

Electronic Thesis and Dissertation Repository

9-16-2013 12:00 AM

Structure-Function Analysis of UDP-Sugar: Polyisoprenyl Phosphate Sugar-1-Phosphate Transferases

Sarah E. Furlong
The University of Western Ontario

Supervisor
Dr. Miguel Valvano
The University of Western Ontario

Graduate Program in Microbiology and Immunology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Sarah E. Furlong 2013

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Medical Microbiology Commons](#)

Recommended Citation

Furlong, Sarah E., "Structure-Function Analysis of UDP-Sugar: Polyisoprenyl Phosphate Sugar-1-Phosphate Transferases" (2013). *Electronic Thesis and Dissertation Repository*. 1629.
<https://ir.lib.uwo.ca/etd/1629>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

**STRUCTURE-FUNCTION ANALYSIS OF UDP-SUGAR:
POLYISOPRENYL PHOSPHATE SUGAR-1-PHOSPHATE
TRANSFERASES**

(Thesis format: Integrated Article)

by

Sarah Ellen Furlong

Graduate Program in Microbiology and Immunology

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

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

© Sarah E. Furlong 2014

Abstract

The synthesis of lipid-linked glycans is a conserved process in eukaryotes and prokaryotes that is initiated by two major enzyme families: the polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPTs) and the polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferases (PNPTs). These enzymes contain multiple membrane domains and transfer a sugar-1-phosphate from a nucleotide sugar precursor to a lipid carrier. The prototypic PNPT member used in this study is the *E. coli* WecA, which initiates the synthesis of O antigen and enterobacterial common antigen in *Enterobacteriaceae* by transferring N-acetylglucosamine-1-P to undecaprenyl phosphate (Und-P). We investigated the topology and function of the highly conserved VFMGD motif. Our results revealed that this motif faces the cytosol and defines a region in PNPTs that contributes to the active site, likely involved in the binding and/or recognition of the nucleotide moiety of the nucleoside phosphate precursor. The PHPT family member used in this study is the *E. coli* WcaJ, which transfers glucose-1-P to Und-P to initiate colanic acid synthesis. We provide the first detailed topological analysis of a PHPT member, which is inverted compared to the *in silico* topological predictions; the N-terminus, C-terminal tail and the large soluble loop all reside in the cytoplasm. We also found that the last membrane domain does not fully span the membrane and is likely ‘pinched in’. We further investigated the role of the N-terminal domain of WcaJ and our data suggest that it likely contributes to the protein folding and/or stability of PHPT family members. Together this work sheds light on the topological and mechanistic differences between these two major enzyme families and will guide further structural studies.

Key words: O antigen, colanic acid, UDP-sugar transferase, membrane protein

Statement of Co-Authorship

In Chapter Three of this thesis, Lorena Albarnez Rodriguez, a visiting student from the University of Valparaíso, Chile created plasmids (pLA1, pLA3, pLA5, pLA6, pLA7, pLA8, pLA9, pLA11, pLA18, pLA19, pLA20, pLA23, pLA24, pLA30, pLA31, pLA32, pLA33) and performed preliminary *in vivo* β -galactosidase and alkaline phosphatase activity assays on the respective plasmids. All other work is by S.E.F.

In Chapter Four, Dr. Mohamad A. Hamad, a postdoctoral fellow in the Valvano laboratory, created the strains MH110 and MH111. All other work is by S.E.F.

Dedication

This work is dedicated to my mother Cathryn, who has always demonstrated courage and strength in the face of tragedy and hardship.

Acknowledgements

I would like to offer my sincerest gratitude to my supervisor Dr. Miguel Valvano for his guidance, support, and motivation during my PhD. I have greatly benefited from his mentorship and I am honored to have been a member of his laboratory. Thank you to my advisory committee members, Dr. David Heinrichs and Dr. Lars Konermann for their insightful comments and suggestions in my projects over the years.

I would like to express my deepest appreciation to my family. My mom Cathryn, my twin sister Wendy, my brother Harrison, and my brother Jason and his wife Rebecca have all provided me with tremendous support. A special thank you to my mom and sister for our trips to Toronto.

I feel extremely fortunate to have made amazing friendships during my time in the lab. I am appreciative for my relationship with my best friend Kina Patel. I would like to thank her for her friendship and for bringing out my sense of adventure during our travels. I would also like to thank the Patel family, who has treated me like part of their family. Thank you for my friends in the lab, Chelsea Clarke, Maha Al-Zayer, Mo Hamad, Lorena Albarnez, Yasmine Fathy, Slade Loutet, Cristina Marolda, Crystal Schmerk, Katie Nurse, Daniel Aubert and Soledad Saldias. Thank you all for making the long days in the lab so much fun and for making me a part of your lives. I am also grateful for my friendships with Piya Lahiry and Tisha Joy. Thank you and Kina for adopting me into your group as “Seema”.

Finally I would like to thank my boyfriend, Graeme Ditner who joined me in the last two years of this journey. I am so grateful for the love, support and encouragement that he has given me during the final stages of my PhD.

Table of Contents

| Section | Page |
|---|-------|
| Title Page | i |
| Abstract | ii |
| Statement of Co-Authorship | iii |
| Dedication | iv |
| Acknowledgements | v |
| Table of Contents | vi |
| List of Tables | x |
| List of Figures | xi |
| List of Abbreviations | xiii |
| Chapter 1- Introduction | 1 |
| 1.1 The prokaryotic cell envelope | 2 |
| 1.1.1 Lipopolysaccharide | 2 |
| 1.1.1.1 Lipid A | 2 |
| 1.1.1.1.1 Lipid A biosynthesis | 4 |
| 1.1.1.2 Core-oligosaccharide | 7 |
| 1.1.1.2.1 Core-oligosaccharide biosynthesis | 10 |
| 1.1.1.3 O antigen | 11 |
| 1.1.1.3.1 O antigen biosynthesis | 12 |
| 1.1.1.3.1.1 Wzy/Wzx-dependent biosynthesis pathway | 12 |
| 1.1.1.3.1.2 ABC transporter-dependent pathway | 16 |
| 1.1.1.3.1.3 Other export pathways: Synthase- and Wxk-dependent pathways | 17 |
| 1.1.1.3.2 O antigen ligation | 18 |
| 1.1.1.3.3 Und-PP synthesis and recycling | 19 |
| 1.1.1.3.4 Lipopolysaccharide export to the outer membrane | 20 |
| 1.2 Initiation of the synthesis of cell surface polysaccharides | 23 |
| 1.2.1 Polyisoprenyl-phosphate <i>N</i> -acetylaminosugar-1-Phosphate Transferases | 23 |
| 1.2.1.1 Bacterial PNPT initiated lipid-linked glycans | 24 |
| 1.2.1.1.1 O antigen | 24 |
| 1.2.1.1.2 <i>Aeromonas hydrophila</i> O Antigen | 24 |
| 1.2.1.1.3 Enterobacterial common antigen | 25 |
| 1.2.1.1.4 Peptidoglycan | 26 |
| 1.2.1.2 Eukaryotic PNPT initiated lipid-linked glycans | 28 |
| 1.2.1.3 Topology of PNPTs | 29 |
| 1.2.1.4 Functional motifs of PNPTs | 30 |
| 1.2.1.5 Proposed catalytic mechanisms of PNPTs | 33 |
| 1.2.2 Polyisoprenyl-phosphate Hexose-1-Phosphate Transferases | 35 |
| 1.2.2.1 PHPT initiated lipid-linked glycans | 35 |
| 1.2.2.1.1 <i>Salmonella enterica</i> O antigen | 35 |
| 1.2.2.1.2 Colanic acid | 35 |
| 1.2.2.1.3 Capsule | 37 |

| | | |
|--|---|-----|
| 1.2.2.1.4 | S-Layer | 43 |
| 1.2.2.2 | Topology of PHPTs | 44 |
| 1.2.2.3 | Functional domains of PHPTs | 46 |
| 1.3 | Research objectives and summary of data obtained | 50 |
| 1.4 | Chapter one references | 52 |
| Chapter 2 - Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase | | 66 |
| 2.1 | Introduction | 67 |
| 2.2 | Materials and Methods | 70 |
| 2.2.1 | Bacterial strains, plasmids, media, and growth conditions | 70 |
| 2.2.2 | Site-directed mutagenesis | 72 |
| 2.2.3 | Growth conditions of cells for protein preparation | 72 |
| 2.2.4 | Total membrane preparation and immunoblotting | 73 |
| 2.2.5 | Sulfhydryl labeling of cysteine residues using biotin maleimide and protein purification | 74 |
| 2.2.6 | LPS analysis | 75 |
| 2.2.7 | <i>In vitro</i> transferase assay | 75 |
| 2.2.8 | Thin layer chromatography | 76 |
| 2.2.9 | Tunicamycin binding activity of WecA | 76 |
| 2.3 | Results | 77 |
| 2.3.1 | The substituted cysteine accessibility method reveals that aspartic acid residue D217 is exposed to the cytosol | 77 |
| 2.3.2 | Substitution of aspartic acid residue D217 by non acidic but polar residues results in a more active enzyme | 78 |
| 2.3.3 | The D217N WecA mutant has a higher V_{max} but no significant change in K_m | 80 |
| 2.3.4 | Alanine replacements of other residues of the VFMGD motif result in diminished enzymatic activity | 83 |
| 2.4 | Discussion | 86 |
| 2.5 | Chapter two references | 91 |
| Chapter 3 - Topological analysis of the <i>Escherichia coli</i> WcaJ reveals a new conserved configuration for the polyisoprenyl-phosphate hexose-1-phosphate transferase family of initiating enzymes | | 95 |
| 3.1 | Introduction | 96 |
| 3.2 | Materials and Methods | 99 |
| 3.2.1 | Bacterial strains, plasmids, media, and growth conditions | 99 |
| 3.2.2 | Cloning | 102 |
| 3.2.2.1 | Construction of pLA1 | 102 |
| 3.2.2.2 | Construction of pLA3 | 102 |
| 3.2.2.3 | Construction of pLA5 | 104 |
| 3.2.2.4 | Construction of Flag-WcaJ-A134-PhoA | 104 |

| | | |
|--|--|-----|
| 3.2.2.5 | Construction of Flag-WcaJ-PhoA at the following positions: V39, G74, D106, L223, N254, L272 | 104 |
| 3.2.2.6 | Construction of Flag-WcaJ-PhoA S120, N143, M160, P180, K301 | 105 |
| 3.2.2.7 | Construction of Flag-WcaJ-LacZ at the following positions: V39, G74, D106, S120, A134, N143, M160, P180, L223, N254, L272, K301, G316 | 105 |
| 3.2.3 | Creation of cysteine-less WcaJ and introduction of novel cysteines using site directed mutagenesis | 106 |
| 3.2.4 | Complementation of colanic acid in <i>E. coli</i> $\Delta wcaJ$, XBF1 | 107 |
| 3.2.5 | Growth conditions of cells for protein preparation | 107 |
| 3.2.6 | Total membrane preparation and immunoblotting | 108 |
| 3.2.7 | <i>In vivo</i> alkaline phosphatase assay | 108 |
| 3.2.8 | <i>In vivo</i> β -galactosidase assay | 109 |
| 3.2.9 | <i>In vitro</i> alkaline phosphatase assay | 109 |
| 3.2.10 | <i>In vitro</i> β -galactosidase assay | 110 |
| 3.2.11 | Labeling of novel cysteines in the Flag-WcaJ _{Cys-Less} protein using Methoxypolyethylene glycol maleimide (PEG-Mal) in total membrane preparations | 111 |
| 3.2.12 | Labeling of novel cysteines in the Flag-WcaJ _{Cys-Less} protein using PEG-Mal in EDTA permeabilized whole cells | 112 |
| 3.3 | Results | 114 |
| 3.3.1 | The last 212 amino acids of the C-terminus of the <i>E. coli</i> WcaJ is sufficient for enzymatic activity <i>in vivo</i> | 114 |
| 3.3.2 | Topological analysis of WcaJ using LacZ and PhoA reporter fusions reveals an unexpected membrane topology | 116 |
| 3.3.3 | Analysis of the borders of TMH-V suggests a helix-break-helix structure | 122 |
| 3.4 | Discussion | 128 |
| 3.5 | Chapter three references | 134 |
| Chapter 4 – Investigating the role of the N-terminal domain of the <i>Escherichia coli</i> UDP-glucose:undecaprenyl phosphate glucose-1-phosphate transferase WcaJ | | 137 |
| 4.1 | Introduction | 138 |
| 4.2 | Materials and Methods | 140 |
| 4.2.1 | Bacterial strains and plasmids | 140 |
| 4.2.2 | Construction of WcaJ cytoplasmic loop domain and α -helix deletions in pLA3 | 142 |
| 4.2.3 | Cloning various WcaJ constructs with N- or C-terminal Flag fusions | 142 |
| 4.2.4 | Cloning plasmids to make chromosomal Flag-tagged <i>wcaJ</i> and Flag-tagged <i>wcaJ_{CT}</i> | 145 |
| 4.2.5 | Construction of Flag-tagged <i>wcaJ</i> and <i>wcaJ_{CT}</i> strains | 146 |
| 4.2.6 | Growth conditions of cells for protein preparation | 146 |
| 4.2.7 | Membrane preparation and immunoblotting | 147 |
| 4.2.8 | <i>In silico</i> secondary structure prediction tools | 148 |

| | | |
|-----------------------------|---|-----|
| 4.2.9 | Colanic acid complementation assay | 148 |
| 4.2.10 | <i>In vitro</i> transferase assay | 148 |
| 4.3 | Results | 149 |
| 4.3.1 | In <i>silico</i> structural prediction programs suggest that the large cytoplasmic loop domain of WcaJ comprises α -helices and β -sheets reminiscent of a Rossmann fold | 149 |
| 4.3.2 | The N-terminal domain of WcaJ may be important for protein folding and/or stability | 152 |
| 4.3.3 | Deletion of the large cytoplasmic loop domain affects protein expression and/or stability of WcaJ while deletion of individual α -helices in the loop domain negatively affect enzymatic activity <i>in vivo</i> and <i>in vitro</i> without compromising protein expression | 154 |
| 4.3.4 | Expression of various N-terminal domain constructs does not restore enzymatic function in MH111 (chromosomal Flag-WcaJ _{CT}) | 162 |
| 4.4 | Discussion | 162 |
| 4.5 | Chapter four references | 169 |
| Chapter 5 – Discussion | | 171 |
| 5.1 | PNPTs and PHPTs are two distinct families of enzymes that initiate lipid-linked glycan synthesis | 172 |
| 5.1.1 | PNPT and PHPT enzymes have distinct membrane protein topologies and catalytic domains | 172 |
| 5.1.2 | The N-terminal domain of PHPTs may contribute to protein stability and binding of UDP | 174 |
| 5.2 | Future studies | 175 |
| 5.2.1 | Identification of the catalytic nucleophile of PNPT members | 175 |
| 5.2.2 | Structural analysis of the PHPT members | 177 |
| 5.3 | Concluding remarks and significance | 178 |
| 5.4 | Chapter five references | 179 |
| Appendix- Copyright release | | 181 |
| Curriculum Vitae | | 182 |

List of Tables

| Table | Title | Page |
|--------------|---|-------------|
| 2.1 | Characteristics of the bacterial strains and plasmids used in this study | 71 |
| 2.2 | Kinetic parameters of UDP-GlcNAc for parental WecA and D217N mutant enzymes | 85 |
| 3.1 | List of strains and plasmids | 100-101 |
| 3.2 | List of primers used for cloning | 103 |
| 3.3 | Alkaline phosphatase and β -galactosidase results of Flag-WcaJ-LacZ/PhoA fusion proteins expressed in DH5 α or CC118 cells | 118 |
| 4.1 | Characteristics of the bacterial strains and plasmids used in this study | 141 |
| 4.2 | Primers used in this study | 143-144 |

