>4×</td>
<td>Glycerol</td>
<td>40% (v/v)</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>10% (v/v)</td>
</tr>
<tr>
<td></td>
<td>SDS</td>
<td>8% (w/v)</td>
</tr>
<tr>
<td></td>
<td>Tris-HCl (pH 6.8)</td>
<td>1.0 M</td>
</tr>
</tbody>
</table>
Figure 3.5: Effect of sample buffer composition on LPS mobility and resolution. Sample 1-3 corresponds to whole-cell lysates from *E. coli* O9a, *E. coli* Top10, *E. coli* Top10 containing pWQ800, respectively. LPS samples were analyzed by A) silver stained SDS-PAGE gel and B) corresponding α-O:54 Western immunoblot.
Figure 3.6: Effect of sample buffer SDS concentration on LPS mobility and resolution. Sample 1-3 corresponds to whole-cell lysates from *E. coli* O9a, *E. coli* Top10, *E. coli* Top10 containing pWQ800, respectively. LPS samples were analyzed by A) silver stained SDS-PAGE gel and B) corresponding α-O:54 Western immunoblot.
that the amount of O:54 antigen produced by the synthase-dependent pathway could also be regulated by growth phase. The quantity of O:54 LPS produced during logarithmic growth and stationary growth phases was examined. Before this experiment was performed, a growth curve was produced for *E. coli* Top10, with and without pWQ800, to determine whether O:54 antigen expression had any effect on bacterial growth (Figure 3.7). The two growth curves were indistinguishable. Cells reached mid-logarithmic phase after 2 hours of growth (OD$_{600}$ of approximately 0.5), and stationary phase after 4.5 hours of growth (OD$_{600}$ of approximately 3.0). Qualitatively, there was no significant difference in the O:54 LPS profiles from cells in either growth phase (Figure 3.8). Therefore, cells were harvested for subsequent experiments at an approximate OD$_{600}$ of 0.5-0.6 for LPS preparations, in order to minimize preparation time.

Sample preparations of O:54 LPS produced in an *E. coli* K-12 derivative strain and *S. enterica* serovars with native O:54 LPS are based on optical density readings. To rule out any effect cell sizes may have on optical density measurements, which would ultimately affect the LPS profiles, the average number of CFU/mL of cells grown to an OD$_{600}$ of 1.0 was calculated for each bacterial species. From this data set, standard deviation and standard error were also calculated using the accepted formulae. The average CFU/mL of *S. enterica* sv. Borreze, *E. coli* Top10, and CWG1167 (strain used in later experiments and is described below) were comparable within standard error (Figure 3.9). Thus, direct comparison of the optical density measurements for each of these strains is valid.

**3.2 Effects of expressing WbbE and WbbF from separate plasmids on O:54 LPS profiles**

Production of the O:54 antigen by WbbE and WbbF proteins expressed from the $wb^{*}_{O:54}$ biosynthetic gene cluster in *E. coli* Top10 was compared with native production of O:54 antigen in *S. enterica* sv. Borreze under conditions where all of the optimization parameters were
Figure 3.7: Growth curve of *E. coli* Top10 and *E. coli* Top10 containing pWQ800. Overnight cultures were subcultured 1:100 into 25 mL fresh LB medium and allowed to grow for a period of time. Optical density (600 nm) was measured every 0.5 hr from inoculation. Experiment was performed in triplicate and OD<sub>600</sub> was averaged for plot. Variation of calculated standard deviation and standard error were insignificant, and so error bars are not visible on plot. Mid-logarithmic and stationary growth phase time points are indicated by arrows on plot.
Figure 3.8: LPS profiles of *E. coli* Top10 cells in mid-logarithmic and stationary growth phases. Whole-cell lysates were from *E. coli* Top10 cells (with and without pWQ800) harvested at an OD$_{600}$ of 0.5 and an OD$_{600}$ of 2.5 when cells. *E. coli* O9a was included as a control. Samples were analyzed by A) silver stained SDS-PAGE and B) anti-O:54 Western immunoblotting.
Figure 3.9: Cell counts (colony-forming units) for *S. enterica* sv. Borreze, *E. coli* Top10 and CWG1167 (ΔwecB) corresponding to an OD$_{600}$ of 1.0. Three cultures of each strain were grown to an optical density (600 nm) of 1.0, then serially-diluted and plated to calculate colony forming units per mL (CFU/mL). The average CFU/mL and standard error for each strain was calculated and plotted.
incorporated. In addition, the O:54 LPS profile produced from cells expressing WbbE and WbbF from separate plasmids was compared. The purpose of this experiment was to determine whether it would be feasible to keep wbbF on a separate plasmid to simplify analysis of the effects of its manipulation. Different O:54 profiles were observed in E. coli Top10 when WbbE and WbbF were expressed from separate plasmids compared to when the two GT enzymes were expressed in the context of wb*O:54mnaA− (pWQ800) (Figure 3.10). Expression of WbbE and WbbF from the plasmid pWQ800 in E. coli Top10 resulted in a reduced amount of O:54 LPS compared to native production of O:54 from pWQ799 in S. enterica sv. Borreze. The average chain length, i.e. the region showing the highest intensity in the Western immunoblot, also decreased. The O:54 LPS produced when WbbE and WbbF were expressed from separate promoters (in constructs pWQ823 and pKB101, respectively), was barely detectable on both the silver stained gel and Western immunoblot. The average chain length was also significantly decreased compared to the O:54 LPS profile observed from E. coli Top10 (pWQ800) and S. enterica sv. Borreze. Because three open reading frames were identified in the wb*O:54 operon, but only one promoter in the upstream sequence of wbbE (Keenleyside and Whitfield, 1996), it is highly probable that the transcription and translation of wbbE and wbbF is coupled.

Although these results confirmed detectable O:54 antigen can be assembled in an E. coli K-12 host with only the presence of WbbE and WbbF from the wb*O:54 operon (pWQ800), the generation of O:54 antigen was significantly decreased compared to native O:54 LPS production in S. enterica sv. Borreze (pWQ799). This suggested that either one or more of the activities of the enzymes expressed from the O:54 biosynthetic gene cluster in pWQ800 is reduced, or there is a substantial difference when O:54 antigen is expressed in E. coli versus S. enterica. Because of this, I sought to further understand the three enzymes from the wb*O:54 operon within the
Figure 3.10: O:54 production in *E. coli* Top10 by WbbE + WbbF expressed from separate constructs and expressed from pWQ800 (wb*O:54ΔmnaA−) compared to *S. enterica* sv. Borreze. A) silver stained SDS-PAGE gel of LPS samples from whole-cell lysates of *E. coli* Top10 containing pWQ572 (negative control), pWQ800, or pWQ823 (WbbE) and pKB101 (WbbF), and *S. enterica* sv. Borreze. B) corresponding anti-O54 Western immunoblot.
context of the synthase-dependent pathway to determine what is truly required for optimum O:54 production.

3.3 Upstream sequence of \textit{wbbE} does not improve O:54 production

Since O:54 LPS production was higher in \textit{E. coli} containing pWQ800, compared to the separate plasmids, elevated expression of WbbEF could potentially maximize O:54 production in \textit{E. coli} Top10. To explore this, a construct containing \textit{wbbEF} was generated under control of an exogenous promoter. Examination of the \textit{wb*}	extsubscript{O:54} operon revealed several potential, but more rarely used start codons that were in frame with \textit{wbbE} and upstream from the proposed start codon (Figure 3.11) (Gualerzi and Pon, 1990). Only one initiation codon, ATG, is preceded by an identifiable ribosome binding site (RBS) candidate, but protein synthesis can occur without an RBS (Keenleyside and Whitfield, 1996; Nakamoto, 2009). Using an incorrect start codon will produce a truncated protein, which may compromise the activity of the protein. Therefore, to resolve this potential issue, two different forward primers were designed to incorporate each potential start codon in two constructs encoding \textit{wbbE}. The constructs were made in the vector pWQ572 because of the tight control over the inducible promoter. Plasmid pBH13 contains \textit{wbbEF} with the expected start codon and the deduced RBS of \textit{wbbE} (Keenleyside and Whitfield, 1996). The second construct, pBH6 contains \textit{wbbEF} plus an additional 108 base pairs upstream from \textit{wbbE} to include the three other potential start codons that were identified. This construct also includes the suggested upstream promoter region, which was described in previous research (Keenleyside and Whitfield, 1996). The expression of O:54 LPS from the constructs pBH6 and pBH13 was examined using different amounts of AhT inducer to vary protein expression levels (Figure 3.12). A similar quantity and quality of O:54 LPS from whole-cell lysates of \textit{E. coli} Top10 cells containing either of the constructs was observed, confirming both versions of WbbE
1- TTGAA TTTGAGCTGTAGT GACATATCTT AAGTTGCTTTA AGTA
51- AATCATTTTTACCTATCTGCA TGCGATGCTGTATAGGA AATTAGGAAT
101- TAGGGAAAT ATGGGCAGCTAGT TACGTGTTGTTGAC ATACGGAAC
151- GATGCAACAGAACATTTCAATGTGCGCTGTATATTGTA AATTGGAAT
201- CAGTAAACCCAATCGAGGTTATTTGATGTGATAATTGGTGA ATTACAGTA
301- GACATTATGAATAATTTGCTAATATTAGTGAATC ACAAGCATTCGCTTG
351- TTGAGAATAGCA CATCTAAGCTCTGTTGGGCAGCGAGAAACATCGC ACT
401- GCAATGCTGCA AAGGTACGTTTCAGTTGGCAGATCAGATTAATTA
451- ATGAGCCTGACCGTTTTATTAAAGCTATTCTCTGAGCTA AACGATAGCGTA
501- AGAATGGTGGGGCATGTTGGTGCAGA AATTAGTCTGTGAAGAAAACGATAG
551- AGGCAATTGTCGTAAGTCTCAATGCGAGATTGATTTGGA ATTATTCGATT TA
601- CAGGATTCGATCACCTGTGTAATAACCCAAACTATCTGTTACTTTAAAAT
651- GATGCACTGAGTGTGAAGGGGTTATAACACGTCTTACTTTAATTTGAGAGA
701- CTATGATTATGTGACGAGAATTTATCAGGAATGGGAAAATTATCGAAACA
751- TTTCCCATTTTATGGAACCTCCGAGTTGGTGAATGCTAAAAAGCTTTAT
801- AAAAAAGAGGTGGGCCCTTTTTATCAACGAGGAATAAATCTGCTAGCG
851- AAGAATTCTGGAAGATATATTAGATTGTCTTCAGTTGTGATTAATG
901- TAATCTAAAGAATTTCTGTTAGGTTATACCTTGCTTGCTTTTAATGTG
951- TTTTATTTTAAGGTTAAGAAGC ATGAATGA

**Figure 3.11:** Nucleotide sequence of *wbbE* including 108 base pairs upstream. The *wbbE* nucleotide sequence obtained from Genbank (accession AAC98401.1) is highlighted in grey with the nucleotides overlapping with *wbbF* highlighted in aqua and italicized. Potential start codons are highlighted in green and the previously determined start codon is underlined. The nucleotides predicted to be part of the ribosome binding site are red.
Figure 3.12: Inclusion of DNA upstream of the potential start codon of WbbE does not improve O:54 production. LPS samples were prepared from whole-cell lysates of *E. coli* Top10 containing pBH6 (containing *wbbE*, *wbbF* and 108 base pairs upstream of *wbbE*) compared to LPS samples of whole-cell lysates of *E. coli* Top10 containing pBH13 (RBS*wbbE*, and *wbbF*) on A) silver stained SDS-PAGE gel and B) corresponding anti-O:54 Western immunoblot. The concentration of anhydrotetracycline (AhT) used to induce LPS samples is indicated above each lane.
are functional (Figure 3.12B). Whether the upstream region is translated into the N-terminal domain of WbbE is unknown from the current data, but these results prove that this region is not necessary for a functional protein. Furthermore, the concentration of inducer (AhT) added to express WbbE (both versions) and WbbF proteins did not affect the amount or quality of O:54 antigen produced, since the profile did not vary for each sample. The amount of O:54 LPS and average chain length of the O:54 polysaccharide was still lower compared to native O:54 generated in S. enterica sv. Borreze, suggesting some other factor was affecting the productivity of the synthase-dependent pathway when expressed in an E. coli background.

3.4 Construction of a wecB deletion mutant in E. coli Top10

The main component missing from the previous experiments is mnaA from the wb*O:54 operon. The sole EcoRI site in the gene mnaA was used to cut the wb*O:54 operon for ligation into the MCS site of pGEM-7Zf (+) to form the hybrid plasmid pWQ800, thereby generating an mnaA mutant (Keenleyside et al., 1994; Keenleyside and Whitfield, 1996). This raised questions concerning the role of MnaA in the O:54 antigen synthase-dependent pathway, particularly whether it is effectively replaced by the functionally redundant epimerase encoded by the chromosomal wecB gene. A requirement for MnaA could reflect some critical protein-protein interaction or a critical quantity of the donor sugar nucleotide pool required for optimal WbbF activity. The efficiency of MnaA and WecB in precursor production could differ, and so WecB may provide lower donor sugar nucleotide stocks which would result in decreased O:54 production. Alternatively, poor O:54 production could be attributed to some unknown feature of the O:54 antigen biosynthesis system that is poorly handled by the strain E. coli Top10, and thus restrains O:54 antigen production.
To test the hypothesis that MnaA influences the GT activity during O:54 antigen production, the chromosomally-encoded \( \text{wecB} \) was first deleted from \( E. \text{coli} \) Top10 to ensure any observations of the O:54 LPS profile can be attributed to only the presence or absence of MnaA from the synthase-dependent pathway. The lambda Red mutagenesis system was utilized to generate the \( \Delta \text{wecB} \) deletion mutant strain by recombining a \( \text{cm} \)-resistance cassette into the target gene (Datsenko and Wanner, 2000). The \( \text{cm} \)-resistance cassette was subsequently removed to eliminate conferred antibiotic resistance. This strategy is summarized in a cartoon diagram shown in Figure 3.13A and the PCR results confirming the mutation is provided in Figure 3.13B.

Potential recombinants were preliminarily screened for insertion of the gene, \( \text{cat} \) encoding the \( \text{cm} \)-resistance cassette by colony PCR. The primers WecBFwdseq and WecBRevseq were designed to adhere to a region of the flanking upstream and downstream sequences of \( \text{wecB} \) to amplify the entire \( \text{cm} \)-resistance cassette. Observation of a band corresponding to the expected size of the \( \text{cat} \) gene (1231 bp) confirmed the successful removal of \( \text{wecB} \) (Figure 3.13B). Additional validation of the insertion of \( \text{cat} \) into \( \text{wecB} \) was achieved by using internal primers Catint5'Rev and Catint3'Fwd specific to the \( \text{cat} \) cassette along with WecBFwdseq and WecBRevseq, in separate PCR reactions. These primer pairs amplify the 5' and 3' junctions where the \( \text{cat} \) cassette connects with the upstream or downstream sequence flanking \( \text{wecB} \), respectively. Approximately, a 286 bp DNA fragment from the 5' PCR reaction and a 583 bp DNA fragment from the 3' junction PCR reaction were observed, reflecting the predicted chromosomal organization in the mutant. As a control, the same PCR reactions were performed using \( E. \text{coli} \) Top10, which produced a DNA fragment corresponding to the expected size (1329 bp) when the external primers WecBFwdseq and WecBRevseq were used. No products for the 5' and 3' junction PCR reactions were observed, as expected. After confirmation
**Figure 3.13:** Generation of CWG1167, an *E. coli* Top10 ΔwecB mutant. **A)** Cartoon of the Lambda Red strategy used to delete *wecB* via insertion of a chloramphenicol-resistance cassette (*cat*), and subsequent removal of *cat* cassette using the flanking FRT sites with the helper plasmid pCP20. **B)** Diagnostic DNA fragments from colony PCR screens of a potential mutant with insertion of *cat* cassette. The “scar region” is the section of DNA left behind after removal of the *cat* gene. *E. coli* Top10 (Wt) was used as a control. The primers used to screen potential ΔwecB deletion mutants by colony PCR are indicated by arrows above the genomic DNA in **A).**
that the cat gene was successfully inserted into wecB, the antibiotic resistance cassette was removed using the plasmid pCP20, leaving behind a DNA “scar region”. Successful removal of cat was confirmed by colony PCR using the primers WecBFwdseq and WecBRevseq, which produced the predicted 277 bp DNA fragment (Figure 3.13B). The verified E. coli Top10 wecB deletion mutant strain was labeled as strain CWG1167.

3.5 Presence of MnaA required for optimal O:54 production in E. coli Top10

To investigate the role of MnaA in the synthase pathway, two constructs were generated: pBH9 containing DNA encoding MnaA with a C-terminal hexa-histidine tag to detect MnaA expression, and pWQ203 containing the complete wb*<sub>O:54</sub> biosynthetic gene cluster to simulate native expression of the synthase-dependent pathway enzymes in an E. coli strain. To verify no other potential sources of the donor nucleotide sugar UDP-ManNAc are available in the strain CWG1167 that could have influenced production of O:54 polysaccharide, WbbE and WbbF were expressed (from pBH13) in the ΔwecB deletion strain (Figure 3.14). As expected, no O:54 antigen was detected on either the silver stained SDS-PAGE gel or the anti-O:54 immunoblot (Figure 3.14). To rule out the possibility that the construct pBH13 was mutated, rendering either WbbE and/ or WbbF non-functional, the plasmid was isolated from a portion of the cells used to prepare the LPS sample and transformed into E. coli Top10 to confirm function. The plasmid pBH13, that originated from ΔwecB strain, conferred O:54 antigen production in E. coli Top10, confirming no mutations were present in this plasmid. Moreover, this result validates the observation that no other source of UDP-ManNAc donor sugars is present in the ΔwecB mutant that would influence O:54 LPS biosynthesis. When mnaA from pBH9, was co-expressed with wbbEF from pBH13 in the strain CWG1167 (ΔwecB), the quantity of O:54 antigen generated was equivalent to that seen in S. enterica sv. Borreze (Figure 3.14). Moreover, chain length
Figure 3.14: Expression of O:54 LPS in CWG1167 (ΔwecB). LPS samples from whole-cell lysates of either E. coli Top10 or CWG1167 containing pBH9 and pBH13 were analyzed by A) silver stained SDS-PAGE gel and B) corresponding Western immunoblot. Induction of the protein expression of MnaA from pBH9 and WbbEF from pBH13 was achieved with the addition of 0.02% arabinose and 2.5 ng/mL anhydrotetracycline, respectively. The presence of MnaA and/or WbbE + WbbF is indicated above the appropriate lanes. The asterisk denotes an LPS sample prepared from E. coli Top10 containing pBH13 which was isolated from CWG1167 to confirm the absence of inactivating mutation. S. enterica sv. Borreze was included as a control.
distribution of the O:54 antigen was similar in these backgrounds. The observation of native production levels of O:54 LPS in *E. coli* Top10 rules out the possibility that *E. coli* K-12 strains are unable to cope with the O:54 production system, and establishes that optimal O:54 biosynthesis is achieved only in the presence of MnaA.

The next question was whether the combined activity of WecB with MnaA would enhance O:54 LPS production beyond the native expression level observed in *Salmonella* serovars. To test this, the complete *wb*<sub>O:54</sub> operon was expressed from the construct pWQ203 in the strain *E. coli* Top10 in the presence of endogenously expressed WecB and production of O:54 antigen was examined. The presence of a chromosomal *wecB* gene had no effect on either the amount or the average chain length of the O:54 antigen (Figure 3.15).

O:54 LPS production from *E. coli* Top10 containing the complete *wb*<sub>O:54</sub> gene cluster (pWQ203) was also compared to O:54 LPS produced from the expression of *wb*<sub>O:54</sub> operon containing a disruption within *mnaA* (pWQ800). As expected, the LPS profile derived from the expression of pWQ800 was considerably different than O:54 LPS profile observed from pWQ203 (Figure 3.15). An overall shift was observed in the O:54 LPS profile generated from the expression of pWQ800 compared to pWQ203, including a lower average chain length, as well as, a lower maximum detectable chain length. Overall, the experimental results from Figure 3.14 and 3.15 establish that the presence of MnaA does influence O:54 production and this function cannot be effectively replaced or enhanced by WecB.