List of Figures

| Figure | Title | Page |
|--------|---|------|
| 1.1 | Illustrations of Gram-positive and Gram-negative bacterial cell envelopes | 3 |
| 1.2 | Lipid A biosynthesis in <i>E. coli</i> K-12 and <i>S. enterica</i> | 5 |
| 1.3 | Structure of the inner and outer core of <i>E. coli</i> K-12 | 8 |
| 1.4 | The core types in <i>E. coli</i> R1, R2, R3, R4, K-12 and <i>S. enterica</i> | 9 |
| 1.5 | O antigen biosynthesis pathways | 13 |
| 1.6 | Model for Und-PP recycling in <i>E. coli</i> | 21 |
| 1.7 | Model of LPS export from the inner membrane to the outer membrane | 22 |
| 1.8 | Proposed biosynthesis pathway of ECA _{PG} | 27 |
| 1.9 | Topological model of the <i>E. coli</i> WecA | 31 |
| 1.10 | Proposed biosynthesis pathway for colanic acid | 38 |
| 1.11 | <i>E. coli</i> K-12 colanic acid repeating structure | 39 |
| 1.12 | Model for the biosynthesis of group 1 and 4 capsules | 41 |
| 1.13 | Model for the biosynthesis of group 2 and 3 capsules | 42 |
| 1.14 | Predicted topology of <i>S. enterica</i> WbaP | 45 |
| 1.15 | Alignment of predicted topologies of PHPT members: <i>S. enterica</i> WbaP, <i>E. coli</i> WcaJ, <i>C. crescentus</i> HfsE, PssY, and PssZ | 47 |
| 2.1 | Topological model of the <i>E. coli</i> WecA | 69 |
| 2.2 | Sulfhydryl labeling accessibility of the cysteine replacement in WecA _{D217C} | 79 |
| 2.3 | <i>In vivo</i> and <i>in vitro</i> complementation of the D217 replacement mutants | 81 |
| 2.4 | The WecA _{D217N} protein mediates increased Und-P-P-GlcNAc production compared to parental WecA | 82 |
| 2.5 | GlcNAc-1-P transferase activity of parental WecA (A) the D217N mutant (B) | 84 |
| 2.6 | Tunicamycin binding assay using parental WecA and D217 replacement mutants | 87 |
| 2.7 | <i>In vivo</i> and <i>in vitro</i> complementation of alanine replacement mutants in the VFMGD motif | 88 |
| 3.1 | Predicted topology of <i>E. coli</i> W3110 WcaJ using TMHMM results from PhoA/LacZ reporter fusions | 98 |
| 3.2 | Membrane protein expression of Flag-WcaJ and Flag-WcaJ _{CT} constructs and complementation of CA in XBF1/pWQ499 | 115 |
| 3.3 | Membrane expression of Flag-WcaJ-PhoA fusion proteins in CC118 cells | 120 |
| 3.4 | Membrane expression of Flag-WcaJ-LacZ fusion proteins in DH5 α cells | 121 |
| 3.5 | Membrane protein expression of Flag-WcaJ _{Cys-Less} and complementation of CA | 123 |
| 3.6 | Complementation of CA in XBF1/pWQ499 by cysteine replacement proteins | 124 |
| 3.7 | Substituted cysteine labeling using PEG-Mal in total membrane preparations expressing Flag-WcaJ _{Cys-Less} cysteine replacement proteins | 125 |

| | | |
|------|--|-----|
| 3.8 | Substituted cysteine labeling using PEG-Mal in EDTA-permeabilized whole cells and lysed cells | 127 |
| 3.9 | <i>In silico</i> analysis of TMH-V of the <i>E. coli</i> WcaJ | 129 |
| 3.10 | ConSeq results of WcaJ _{CT} | 130 |
| 3.11 | Refined topological model of the <i>E. coli</i> WcaJ after using reporter fusions and substituted cysteine labeling with PEG-Mal | 131 |
| 4.1 | Topological alignments of various PHPTs | 139 |
| 4.2 | Secondary structure of WcaJ loop domain (R131-D278) using PSIPRED | 150 |
| 4.3 | Structural prediction of the WcaJ large cytoplasmic loop domain using HHpred | 151 |
| 4.4 | Colanic acid production in <i>E. coli</i> strains W3110, XBF1 (<i>wcaJ</i> deletion strain), MH110 and MH111 (chromosomally tagged Flag-WcaJ and Flag-WcaJ _{CT}) co-expressing RcsA | 153 |
| 4.5 | Complementation of colanic acid production in XBF1/pWQ499 by Flag-WcaJ and Flag-WcaJ _{CT} with increasing concentrations of arabinose | 155 |
| 4.6 | Immunoblotting of total membranes expressing Flag-WcaJ and Flag-WcaJ _{CT} with increasing concentrations of arabinose | 156 |
| 4.7 | Complementation of colanic acid synthesis in XBF1/pWQ499 expressing the wild-type Flag-WcaJ and the loop deletion construct Flag-WcaJ Δ N143-L272 | 158 |
| 4.8 | <i>In vitro</i> transferase assay of the wild-type Flag-WcaJ and the loop deletion construct Flag-WcaJ Δ N143-L272 | 159 |
| 4.9 | Complementation of colanic acid synthesis in XBF1/pWQ499 expressing the wild-type Flag-WcaJ and the α -helix deletion constructs | 160 |
| 4.10 | <i>In vitro</i> transferase assay of the wild-type Flag-WcaJ and the Flag-WcaJ α -helix deletion constructs (Δ L152-N164, Δ N192-A202, Δ D216-D229) | 161 |
| 4.11 | Diagram of the various WcaJ constructs cloned into pBAD vectors | 163 |
| 4.12 | Expression of various WcaJ constructs in XBF1/pWQ499 | 164 |
| 4.13 | Expression of various WcaJ constructs in MH111 (Flag- <i>wcaJ</i> _{CT})/pWQ499 | 165 |
| 4.14 | Immunoblotting of total membrane expressing various WcaJ constructs | 166 |

List of Abbreviations

| | |
|-------------------------|--|
| ABC | ATP-binding cassette |
| Amp | Ampicillin |
| Ara4N | 4-amino-4-deoxy- α L-arabinose |
| CA | Colanic acid |
| CA_{LPS} | Colanic acid attached to lipid A-core OS |
| CAP | Covalently attached protein |
| Cm | Chloramphenicol |
| Core OS | Core oligosaccharide |
| CPS | Capsular polysaccharide |
| CT | C-terminal |
| ECA | Enterobacterial common antigen |
| EPS | Exopolysaccharide |
| Fuc | Fucose |
| Gal | Galactose |
| Glc | Glucose |
| GlcNAc | N-acetylglucosamine |
| IMP | Integral membrane protein |
| Kdo | 3- <i>deoxy</i> -D- <i>manno</i> -oct-2-ulosonic acid |
| Kn | Kanamycin |
| Ko | D- <i>glycero</i> - α -D- <i>talo</i> -oct-2-ulosonic acid |
| LB | Luria-Bertani |
| LP | Lipoprotein |
| LPS | Lipopolysaccharide |
| LTA | Lipoteichoic acid |
| O Ag | O antigen |
| OMP | Outer membrane protein |
| PCP | Polysaccharide co-polymerases |
| PCR | Polymerase chain reaction |
| PEG-Mal | Methoxypolyethylene glycol maleimide |
| PHPT | Polyisoprenyl-phosphate Hexose-1-Phosphate Transferase |
| PNPT | Polyisoprenyl-phosphate <i>N</i> -acetylaminosugar-1-Phosphate Transferase |
| Rha | Rhamnose |
| SDS-PAGE | Sodium dodecylsulfate polyacrylamide gel electrophoresis |
| Tc | Tetracycline |
| TLC | Thin layer chromatography |
| TM | Transmembrane |
| TMH | Transmembrane helix |
| Tp | Trimethoprim |
| UDP | Uridine diphosphate |
| UMP | Uridine monophosphate |
| Und-P | Undecaprenyl phosphate |
| WTA | Wall teichoic acid |

Chapter 1

Introduction

1.1 The prokaryotic cell envelope

The prokaryotic cell envelope is highly complex and helps bacteria resist environmental stresses, maintain cell shape, remove waste products from the cell and selectively pass nutrients into the cell. Bacteria can be separated into two groups: Gram positive and Gram negative. Both groups have an inner membrane phospholipid bilayer. Gram-negatives also contain an outer membrane forming an asymmetric bilayer made of phospholipids and lipopolysaccharides, and a thin layer of cell wall peptidoglycan in the area between the inner and outer membrane, called the periplasm. Gram-positive bacteria lack outer membrane and instead have a thicker peptidoglycan cell wall, which is decorated with teichoic acids (Fig. 1.1) (Silhavy *et al.*, 2010).

1.1.1 Lipopolysaccharide

Lipopolysaccharide (LPS) is major component of the outer leaflet of the outer membrane of Gram-negative bacteria. LPS has three major components: a hydrophobic domain Lipid A (see section 1.1.1.1.), core oligosaccharide, (see section 1.1.1.2.) and O antigen (see section 1.1.1.3). LPS plays an important role in the barrier function by contributing to membrane permeability and stability, resistance against antimicrobial peptides, and protection against host immune responses (Silhavy *et al.*, 2010, Raetz and Whitfield, 2002).

1.1.1.1 Lipid A

Lipid A is the most highly conserved component of LPS molecules and serves as the hydrophobic anchor for LPS. Lipid A consists of a glucosamine-based glycolipid

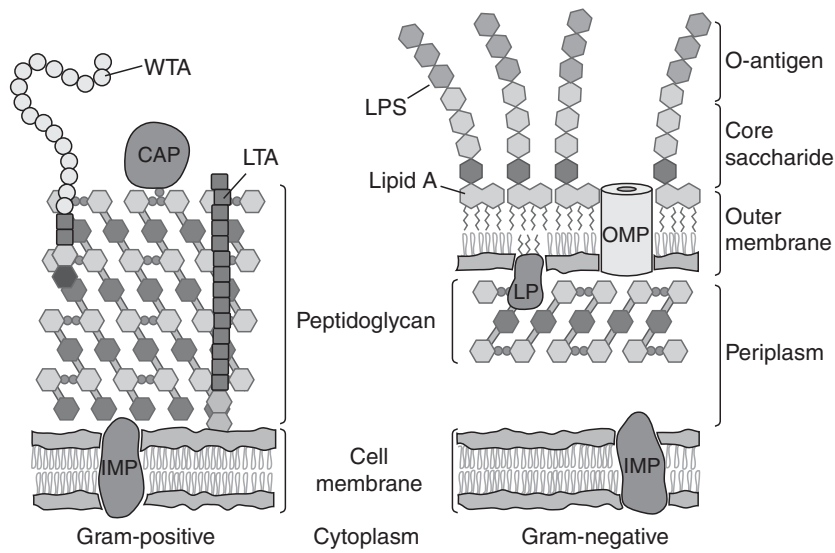


Figure 1.1. Illustrations of Gram-positive and Gram-negative bacterial cell envelopes. Figure from (Silhavy *et al.*, 2010). Abbreviations: CAP, covalently attached protein; IMP, integral membrane protein; LP, lipoprotein; LPS, lipopolysaccharide; LTA, lipoteichoic acid; OMP, outer membrane protein; WTA, wall teichoic acid.

(Raetz *et al.*, 2007). Lipid A usually contains six or seven saturated acyl chains, although this can vary depending on the specific bacterium. These acyl chains contribute to the tight packing of the outer leaflet of the outer membrane, which prevents entry of hydrophobic molecules (Silhavy *et al.*, 2010). Lipid A is a potent endotoxin in mammals. Lipid A activates the TLR4 receptor of the mammalian innate immune system, which stimulates the production of factors that promote clearing of the infection. However, overproduction of these factors can cause septic shock (Raetz and Whitfield, 2002).

1.1.1.1.1 Lipid A biosynthesis

Lipid A biosynthesis begins in the cytoplasm with the acylation of UDP-GlcNAc by LpxA (Anderson and Raetz, 1987) (Fig. 1.2). The resulting UDP-3-O-(R-3-hydroxymyristoyl)-GlcNAc is then deacetylated by LpxC, a bacterial zinc metalloamidase to form UDP-3-O-(acyl)-GlcN, the first committed step of lipid A biosynthesis (Jackman *et al.*, 2001). LpxD then incorporates a second acyl chain, a β -hydroxymyristate moiety, onto the substrate to produce UDP-2,3-diacylglucosamine (Kelly *et al.*, 1993). LpxH hydrolyzes the pyrophosphate bond of UDP-2,3-diacylglucosamine to produce 2,3-diacylglucosamine-1-phosphate (also referred to as lipid X) and releasing UMP (Babinski *et al.*, 2002). LpxB, the disaccharide synthase, condenses with another molecule of lipid X to form the characteristic β -1',6-glycosidic bond of lipid A (Metzger and Raetz, 2009).

Some Gram-negative bacteria lack the *lpxH* gene and produce 2,3-diacylglucosamine-1-phosphate with a different enzyme called LpxI. LpxI was discovered in *Caulobacter crescentus* and generates this product by the attack of water on the β -phosphate of UDP, instead of the α -phosphate (Metzger and Raetz, 2010). Then the

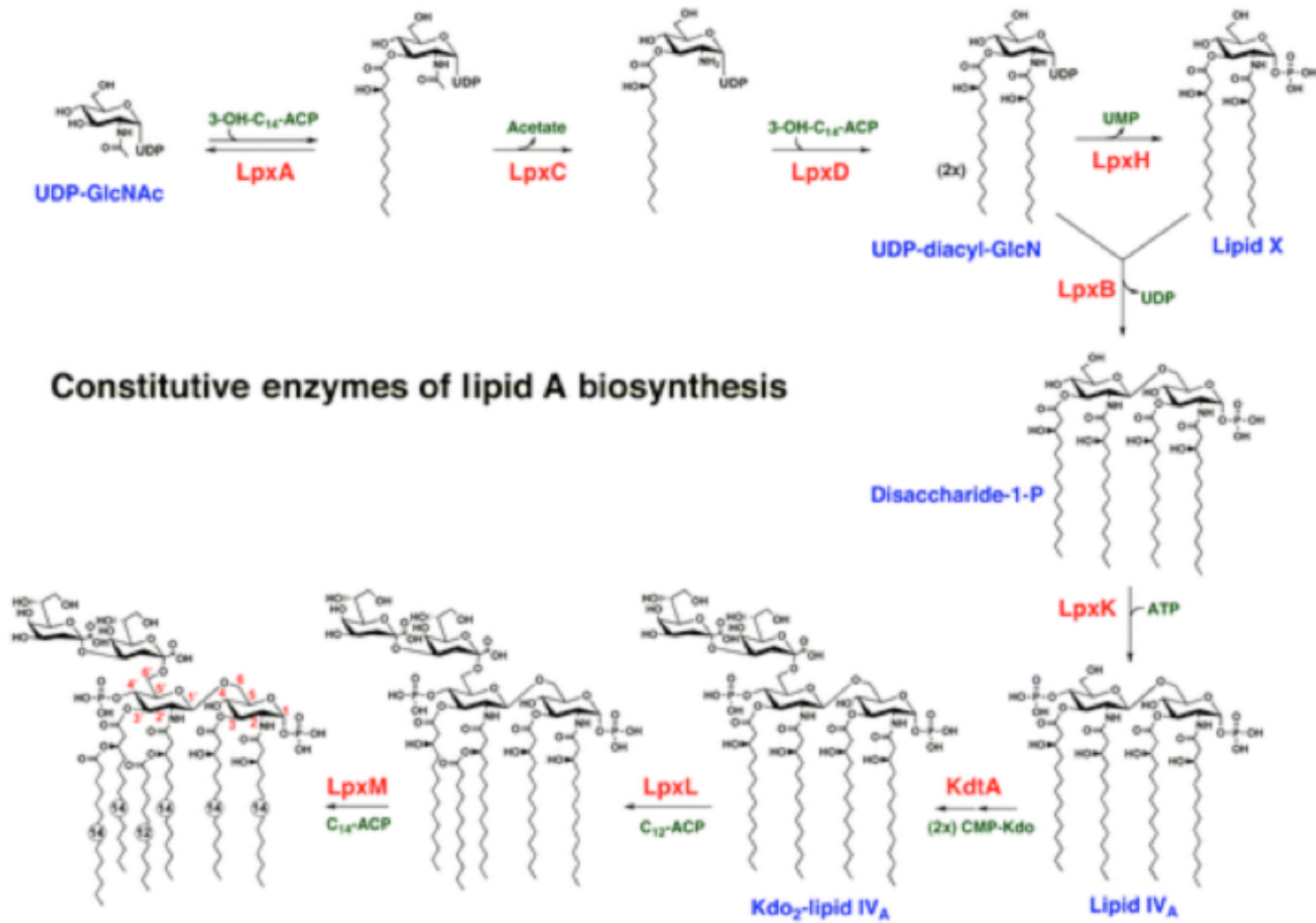


Figure 1.2. Lipid A biosynthesis in *E. coli* K-12 and *S. enterica*. Figure from (Stead *et al.*, 2011).

tetraacyldisaccharide-1-phosphate intermediate is phosphorylated by LpxK to form lipid IV_A (Emptage *et al.*, 2013). WaaA, formerly KdtA, transfers two 3-*deoxy-D-manno*-oct-2-ulosonic acid (Kdo) moieties to the distal glucosamine of lipid IV_A, which constitutes the link between lipid A and core oligosaccharide (Clementz, 1992, Stead *et al.*, 2011). Some homologs of WaaA have the ability to transfer between one to four Kdo residues. After this, secondary acyl transferases add acyl chains to Kdo₂- lipid IV_A. The hexaacylated lipid A species in *Escherichia coli* and *Salmonella* are produced by acyl carrier protein (ACP)-dependent acyltransferases LpxL and LpxM. LpxL first transfers the lauroyl group from lauroyl-ACP to the 2'-position of Kdo₂- lipid IV_A (Carty *et al.*, 1999, Clementz *et al.*, 1996). Then LpxM transfers a myristoyl group from myristoyl-ACP to the 3'-position of Kdo₂- lipid IV_A (Carty *et al.*, 1999). In other Gram negative bacteria, the secondary acyl chain placements can vary depending on the types of secondary acyltransferases expressed, which contribute to lipid A diversity (Stead *et al.*, 2011).