**3.6 Effect of MnaA, WbbE and WbbF protein levels on O:54 LPS production**

To further understand the involvement of MnaA in O:54 biosynthesis, a range of concentrations of MnaA were tested to see if the amount of the protein MnaA influences the O:54 LPS profile. To achieve this, the expression of the protein MnaA-His<sub>6</sub> from *mnaA* encoded
Figure 3.15: Examination of effect of different sources of UDP-GlcNAc 2-epimerase on O:54 antigen production. Whole cell lysates of *E. coli* Top10 containing either pWQ800 (*wb*\*O:54 *mnaA*), pWQ203 (*wb*\*O:54), or pBH13 (WbbEF) were examined, along with CWG1167 containing pWQ203. A) silver stained SDS-PAGE gel and B) corresponding α-O:54 Western immunoblot. All samples containing either pWQ203 or pBH13 were induced with 2.5 ng/mL anhydrotetracycline.
in pBH9 was incrementally titrated using increasing concentrations of arabinose inducer in CWG1167 (ΔwecB). WbbEF was provided by pBH13 and the induction of wbbEF expression was kept constant with 2.5 ng/mL AhT. LPS and protein samples were prepared for all samples from whole-cell lysates, and the amount of MnaA and the quantity of O:54 LPS were examined. As expected, the amount of a protein band with the apparent molecular weight of MnaA (predicted to be 43.3 kDa) increased with increasing concentrations of arabinose (Figure 3.16A). Under repressed conditions (0.4% glucose) or in the absence of the inducer, arabinose, no MnaA was detected. In all samples expressing MnaA, an additional protein was detected with a calculated apparent molecular weight of 23.5 kDa (determined from calculated R<sub>f</sub> value). The intensity of the low molecular weight protein band changed in parallel with the high molecular weight protein band. Thus, it was presumed the low molecular weight protein represented a truncated version of MnaA, retaining the hexa-histidine tag. Any other bands were of low intensity and were also visible in the negative control sample expressing only the vector pBAD24. These were assumed to reflect promiscuous binding of the anti-His<sub>5</sub> antibody.

Production of O:54 LPS increased as the amount of MnaA protein increased (Figure 3.16A). However, the quantity of O:54 antigen did not achieve the levels produced in cells containing pWQ203 (Figure 3.16B, C). Remarkably, where detectable O:54 antigen was evident, the chain length distribution of the polysaccharide remained quite uniform at all concentrations of inducer. Only a minor amount of O:54 antigen from the sample induced with 0.0005% arabinose was detectable on the Western immunoblot, signifying the amount of O:54 antigen produced was near the limit of detection. No O:54 LPS could be detected in samples from cells that lacked the inducer, arabinose, or with the addition of 0.4% glucose, which was expected given the absence of detectable MnaA protein in these samples.
Figure 3.16: Titration of MnaA protein expression level affects quantity of O:54 LPS. LPS and protein samples were prepared from whole-cell lysates of CWG1167 containing pBH9 and pBH13. MnaA expression from pBH9 was induced using a range of arabinose concentrations (as indicated), but *wbbEF* expression from pBH13 was kept constant by inducing with 2.5 ng/mL of anhydrotetracycline. A) Protein level of MnaA-His\textsubscript{6} detected by Western immunoblot using anti-His\textsubscript{6} serum. O:54 LPS in the corresponding whole-cell lysates were examined by B) silver stained SDS-PAGE gel and C) anti-O:54 Western immunoblot.
Before investigating the effect of WbbE and WbbF on O:54 antigen production, the protein preparation method for WbbF was examined. After resuspending pelleted cells in a lysis buffer that includes SDS, the samples are typically boiled to solubilize the cells. However, during initial protein experiments, only degraded WbbF products were observed when WbbF protein samples were boiled, even when optimized to express in *E. coli* strains (Figure 3.17). This phenomenon has been observed with other membrane proteins, where the bulk of the protein aggregates in boiled samples (Hennessey Jr and Scarborough, 1989; Nakamura and Mizushima, 1976). To prevent these problems, cells were resuspended in sample buffer, and warmed at 45°C for 30 min. A 26 G1/2 needle was then used to physically shear cells to disrupt DNA and further homogenize cells. Full length WbbF protein was observed when samples were treated in this manner (Figure 3.17).

To determine whether the amount of WbbE and WbbF also influences O:54 antigen production, the expression of *wbbEF* from the pBH13 construct was titrated using increasing concentrations of AhT in the strain CWG1167. The amount of the protein MnaA-His₆ expressed from *mnaA* in pBH9 was kept constant in each sample as shown in Figure 3.18A. The molecular weight of WbbF is predicted to be 53.3 kDa, but the protein band is observed at approximately the 50 kDa marker in SDS-PAGE gels. Because the protein band for MnaA (predicted molecular weight of 43.3 kDa) is close to the protein band observed for WbbF in SDS-PAGE gels, protein samples were separated on an 18% SDS-PAGE gel to yield greater separation of WbbF and MnaA. As expected the expression level of WbbF increased with the inducer concentration. WbbF was undetectable at low inducer concentrations (0.1 ng/mL, 1 ng/mL). However, this may not necessarily signify that no protein was expressed, but rather the protein level was beyond the limits of detection by chemiluminescence using horseradish peroxidase.
Figure 3.17: WbbF protein is not detected in boiled samples. Overnight cultures of *E. coli* Top10 containing either pKB7bc or pKB101 were subcultured, grown until an OD\textsubscript{600} of 0.6 was reached and then protein expression was induced by 0.5 mM IPTG. Samples were grown for 2 hr after induction and then protein samples prepared from whole-cell lysates. During protein preparation, samples were subjected to either warming at 45°C then sheared with a needle or were boiled for 10 min after resuspending cells in sample buffer. WbbF was observed on a A) coomassie blue stained SDS-PAGE gel and B) anti-His\textsubscript{5} Western immunoblot. A protein sample was also prepared from *E. coli* Top10 cells as a negative control, and the relevant lane is labeled with a minus sign.
Figure 3.18: Titration of WbbE and WbbF protein expression levels affects O:54 production. LPS and protein samples were prepared from whole-cell lysates of CWG1167 containing pBH13 and pBH9. WbbEF expression from pBH13 was induced using a range of anhydrotetracycline concentrations (as indicated), but MnaA expression from pBH9 was kept constant by inducing with 0.02% arabinose. A) The amount WbbF-His\textsubscript{6} and MnaA-His\textsubscript{6} detected by Western immunoblot using anti-His\textsubscript{5} serum. O:54 LPS in the corresponding whole-cell lysates were examined by B) silver stained SDS-PAGE gel and C) corresponding anti-O:54 Western immunoblot.
The amount and chain length distribution of O:54 LPS in the corresponding whole-cell lysates were examined. The quantity of O:54 LPS was significantly reduced in samples from cells with the lowest amount of AhT (0.1 ng/mL), and a similar profile was observed in uninduced cells (Figure 3.18B, C). Only high modal chain O:54 LPS can be detected in these samples, suggesting a preference for WbbF to generate polysaccharide with the largest chain length. The quantity of long chain-length O:54 LPS increased significantly in cells induced with higher concentrations of inducer. Moreover, the amount of lower length O:54 chains also increased. The O:54 LPS profile of the samples with high concentrations of AhT (2.5-25 ng/mL) appeared similar in both quantity and in O:54 chain length.

3.7 Deletion of waaL from E. coli Top10

The O-Ag polymer is assembled in a separate pathway from the lipid A-core portion of LPS and the two parts are subsequently joined together by the ligase, WaaL, to form the complete LPS molecule (Han et al., 2012; Kaniuk et al., 2002; Mulford and Osborn, 1983). Salmonella species are known to produce O-Ag with heterogeneous chain lengths from chromosomally-encoded O-Ag biosynthesis genes (Palva and Mäkelä, 1980). However, some E. coli isolates can also produce and export O-Ag that lacks attachment to the lipid A-core. This form is called an O-Ag capsule (Goldman et al., 1982; Peleg et al., 2005). In this case, the glycan is exported to the cell surface via the capsule export pathway (Whitfield, 2006). The possibility that the O:54 polysaccharide is exported in the absence of ligation to the lipid A-core was considered. To determine the extent of O:54 antigen attachment to lipid A-core, the chromosomally-encoded waaL was deleted via insertion of a cm-resistance cassette using again the Lambda Red system (Datsenko and Wanner, 2000). The cm-resistance cassette was inserted
into \textit{waaL} on the chromosome, and subsequently removed using the helper plasmid pCP20. A schematic diagram of the process is shown in Figure 3.19A.

Potential mutant colonies were initially screened via colony PCR. The primers WaaLseqFwd and WaaLseqRev, specific to the upstream and downstream sequence regions just outside of the target gene \textit{waaL}, were used to amplify the Cm-resistance cassette. The colony PCR reaction generated a DNA fragment approximately the size of 1266 bp, as predicted (Figure 3.19B). The same colony PCR reaction was performed with wildtype \textit{E. coli} Top10 colonies to amplify \textit{waaL}, producing a DNA fragment that is approximately the expected 1493 bp size. Two additional colony PCR reactions were done to provide further validation of the successful replacement of \textit{waaL} with \textit{cat}. The same primers, WaaLseqFwd and WaaLseqRev were used with the internal primers Catint5′Rev and Catint3′Fwd, which are specific to \textit{cat} to amplify the 5′ and 3′ junction regions between the upstream and downstream flanking sequences and the \textit{cat} gene. Positive reactions generated DNA bands consistent with the expected sizes of 267 bp DNA fragment for the 5′ junction and a 592 bp DNA fragment for the 3′ junction. No DNA fragments were observed when the colony PCR reactions were carried out with \textit{E. coli} Top10 DNA, as expected. Afterwards, the cm-resistance cassette was removed using the helper plasmid pCP20 and a "scar region" of the \textit{cat} gene was left behind. Removal of the cm-resistance cassette was confirmed by colony PCR in which the scar region was amplified using the WaaLseqFwd and WaaLseqRev primers. This reaction generated a band at approximately 312 bp, which was the expected size. The Δ\textit{waaL} deletion strain was named CWG1168.