Lipid A diversity occurs through lipid A modifications. These modifications arise in response to certain environmental stressors such as chemicals, antibiotics, changes in pH, and antimicrobial peptides (Gunn, 2008). *Salmonella enterica* Typhimurium, modifies LPS in response to changes in the outer membrane, such as exposure to cationic peptides or in reduced divalent cation environments. These changes are sensed by the PhoPQ two-component system and result in the increased transcription of genes *pagP*, *lpxO*, and genes regulated by the two-component system PmrA/B (Murata *et al.*, 2007). Some of these modifications include the PagP-catalyzed palmitoylation of lipid A and incorporating phosphoethanolamine and L-4-aminoarabinose (L-Ara4N) at the 1 and 4' phosphates of lipid A. Addition of L-Ara4N contributes to antimicrobial peptide and toxic

metals resistance by neutralizing the phosphates on lipid A (Kato *et al.*, 2012, Gunn, 2008). Interestingly, addition of L-Ara4N on lipid A in *Burkholderia cenocepacia* is required for bacterial survival and also contributes to its high level of antimicrobial peptide resistance (Hamad *et al.*, 2012).

Although lipid A is the most conserved component of lipopolysaccharides, these acylation steps and lipid A modifications contribute to its diversity amongst various bacteria.

1.1.1.2 Core oligosaccharide

The core oligosaccharide comprises sugars that are added to the lipid A molecule. The core OS can be divided into two regions: the inner core, which is proximal to lipid A, and the outer core, which is connected to the O antigen polysaccharide (Fig. 1.3) (Friedrich and Whitfield, 2005, Raetz and Whitfield, 2002). The outer core is more diverse than the inner core, which is conserved even in distantly related bacteria. All known inner cores contain 3-*deoxy-D-manno*-oct-2-ulosonic acid (Kdo) or its derivative D-*glycero- α -D-talo*-oct-2-ulosonic acid (Ko), and most contain L-*glycero-D-manno*-heptose (L,D-Hep) (Raetz and Whitfield, 2002). The inner core can also be decorated with other sugars, such as rhamnose (Rha), galactose (Gal), N-acetylglucosamine (GlcNAc), Kdo, and groups such as phosphate, phosphorylcholine, and pyrophosphorylethanolamine, which contribute to diversity even within a single species. In *E. coli*, there are five known core structures: R1, R2, R3, R4, and K-12 (Fig. 1.4) (Friedrich and Whitfield, 2005, Raetz and Whitfield, 2002). The outer core is more diverse, with five known outer cores identified in *E. coli*. This diversity is likely attributed to selective pressures from host and environmental stressors. The core OS plays an important role in maintaining membrane

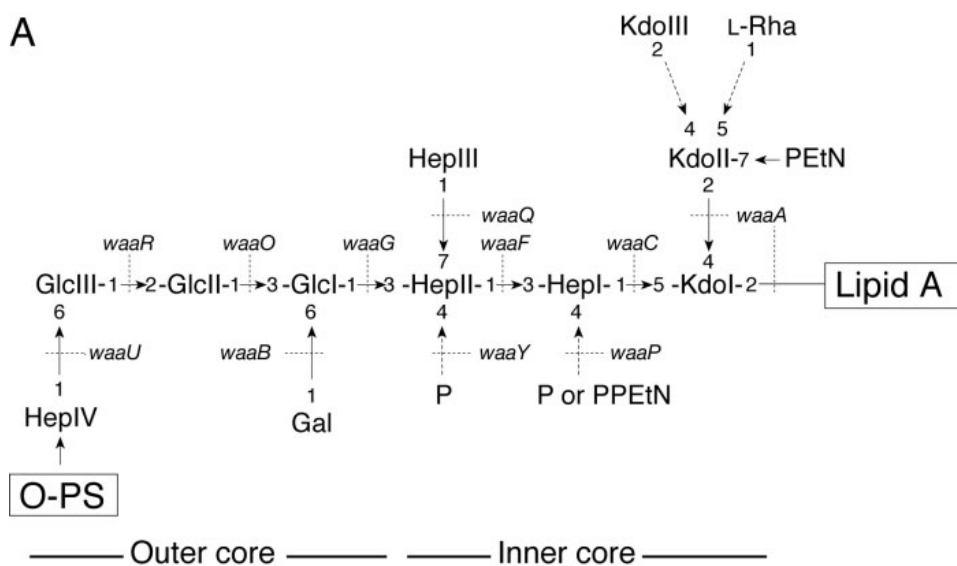


Figure 1.3. Structure of the inner and outer core of *E. coli* K-12. Non-stoichiometric substitutions and are shown with dashed arrows and known or predicted genes involved in the linkages are shown in dotted lines. Figure is from (Friedrich and Whitfield, 2005).

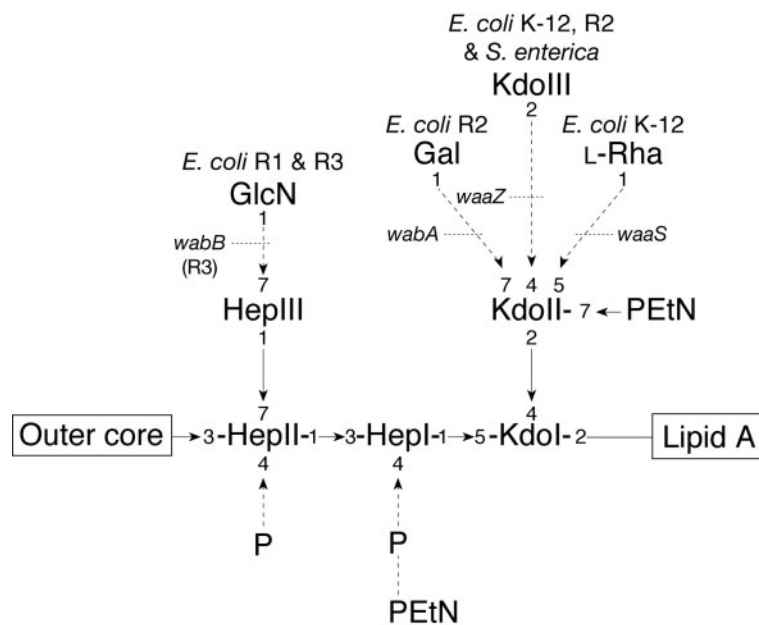


Figure 1.4. The core types in *S. enterica* and *E. coli* R1, R2, R3, R4, K-12. Non-stoichiometric substitutions are represented by dashed arrows and genes involved in these linkages are shown by dotted lines. Figure from (Frirdich and Whitfield, 2005).

stability by providing resistance against dyes, hydrophobic antibiotics, fatty acids, and detergents (Raetz and Whitfield, 2002). In *E. coli*, this resistance is maintained by the addition of phosphate groups to the core (Yethon *et al.*, 2000), as negatively charged groups on LPS bind divalent cations and this is presumed to stabilize adjacent LPS molecules (Nikaido and Vaara, 1985, Whitfield *et al.*, 2003). Although the core OS is important for membrane integrity, it has recently been found that the core is not essential for viability as some isolates can only synthesize lipid A or a lipid A precursor (Meredith *et al.*, 2006, Mamat *et al.*, 2008).

1.1.1.2.1 Core oligosaccharide biosynthesis

The core oligosaccharide biosynthesis genes reside in the *waa* locus, which is generally conserved amongst *E. coli*, *Salmonella* and *Shigella* core types. In *E. coli*, the *waa* locus comprises three operons: *hldD*, *waaQ*, and *waaA* (Friedrich and Whitfield, 2005).

The link between lipid A and inner core OS is through a Kdo moiety. The *waaA* operon contains the gene for the WaaA enzyme, formerly KdtA, which transfers two Kdo moieties to the distal glucosamine of lipid IV_A (Clementz, 1992, Stead *et al.*, 2011). Then, the acylation of lipid A is completed prior to the addition of inner core sugars to the lipid A-Kdo₂ acceptor (Friedrich and Whitfield, 2005). The inner core sugar ADP-L-glycero-D-manno-heptose is synthesized by HldD, the ADP-L-glycero-D-manno-heptose 6-epimerase (Morrison and Tanner, 2007). These heptoses (Hep I and Hep II) are subsequently added to Kdo by the heptosyltransferases WaaC and WaaF (Fig. 1.3) (Gronow *et al.*, 2000). Genes encoded in the *waaQ* operon synthesize the outer core and modifications to inner core. Generally, the *E. coli* outer core is synthesized by

glucosyltransferases WaaG, WaaO, WaaR, and the galactosyltransferase WaaB (Whitfield *et al.*, 2003) (Fig. 1.3). Modifications to the inner core can be attributed to gene products from *waaS* (K-12 core), *wabA* (R2 core), *waaZ* (K-12, R2), and *wabB* (R3) (Whitfield *et al.*, 2003, Frirdich and Whitfield, 2005) (Fig. 1.4). After the lipid A-core OS is complete, MsbA, a lipid activated ATP-binding cassette (ABC) transporter (Doerrler and Raetz, 2002) will translocate the molecule across the inner membrane from the cytoplasmic side to the periplasmic space (Doerrler *et al.*, 2004), where the completed O Ag chain is ligated to lipid A-core OS (see section 1.1.1.3.2 for O antigen Ligation).

1.1.1.3 O antigen

O antigen is the most surface exposed component of the LPS molecule and is also the most diverse (Valvano, 2003). O Ag plays an important role in bacterial survival against host defenses. O Ag chain length is crucial for resistance against complement-mediated cell lysis (Burns and Hull, 1998, Bravo *et al.*, 2008) and influences macrophage invasion or uptake (Murray *et al.*, 2006, Saldías *et al.*, 2009).

The O Ag chain extends from the core OS and is comprised of either repeating subunits to form a heteropolymer or a single repeating sugar to form a homopolymer. These chains can vary in the types of sugars, sugar branching, and presence of nonstoichiometric modifications, such as O-acetylation and glycosylation. This creates O Ag diversity even within a single species, known as O-serological specificity. For example, *E. coli* has approximately 170 different serotypes (Raetz and Whitfield, 2002).

1.1.1.3.1 O antigen biosynthesis

The biosynthesis of O Ag is highly complex and involves several enzymes in the cytoplasmic space and inner membrane. Initiation enzymes commence the synthesis of O Ag by transferring the first sugar of the O Ag chain to the lipid carrier, undecaprenyl-phosphate (Und-P), on the cytoplasmic face of the inner membrane. The initiation enzymes are grouped into two families: the Polyisoprenyl-phosphate *N*-acetylaminosugar-1-Phosphate Transferases (PNPTs) and the Polyisoprenyl-phosphate Hexose-1-Phosphate Transferases (PHPTs) (see sections 1.2.1 and 1.2.2, respectively) (Valvano, 2003). The initiation enzymes transfer uridine diphosphate (UDP)-sugar to Und-P, to form a phosphoanhydride bond and releasing uridine monophosphate (UMP). The Und-P linked sugar becomes an acceptor molecule for one of the four known O Ag biosynthetic pathways: Wzy/Wzx-, ABC transporter-, Synthase- and Wzk-dependent pathways (Valvano, 2003, Valvano *et al.*, 2011) (Fig. 1.5).

1.1.1.3.1.1 Wzy/Wzx-dependent biosynthesis pathway

Bacteria that produce a heteropolymeric chain of repeating O Ag subunits utilize the Wzy/Wzx dependent pathway (Fig. 1.5). After the initiation reaction, glycosyltransferases on the cytoplasmic face of the inner membrane add sugars to the Und-P-P-sugar acceptor molecule to complete the O Ag subunit (Samuel and Reeves, 2003). The O Ag unit is then translocated across the inner membrane to the periplasmic

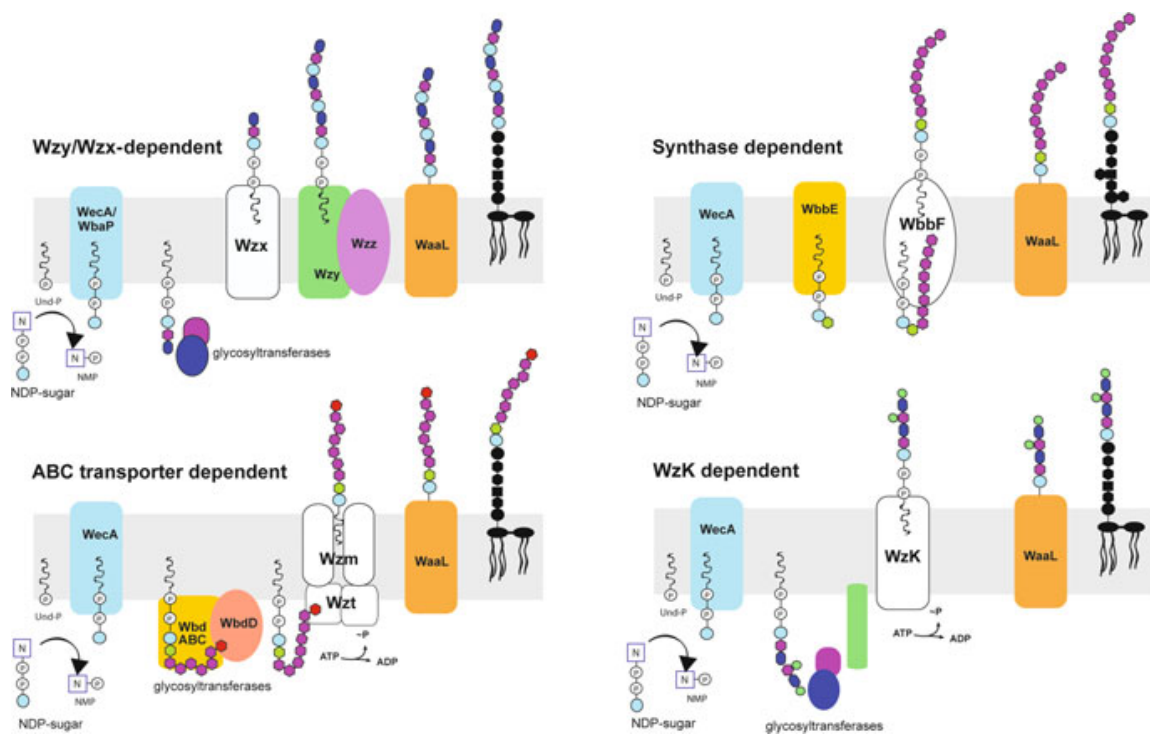


Figure 1.5. O antigen biosynthesis pathways. Figure unmodified from (Valvano *et al.*, 2011).

space by the Wzx integral membrane enzyme (Liu *et al.*, 1996). The topology of Wzx, consisting of 12 predicted transmembrane domains, has been experimentally refined for *S. enterica* (Cunneen and Reeves, 2008), *Pseudomonas aeruginosa* (Islam *et al.*, 2010), *Rhizobium leguminosarum* (Mazur *et al.*, 2005), and *E. coli* O157 (Marolda *et al.*, 2010). The role of Wzx was first elucidated during *in vivo* experiments when a *wzx* deletion strain of *S. enterica* accumulated radiolabeled Und-P-P-O Ag on the cytoplasmic surface of the inner membrane (Liu *et al.*, 1996). Wzx only requires recognition of the first sugar of the O Ag unit to translocate the Und-P-P-O Ag (Marolda *et al.*, 2004). It is also known that four charged amino acids in Wzx are required for function, two of which reside in transmembrane helices II and VI. The authors suggest that Wzx transmembrane helices may be interacting with one another, similar to the LacY permease, and/or the Und-P-P-sugars (Marolda *et al.*, 2010).

An *in vitro* “flipping” assay has not been developed for Wzx, which has hampered the ability to identify the exact flipping mechanism. Despite this, Islam *et al* have suggested an antiport model for the flipping mechanism. Using tertiary structure homology modeling based on the known structure of NorM, an exporter from the same superfamily, these authors identified a central cationic lumen between transmembrane segments. These authors also identified a number of charged and aromatic residues in the transmembrane segments, which are essential for enzyme function and may contribute to an antiport flipping mechanism. The anionic amino acids in the outward-facing conformation of Wzx may bind ions from the periplasmic space, causing the protein to have an inward-facing conformation, which would allow the cationic channel lumen to bind the O Ag substrate (Islam *et al.*, 2012). This results in a change to the outward-

facing conformation (Wzx open towards the periplasmic space) and the translocation of the Und-P-P-O Ag (Islam and Lam, 2013). This model provides an interesting explanation for the flipping mechanism, but still needs to be experimentally validated (Islam *et al.*, 2012).

After the O Ag subunits are translocated to the periplasmic space, the Wzy protein polymerizes the O Ag chain. The chain grows from the reducing end, a biosynthesis mechanism that is novel from other polysaccharides that typically grow from the non-reducing end (Bray and Robbins, 1967). Polysaccharide co-polymerases (PCP) determine the length of the O Ag chain. Using purified Wzy and two different PCPs, known as Wzz proteins, *in vitro* experiments reveal that Wzz determines strain specific chain lengths (Woodward *et al.*, 2010). PCPs share a common protomer structure with an alpha helix that extends into the periplasmic space, suggesting that the top of PCPs may contribute to chain length regulation; however, it is still unclear how these proteins control the polysaccharide length. Also, the number of protomers in the oligomer is under discussion (Tocij *et al.*, 2008, Larue *et al.*, 2009, Kalynych *et al.*, 2012b). With knowledge of 3D PCP structures, Kalynych *et al* have proposed a new mechanism for the chain length determination. These authors suggested that PCPs act to stabilize the growing polysaccharide and prevent the release of the growing chain from the Wzy polymerase. When the chain becomes too long for the PCPs to stabilize, the chain is released. This model explains why there are not a precise number of O Ag units associated with PCPs and why there have been no observed physical interactions observed between Wzy and Wzz (Kalynych *et al.*, 2012a).

1.1.1.3.1.2 ABC transporter-dependent pathway

The ABC transporter-dependent pathway has been well described for *E. coli* O8 and O9 Ag biosynthesis (Fig 1.5). This pathway resembles the biosynthesis of group II polysaccharides (Bronner *et al.*, 1994). O Ag synthesis commences after WecA transfers GlcNAc-1-P to Und-P to form Und-P-P-GlcNAc (Rick *et al.*, 1994, Jann *et al.*, 1979). Using purified enzymes and *in vitro* assays, Greenfield *et al* demonstrated the functions of mannosyltransferases involved in O Ag production. WbdC and WbdB are two conserved mannosyltransferases that are responsible for producing the tetrasaccharide adaptor region found in O8 and O9 antigens. WbdA enzymes are serotype specific polymerizing multi-domain mannosyltransferases that build the O Ag repeat unit (Greenfield *et al.*, 2012b, Greenfield *et al.*, 2012a). Chain elongation continues until it is terminated by the addition of a methyl (O8) or methyl phosphate (O9a) group by WbbD (Clarke *et al.*, 2004, Clarke *et al.*, 2011). The completed O Ag chain is exported by the Wzm-Wzt ABC transporter (Bronner *et al.*, 1994). Wzm is a transmembrane component and Wzt is a nucleotide-binding component. Wzt has two domains, an N- and a C-terminal domain, both of which are required for export but can complement function if expressed separately. The N-terminal domain resembles typical ABC transporters that can hydrolyze ATP, however, the C-terminal domain has an O polysaccharide-binding pocket that is strain specific. This provides insight into possible strain specific ABC transporter-dependent mechanisms of O Ag biosynthesis and export (Cuthbertson *et al.*, 2007). A variety of other bacteria that synthesize O Ag polysaccharides using the ABC transporter-dependent pathway also utilize terminal capping and export (Greenfield and Whitfield, 2012).

1.1.1.3.1.3 Other export pathways: Synthase- and Wzk-dependent pathways

The Synthase-dependent mechanism has only been observed to produce O Ag in *Salmonella enterica* serovar Borreze (Keenleyside and Whitfield, 1996) (Fig. 1.5). *S. enterica* serovar Borreze carries a 6.9-kb plasmid pWQ799 that encodes the entire O:54 biosynthetic gene cluster (Keenleyside and Whitfield, 1995). The synthesis of the O:54 homopolymer is initiated by the transfer of UDP-GlcNAc to Und-P by WecA, which serves a priming sugar for synthesis. The Rfb_{O:54} is a processive *N*-acetylmannosaminyltransferase that adds ManNAc to Und-P-P-GlcNAc to produce a polymeric O Ag chain in the cytoplasmic space. The synthase itself also serves to export the polysaccharide across the inner membrane, however the mechanism is unknown (Keenleyside and Whitfield, 1996). The polymer is then ligated to lipid A-core OS.

The Wzk-dependent O Ag synthesis pathway has only been identified in *Helicobacter pylori* (Fig. 1.5). The O Ag synthesis is initiated by WecA, which results in the formation of Und-P-P-GlcNAc. Glycosyltransferases alternate the addition of GalNAc and Gal residues to form the O Ag chain, which is then decorated by fucosyltransferases to create Lewis antigens. Afterwards, the Wzk flippase translocates this O Ag chain to the periplasmic space where WaaL will ligate it to the lipid A-core OS. Wzk is not related to any known O Ag flippases but has homology to a flippase involved in protein N-glycosylation in *Campylobacter jejuni*, PglK. Interestingly, Wzk has relaxed substrate specificity because it can translocate Und-P linked substrates of various sugar compositions and length such as, Lewis antigens, *E. coli* O16 antigen, and *C. jejuni* and *N*-glycosylation heptasaccharide (Hug *et al.*, 2010).

1.1.1.3.2 O antigen ligation

Irrespective of the export pathway, after formation of the O Ag chain, the chain is ligated to the lipid A-core OS in the periplasmic space (Mulford and Osborn, 1983, McGrath and Osborn, 1991, Abeyrathne *et al.*, 2005). This step is catalyzed by the O Ag ligase WaaL, which denotes a family of enzymes with multiple transmembrane domains and a large periplasmic loop region. In the *E. coli* K-12 WaaL, the large periplasmic loop domain has two predicted alpha helices, each with a functionally important basic amino acid, and a third periplasmic loop with a basic amino acid critical for ligase activity (Perez *et al.*, 2008). It was unclear if WaaL was ATP-dependent, as two groups published contradictory results. Abeyrathne and Lam concluded that the *Pseudomonas aeruginosa* WaaL hydrolyzes ATP and that ATP hydrolysis is essential for the *in vitro* ligation reaction (Abeyrathne and Lam, 2007). Hug *et al* provided evidence that the *Helicobacter pylori* WaaL enzyme did not require ATP (Hug *et al.*, 2010). Ruan *et al* purified the *E. coli* K-12 and *P. aeruginosa* WaaL ligases and demonstrated that these enzymes do not hydrolyze ATP and the reaction can proceed in the complete absence of ATP. More importantly, these authors also provided evidence that WaaL is a metal ion-independent inverting glycosyltransferase as the attachment of proximal sugar to the lipid A-core OS is in the β -conformation, despite that diphosphate nucleotide sugar precursors are in the α -conformation. These authors suggest that there is a nucleophilic attack by the acceptor hydroxyl group and that the basic amino acids in the periplasmic loop of the ligase may stabilize the Und-P-P leaving group. This would invert the glycosyl linkage with the proximal sugar and lipid A-core OS. Although this mechanism has yet to be clearly demonstrated, these authors provide compelling evidence WaaL ligases are metal-ion

ATP-independent inverting glycosyltransferases (Ruan *et al.*, 2012).

1.1.1.3.3 Und-PP synthesis and recycling

Und-P is used as a lipid carrier in the synthesis of multiple polysaccharide polymers, such as peptidoglycan and O Ag. Und-P is synthesized by the addition of isoprene units onto farnesyl pyrophosphate, which results in the formation of the C55-PP (Und-PP) lipid substrate. The Und-PP is dephosphorylated to make Und-P available for use in the cell. This dephosphorylation step is performed mainly by UppP (formerly BacA), which is responsible for 75% of the dephosphorylation activity in cells, demonstrated by the remaining 25% activity in a *uppP* mutant strain (El Ghachi *et al.*, 2004). However, the remaining dephosphorylation activity occurs by the integral membrane proteins YbjG and PgpB, because deletion of all three genes results in a lethal phenotype, consistent with the notion that the generation of Und-P is essential for cell survival (El Ghachi *et al.*, 2005).

Und-P is synthesized *de novo* in very small amounts, therefore Und-P is a limiting factor in the biogenesis of lipid-linked glycans and must be recycled by the cell. Und-P is regenerated from Und-P-P linked glycans after the glycan has been transferred onto a specific acceptor molecule in the periplasmic space, such as the transfer of an O Ag chain onto the lipid A-core OS acceptor (Valvano, 2008). There are two known Und-P-P recycling mechanisms in bacteria: the dephosphorylation of Und-P-P and the transfer of the phosphate from Und-P-P to lipid A (Tatar *et al.*, 2007, Touze *et al.*, 2008) (Fig. 1.6). Tatar *et al* identified two enzymes in *E. coli*, YbjG and YeiU, which dephosphorylate Und-P-P on the periplasmic face of the inner membrane. These enzymes bear resemblance to the eukaryotic dolichol-P-P recycling enzyme, Cwh8 (Tatar *et al.*, 2007).

Touzé *et al* demonstrated LpxT transfers the phosphate from Und-P-P to lipid A to form lipid A 1-diphosphate. A thermosensitive MsbA variant, that cannot transport the lipid A across the inner membrane, could not receive this phosphate, suggesting that this transfer occurs on the periplasmic face of the inner membrane (Touze *et al.*, 2008).

1.1.1.3.4 Lipopolysaccharide export to the outer membrane

The complete LPS molecule is exported from the inner membrane to the outer membrane by the Lpt pathway. There are seven proteins that are essential for the export of LPS, LptABCDEFG, which form a trans-envelope bridge between the inner membrane and outer membrane (Fig. 1.7) (Chng *et al.*, 2010a, Sperandio *et al.*, 2008).

LptD is an outer membrane protein that forms a stable complex with LptE, which interacts specifically with LPS. It is thought that LptE receives LPS and LptD is required for LPS assembly in the outer membrane (Chng *et al.*, 2010b). The proper assembly of the LptD/E outer membrane complex is required for the disulphide bond formation in LptD and the subsequent assembly of the trans-envelope bridge. After the LptD/E complex forms, the N-terminal domain of LptD interacts with C-terminal domain of LptA, a periplasmic protein that specifically binds to the lipid A domain. It is thought that LptA may shield the hydrophobic lipid A from the aqueous environment of the periplasm, facilitating the transfer from the inner to the outer membrane (Tran *et al.*, 2008). The N-terminal domain of LptA also contacts the C-terminus of LptC to create a continuous bridge of anti-parallel β -barrel sheets that connect the inner and outer membrane (Freinkman *et al.*, 2012). Interestingly, only the periplasmic region of LptC is required to associate with the inner membrane LptBFG complex (Villa *et al.*, 2013), an

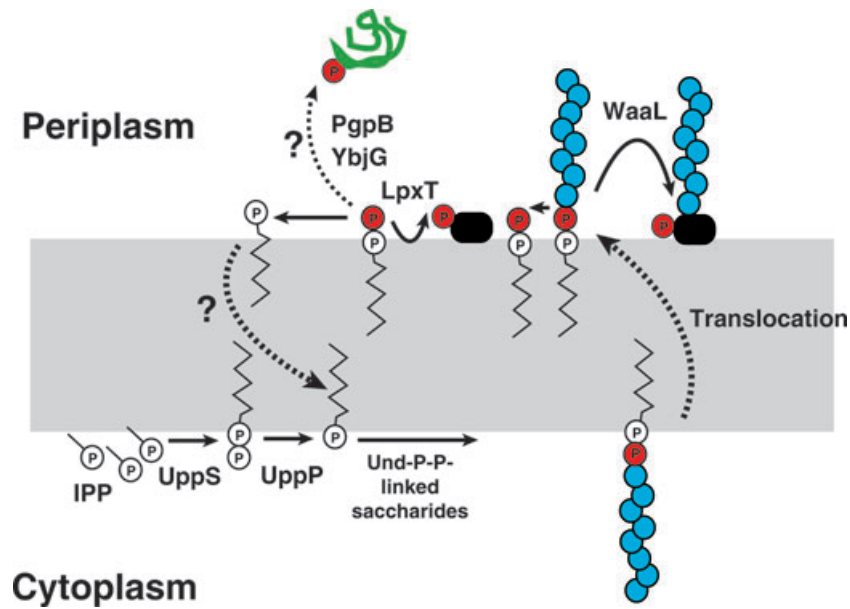


Figure 1.6. Model for Und-PP recycling in *E. coli*. The blue circles represent O Ag units and the black square represents lipid A-core OS. Phosphates are shown in red. Figure from (Valvano, 2008).

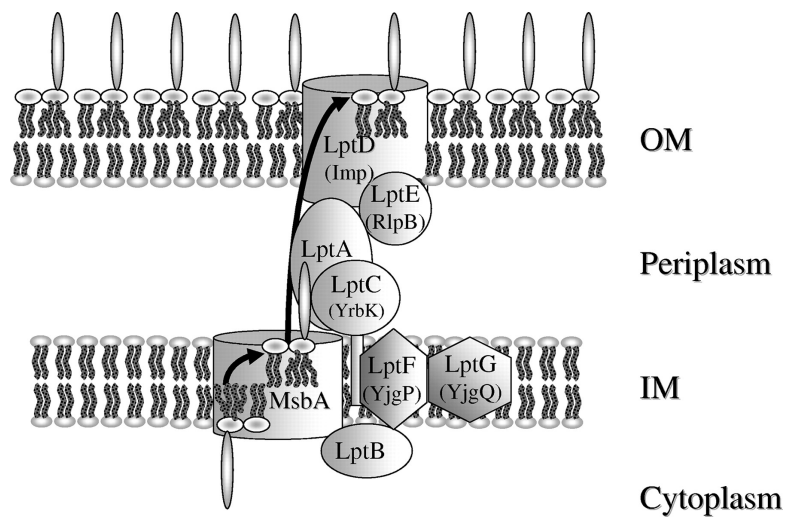


Figure 1.7. Model of LPS export from the inner membrane to the outer membrane.

Figure from (Sperandeo *et al.*, 2008).

ABC transporter complex that provides energy to extract the LPS from the inner membrane and transport it across the membrane (Narita and Tokuda, 2009). It is critical that the structure of LptC is maintained to have an interaction with both the LptBFG complex and LptA (Villa *et al.*, 2013).

Prediction modeling programs reveal that the periplasmic loops of LptF and LptG, LptA, LptC, and the terminal region of LptD all contain a jellyroll fold, which Villa *et al* call an “lpt fold”, a fold shared by LPS export machinery proteins. This common fold may serve as a structural element for protein-protein interactions in the trans-envelope bridge. These studies provide evidence that there is regulation of the assembly of LPS export machinery and that there may be a continuous path between the inner and outer membrane for export of LPS to the outer membrane (Villa *et al.*, 2013).