\textbf{3.8 Production of O:54 in Δ\textit{waaL} deletion mutant causes a growth defect that is recovered via endogenous inhibition of O:54 production}

Transformation of the plasmid pWQ800 into the Δ\textit{waaL} deletion strain, CWG1168 proved difficult. Each attempt generated either only a few (1-3) colonies, or no transformants at
Figure 3.19: Production of CWG1168, an *E. coli* Top10 Δ*waal* deletion mutant. A) Cartoon of the Lambda Red system strategy used to delete *waal* from *E. coli* Top10 via insertion of a chloramphenicol-resistance cassette (*cat*), and subsequent removal of *cat* cassette using the flanking FRT sites with the helper plasmid pCP20. B) Diagnostic DNA fragments from colony PCR screens of a potential mutant with insertion of *cat* cassette. The “scar region” is the section of DNA left behind after removal of the *cat* gene. *E. coli* Top10 (Wt) was used as a control. The primers used to screen potential Δ*weCB* deletion mutants by colony PCR are indicated by arrows above the genomic DNA in A).
all. Test transformations were carried out using the vector pBAD24, which generated approximately 100-150 colonies each test transformation, indicating no issues with the electrocompetent cells. The plasmid, pWQ800, contains the promoter region of \( wb^{*}_{O:54} \), and so induction of the expression of the O:54 biosynthetic gene cluster is achieved endogenously. Early induction of the pathway could have stressed newly transformed cells, resulting in no recovered transformants. To test this, pWQ203, containing the \( wb^{*}_{O:54} \) operon behind the AhT inducible promoter was also transformed into \( \Delta waaL \) cells. The test transformation (empty plasmid) and the transformation of pWQ203 both produced over 100 colonies when the cells were recovered in the absence of inducer.

Whole-cell lysates were made from CWG1168 transformants containing pWQ800 or pWQ203 (induced with 2.5 ng/mL AhT) and examined for the presence of O:54 LPS. Cultures of these cells grew very slowly, such that neither sample reached an \( OD_{600} \) of greater than 0.02 after 4 hr of incubation. Recovery of growth was observed after 5 hr of incubation, so samples were left overnight and LPS samples were prepared the following day. All LPS samples prepared from CWG1168 whole cell lysates lacked O:54 LPS on the silver stained SDS-PAGE gel (Figure 3.20A), but a trace amount of O:54 antigen with a high modality (very long O-Ag) was detectable on the Western immunoblot using O:54 antiserum for the samples containing pWQ203 (Figure 3.20B). To test whether the absence of O:54 antigen was due to a defect in pWQ203 or pWQ800 in these transformants, the plasmids were isolated from a portion of the \( \Delta waaL \) cells used for the LPS preparations and transformed into \( E. \ coli \) Top10 cells to test for function. Only one pWQ203 sample (sample #1) recovered from CWG1168 retained function. The pWQ800 sample directed production of a trace amount of very long O:54 antigen in \( E. \ coli \) Top10. These
Figure 3.20: Production of O:54 in CWG1168 (ΔwaaL). Whole-cell lysates of CGW1168 containing pWQ203 or pWQ800 were prepared. From the CWG1168 transformants, the pWQ203 and pWQ800 constructs were isolated then transformed into E. coli Top10. LPS preparations were made to test for function of construct. A) silver stained SDS-PAGE gel and B) corresponding anti-O:54 Western immunoblot.
experiments indicate that growth of CWG1168 ΔwaaL is severely impaired when O:54 antigen is produced and cells only recover by mutations that turn-down or turn-off O:54 antigen synthesis.

Unfortunately, the results from this experiment do not address the original question of whether or not the entire O:54 polysaccharide produced in E. coli Top10 is ligated to the lipid A-core portion of LPS. To examine this issue, an agglutination test using anti-O:54 serum was performed to test for presence of O:54 on the external surface of the ΔwaaL cells after activating O:54 antigen production by induction of pWQ203 with AhT. The cell density was decreasing at this time due to the stress on the cells. No evidence of agglutination was observed, signifying no O:54 polysaccharide was exported to the external surface of the cell. The same agglutination test was performed on E. coli Top10 cells containing pWQ203 induced to express the O:54 biosynthetic gene cluster. These positive control cells did agglutinate over time, but the process was relatively slow compared to a comparable agglutination test performed on E. coli O9a cells using anti-O9a serum. Thus, the negative result with CWG1168 (pWQ203) could just reflect an inefficient agglutination reaction.

The growth-defect in CWG1168 with pWQ203 during O:54 antigen production was unanticipated. waaL deletions have been made in many bacteria including Salmonella and E. coli (Klena et al., 1992; Kong et al., 2011; Woodward et al., 2010) with no apparent defects, other than the lack of O-Ag substituted LPS. To investigate this peculiar growth pattern, a growth curve was generated. Subcultured samples of CWG1168 containing pWQ203 were divided into two equal aliquots once an OD$_{600}$ of 0.2 was reached. One of the cultures was then induced with 2.5 ng/mL of AhT to express the O:54 biosynthetic gene cluster, while the other culture remained uninduced. Growth of the sample without inducer was similar to the growth of CWG1168 ΔwaaL strain lacking pWQ203, as expected (Figure 3.21). The optical density
Figure 3.21: Growth curve of CWG1168 (ΔwaaL mutant) containing pWQ203 with, and without induction of O:54 biosynthesis. Subcultured samples of CWG1168 containing pWQ203 were grown until an OD$_{600}$ of 0.2 was attained, and then divided into two samples. One sample remained uninduced and the other sample received 2.5 ng/mL AhT for induction. OD$_{600}$ was measured every 0.5 hr for 5 hr.
increased exponentially for samples lacking O:54 LPS synthesis until the culture entered the stationary growth phase. One hour after activation of O:54 production, the OD$_{600}$ of the culture containing pWQ203 dropped dramatically, declining to an OD$_{600}$ of 0.12, 2.5 hr after induction. Beyond this point, the optical density of the culture began to increase slowly. After overnight growth, the OD of the culture reached an OD$_{600}$ of 1.6, typical of other stationary phase E. coli cultures. Samples were examined under the microscope during the growth experiment to determine whether the growth defect caused any visible defects to the cell shape or size. At 0 hr, cells in all samples appeared similar in size and were rod shaped, as expected (Figure 3.22).

After the sample was divided into two, there was a clear divergence in appearance of cells producing O:54 antigen. One hr after induction, when the growth-defect became apparent, the cells appeared shorter than those lacking O:54 antigen, and there was a tendency for cells to aggregate into small groups of 2 or 3 cells. When the cultures producing O:54 antigen began to recover growth at 4.5 hr after induction, a mix of elongated rod shaped cells, rounded cells and regular-length rod shaped cells was observed. There were no longer any visible aggregates, but debris was observed, presumably due to lysed cells.

Whole-cell lysates were prepared during the growth experiment to examine how O:54 production was affected during the experimental time course. As expected, no O:54 antigen was detected before induction of O:54 production (Figure 3.23). Although the CWG1168 cells were still growing at 0.5 hr after induction, significantly less O:54 antigen was produced relative to the profile in the ligation-proficient E. coli Top10 sample, and the average chain length was high. At 1 hr after induction, the quantity of O:54 antigen increased, but only higher molecular weight product was observed. The amount and size of O:54 antigen did not vary up to 4.5 hr after
Figure 3.22: Brightfield microscopy using a differential interference contrast filter of CWG1168 ($\Delta\text{waaL}$) cells before and during O:54 production. A CWG1168 culture containing pWQ203 was split and one portion was induced for O:54 LPS production. Cells were viewed under a 63× objective lens using a Zeiss Axiovert 200M microscope at 0 hr, 1 hr and 4.5 hr after induction. CWG1168 cells lacking pWQ203 were included as a control. Expected shape and size of cell indicated by a black circle outline. Unusual cells observed are indicated by a black triangle.
Figure 3.23: Production of O:54 over time in CWG1168 (Δwaal). Two samples of whole-cell lysates of CWG1168 containing pWQ203 from two different cultures were prepared at the noted time points and O:54 antigen was detected by A) silver stained SDS-PAGE gel and B) corresponding anti-O:54 Western immunoblot. A sample of whole-cell lysate of *E. coli* Top10 containing pWQ203 induced with 2.5 ng/mL AhT was also examined as a control.
induction. At 24 hours after induction, the quantity of the high modal length O:54 decreased significantly. Surprisingly, the core-lipid A portion of O:54 LPS on the silver stained gel appeared to increase from 1-24 hr after induction. This observation is perplexing since there is no precedent as to why the amount of lipid A-core LPS would increase.

Given the change in cell morphology, changes in O:54 LPS may be a result of an inaccurate OD\textsubscript{600} measurement. To address this, cells equivalent to an OD\textsubscript{600} of 1.0 from each sample (induced and uninduced) were serially diluted then plated to determine the number of CFU/mL. The uninduced samples had an average of CFU/mL of 4.0 \times 10^8 and the induced samples had an average CFU/mL of 2.5 \times 10^7. The number of colony forming units was significantly lower for the induced samples than the uninduced samples. Therefore, the OD\textsubscript{600} measurements for the induced samples are inaccurate and the amount of O:54 antigen in \textit{E. coli} Top10 cannot be compared directly to that in a ligase-proficient strain.

3.9 Development of a functional cysteine-free WbbF

The structure and function of WbbF and related synthases are poorly understood. In order to begin a detailed investigation of the enzymatic and export functions of WbbF, the topology must first be elucidated. This will provide insight into the membrane-spanning/membrane-associated regions and the potential organization of the functional domains. The strategy chosen to decipher the topology of WbbF is the cysteine-accessibility method, which requires the construction of WbbF derivatives with single Cys residues at precise reporter sites. WbbF contains seven Cys residues, which therefore must be removed before reporter-specific single Cys constructs can be made. In this method, confounding effects of Cys replacements on protein structure can be ruled out if those mutants retain function.
To replace the seven Cys residues within WbbF, single Cys to Ala substitutions were made in pWQ203 (\(wb^*_{O:54}\) operon) by site-directed mutagenesis to create the Cys → Ala changes (Figure 3.24). Initially, each Cys was individually substituted in order to determine if any Cys residues are essential for enzymatic activity of WbbF. O:54 LPS was detected in whole-cell lysates from cells containing constructs with any one of the seven Cys residues mutated in \(wbbF\) in the \(wb^*_{O:54}\) operon. Therefore, no native Cys residue was required for WbbF function (Figure 3.25).