1.2 Initiation of the synthesis of cell surface polysaccharides

Two major families of enzymes initiate the synthesis of cell surface polysaccharides: Polyisoprenyl-phosphate *N*-acetylaminosugar-1-Phosphate Transferases (PNPTs) and Polyisoprenyl-phosphate Hexose-1-Phosphate Transferases (PHPTs). These initiation enzymes catalyze the formation of a phosphoanhydride bond between sugar-1-phosphate, from a uridine diphosphate (UDP)-sugar, and an isoprenoid lipid carrier resulting in the release of uridine monophosphate (UMP) (Valvano, 2003, Valvano *et al.*, 2011).

1.2.1 Polyisoprenyl-phosphate *N*-acetylaminosugar-1-Phosphate Transferases

Polyisoprenyl-phosphate *N*-acetylaminosugar-1-Phosphate Transferases initiate the synthesis of bacterial polysaccharides and eukaryotic N-linked protein glycosylation by

transferring the nucleotide-sugar UDP-GlcNAc onto a lipid carrier. This is a highly conserved process shared between eukaryotes and prokaryotes (Valvano, 2003). These enzymes are integral membrane proteins that share a number of conserved amino acids and functional motifs in their cytosolic loops (see 1.2.1.4 Functional motifs of PHPTs).

1.2.1.1 Bacterial PNPT initiated lipid-linked glycans

The bacterial PNPTs initiate a number of lipid-linked glycans that are involved in the synthesis of O Ag, enterobacterial common antigen, and peptidoglycan. Bacterial PNPTs transfer *N*-acetylaminosugars to the bacterial lipid carrier, Und-P (Valvano, 2003).

1.2.1.1.1 O antigen

The PNPT family member WecA is an inner membrane enzyme that is encoded by *rfe* in the enterobacterial gene cluster (Meier and Mayer, 1985). WecA initiates the synthesis of homopolymeric and heteropolymeric O Ag in a number of *Enterobacteriaceae* members, such as *E. coli*, *Shigella*, and *Klebsiella* (Kido *et al.*, 1995, Clarke *et al.*, 1995, Klena and Schnaitman, 1993, Alexander and Valvano, 1994). WecA transfers GlcNAc-1-P to Und-P to form Und-P-P-GlcNAc, which serves as the first step for the growing O Ag chain (see 1.1.1.3.1 O antigen biosynthesis) and enterobacterial common antigen unit (see 1.2.1.1.3 Enterobacterial common antigen) (Alexander and Valvano, 1994).

1.2.1.1.2 *Aeromonas hydrophila* O Antigen

The biosynthesis of *Aeromonas hydrophila* AH-3 O Ag is Wzy-dependent; however, a unique PNPT family member, WecP, initiates the biosynthesis. WecP transfers GalNAc-1-P from UDP-GalNAc to the lipid carrier, Und-P. This enzyme was named WecP as it is

an *N*-acetylaminosugar-1-P transferase, like WecA, but has a similar predicted topology to the *S. enterica*, WbaP. WecP is the first known PNPT family member that can transfer GalNAc-1-P (Merino *et al.*, 2011).

1.2.1.1.3 Enterobacterial common antigen

Enterobacterial common antigen (ECA) is a cell surface polysaccharide that is localized in the outer membrane of most members of the *Enterobacteriaceae* family (Lugowski and Romanowska, 1983, Rinno *et al.*, 1980). ECA is a linear chain comprised of a trisaccharide repeating unit of *N*-acetyl-D-mannosaminuronicacid (ManNAcA), *N*-acetyl-D-glucosamine (GlcNAc), 4-acetamido-4,6-dideoxy-D-galactose (Fuc4NAc) (Lugowski and Romanowska, 1983). There are three major forms of ECA: ECA_{PG}, ECA_{LPS}, ECA_{CYC}, (Dell *et al.*, 1984, Kajimura *et al.*, 2005). Phosphoglyceride linked ECA, ECA_{PG}, is produced by all *Enterobacteriaceae*. The ECA polysaccharide chain is covalently linked to diacylglycerol through a phosphodiester bond with the GlcNAc residue of the first ECA unit (Kuhn *et al.*, 1983). ECA_{LPS} is immunogenic and is covalently attached to the LPS core OS that lack the O Ag chain (Rick *et al.*, 1985). Some members of the *Enterobacteriaceae* family, including *E. coli* K-12, also produce a cyclic form of ECA (ECA_{CYC}) found in the periplasmic space (Kajimura *et al.*, 2005, Dell *et al.*, 1984). The repeating units of ECA_{CYC} are synthesized in the same pathway as immunogenic ECA, which also requires the modulator of chain length, WzzE. However details of the rest of synthesis and the function of ECA_{CYC} have yet to be elucidated (Kajimura *et al.*, 2005).

The biosynthetic pathway for ECA_{PG} is illustrated in Figure 1.8 (Kajimura *et al.*, 2005). The synthesis of ECA is initiated by WecA, which transfers GlcNAc-1-P to Und-P

with the release of UMP to form lipid I (Meier-Dieter *et al.*, 1990, Barr and Rick, 1987). WecG adds ManNAcA from a nucleotide precursor to Und-P-P-GlcNAc with the release of UDP to form lipid II (Barr and Rick, 1987). WecF transfers Fuc4NAc to lipid II from TDP-Fuc4NAc to form lipid III (Rahman *et al.*, 2001). It is not well understood, but genetic evidence suggests that lipid III may be translocated by WzxE “flippase” and polymerized into a polymeric chain by WzyE, with chain length regulation by WzzE. The ECA chain may be transferred to a lipid acceptor and exported to the outer membrane (Erbel *et al.*, 2003, Rick *et al.*, 1998). However, there is no concrete evidence that this is the synthesis pathway and it must be validated experimentally.

1.2.1.1.4 Peptidoglycan

Peptidoglycan is present in both Gram-positive and Gram-negative bacteria and plays an important role in maintaining cell shape and structural rigidity (Silhavy *et al.*, 2010, Lovering *et al.*, 2012). The peptidoglycan cell wall is composed of alternating *N*-acetylmuramic acid (MurNAc) and GlcNAc residues with peptide crosslinking between them, which provides rigidity and resistance against osmotic pressure. The synthesis of peptidoglycan is quite complex, with over 20 different enzyme reactions involved.

Peptidoglycan synthesis commences with the synthesis of a pentapeptide on UDP-MurNAc by a series of enzymes in the cytoplasm (Lovering *et al.*, 2012). Then the PNPT enzyme *MraY* commits the first membrane associated step by transferring phospho-MurNAc-pentapeptide to the lipid carrier, Und-P to form Und-P-P-MurNAc-pentapeptide

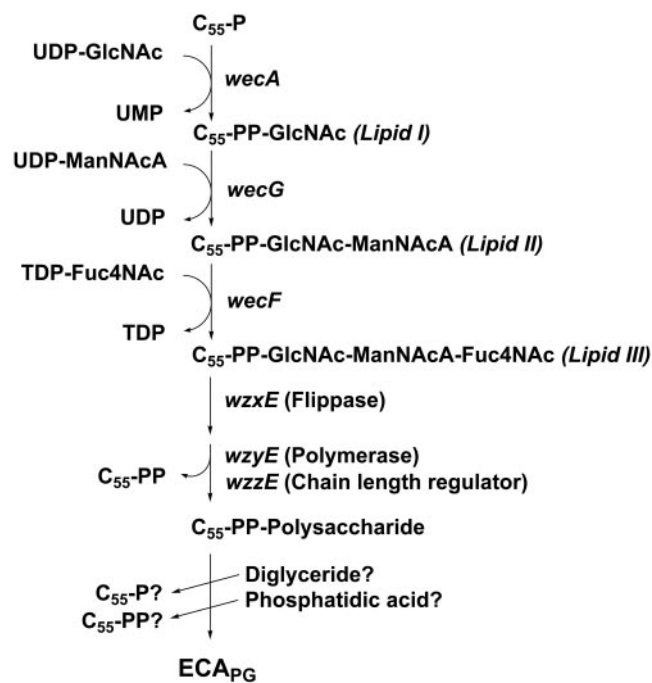


Figure 1.8. Proposed biosynthesis pathway of ECA_{PG}. Figure is unmodified from (Kajimura *et al.*, 2005).

to form lipid intermediate I (Ikeda *et al.*, 1991). The reaction mechanisms for MraY have been disputed with conflicting evidence between a one-step and two-step mechanism (see 1.2.1.5. Proposed catalytic mechanisms) (Heydanek *et al.*, 1969, Lloyd *et al.*, 2004, Al-Dabbagh *et al.*, 2008). Afterwards, MurG forms the lipid intermediate II (Und-P-P-MurNAc(pentapeptide)GlcNAc) by adding GlcNAc from UDP-GlcNAc to lipid I (Mengin-Lecreulx *et al.*, 1991). Using a bioinformatics approach, a putative lipid II flippase was identified, MurJ (MviN). This enzyme is thought to translocate the lipid II from the cytoplasmic face of the inner membrane to the periplasmic space (Ruiz, 2008). Penicillin binding proteins (PBPs) then crosslink the peptide side chains to form a three-dimensional structure (Ruiz, 2008, Popham and Young, 2003).

1.2.1.2 Eukaryotic PNPT initiated lipid-linked glycans

The biogenesis of lipid-linked glycans is a conserved process in bacteria and eukaryotes. Eukaryotes utilize lipid-linked glycans for asparagine (N)-linked protein glycosylation (Valvano, 2003). N-linked protein glycosylation involves the transfer of a 14-residue oligosaccharide to asparagines on nascent polypeptides in the rough endoplasmic reticulum (Lehrman, 1991). GlcNAc-1-P transferases (GPTs) are multiple transmembrane domain eukaryotic PNPTs that initiate the synthesis of the 14-residue oligosaccharide in the RER (Lehrman, 1991, Dan *et al.*, 1996). Like other PNPTs, GPTs are also sensitive to the antibiotic tunicamycin (Lehrman, 1991). GPTs transfer GlcNAc-1-P to the lipid carrier, dolichol-phosphate (Dol-P) to form Dol-P-P-GlcNAc, the first committed step of N-linked protein glycosylation. Afterwards, glycosyltransferases complete the 14-residue oligosaccharide, comprised of two GlcNAc,

nine mannoses (Man), and three glucoses (Glc) (Lehrman, 1991, Elbein, 1987).

1.2.1.3 Topology of PNPTs

In silico prediction programs suggest that PNPTs contain ten to eleven transmembrane domains, with five cytoplasmic loops that contain several highly conserved amino acids (Bouhss *et al.*, 1999, Lehrer *et al.*, 2007, Valvano, 2003). Topological analysis of the *E. coli* and *S. aureus* MraYs, enzymes that initiate the synthesis of peptidoglycan, were studied using β -lactamase fusions. Ten β -lactamase fusions were constructed in each enzyme and expressed in cells to assess for ampicillin resistance. Expression of β -lactamase in the periplasmic space would hydrolyze the antibiotic, rendering the cells resistant. The topological results were relatively similar to the *in silico* predicted topologies with ten transmembrane domains and five cytosolic loops (Bouhss *et al.*, 1999). The topology of another PNPT family member, the *E. coli* WecA that initiates the synthesis of O Ag, was studied using a substituted cysteine labeling technique. Novel cysteines were introduced into a cysteine-less WecA protein and labeled with a membrane permeable sulfhydryl reactive reagent with or without the pre-treatment of a membrane impermeable blocking reagent. The topological analysis revealed that the *E. coli* WecA has eleven transmembrane domains and five cytosolic loop domains (Fig. 1.9). A green fluorescent protein (GFP) was fused to the C-terminus of WecA to independently confirm that the C-terminus was cytoplasmic (Lehrer *et al.*, 2007). Although the topology of the *E. coli* WecA has been experimentally refined (Lehrer *et al.*, 2007), there are still some regions that need to be investigated. This includes the highly conserved V/IFMGD motif, which is conserved in prokaryotes and eukaryotes (Price and Momany, 2005).

1.2.1.4 Functional motifs of PNPTs

A number of highly conserved functional motifs have been identified in the cytosolic loop domains II, III, IV, and V of PNPTs: DDxxD, NxxNxxDGIDGL, V/IFMGD, and the carbohydrate recognition domain (Fig. 1.9). These motifs have roles in coordinating the metal cofactors, nucleotide sugar substrates, or serve as potential catalytic nucleophiles.

DDxxD motif

The DDxxD motif resides in the cytosolic loop region II of PNPT members. This domain was originally hypothesized to contribute to coordination of a metal cofactor because of its similarity to the aspartate-rich motif found in prenyl transferases, which bind to diphosphate-containing substrates via Mg^{2+} bridges (Ashby and Edwards, 1990, Tarshis *et al.*, 1994). Mutational analysis of the DD residues in the *E. coli* WecA, MraY, and *B. subtilis* MraY resulted in decreased enzymatic activity but only the D₉₀ of D₉₀D₉₁xxD₉₄ in the *E. coli* WecA was found to contribute to Mg^{2+} cofactor binding (Lehrer *et al.*, 2007).

V/IFMGD motif

The V/IFMGD motif is a highly conserved motif in eukaryotic and prokaryotic PNPTs (Price and Momany, 2005). Although this motif's predicted location is inside the membrane domain after cytosolic loop IV, it has been speculated that this aspartic acid residue may contribute to catalysis. Lloyd *et al* suggested that this residue in MraY

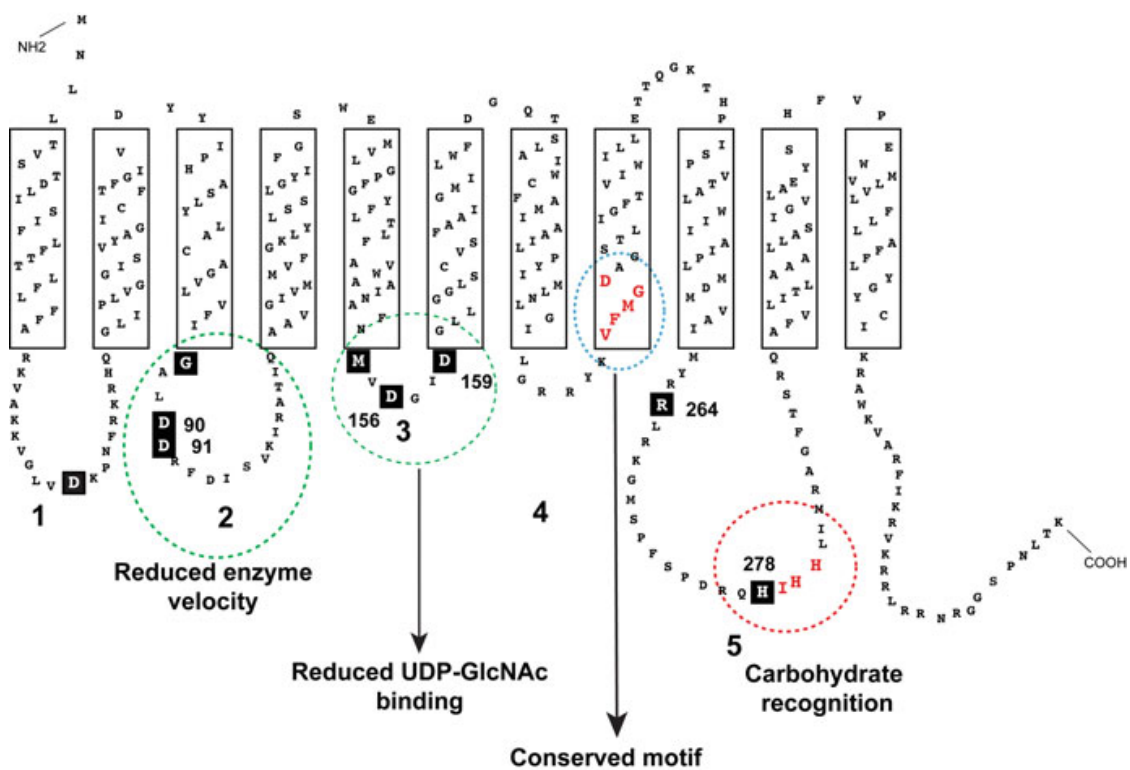


Figure 1.9. Topological model of the *E. coli* WecA. Residues important for function are shown in black boxes. The functional motifs are shown in dotted circles. Figure from (Valvano *et al.*, 2011).

contributes as a catalytic nucleophile (see 1.2.1.5. Proposed catalytic mechanisms) (Lloyd *et al.*, 2004), however no mass spectral intermediate-trapping experiments have been performed to confirm this and there has been no topological evidence to confirm that this motif does not reside in the membrane. Due to the high conservation of this motif, it is tempting to speculate that this region may contribute to enzyme function. Therefore it is important to elucidate the role and refine the topology of this region.

Carbohydrate Recognition Domain

PNPTs contain a proposed carbohydrate recognition (CR) domain in cytosolic loop V. This domain is highly basic, with a *pI* of 11.0, which suggests that this region may interact with an acidic ligand, such as a sugar substrate (Anderson *et al.*, 2000). This domain also resembles the HIGH motif that is found in class I aminoacyl-tRNA synthetases (Sekine *et al.*, 2001) and another superfamily of related nucleotidyltransferases that contributes to phosphodiesterase activity and is postulated to contribute to the structure of active site and nucleotide binding (Venkatachalam *et al.*, 1999). Amer and Valvano postulated that the HHH motif in the *E. coli* WecA contributes to the binding of the nucleotide substrate, UDP-GlcNAc. When the highly conserved R₂₆₅ and H₂₇₉HH₂₈₂ motif in the *E. coli* WecA was replaced with non-charged amino acids, the enzyme could not complement O7 antigen synthesis *in vivo* and had no *in vitro* transferase activity. These authors further support the notion that R₂₆₅ and the HHH motif contributes to nucleotide substrate binding by showing that amino acid replacements of H₂₇₉, R₂₆₅, and the HHH motif lose the ability to bind tunicamycin (Amer and Valvano, 2001), the nucleoside antibiotic that resembles the UDP-GlcNAc-polyisoprenoid lipid reaction intermediate (Heifetz *et al.*, 1979).

1.2.1.5 Proposed catalytic mechanisms of PNPTs

The PNPTs share a number of conserved domains and residues and therefore it is conceivable that they share a common enzymatic mechanism. There are two mechanisms proposed for the *Bacillus subtilis* and *E. coli* *MraY* enzymes: the two-step and the one-step.

The two-step (double displacement) mechanism suggests that the aspartic acid residue of the highly conserved V/IFMGD motif acts as a catalytic nucleophile and forms a covalent acyl-enzyme intermediate with the UDP-sugar substrate, which results in the release of UMP (Heydanek *et al.*, 1969, Lloyd *et al.*, 2004). Lloyd *et al* proposed that the aspartic acid residue in this motif is likely the catalytic nucleophile because it is strictly conserved amongst prokaryotic and eukaryotic homologs (Lloyd *et al.*, 2004), however prediction programs place this motif inside transmembrane domain VIII (Price and Momany, 2005). Lloyd *et al* observed the formation of a covalent intermediate after gel filtration of purified *MraY* protein and radiolabeled substrate, MurNAc-pentapeptide (Lloyd *et al.*, 2004). The second step of this reaction involves the nucleophilic attack on the covalent intermediate by the phosphate group on Und-P, resulting in the formation of phosphoanhydride bond between Und-P and the sugar-1-P. Despite the observation of a covalent intermediate, the catalytic nucleophile has not been unequivocally determined (Lloyd *et al.*, 2004, Heydanek *et al.*, 1969).

The one-step mechanism suggests that a single conserved aspartic acid residue in the DDxxD motif (in cytosolic loop II) in *B. subtilis* *MraY* removes a hydrogen atom from the phosphate group of Und-P, which would allow the subsequent nucleophilic

attack on the UDP-sugar substrate. This residue was postulated to contribute to the removal of the hydrogen atom because after mutational analysis of *B. subtilis* MraY, this was the only mutant that restored some enzymatic activity after increasing the pH. These data suggest that this aspartic acid residue serves to deprotonate the Und-P (Al-Dabbagh *et al.*, 2008). However, this model is inconsistent with the notion that this highly conserved aspartic acid residue in the DDxxD motif contributes to coordinating the Mg²⁺ cofactor in another PNPT family member, WecA (Lehrer *et al.*, 2007). It is also inconsistent with the previously described two-step mechanism identified in the *E. coli* MraY enzyme (Lloyd *et al.*, 2004).

To date there are two potential aspartic acid residues that could serve as catalytic nucleophiles: the first aspartic acid residue of the NxxNxxDGIDGL motif, in predicted cytosolic loop III of the bacterial members of the family, and the aspartic acid residue of the V/IFMGD, in the eighth transmembrane domain of PNPTs. It is tempting to speculate that the residue from the NxxNxxDGIDGL motif acts as a nucleophile because in the *E. coli* WecA, D156 is the only conserved, cytosolic exposed aspartic acid residue that when replaced has abolished enzymatic activity *in vivo* and *in vitro* and loses the ability to bind tunicamycin (Amer and Valvano, 2002, Lehrer *et al.*, 2007). It is possible that the aspartic acid residue in the V/IFMGD motif may serve this function as replacement of this residue in the *E. coli* MraY lost *in vitro* enzymatic activity after replacement with asparagine (Lloyd *et al.*, 2004). However, to identify the catalytic nucleophile, it will be important to further examine these motifs in other PNPTs.

1.2.2 Polyisoprenyl-phosphate Hexose-1-Phosphate Transferases

Polyisoprenyl-phosphate Hexose-1-Phosphate Transferases (PHPTs) are generally large, basic proteins with multiple transmembrane spanning domains and a highly conserved C-terminal catalytic domain (Valvano, 2003, Saldías *et al.*, 2008, Patel *et al.*, 2010, Patel *et al.*, 2012b). These enzymes are involved in the synthesis of various bacterial polysaccharides, including O Ag, LPS, exopolysaccharides, capsules and N-linked protein glycosylation (Valvano, 2003, Saldías *et al.*, 2008).

1.2.2.1 PHPT initiated lipid-linked glycans

1.2.2.1.1 *Salmonella enterica* O antigen

The *S. enterica* O Ag unit contains three to six sugar residues and the number and nature of sugars vary depending on the O serotype. For example, serotypes B and E1 share the backbone mannose-rhamnose-galactose, with differences in the addition of abequose in serotype B and the linkage between mannose and rhamnose. This variation is given rise by the different types of nucleotide sugar biosynthesis pathways and glycosyltransferases. However, all *S. enterica* serotypes have a gene encoding WbaP, which transfers Gal-1-P from UDP-Gal to Und-P (Liu *et al.*, 1993). This serves as an acceptor substrate for subsequent glycosyltransferases to complete the O Ag unit (see 1.1.1.3.1 O antigen biosynthesis).

1.2.2.1.2 Colanic acid

Colanic acid (CA) or M antigen, is an exopolysaccharide produced by many strains of *Enterobacteriaceae* (Grant *et al.*, 1969). Colanic acid (CA) or M antigen is a high

molecular weight extracellular polysaccharide is critical for the formation of *E. coli* biofilm, resistance to desiccation, cell stress responses, protection against acid and heat (Whitfield, 2006, Danese *et al.*, 2000, Grant *et al.*, 1969, Mao *et al.*, 2001). CA has a common structure amongst members that produce it and the structure resembles group 1-capsules (Grant *et al.*, 1969, Whitfield, 2006). However, the expression of CA differs from group 1 capsule as it cannot be produced at 37°C, which is likely related to its regulation. Unlike capsular polysaccharides, CA is loosely associated with the outer membrane, but under specific environmental stimuli units of CA can be attached to the LPS core OS, called M_{LPS} (Meredith *et al.*, 2007).

CA is produced under conditions of osmotic shock and stress or damage to the cell envelope. The production of colanic acid is regulated by the *rca* (regulator of capsule synthesis) genes *rcaA*, *rcaB*, and *rcaC*. RcsA and RcsB positive transcriptional regulators and are required for maximal CA production (Gottesman and Stout, 1991). RcsA is very unstable *in vivo* as it is degraded by Lon protease (Stout *et al.*, 1991).

The CA biosynthesis genes reside in the CA gene cluster, which is composed of 19 genes, including genes for the synthesis of nucleotide-sugar precursors required for CA production such as GDP-L-fucose. Stevenson *et al* propose that CA is synthesized in a pathway similar to the Wzy-dependent O Ag synthesis pathway (Fig. 1.10). CA is produced by assembling repeating E units of D-glucose, D-fucose, D-galactose, D-glucuronic acid, and pyruvate linked to galactose, and O-acetylation to the first fucose (Fig. 1.11) (Stevenson *et al.*, 1996, Sutherland, 1969). Like O Ag units, E units are synthesized on the lipid carrier Und-P (Stevenson *et al.*, 1996). WcaJ initiates the synthesis of E units by transferring Glc-1-P to Und-P to form Und-P-P-Glc (Patel *et al.*,

2012b). It is thought that WcaA, WcaC, WcaE, WcaL, and WcaI serve as glycosyltransferases to complete the E unit. WcaB and WcaF both encode acetyltransferases, although it is unclear which one adds the acetyl group to the first fucose of the E unit. Stevenson *et al* have called *orf17* Wzx and propose that it acts as a flippase to translocate the E unit across the inner membrane. These authors have also proposed that WcaD acts as an E unit polymerase. However, despite the annotations, these functions have to be experimentally confirmed (Stevenson *et al.*, 1996). Gene products encoded by *wza*, *wzb*, and *wzc* most likely export CA since they are homologous to similar genes associated with the export of EPS in *Pseudomonas spp.* and capsular polysaccharide in *Klebsiella pneumoniae* (Stevenson *et al.*, 1996).

1.2.2.1.3 Capsule

Capsules, also known as K antigen, are high molecular weight polysaccharides that are firmly attached to the cell surface. Capsules are important virulence factors as they contribute to resistance against complement-mediated killing and opsonophagocytosis. There are over 80 types of capsules in *E. coli* and are divided into four major groups: 1, 2, 3, and 4 (Whitfield, 2006).

Group 1 and 4 capsules are found in *E. coli* isolates that cause intestinal infections. *E. coli* group 1 capsules are similar to *K. pneumoniae* capsules and CA (see 1.2.2.1.2 Colanic acid). Group 4 capsules have a more diverse structure with no known similarities to other capsules. Group 1 and 4 capsules can be attached to the cell surface in two forms: K_{LPS} and capsular K antigen (high molecular weight capsule). K_{LPS} is attached to the lipid A-core OS and it is unknown how capsular K antigen is attached to

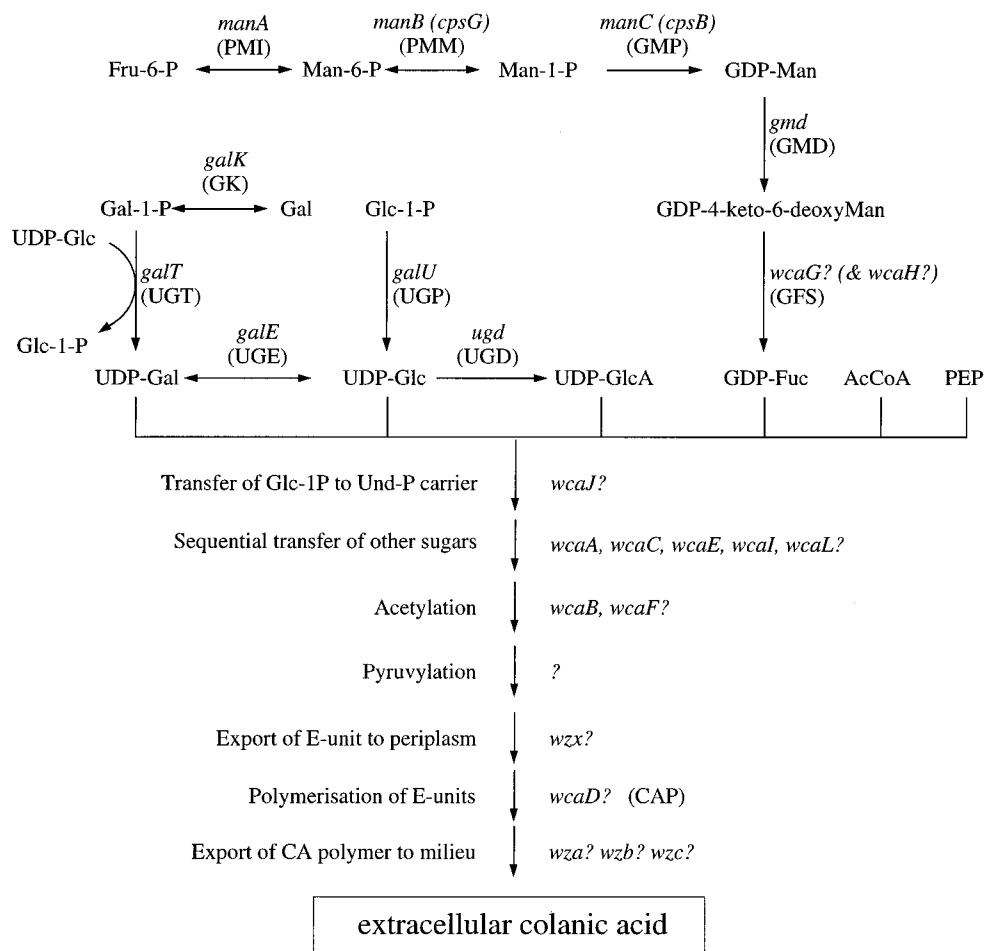


Figure 1.10. Proposed biosynthesis pathway for colanic acid. Abbreviations: AcCoA, acetyl coenzyme A; CAP, colanic acid polymerase; Fru, fructose; GFS, GDP-L-fucose synthetase; GK, galactokinase; GLK, glucokinase; GMP, GDP-D-mannose pyrophosphorylase; GMD, GDP-D-mannose dehydratase; Man, D-mannose; PEP, phosphoenolpyruvate; PGI, phosphoglucose isomerase; PGM, phosphoglucomutase; PMI phosphomannose isomerase; PMM, phosphomannomutase; UGD, UDP-D-glucose dehydrogenase; UGE, UDP-D-galactose-4-epimerase; UGP, UDP-D-glucose pyrophosphorylase; UGT, UTP-D-galactose-1-phosphate uridylyl transferase. Figure unmodified from (Stevenson *et al.*, 1996).

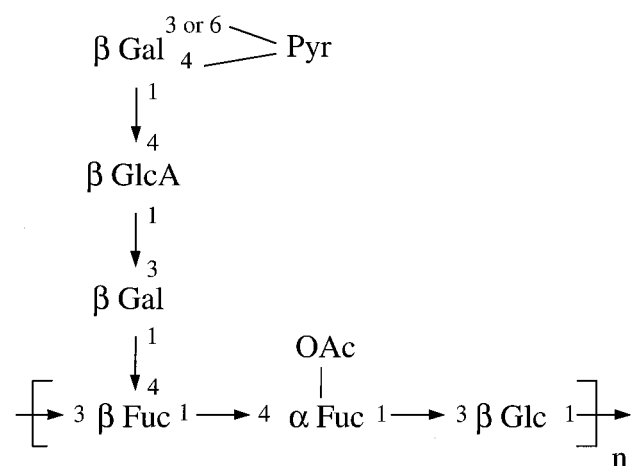
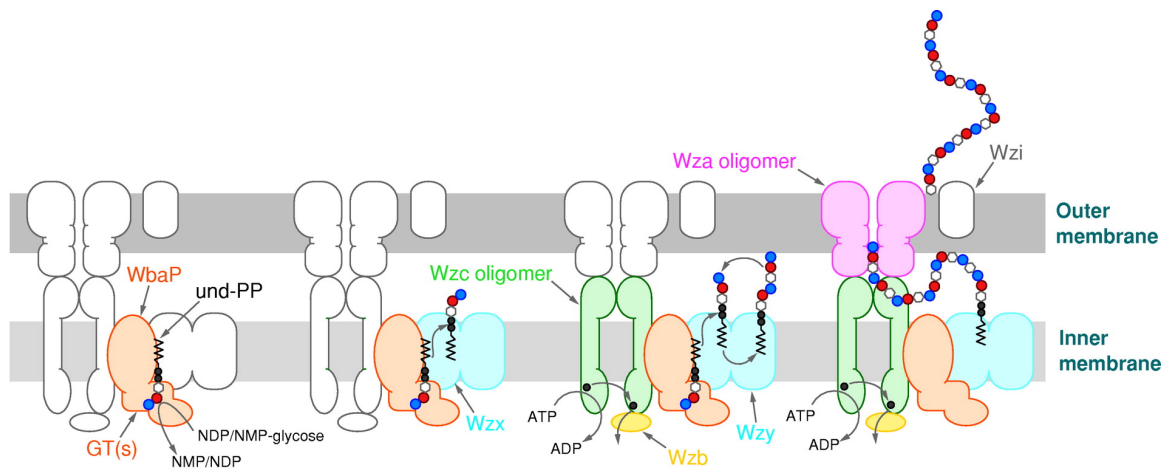


Figure 1.11. *E. coli* K-12 colanic acid repeating structure. Figure from (Stevenson *et al.*, 1996).

the cell surface (MacLachlan *et al.*, 1993). Group 2 and 3 K antigens are found in bacteria that produce extraintestinal infections. These antigens are structurally diverse with some similarities to capsules in *Neisseria meningitidis* and *Haemophilus influenza* (Whitfield, 2006).