All seven Cys residues were sequentially mutated in the same pWQ203 construct, starting at the N-terminus with Cys49Ala and progressing to the C-terminus, with the exception of C183 since function of WbbF with the C183A change was still being questioned at the time (see below). After each cycle of mutation, the ability of the construct to direct production of O:54 LPS was examined to ensure WbbF remained enzymatically active. WbbF was fully functional after the removal of six of the seven Cys residues, based on observation of similar quantity and size distribution of O:54 LPS prepared from whole-cell lysate containing each mutant construct (Figure 3.26). Though WbbF was found to be still functional after the removal C183, a small decrease in activity was evident by the reduced quantity of O:54 antigen produced. This was confirmed in a second, independent WbbF mutant with all seven Cys residues removed. Sequencing revealed no (additional) mutations in the \(wb^*_{O:54}\) operon (data not shown).

Nonetheless, a functional, cysteine-free WbbF was successfully generated, and can now be used to make the reporter-specific single Cys constructs required to map the topology of WbbF.

### 3.10 Preliminary protein experiments

Preliminary analysis was performed to determine the most suitable way of detecting WbbF protein expression for future experiments. WbbF protein expression from a high expression vector and low expression vectors was examined to compare how well WbbF is
Figure 3.24: Amino acid sequence of WbbF with cysteine residues highlighted in green and underlined. The signature DXD and QXXRW motifs are highlighted red and pink, respectively.
Figure 3.25: Analysis of WbbF Cys → Ala mutants. LPS samples were prepared from whole cell lysates of *E. coli* Top10 cells containing pWQ203, pWQ572 or pBH68, -69, -70, -71, -72, -73, -74 (containing one of the seven cysteines substituted to alanine). O:54 LPS was detected by A) silver stained SDS-PAGE gel and B) corresponding anti-O:54 Western immunoblot.
Figure 3.26: Analysis of function of WbbF Cys → Ala mutant after the sequential substitution of each cysteine residue to alanine. Whole cell lysates of *E. coli* Top10 cells containing pWQ572 and pWQ203 containing one, two, three, four, five, six and seven of the seven cysteines substituted to alanine (denoted from CWbbFA1-7) were prepared. O:54 LPS was detected by A) silver stained SDS-PAGE gel and B) corresponding anti-O:54 Western immunoblot.
detected under these conditions. The arabinose-inducible promoter from pBAD24 potentially yields high levels of protein, so the construct pKB101 containing the pBAD24 promoter was used for WbbF expression (Guzman et al., 1995). To achieve a low level of expression of WbbF, the constructs pBH6 and pBH13 containing the low expression P_\text{Tet} promoter were used (Hillen and Berens, 1994). As expected, the high expression plasmid, pKB101 resulted in considerably more detectable WbbF protein compared to plasmids, pBH6 and pBH13 (Figure 3.27). In the sample with a high level of WbbF expression, a less dense band was also observed around the 36 kDa marker. Because WbbF is a membrane protein, this band could represent full length WbbF with aberrant migration due to SDS binding (Rath and Deber, 2013; Rath et al., 2009; Reynolds and Tanford, 1970) or this band could represent degraded WbbF protein.

The subcellular location of WbbF was examined under conditions where expression was modulated (based on the amount of inducer) to determine how much protein is actually embedded in the IM. Cell fractions were separated by centrifugation. When WbbF was expressed at a high level from pKB101, protein was observed in both the pellet and supernatant samples taken after centrifugation at 4000 x g and 12 000 x g, as well as in the pellet sample taken after centrifugation at 100 000 x g which contains the membrane fraction of the cells (Figure 3.28). The low speed pellet contained cell debris, unbroken cells, and inclusion body contamination (Thein et al., 2010). Under low protein expression conditions, a small amount of WbbF was detected in the 12000 x g pellet sample, but the majority resided in the 100 000 x g pellet samples (Figure 3.29). Based on these results, it can be concluded WbbF is localized in the membrane fraction. However, WbbF can also be found in the other fractions when expressed at a high level. Therefore, a lower level of expression is required for cysteine-accessibility experiments.
Figure 3.27: Examination of WbbF expression from different plasmids. Overnight cultures of *E. coli* Top10 containing either pKB101, pBH6 or pBH13 were subcultured, grown until an OD$_{600}$ of 0.6 was reached and then protein expression was induced by 0.5 mM IPTG (pKB101) or 5 ng/mL AhT (pBH6 and pBH13). Samples were grown for 2 hr after induction and then protein samples were prepared from whole-cell lysates. A) coomassie blue stained SDS-PAGE gel and B) anti-His$_5$ Western immunoblot. A protein sample was also prepared from *E. coli* Top10 cells as a negative control, and is the relevant lane labeled with a minus sign.
Figure 3.28: Subcellular localization of highly expressed WbbF. Overnight cultures of *E. coli* Top10 containing pKB101 were subcultured and induced with 0.5 mM IPTG to express WbbF. Once an OD$_{600}$ of 0.6 was reached, cells were collected, resuspended in buffer, and lysed by sonication. Subcellular fractions were separated by three steps of centrifugation and resuspension cell pellet until membrane fraction was collected. A protein sample was taken at each step and then examined on a A) coomassie blue stained SDS-PAGE gel and B) an anti-His$_5$ Western immunoblot. The original cell-free lysate, supernatant (S) and pellet (P) samples from centrifugation at 4 000, 12 000 and 100 000 x g are indicated.
Figure 3.29: Subcellular localization WbbF under low-level expression conditions. Overnight cultures of *E. coli* Top10 containing pBH13 were subcultured and induced with 5 ng/mL AhT to express WbbF. Once an OD$_{600}$ of 0.6 was reached, cells were collected, resuspended in buffer, and lysed by sonication. Subcellular fractions were separated by three steps of centrifugation and resuspension cell pellet until membrane fraction was collected. A protein sample was taken at each step and then examined on a A) coomassie blue stained SDS-PAGE gel and B) an anti- His$_5$ Western immunoblot. The original cell-free lysate, supernatant (S) and pellet (P) samples from centrifugation at 4 000, 12 000 and 100 000 x g are indicated.
4.0 DISCUSSION

Difficulties encountered in visualizing O:54 LPS were problematic for interpreting experimental data, so optimization of the techniques used to detect O:54 LPS was the first objective of this thesis. An important technique developed for examining LPS is SDS-PAGE gel electrophoresis. The LPS profiles observed on an SDS-PAGE gel are complex since a single strain produces R-LPS and S-LPS with a mixture of O-Ag chain lengths (Raetz and Whitfield, 2002). SDS-PAGE gel separates the LPS mixture into a ladder-like banding pattern based on molecular weight (Jann et al., 1975). The fastest migrating band is R-LPS lacking the O-Ag portion, and S-LPS with the largest O-Ag attached migrates the slowest in the gel. Each consecutive band or “rung” in the ladder that runs slightly slower than the previous band represents LPS containing an extra O-Ag repeat unit (Palva and Mäkelä, 1980; Peterson and McGroarty, 1985). The ladder-like pattern observed from a single variant of LPS is reproducible, so this technique is advantageous for characterizing LPS (Hitchcock and Brown, 1983).

The development of the silver stain protocol by Tsai and Frasch in 1982 was crucial for LPS visualization since LPS molecules did not stain well using standard stains for carbohydrates (Tsai and Frasch, 1982; Hitchcock and Brown, 1983). The chemistry behind the silver staining technique, although not proven, is believed to work by first oxidizing the glycol chains to aldehydes. These then react with silver ions to deposit metallic silver allowing visualization of the LPS bands (Fomsgaard et al., 1990; Zhu et al., 2012). This sensitive staining method has worked for many LPS structures with varying structures (Hitchcock and Brown, 1983). Thus, it was surprising that the staining of O:54 LPS was significantly poorer than the positive controls, *E. coli* O8 and O9a LPS. All three O-Ags are linear homopolymers, but differ in type and linkage of the sugar residues in the repeat unit. The O8 and O9a O-Ags consist of repeat units of mannose sugar residues with different linkages (Stenutz et al., 2006), whereas the O:54 antigen
consists of ManNAc sugar residues (Keenleyside and Whitfield, 1994). The difference in staining sensitivity between *S. enterica* O:54 and *E. coli* O8 and O9a can likely be attributed to the differences in the chemical properties of the sugar residues in the O-Ags. In particular, the chemical reactions involved in staining O:54 silver may be influenced by the *N*-acetyl side group of the ManNAc sugar in the O:54 polysaccharide, although there is no evidence in the literature that specifies acetamidohexoses stain more poorly than hexoses. Previous research reported that the efficiency of silver staining of O:54 LPS from some *Salmonella* serovars containing pWQ799 or pWQ800 was weaker than others (Keenleyside and Whitfield, 1994). The poor staining was presumed to be due to a decreased amount of smooth LPS produced, suggesting a more sensitive staining protocol is necessary to increase staining intensity.

Some factors that affect the staining sensitivity of LPS can be easily overcome, like preventing the use of ammonium hydroxide whose concentration has been reduced by repeated exposure to air. This causes a brown precipitate to form in the staining solution, which significantly decreases the staining intensity of LPS (Hitchcock and Brown, 1983). Other factors contributing to poor LPS detection are more complex, and are not easily resolved with a simple adjustment to either the SDS-PAGE gel electrophoresis protocol or the silver staining protocol. Indeed, researchers have found that the use of one silver staining protocol versus another produces such variable results for the detection of LPS from different bacterial species that there is no definitive, optimal protocol (Fomsgaard et al., 1990; Guard-Petter et al., 1995; Tan and Grewal, 2002).

The modified silver staining technique by Fomsgaard and colleagues (Fomsgaard et al., 1990) proved to be the best method for detecting O:54 LPS. This method was specifically developed to increase the staining intensity of poorly stained LPS profiles. Moreover, the
Fomsgaard method is advantageous because it is faster and has simpler steps than the original silver staining method by Tsai and Frasch (1982). Fomsgaard and colleagues also recommended that Western immunoblotting using O-specific antisera should be performed for thorough characterization of the LPS profile since the researchers had previously observed certain bands present on the immunoblot that were not detected via silver staining (Fomsgaard et al., 1990; Fomsgaard et al., 1988). In this thesis, the visualization of O:54 LPS by silver stained SDS-PAGE gel and Western immunoblotting was improved, although unequivocal detection of the O:54 antigen profile was still best done using an anti-O:54 immunoblot.

Re-examination of the O:54 LPS profile from *S. enterica* sv. Borreze after optimization experiments revealed that O:54 LPS contained a heterogeneous mixture of O-Ag chain lengths. Similar observations have been made in other *Salmonella* serovars including the well-studied *S. enterica* sv. Typhimurium, as previously described. However, no distinct tight clusters of bands with measurable modal lengths were observed for O:54 antigen like other *Salmonella* serovars (Peterson and McGroarty, 1985; Goldman and Leive, 1980; Munford et al., 1980). The migration pattern of O:54 LPS produced in *S. enterica* sv. Borreze containing the naturally-occurring plasmid pWQ799 was consistently observed. Based on qualitative observations, a greater amount of higher molecular weight O:54 antigen was generated in *S. enterica* sv. Borreze, compared to the amount of lower molecular weight material.

The quantity and average chain length of O:54 antigen produced was higher when the overlapping genes *wbbE* and *wbbF* were expressed together from one inducible promoter, mimicking the expression of these proteins from *wbb*<sup>®</sup> <sub>O:54</sub> operon, compared to the O:54 LPS produced from the expression of *wbbE* and *wbbF* from separate plasmids. Because the LPS profile of a single variant on SDS-PAGE is reproducible, any deviations in the profile reflect
changes in expression of that LPS. Therefore, the separate expression of each GT was not optimal for the O:54 biosynthesis system to work efficiently. This could be due to poor expression of one enzyme, or could reflect an imbalance in the necessary stoichiometry of the enzymes. In the native pWQ799, each gene encoded in the \( wb^{*}_{O:54} \) operon overlaps with the preceding gene, which led Keenleyside and Whitfield (1996) to propose that translation of all three genes in the \( wb^{*}_{O:54} \) operon was coupled.

Although O:54 LPS production was improved by expression of \( wbbEF \) from a single promoter, the LPS profile in \( E. \) coli Top10 was still different than the native O:54 LPS profile in \( Salmonella \). Clues to the basis for different SDS-PAGE profiles of O:54 LPS in \( E. \) coli and \( Salmonella \) may be obtained by examining pWQ799. The O:54 antigen is the only described O-factor in \( Salmonella \) spp for which the entire biosynthetic gene cluster is contained on a naturally-occurring plasmid, although there are examples within the literature where one or more O-Ag biosynthesis genes are plasmid-borne. All of the other examples involve a Wzy-dependent O-Ag biosynthesis pathway. For example, expression of \( Shigella \) sonnei form I O-Ag requires a large virulence plasmid (Kopecko et al., 1980; Viret et al., 1993; Yoshida et al., 1991). A plasmid-encoded galactosyltransferase is required for the synthesis of the O:1 factor in \( Shigella dysenteriae \) type I. In these systems, the remaining functions are encoded by chromosomal \( wb^{*} \) genes (Haider et al., 1990; Sturm et al., 1986). Interestingly, the absence of the virulence plasmid from \( S. \) enterica sv Dublin, or the small pHS-2 plasmid from \( Shigella \) flexneri carrying the chain-length regulator \( (cld_{pHS-2}) \) gene does not prevent O-Ag biosynthesis, but only short O-Ag polymers are produced, affecting O-Ag function (Hong and Payne, 1997; Kawahara et al., 1989). The \( wb^{*}_{O:54} \) operon encoded on pWQ799 is an extreme example of plasmid-borne O-Ag biosynthesis genes. All genes encoded by pWQ799 were believed to be confined to those
absolutely necessary for the mobilization and transfer of the plasmid, and the biosynthesis of the O:54 antigen. These details prompted questions regarding the role of MnaA in the O:54 biosynthesis pathway.

Surprisingly, MnaA (with, and without WecB) restored O:54 LPS production to a similar level and quality as the O:54 LPS produced in S. enterica sv. Borreze. Thus, the function of MnaA cannot be effectively replaced by WecB, nor can the presence of WecB enhance the effect of MnaA on the production of O:54 antigen by increasing the size of the donor pool. MnaA may be the preferred component of the O:54 biosynthesis pathway, providing insight into why mnaA was retained in the naturally occurring plasmid-encoded wb*O:54 operon. The only described role for MnaA is to provide the donor, UDP-ManNAc. It is conceivable that MnaA may play a more involved role in O:54 biosynthesis, such as regulation of the activity of one or both of the GTs or influencing the processivity of WbbF. However, the most likely explanation is simply that MnaA provides a larger donor pool than WecB for the assembly of O:54 antigen when the enzymes are endogenously expressed during O:54 production. To unequivocally resolve this, the protein expression level of WecB would have to be titrated in the ΔwecB mutant (similar to the titration of MnaA experiment) in the presence of WbbE and WbbF, and O:54 LPS profiles examined.

MnaA is a uridine diphosphate-\(N\)-acetylglucosamine (UDP-GlcNAc) 2-epimerase. These enzymes convert UDP-GlcNAc to ManNAc intermediates (attached to UDP or free ManNAc where UDP is lost during the epimerization reaction), which are used to assemble a variety of cell surface glycoconjugates. Hydrolyzing UDP-GlcNAc 2-epimerases are involved in the sialic acid biosynthesis pathway and are found almost exclusively in mammals, with the exception of SiaA from Neisseria meningitidis and NeuC from E. coli K1 (Murkin et al., 2004; Tanner, 2005; Vann et al., 2004). Mammalian UDP-GlcNAc 2-epimerases possess irreversible epimerase
activity (thus the enzyme is not considered a ‘true’ epimerase) that produces free ManNAc sugars and UDP from UDP-GlcNAc. However, they also possess kinase activity that phosphorylates ManNAc to generate ManNAc-6-P, a precursor for sialic acid biosynthesis (Hinderlich et al., 1997; Stäsche et al., 1997). The N-terminal domain containing the epimerase activity functions independently of the C-terminal kinase domain. The N-terminal domain is therefore considered to be related evolutionarily and mechanistically to the monofunctional bacterial UDP-GlcNAc 2-epimerases (Effertz et al., 1999; Tanner, 2005). The precursor, ManNAc generated by the mammalian enzymes is the only precursor in the sialic acid biosynthesis pathway whose quantity affects the degree of cell surface sialylation (Keppler et al., 1999). Therefore, the mammalian UDP-GlcNAc 2-epimerase catalyzes the rate-limiting step in the sialic acid biosynthesis pathway and is considered a key regulator of this pathway.

Bacterial UDP-GlcNAc 2-epimerases catalyze the reversible, interconversion of UDP-GlcNAc and UDP-ManNAc, to be used as donors for the synthesis of an assortment of polysaccharides (Kawamura et al., 1979; Kawamura et al., 1978; Morgan et al., 1997; Sala et al., 1996). Many putative and confirmed bacterial UDP-GlcNAc 2-epimerases that have been identified via sequence homology including WecB involved in the assembly of the enterobacterial common antigen (ECA) found in all members of the Enterobacteriaceae family including E. coli and S. enterica (Kuhn et al., 1988; Lew et al., 1978).

In some circumstances, the use of sequence homology has led to the misidentification of candidate UDP-GlcNAc 2-epimerases. For instance, the proposed UDP-GlcNAc 2-epimerase, WbpI was not functionally interchangeable with WecB via cross-complementation experiments in Pseudomonas aeruginosa serotype O5 (wbpI::Gm) and E. coli 21566 (wecB::Tn10) (Burrows et al., 2000). A subsequent study revealed that WbpI was a UDP- α-D-ManNAc3NAcA 2-
epimerase, rather than a UDP-GlcNAc 2-epimerase (Westman et al., 2007). Previous work in our laboratory identified mnaA in the wb*O:54 operon and proposed it encodes a functional homolog of WecB. However, definitive proof of UDP-GlcNAc 2-epimerase activity was not provided (Keenleyside and Whitfield, 1996). The results presented in this thesis confirm MnaA is a UDP-GlcNAc 2-epimerase since the presence of mnaA was able to restore the loss of O:54 antigen expression in a ΔwecB host strain expressing WbbE and WbbF.

Recently, researchers were able to identify another UDP-GlcNAc 2-epimerase in Bacillus anthracis using the bacteriophage lysin, PlyG, which cleaves peptidoglycan-linked cell wall glycopolymer (Schuch et al., 2013). A receptor in the cell wall glycopolymer that PlyG binds, and its cognate biosynthetic pathway were identified in this study, which led to the identification of a UDP-GlcNAc 2-epimerase that was determined to be required in the biosynthesis pathway and for cell growth of B. anthracis. Because of this, a potential antimicrobial inhibitor known as epimerox was designed and subsequently shown to be a potent inhibitor of the UDP-GlcNAc 2-epimerase in B. anthracis. The bacteria did not develop resistance to epimerox over time, and allowed 100% survival of mice 14 days post-infection of B. anthracis. Epimerox is therefore a viable lead for drug discovery.

An X-ray crystallographic structure of the E. coli WecB with bound UDP was solved at 2.5 Å resolution. The structure revealed a homodimer, with each monomer containing a Rossman fold, linked by α-helices (Campbell et al., 2000). The crystal structure of the UDP-GlcNAc 2-epimerase from B. anthracis was also solved with UDP and UDP-GlcNAc present in the active site (Velloso et al., 2008). This structure resembled the E. coli crystal structure. An interdomain rotation of each monomer was identified and found to allosterically regulate enzymatic activity by “opening” or “closing” the active site cleft (Campbell et al., 2000). Velloso and colleagues
(2008) determined that the binding of UDP-GlcNAc induces major, localized conformational changes of the protein, directing it into a catalytically active form. Moreover, in the absence of UDP-GlcNAc, most UDP-ManNAc did not epimerize in the reverse reaction to form UDP-GlcNAc, thus UDP-GlcNAc is considered the allosteric regulator of bacterial UDP-GlcNAc 2-epimerases (Kawamura et al., 1979; Morgan et al., 1997; Samuel and Tanner, 2004).