The biosynthesis of Group 1 and 4 K antigens is Wzy-dependent, which is initiated by the transfer of a sugar-1-P to the lipid carrier, Und-P (Fig. 1.12). The PHPT family member, WbaP, initiates the synthesis of group 1 K antigens by transferring Gal-1-P to Und-P (Whitfield and Roberts, 1999) and group 4 K antigens are initiated by WecA, which transfers GlcNAc-1-P to Und-P (Amor and Whitfield, 1997, Meier-Dieter *et al.*, 1990). The K antigen unit is polymerized on the lipid-linked sugar on the cytoplasmic face of the inner membrane. Wzx translocates this unit across the inner membrane and Wzy polymerizes the units into a polymeric chain. At this point, K antigen can be ligated to lipid A-core OS by WaaL to form K_{LPS}, or continue to produce high molecular weight K antigen. In group 4 K_{LPS}, the length of the chain is regulated by Wzz, which is absent in group 1 K_{LPS} (Amor and Whitfield, 1997, Whitfield, 2006). Wzc regulates group 1 and 4 capsule length (Whitfield, 2006). Wza, an outer membrane lipoprotein, is essential for the surface expression of group 1 capsules. Wzi is unique to group 1 capsule biosynthesis and is postulated to play a role in attaching the capsule to the cell surface (Rahn *et al.*, 2003).

Group 2 and 3 K antigens are synthesized in an ABC transporter-dependent pathway (Fig. 1.13) (Whitfield, 2006). Processive glycosyltransferases build the polysaccharide chain, which is translocated across the inner membrane by an ABC




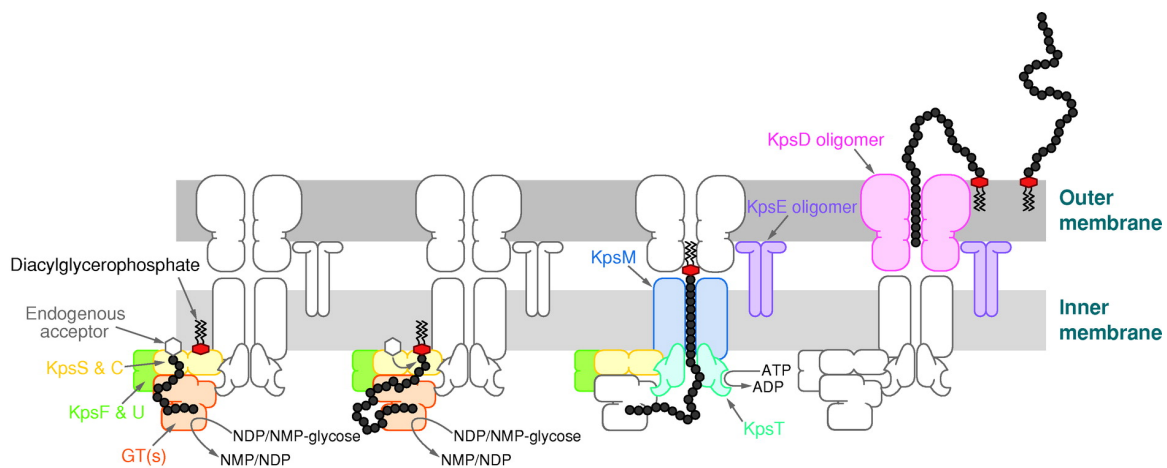
 Whitfield C. 2006.
Annu. Rev. Biochem. 75:39–68

Figure 1.12. Model for the biosynthesis of group 1 and 4 capsules. The Und-P-P-sugar repeat units are synthesized on the cytoplasmic face of the inner membrane, which are translocated across the membrane by Wzx. The repeat units are polymerized into a chain on Und-P by the Wzy polymerase. Wzc controls the high-level polymerization of group 1 capsular polysaccharides and colanic acid. Wza translocates group 1 capsular polysaccharides across the outer membrane. Wzi is unique to group 1 capsule biosynthesis and contributes to the final stages of capsule assembly on the outer membrane. Figure from (Whitfield, 2006).



R Whitfield C. 2006.
Annu. Rev. Biochem. 75:39–68

Figure 1.13. Model for the biosynthesis of group 2 and 3 capsules. The capsule polymer is synthesized on an unknown acceptor (open hexagon) by glycosyltransferases on the cytoplasmic face of the inner membrane. The polymeric chain is ligated to a diacylglycerophosphate (or diacylglycerophosphate-Kdo). The polymer is exported to the outer membrane by the ABC-transporter, KpsM and T. KpsS,C, F, and U also participate in the export of the polymer, but their functions remain unclear. KpsE and D provide the membrane-fusion (adaptor) protein and OMA protein functions. Figure from (Whitfield, 2006).

transporter (KpsM and KpsT) (Nsahlai and Silver, 2003, Pavelka *et al.*, 1991, Bliss and Silver, 1996). The export of the polymer to the cell surface is mediated by KpsD and KpsE (Arrecubieta *et al.*, 2001).

1.2.2.1.4 S-layer

S-layers are planar, ordered, crystalline, lattice-like structures found on archaeal and bacterial (both Gram-positive and Gram-negative) cell surfaces (Sleytr and Beveridge, 1999). The homogenous, proteinaceous subunits (protein or glycoprotein) of S-layers self-assemble on the cell surface in an entropy driven process (Sleytr and Messner, 1983, Sleytr and Beveridge, 1999). S-layers can contribute to adhesion, protect the bacteria, and act as a scaffold to hold virulence factors and enzymes (Sleytr and Beveridge, 1999).

The synthesis of S-layer glycoproteins is a highly complex process and has been characterized in *Geobacillus stearothermophilus*. The S-layer protein glycosylation pathway is thought to be analogous to ABC transporter-dependent O Ag biosynthesis, as many S-layer producing bacteria have an ABC transporter in their S-layer protein glycosylation gene clusters. The glycosylation pathway is initiated by the PHPT family member, WsaP, which transfers Gal-1-P from UDP-Gal to a lipid carrier in the inner membrane (Steiner *et al.*, 2007). L-rhamnose is then added to the lipid-linked intermediate by rhamnosyltransferases WsaC, WsaD, and WsaF to form a chain. The methylrhamnosyltransferase, WsaE terminates chain growth by adding an O-methyl group to the terminal rhamnose. The ABC transporter exports the polymer, which is linked to the S-layer protein by WsaB, an oligosaccharyl::protein transferase (Steiner *et al.*, 2008).

1.2.2.2 Topology of PHPTs

PHPT family members can vary in amino acid length and the number of predicted transmembrane domains, however they all share at least one membrane domain, presumably for the enzyme to associate with the lipid acceptor and a cytoplasmic C-terminal domain, where the enzymatic activity resides (Patel *et al.*, 2012b, Wang *et al.*, 1996, Patel *et al.*, 2012a, Saldías *et al.*, 2008).

Larger PHPTs, such as *S. enterica*, WbaP, contain three distinct domains: an N-terminal domain, containing four of the five predicted transmembrane domains, a large predicted periplasmic loop domain, between TM IV and TM V and a C-terminal catalytic domain, containing TM V and the long C-terminal tail (Saldías *et al.*, 2008) (Fig. 1.14). The N-terminal and periplasmic loop domains are conserved in some *E. coli*, *Salmonella*, *Acinetobacter*, and *Haemophilus influenzae* PHPTs that are involved in the biosynthesis of LPS and O Ag. The C-terminal catalytic domain is located in the cytoplasmic space, as confirmed by Green fluorescent protein (GFP) fusion to the tail of WbaP_{CT}, and is highly conserved amongst a large number of PHPTs, providing a hallmark trait for this family of enzymes (Valvano *et al.*, 2011, Patel *et al.*, 2010). Some shorter PHPTs only contain the last transmembrane domain and C-terminal tail, which include the holdfast synthesis enzymes in *C. crescentus* PssY and PssZ, and the N-linked protein glycosylation enzymes in *Campylobacter jejuni* PglB and PglC (Patel *et al.*, 2012b, Valvano *et al.*, 2011) (Fig. 1.15).

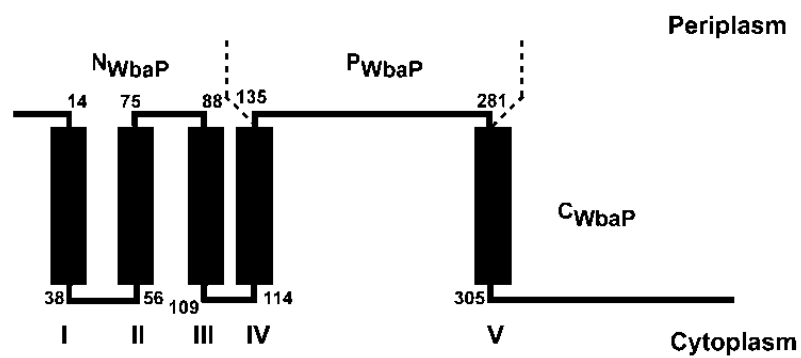


Figure 1.14. Predicted topology of *S. enterica* WbaP. WbaP has three distinct domains, the N-terminal (TMs I-IV), the larger periplasmic loop domain, and the C-terminal domain (TM V and cytosolic tail). Figure from (Saldías *et al.*, 2008).

To date there is only one topological analysis of a PHPT family member, BceB, an enzyme that initiates the synthesis of Cepacian EPS in *Burkholderia cenocepacia*. BceB has 7 transmembrane domains and lacks a distinct large periplasmic loop region and instead has two major cytosolic loop domains, which is inconsistent with previous topological predictions. However, a detailed analysis was not performed, as only 5 C-terminal LacZ and PhoA fusions were studied (Videira *et al.*, 2005). Interestingly, Patel *et al* suggested that the large periplasmic loop domain may reside in the cytoplasm as the N-terminal His₆-TrxA fusion partner on WbaP_{CT}(F258-Y476), was found to be cytoplasmic as it was inaccessible to the TEV protease in spheroplasts. However, it was unclear if this was due to the rapid folding of the TrxA fusion partner, or due to the topology of the C-terminal WbaP protein itself (Patel *et al.*, 2010). The notion that the periplasmic loop domain may indeed be cytoplasmic may be consistent with the fact that the loop contributes to regulating enzymatic activity *in vivo*, as it has been demonstrated that the loop domain in WbaP plays a role in regulating the O Ag chain length (Saldías *et al.*, 2008). It is clear that topology of PHPTs is not well understood and a more detailed analysis of these enzymes needs to be conducted.

1.2.2.3 Functional domains of PHPTs

Previous studies suggested that there were two distinct functional domains of the *Salmonella enterica* WbaP, a N-terminal domain (T block) and C-terminal domain (GT block) that was involved in flipping of the Und-P linked O Ag unit and the C-terminal domain of WbaP contained Gal-1-P transferase activity *in vitro*, as measured by the incorporation of radiolabeled Gal in lipid extractions (Wang *et al.*, 1996). Using pulse-chase experiments and chromatography, these authors demonstrated that T-block WbaP

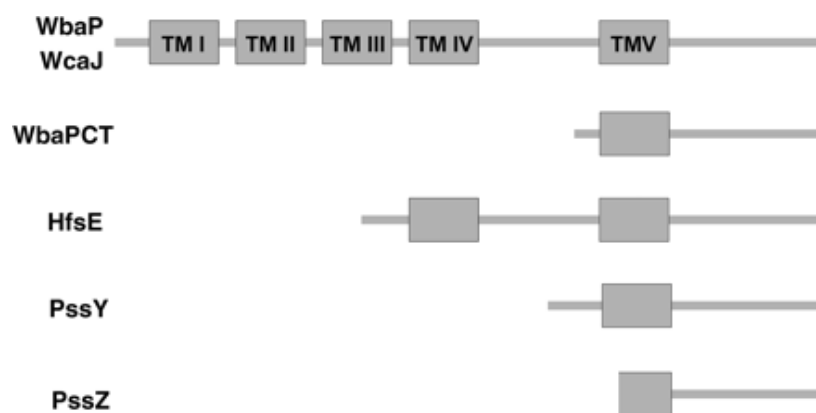


Figure 1.15. Alignment of predicted topologies of PHPT members: *S. enterica* WbaP, *E. coli* WcaJ, *C. crescentus* HfsE, PssY, and PssZ. Figure from (Patel *et al.*, 2012b).

mutants accumulated low-molecular weight Und-P-P linked to incomplete O units, rather than Und-P-P linked to full-length O Ag chains or single O units, which would suggest a block in either ligation or polymerization, respectively. Instead, Wang and Reeves suggested that the T-block (N-terminus) of the WbaP was concerned with flipping the Und-P-P-O unit from the cytoplasmic space to periplasmic space, where ligation to the lipid A-core OS would occur (Wang and Reeves, 1994). The authors admit that this notion required further analysis (Wang and Reeves, 1994) and later work by Saldías *et al* demonstrated that the N-terminal and large periplasmic loop domain of the *S. enterica* WbaP were dispensable for *in vivo* function as C-terminus alone could complement O Ag synthesis in a $\Delta wbaP$ strain, MSS2. Also, Saldías *et al* made the WbaP T block mutants referred to by Wang and Reeves and found that these mutations did not affect O Ag complementation in MSS2, suggesting that there may have been a deletion or deletions in other O Ag genes (Saldías *et al.*, 2008).

Saldías *et al* found that the large periplasmic loop domain was required for normal O Ag chain length distribution. When the WbaP_{M1-I144/F258-Y476} construct, lacking the loop domain, and WbaP_{Y250-Y476} protein, lacking the N-terminal domain and most of the periplasmic loop, was expressed in MSS2 ($\Delta wbaP$) strain, different O Ag chain length distributions were observed compared to the wild-type *Salmonella* Typhimurium strain LT2. A non-functional WbaP construct containing the periplasmic loop domain, domain V, and a truncated C-terminal tail, WbaP_{M170-R354}, was expressed in the wild-type *Salmonella* strain LT2. Expression of this construct reduced the amount of high molecular weight O Ag molecules, suggesting that the large periplasmic loop domain may modulate O Ag chain length by potentially interacting with enzymes downstream of

the O Ag biosynthesis assembly. This work demonstrated that the *S. enterica* WbaP has three distinct functional domains, N-terminal (TMs I-IV), the large periplasmic loop domain (between TMs IV and V), and the C-terminal domain (TM V and cytosolic tail) (Fig. 1.14) (Saldías *et al.*, 2008).

The C-terminal domain of the *S. enterica* WbaP is sufficient for enzymatic activity *in vivo* and *in vitro* (Saldías *et al.*, 2008, Wang *et al.*, 1996, Wang and Reeves, 1994) and sequence analysis of hundreds of WbaP homologs reveals that there are nine highly conserved, charged amino acids: three arginines (R319, R377, R401), two aspartic acids (D382 and D458), one lysine (K331) and one glutamic acid (E383). Alanine replacements of these residues reveal that seven of them are critical for enzymatic activity *in vivo* and *in vitro*. Four of these amino acids (R377, D331, E383, and D458) are predicted to be in three α -helices that are conserved in other WbaP homologs, suggesting that they may contribute to a catalytic center (Patel *et al.*, 2010).