Mechanistic studies of the bacterial epimerases exposed a unique sugar nucleotide epimerase reaction mechanism (Morgan et al., 1997; Sala et al., 1996). The reaction scheme involves an anti-elimination of UDP from UDP-GlcNAc to form the intermediate 2-acetamidoglucal, as detected in vitro (Vann et al., 2004). A subsequent syn addition of UDP generates the UDP-ManNAc product (Morgan et al., 1997). The reaction mechanism occurs via an oxocarbenium ion-like transition state. Three residues (Asp95, Glu117 and Glu131) in WecB from E. coli were determined to play key roles in catalysis and/or regulation (Samuel and Tanner, 2004). His213 was determined to be important for UDP binding and possibly catalysis and/or regulation. Lys15 was also believed to be important for binding.

To investigate why MnaA, but not WecB is optimal for O:54 antigen production, their amino acid sequences were compared. Using the Clustal Omega online alignment tool, MnaA was found to be 53% identical and 15% similar to WecB (Figure 4.1). The significant portion of identical and similar regions in the amino acid sequences of these two enzymes is indicative of structurally or functionally conserved regions. The three residues important for catalysis and/or regulation previously identified in WecB (Samuel and Tanner, 2004) are conserved in MnaA, and the His and Lys residues that were associated with binding. The alignment offers no insight into the functional or structural difference between MnaA and WecB that would explain why there is a difference in O:54 antigen production depending on which enzyme is present. A
Figure 4.1: Clustal Omega amino acid sequence alignments of WecB from *E. coli* and MnaA from pWQ799 (Sievers et al., 2011; Goujon et al., 2010). The black square boxes surround residues that are important for catalytic activity, regulation or binding. The webserver program marks a conserved amino acid residue in all sequences with an asterisk (*). Conservation of residues with strongly similar properties is marked with a colon (:) and weakly similar properties conserved are marked with a period (.) underneath. Colour of residue indicates the physiochemical property of that amino acid. Red designates small and hydrophobic residues. Blue, magenta, and grey designate acidic, basic, unusual amino acid residues, respectively. Green designates hydroxyl, and sulfhydryl.
detailed analysis of the kinetic parameters for these enzymes may highlight differences, which could result from even just a single amino acid change.

WecB, but not MnaA from \( wb^{*\_0.54} \), was able to restore capsule synthesis in a strain with a partial deletion of \( neuC \) (Vann et al., 2004). This was surprising, since MnaA shares high sequence similarity to NeuC. The reason why MnaA was unable to complement a \( \Delta neuC \) defect was not resolved. However, they did hypothesize that this result occurred because NeuC was shown to cleave UDP-GlcNAc forming ManNAc via a 2-acetamidoglucal intermediate during catalysis. This reasoning does not make sense since the catalytic reaction mechanism for NeuC involves hydrolysis of UDP from GlcNAc, producing free ManNAc instead of UDP-ManNAc, whereas, WecB and MnaA do not hydrolyze the UDP from GlcNAc.

*Staphylococcus aureus* serotype CP5 also encodes a functionally redundant homolog of the UDP-GlcNAc 2- epimerase, Cap5P, called Mna\( \text{SA}_S \) (Kiser et al., 1999; Keenleyside and Whitfield, 1996). Mna\( \text{SA}_S \) shares 61% amino acid sequence identity with Cap5P, and is considered functionally redundant since Mna\( \text{SA}_S \) and Cap5P were able to restore ECA biosynthesis when expressed in an *E. coli* \( \text{wecB}::\text{Tn}_{10} \) strain (Kiser et al., 1999). The gene encoding Mna\( \text{SA}_S \) was also identified in *S. aureus* serotype CP8 by Kiser and colleagues (1999), which explained why the mutation of \( \text{cap8P} \) (encodes a homolog of Cap5P) in a previous study did not abrogate CP8 production (Sau et al., 1997b). A functional role has not been determined for Mna\( \text{SA}_S \), although it was postulated to be involved in teichoic acid biosynthesis like Mna\( \text{BS} \) from *Bacillus subtilis* (Kiser et al., 1999; Soldo et al., 2002). The \( \text{cap5} \) and \( \text{cap8} \) biosynthetic gene clusters encode common genes flanking a central serotype-specific gene region (Sau et al., 1997a). It is likely mna\( \text{SA}_S \) was already encoded in the chromosome of *S. auerus* as a housekeeping gene for teichoic acid biosynthesis (required for viability) when the \( \text{cap5} \) and \( \text{cap8} \)
biosynthetic gene clusters were introduced into the chromosome. This event would have resulted in two genes in the chromosome coincidentally encoding functionally redundant proteins. Incidences of functional redundancy have been observed for other components in the capsule biosynthetic gene cluster in *S. aureus* serotype CP8 (Sau et al., 1997b). A related scenario is reiterated in the case of O:54 biosynthesis with *wecB* and *mnaA*, and so it was likely *mnaA* was retained on pWQ799 as a way to ensure expression the O:54 factor in any strain pWQ799 is transmitted to, and the presence of the *wecB* was just coincidental.

The production of a reduced amount of O:54 polymer from the expression of WbbE and WbbF in *E. coli* Top10 could simply be attributed to a shortage of donor sugar nucleotides. Because WecB is involved in the ECA biosynthesis pathway, it is possible that WecB only generates the necessary amount of UDP-ManNAc sugars to supplement the ECA pathway. Consequently, the demand to provide UDP-ManNAc could have been too great for WecB with both O:54 biosynthesis and ECA pathways activated, resulting in a decreased donor pool for O:54 biosynthesis. In order to determine whether donor availability explains the altered quantity and/or the average chain length of the O:54 polymer, the expression of MnaA, as well as WbbE and WbbF were examined. Titration of the expression levels of WbbE, WbbF and MnaA produced some interesting and unexpected results. Increasing the protein expression levels of MnaA resulted in a marked increase in O:54 antigen production, but no change in the average chain length distribution in the O:54 LPS profile. This supports the idea that the quantity of the donor sugar nucleotide pool determines the intensity of the SDS-PAGE profiles.

Increasing the amount of *wbbEF* by adding inducer at concentrations ≥2.5 ng/mL AhT did not alter the O:54 LPS profile. In these samples, the LPS profile was similar to *S. enterica* sv, Borreze. However, below this induction concentration, there was a drastic decline in the quantity
of lower chain-length O:54 antigen, until none could be detected. In contrast, only a moderate
decrease was observed in the quantity of higher modal chain length O:54 antigen. WbbE and
WbbF are membrane proteins and membrane availability may limit the amount of integrated, and
thus functional, protein (Gubellini et al., 2011; Wagner et al., 2007; Wagner et al., 2006).
Therefore, further increases beyond 2.5 ng/mL inducer may not increase the number of
functional proteins. In addition, the essential und-P acceptor may be limiting at higher
concentrations of WbbEF. What is not clear from these results is why the quantity of lower
modal chain length O:54 antigen was more sensitive to changes in inducer. Alterations of the
average modal length distribution suggest some additional (unknown) factor has altered the chain
extension/termination in the system.

Currently, there is no known or apparent mechanism to explain how length of the
polymer is regulated, or how chain elongation is terminated in the synthase pathway. The other
two major O-Ag synthesis pathways, the ABC transporter-dependent pathway and the Wzy-
dependent pathway, utilize specific proteins to determine the chain length of the O-Ags. For the
ABC transporter-dependent pathway in E. coli O9a, chain termination occurs via the transfer of a
methyl or a methyl-phosphate group to the non-reducing terminal of the growing polymer by the
enzyme WbdD (Clarke et al., 2004; Clarke et al., 2011). Overexpression of WbdD causes
premature chain termination, resulting in a lower average modal chain length distribution. In the
Klebsiella pneumoniae O2a ABC-transporter pathway, chain termination of O2a does not occur
via the addition of a capping residue (Kos et al., 2009). Chain elongation and export of the
polysaccharide must be coupled for the assembly of the O2a polysaccharide. Furthermore, the
chain length of K. pneumoniae O2a is dependent on the stoichiometry between the ABC-
transporter and the galactosyltransferases involved with O2a assembly. Greenfield and Whitfield
(2012) have suggested chain termination could occur when the ABC-transporter out-competes the galactosyltransferases for the nascent O2a polymer substrate (Greenfield and Whitfield, 2012). What is noticeably different about the *E. coli* O9a ABC-transporter dependent pathway compared to the O2a system and the O:54 synthase pathway is that proper termination of the growing O-Ag polymer is required for export of the nascent chain (Clarke et al., 2004; Cuthbertson et al., 2005; Kos et al., 2009).

In the Wzy-dependent pathway, Wzz homologs are responsible for the regulation of the modal chain length distribution of the O-Ag polymer via an as yet undetermined mechanism (Morona et al., 2009). Previously, it was believed that regulation of the modal length was determined by the oligomerization state of the Wzz proteins (Tocilj et al., 2008). However, recent research suggests that the Wzz proteins likely act as a scaffold upon which O-Ag is assembled since the Wzz proteins were shown to bind the growing polymer to stabilize it for continued polymerization (Kalynych et al., 2012). It was suggested once the polymer reached a specific length, the Wzz protein can no longer bind the polymer, and the O-Ag is released. Researchers believe that the modal length distribution depends primarily on the surface-exposed amino acids in specific regions, which is supported by the identification of a conserved residue and certain regions of the Wzz protein that are critical for modal chain length modulation (Kalynych et al., 2011; Kalynych et al., 2012; Kintz and Goldberg, 2011). Unlike the *E. coli* O9a ABC transporter- and the Wzy-dependent pathways, the O:54 biosynthesis pathway does not seem to contain a specific enzyme that is responsible for regulating the chain length. Therefore, the chain length must be regulated by the enzymes already involved in the O:54 biosynthesis pathway, possibly in a manner similar to the *K. pneumoniae* O2a system.
The mechanism(s) of chain regulation of other enzymes in the HasA sub-family of GTs is
not well understood. Thus far, only the regulation mechanism of HA assembled by Streptococcus
zooepidemicus HAS (szHAS) and type 3 capsule assembled by the Cps3S from S. pneumoniae
has been determined. In both cases, chain length of the polymer is dependent on the donor pool
size and the ratio of each donor in that pool (Chen et al., 2009; Forsee et al., 2009; Ventura et al.,
2006). Chain termination of the type 3 capsule occurs via ejection of the polymer by an abortive
translocation process caused by the insufficient supply of one of the donor sugars, UDP-GlcUA
(Forsee et al., 2000; Yother, 2011). Because synthesis of a new capsule chain depends on
termination of the previous chain, the donor concentration also likely controls the number of
chains produced.

The data for chain termination by HAS synthesis is not substantial, but preliminary
results suggest chain regulation is similar to Cps3S. In one study, researchers found that the
overexpression of the genes encoding the enzymes responsible for the biosynthesis of UDP-
GlcNAc (a precursor for HA), resulted in an increase in the molecular weight of HA product
(Chen et al., 2009). Conversely, the same study discovered that overexpression of the genes
encoding enzymes responsible for the production of UDP-GlcUA (the other precursor), resulted
in a decrease in the molecular weight of the product. Unlike the type 3 capsule, the yield of HA
was not influenced by the donor pool. The authors suggested that termination may occur when
the active site for the one donor is blocked for too long by the other donor sugar nucleotide,
although it is still not clear if there are two separate binding sites for the donor sugar nucleotides.
Modulation of polymer length by a lingering donor would likely be influenced by the relative
levels of each donor and their binding affinities for the active site in HAS. Thus, the molecular
weight of the HA polymer would be determined through the competition between the two donor
sugars. A similar effect has also been observed for spHAS, and the Class II pmHAS, where HA possessing a narrow range of molecular weight was produced when the molar ratio of the precursor sugars and the acceptor molecules was controlled (Jing and DeAngelis, 2000; Tlapak-Simmons et al., 1999).

Since a single donor is used for O:54 biosynthesis, modulating chain termination by the competition between donor sugars can be ruled out. Based on data obtained in this thesis, it is apparent that the UDP-ManNAc pool controls the number of O:54 polysaccharide chains produced (based on SDS-PAGE staining intensity), similar to the type 3 capsule, but does not influence modality of the polymer. Therefore, the control of O:54 polymer chain length does not strictly follow any of the established models, suggesting a novel mechanism. One possibility to consider is that a region within WbbF itself modulates the GT activity of WbbF. For example, the affinity of the polymer binding region could regulate the chain length of O:54. The precise mechanism requires further investigation.

When the O:54 biosynthesis pathway was induced in strain CWG1168 (ΔwaaL), only a small quantity of very long O-Ag was observed. This material was not believed to be surface-located since the cells were not agglutinated by anti-O:54 serum. However, the positive control (E. coli Top10 cells expressing O:54 antigen on the cell surface) agglutinated slowly, so further investigation using immunofluorescence microscopy is required for an unequivocal conclusion.

The growth defect phenotype of the ΔwaaL mutant strain during O:54 antigen production was unexpected since there have been no reports of similar growth impairment during the production of other O-Ags in ΔwaaL mutants. Why this growth defect occurred for O:54 antigen production and not for other O-Ags is unknown. WaaL is an integral membrane protein found in the IM of Gram-negative bacteria and is the only enzyme required for the ligation of O-Ag to the
lipid A core portion to form smooth LPS. Definitive demonstration of this activity was only recently achieved by *in vitro* analysis (Abeyrathne and Lam, 2007; Gebhart et al., 2012; Han et al., 2012). WaaL has a very relaxed specificity for the Und-PP linked O-Ag, but is specific for the lipid A-core acceptor structure (Han et al., 2012; Kaniuk et al., 2002). Topology mapping and site-directed mutagenesis of different WaaL homologues revealed a large periplasmic loop with conserved residues. This region was implicated as the catalytic domain since it contains an H/RX_{10}G motif found also in Wzy, another enzyme that recognizes Und-PP-linked glycan (Han et al., 2012; Islam et al., 2010; Pérez et al., 2008; Schild et al., 2005). Deletion of WaaL in *S. enterica* sv Typhimurium causes a stable build-up of Und-PP-O-Ag in the periplasm, limited by the low concentration of lipid carrier Und-P that is unable to be recycled to generate more O-Ag chains (Mulford and Osborn, 1983). The low quantity of O:54 antigen observed indicates that the limited amount of Und-P lipid carrier likely restricted the amount of O:54 antigen generated in this case as well, making the apparent toxicity even more surprising.

The observation that cells lacking WaaL produced only very long O:54 antigen was unexpected and intriguing. The complete absence of lower modal length O:54 antigen in CWG1168 suggests WaaL may have a role in regulating the chain length. Since WbbF is believed to simultaneously synthesize O:54 antigen while exporting the nascent polymer to the cell surface, the polymer length may depend on the how efficiently WaaL recognizes, and binds the Und-PP linked O:54 polymer, before ligating it to the lipid A-core acceptor. Therefore, the length of each O:54 chain may depend on the kinetics of both the GT and exporter activities of WbbF, as well as the kinetics of WaaL. Even the availability of the lipid A-core acceptor may come into play. To decipher if this is a possibility, the O:54 antigen profile would need to be
compared under conditions where each of these factors was varied. This is currently not possible given the limited understanding of the mechanism of each process.

The construction of a functional cysteine-free WbbF was successfully achieved. Although the function of WbbF marginally decreased with the removal of all seven Cys residues, this does not compromise experiments designed to establish the topology map of WbbF via cysteine-accessibility experiments, performed on single Cys WbbF mutants. Currently in the literature, there are only two examples of important Cys residues in the HasA family of enzymes, which are found in the BcsA and SeHAS enzymes. For BcsA, the recent crystallographic snapshot illuminated a FFCGS motif that interacts with the terminal disaccharide of the product through Cys and Gly. However, neither residue was essential for function (Morgan et al., 2013). Four generally conserved Cys residues were identified in SeHAS. These are found clustered close together near the UDP-sugar binding site on the cytoplasmic side of the IM (Kumari et al., 2002; Kumari and Weigel, 2005). None of the four Cys residues form disulfide bridges, and the Cys-null mutant SeHAS is still active, although activity was significantly reduced (Kumari et al., 2002). In the amino acid sequence alignment (Introduction Figure 1.10), none of the four Cys residues of SeHAS aligned with any of the seven WbbF Cys residues, and WbbF does not contain the FFCGS motif observed in BcsA.

Now that a functional, Cys-free WbbF has been developed, single Cys mutants can be generated and cysteine-accessibility experiments can be executed using these mutants. Preliminary *in silico* prediction models of the topology of WbbF were generated by three different web-server programs, yielding variable results (Figure 4.3 A, B, C). Two of the *in silico* models (Figure 4.3 A, B) predict very similar topologies with five transmembrane segments, a
**Figure 4.2:** *In silico* predictions of the membrane topology of WbbF. A) TMHMM web server prediction outcome using a hidden Markov model approach (Krogh et al., 2001). B) SVMtop web server prediction result using support vector machines (SVM) in a hierarchical structure. Helix prediction is performed in the first stage, followed by topology prediction in the second (Lo et al., 2007). C) HMMTOP web server prediction result using an altered hidden Markov model approach (Tusnády and Simon, 1998, 2001). D) original WbbF topology prediction proposed using hydropathy analysis (Keenleyside and Whitfield, 1996). The pink triangle designates the PhoA fusion made in order to confirm presence of a periplasmic loop.
large cytoplasmic loop near the N-terminus, and a cytoplasmic C-terminal tail. In contrast, the third model (HMMTOP) predicts six transmembrane helices with a large periplasmic loop near the N-terminus. These models, along with the hydropathy plot model generated by Keenleyside and Whitfield (1996) were assessed for potentially informative places to reintroduce (individually) Cys residues for the topology experiments.

During the construction of a Cys-free WbbF, Cys residues were sequentially substituted ending with residue 183, and so one single Cys reporter is already available. This residue is predicted to reside in the large cytoplasmic loop in three out of the four prediction models (Figure 4.3 A, B, D), which may contain the glycosyltransferase activity. This prediction is based on the location of the archetypal DXD motif involved in enzymatic activity within many GTs (Saxena et al., 1995). MnaA and WecB are soluble proteins in the cytoplasm, and so the donor sugar nucleotides are located in the cytoplasm as well. Therefore, the DXD motif must be located in the cytoplasm for GT function. There is no precedent for the donor sugar nucleotides to move into the periplasm, and so the periplasmic location of the large loop containing this motif predicted by the fourth prediction program (HMMTOP) is inconceivable. The other locations recommended for single Cys substitutions are based on the predicted locations of the other, smaller loops in the cytoplasm and periplasm. This strategy will help resolve the number and location of the transmembrane segments within WbbF. The recommended residues to change are I8, T242, V312, K409 and V447. Introducing another single Cys mutant at residue 368 would help confirm the previously established periplasmic location as determined by a PhoA-fusion (Keenleyside and Whitfield, 1996). If any of these substitutions render WbbF non-functional, other residues near the suggested sites should be attempted.
Overall, the research presented here has established that MnaA is required for optimal O:54 antigen production, presumably by increasing the amount of donor sugar nucleotides. The availability of the precursor had no significant effect on the chain length distribution. Influence on chain length distribution, however, was observed at low expression levels of WbbF and in a ΔwaaL mutant. O:54 antigen that is unable to be ligated to the lipid A-core portion of LPS causes growth impairment, which can be reversed over time by turning off (or down) O:54 production. This is the first observation of a significant growth defect phenotype in a ΔwaaL mutant during O-Ag production. These results establish the basic parameter for a detailed investigation of the structure and function of WbbF.
REFERENCES


Whitfield, C., Kaniuk, N., and Friridich, E. (2003). Molecular insights into the assembly and diversity of the outer core oligosaccharide in lipopolysaccharides from *Escherichia coli* and *Salmonella*. J Endotoxin Res 9, 244-249.