The WbaP_{CT} construct has been purified using an N-terminal His₆-TrxA fusion partner, which greatly increased the stability and solubility of the protein. This further allowed the characterization of the enzyme, demonstrating that the C-terminal domain is highly specific for polyisoprenyl phosphate lipid acceptor, Und-P (C55-P) (Patel *et al.*, 2012a). Interestingly, the sequence in the transmembrane domain V resembles a 13-amino acid consensus sequence identified in other prokaryotic and eukaryotic glycosyltransferases (Patel *et al.*, 2012a, Zhou and Troy, 2003, Albright *et al.*, 1989).

1.3 Research objectives and summary of data obtained

The synthesis of lipid-linked glycans is a conserved process in eukaryotes and prokaryotes that is initiated by two major enzyme families: the polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPTs) and the polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferases (PNPTs). These enzymes contain multiple membrane domains and transfer a sugar-1-phosphate from a nucleotide sugar precursor to a lipid carrier. Although these enzyme families perform a similar enzymatic reaction, they have no amino acid sequence similarity and have distinct membrane topologies and catalytic domains. The goal of this work was to demonstrate that the PNPT and PHPT families of enzymes are distinct in their membrane protein topologies and their functional domains. **The objectives were to investigate the function and topology of a highly conserved functional domain in PNPTs, characterize the topology and investigate the role of the N-terminal domain of PHPTs.**

The prototypic PNPT member used in this study is the *E. coli* WecA, which initiates the synthesis of O Ag and ECA in *Enterobacteriaceae* by transferring GlcNAc-1-P to undecaprenyl phosphate (Und-P). The PHPT family member used in this study is the *E. coli* WcaJ, which transfers Glc-1-P to Und-P to initiate colanic acid synthesis. This work has three results chapters in alignment with the three proposed objectives:

1) Chapter Two demonstrates that the highly conserved VFMGD motif in PNPTs faces the cytosol and defines a region in PNPTs that contributes to the active site, likely involved in the binding and/or recognition of the nucleotide moiety of the nucleoside phosphate precursor.

2) Chapter Three provides the first detailed topological analysis of a PHPT member,

which is inverted compared to the *in silico* topological predictions; the N-terminus, C-terminal tail and the large soluble loop all reside in the cytoplasm. We also found that the last membrane domain does not fully span the membrane and is likely 'pinched in'.

3) Chapter Four shows evidence that the N-terminal domain (including the large cytoplasmic loop domain) likely contributes to the protein folding and/or stability of PHPT family members.

1.4 Chapter one references

- Abeyrathne, P. D., Daniels, C., Poon, K. K., Matewish, M. J., and Lam, J. S. (2005) Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. *J Bacteriol* **187**: 3002-3012.
- Abeyrathne, P. D., and Lam, J. S. (2007) WaaL of *Pseudomonas aeruginosa* utilizes ATP in in vitro ligation of O antigen onto lipid A-core. *Mol Microbiol* **65**: 1345-1359.
- Al-Dabbagh, B., Henry, X., El Ghachi, M., Auger, G., Blanot, D., Parquet, C., Mengin-Lecreulx, D., and Bouhss, A. (2008) Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. *Biochemistry* **47**: 8919-8928.
- Albright, C. F., Orlean, P., and Robbins, P. W. (1989) A 13-amino acid peptide in three yeast glycosyltransferases may be involved in dolichol recognition. *Proc Natl Acad Sci U S A* **86**: 7366-7369.
- Alexander, D. C., and Valvano, M. A. (1994) Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing N-acetylglucosamine. *J Bacteriol* **176**: 7079-7084.
- Amer, A. O., and Valvano, M. A. (2001) Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine. *Microbiology* **147**: 3015-3025.
- Amer, A. O., and Valvano, M. A. (2002) Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in *Escherichia coli*. *Microbiology* **148**: 571-582.
- Amor, P. A., and Whitfield, C. (1997) Molecular and functional analysis of genes required for expression of group IB K antigens in *Escherichia coli*: characterization of the his-region containing gene clusters for multiple cell-surface polysaccharides. *Mol Microbiol* **26**: 145-161.
- Anderson, M. S., Eveland, S. S., and Price, N. P. (2000) Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiol Lett* **191**: 169-175.
- Anderson, M. S., and Raetz, C. R. (1987) Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyltransferase that converts UDP-N-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *J Biol Chem* **262**: 5159-5169.

- Arrecubieta, C., Hammarton, T. C., Barrett, B., Chareonsudjai, S., Hodson, N., Rainey, D., and Roberts, I. S. (2001) The transport of group 2 capsular polysaccharides across the periplasmic space in *Escherichia coli*. Roles for the KpsE and KpsD proteins. *J Biol Chem* **276**: 4245-4250.
- Ashby, M. N., and Edwards, P. A. (1990) Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase. *J Biol Chem* **265**: 13157-13164.
- Babinski, K. J., Ribeiro, A. A., and Raetz, C. R. (2002) The *Escherichia coli* gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *J Biol Chem* **277**: 25937-25946.
- Barr, K., and Rick, P. D. (1987) Biosynthesis of enterobacterial common antigen in *Escherichia coli*. *In vitro* synthesis of lipid-linked intermediates. *J Biol Chem* **262**: 7142-7150.
- Bliss, J. M., and Silver, R. P. (1996) Coating the surface: a model for expression of capsular polysialic acid in *Escherichia coli* K1. *Mol Microbiol* **21**: 221-231.
- Bouhss, A., Mengin-Lecreulx, D., Le Beller, D., and Van Heijenoort, J. (1999) Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol* **34**: 576-585.
- Bravo, D., Silva, C., Carter, J. A., Hoare, A., Alvarez, S. A., Blondel, C. J., Zaldivar, M., Valvano, M. A., and Contreras, I. (2008) Growth-phase regulation of lipopolysaccharide O-antigen chain length influences serum resistance in serovars of *Salmonella*. *J Med Microbiol* **57**: 938-946.
- Bray, D., and Robbins, P. W. (1967) The direction of chain growth in *Salmonella anatum* O-antigen biosynthesis. *Biochem Biophys Res Commun* **28**: 334-339.
- Bronner, D., Clarke, B. R., and Whitfield, C. (1994) Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. *Mol Microbiol* **14**: 505-519.
- Burns, S. M., and Hull, S. I. (1998) Comparison of loss of serum resistance by defined lipopolysaccharide mutants and an acapsular mutant of uropathogenic *Escherichia coli* O75:K5. *Infect Immun* **66**: 4244-4253.
- Carty, S. M., Sreekumar, K. R., and Raetz, C. R. (1999) Effect of cold shock on lipid A biosynthesis in *Escherichia coli*. Induction At 12 degrees C of an acyltransferase specific for palmitoleoyl-acyl carrier protein. *J Biol Chem* **274**: 9677-9685.
- Chng, S. S., Gronenberg, L. S., and Kahne, D. (2010a) Proteins required for lipopolysaccharide assembly in *Escherichia coli* form a transenvelope complex. *Biochemistry* **49**: 4565-4567.

- Chng, S. S., Ruiz, N., Chimalakonda, G., Silhavy, T. J., and Kahne, D. (2010b) Characterization of the two-protein complex in *Escherichia coli* responsible for lipopolysaccharide assembly at the outer membrane. *Proc Natl Acad Sci U S A* **107**: 5363-5368.
- Clarke, B. R., Bronner, D., Keenleyside, W. J., Severn, W. B., Richards, J. C., and Whitfield, C. (1995) Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from *Klebsiella pneumoniae* serotype O1. *J Bacteriol* **177**: 5411-5418.
- Clarke, B. R., Cuthbertson, L., and Whitfield, C. (2004) Nonreducing terminal modifications determine the chain length of polymannose O antigens of *Escherichia coli* and couple chain termination to polymer export via an ATP-binding cassette transporter. *J Biol Chem* **279**: 35709-35718.
- Clarke, B. R., Richards, M. R., Greenfield, L. K., Hou, D., Lowary, T. L., and Whitfield, C. (2011) *In vitro* reconstruction of the chain termination reaction in biosynthesis of the *Escherichia coli* O9a O-polysaccharide: the chain-length regulator, WbdD, catalyzes the addition of methyl phosphate to the non-reducing terminus of the growing glycan. *J Biol Chem* **286**: 41391-41401.
- Clementz, T. (1992) The gene coding for 3-deoxy-manno-octulosonic acid transferase and the *rfaQ* gene are transcribed from divergently arranged promoters in *Escherichia coli*. *J Bacteriol* **174**: 7750-7756.
- Clementz, T., Bednarski, J. J., and Raetz, C. R. (1996) Function of the *htrB* high temperature requirement gene of *Escherichia coli* in the acylation of lipid A: HtrB catalyzed incorporation of laurate. *J Biol Chem* **271**: 12095-12102.
- Cunneen, M. M., and Reeves, P. R. (2008) Membrane topology of the *Salmonella enterica* serovar Typhimurium Group B O-antigen translocase Wzx. *FEMS Microbiol Lett* **287**: 76-84.
- Cuthbertson, L., Kimber, M. S., and Whitfield, C. (2007) Substrate binding by a bacterial ABC transporter involved in polysaccharide export. *Proc Natl Acad Sci U S A* **104**: 19529-19534.
- Dan, N., Middleton, R. B., and Lehrman, M. A. (1996) Hamster UDP-N-acetylglucosamine:dolichol-P N-acetylglucosamine-1-P transferase has multiple transmembrane spans and a critical cytosolic loop. *J Biol Chem* **271**: 30717-30724.
- Danese, P. N., Pratt, L. A., and Kolter, R. (2000) Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J Bacteriol* **182**: 3593-3596.

- Dell, A., Oates, J., Lugowski, C., Romanowska, E., Kenne, L., and Lindberg, B. (1984) The enterobacterial common-antigen, a cyclic polysaccharide. *Carbohydr Res* **133**: 95-104.
- Doerrler, W. T., Gibbons, H. S., and Raetz, C. R. (2004) MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J Biol Chem* **279**: 45102-45109.
- Doerrler, W. T., and Raetz, C. R. (2002) ATPase activity of the MsbA lipid flippase of *Escherichia coli*. *J Biol Chem* **277**: 36697-36705.
- El Ghachi, M., Bouhss, A., Blanot, D., and Mengin-Lecreulx, D. (2004) The *bacA* gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. *J Biol Chem* **279**: 30106-30113.
- El Ghachi, M., Derbise, A., Bouhss, A., and Mengin-Lecreulx, D. (2005) Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in *Escherichia coli*. *J Biol Chem* **280**: 18689-18695.
- Elbein, A. D. (1987) Inhibitors of the biosynthesis and processing of N-linked oligosaccharide chains. *Annu Rev Biochem* **56**: 497-534.
- Emptage, R. P., Pemble, C. W. t., York, J. D., Raetz, C. R., and Zhou, P. (2013) Mechanistic Characterization of the Tetraacyldisaccharide-1-phosphate 4'-Kinase LpxK Involved in Lipid A Biosynthesis. *Biochemistry* **52**: 2280-2290.
- Erbel, P. J., Barr, K., Gao, N., Gerwig, G. J., Rick, P. D., and Gardner, K. H. (2003) Identification and biosynthesis of cyclic enterobacterial common antigen in *Escherichia coli*. *J Bacteriol* **185**: 1995-2004.
- Freinkman, E., Okuda, S., Ruiz, N., and Kahne, D. (2012) Regulated assembly of the transenvelope protein complex required for lipopolysaccharide export. *Biochemistry* **51**: 4800-4806.
- Frirdich, E., and Whitfield, C. (2005) Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the Enterobacteriaceae. *J Endotoxin Res* **11**: 133-144.
- Gottesman, S., and Stout, V. (1991) Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. *Mol Microbiol* **5**: 1599-1606.
- Grant, W. D., Sutherland, I. W., and Wilkinson, J. F. (1969) Exopolysaccharide colanic acid and its occurrence in the Enterobacteriaceae. *J Bacteriol* **100**: 1187-1193.
- Greenfield, L. K., Richards, M. R., Li, J., Wakarchuk, W. W., Lowary, T. L., and Whitfield, C. (2012a) Biosynthesis of the polymannose lipopolysaccharide O-antigens from *Escherichia coli* serotypes O8 and O9a requires a unique

- combination of single- and multiple-active site mannosyltransferases. *J Biol Chem* **287**: 35078-35091.
- Greenfield, L. K., Richards, M. R., Vinogradov, E., Wakarchuk, W. W., Lowary, T. L., and Whitfield, C. (2012b) Domain organization of the polymerizing mannosyltransferases involved in synthesis of the *Escherichia coli* O8 and O9a lipopolysaccharide O-antigens. *J Biol Chem* **287**: 38135-38149.
- Greenfield, L. K., and Whitfield, C. (2012) Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways. *Carbohydr Res* **356**: 12-24.
- Gronow, S., Brabetz, W., and Brade, H. (2000) Comparative functional characterization *in vitro* of heptosyltransferase I (WaaC) and II (WaaF) from *Escherichia coli*. *Eur J Biochem* **267**: 6602-6611.
- Gunn, J. S. (2008) The *Salmonella* PmrAB regulon: lipopolysaccharide modifications, antimicrobial peptide resistance and more. *Trends Microbiol* **16**: 284-290.
- Hamad, M. A., Di Lorenzo, F., Molinaro, A., and Valvano, M. A. (2012) Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. *Mol Microbiol* **85**: 962-974.
- Heifetz, A., Keenan, R. W., and Elbein, A. D. (1979) Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate GlcNAc-1-phosphate transferase. *Biochemistry* **18**: 2186-2192.
- Heydanek, M. G., Jr., Struve, W. G., and Neuhaus, F. C. (1969) On the initial stage in peptidoglycan synthesis. 3. Kinetics and uncoupling of phospho-N-acetylmuramyl-pentapeptide translocase (uridine 5'-phosphate). *Biochemistry* **8**: 1214-1221.
- Hug, I., Couturier, M. R., Rooker, M. M., Taylor, D. E., Stein, M., and Feldman, M. F. (2010) *Helicobacter pylori* lipopolysaccharide is synthesized via a novel pathway with an evolutionary connection to protein N-glycosylation. *PLoS Pathog* **6**: e1000819.
- Ikeda, M., Wachi, M., Jung, H. K., Ishino, F., and Matsushashi, M. (1991) The *Escherichia coli* *mraY* gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-pentapeptide transferase. *J Bacteriol* **173**: 1021-1026.
- Islam, S. T., Fieldhouse, R. J., Anderson, E. M., Taylor, V. L., Keates, R. A., Ford, R. C., and Lam, J. S. (2012) A cationic lumen in the Wzx flippase mediates anionic O-antigen subunit translocation in *Pseudomonas aeruginosa* PAO1. *Mol Microbiol* **84**: 1165-1176.

- Islam, S. T., and Lam, J. S. (2013) Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria. *Environ Microbiol* **15**: 1001-1015.
- Islam, S. T., Taylor, V. L., Qi, M., and Lam, J. S. (2010) Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *MBio* **1**: e00189-00110.
- Jackman, J. E., Raetz, C. R., and Fierke, C. A. (2001) Site-directed mutagenesis of the bacterial metalloamidase UDP-(3-O-acyl)-N-acetylglucosamine deacetylase (LpxC). Identification of the zinc binding site. *Biochemistry* **40**: 514-523.
- Jann, K., Kanegasaki, S., Goldemann, G., and Makela, P. H. (1979) On the effect of *rfe* mutation on the biosynthesis of the O8 and O9 antigens of *E. coli*. *Biochem Biophys Res Commun* **86**: 1185-1191.
- Kajimura, J., Rahman, A., and Rick, P. D. (2005) Assembly of cyclic enterobacterial common antigen in *Escherichia coli* K-12. *J Bacteriol* **187**: 6917-6927.
- Kalynych, S., Valvano, M. A., and Cygler, M. (2012a) Polysaccharide co-polymerases: the enigmatic conductors of the O-antigen assembly orchestra. *Protein Eng Des Sel* **25**: 797-802.
- Kalynych, S., Yao, D., Magee, J., and Cygler, M. (2012b) Structural characterization of closely related O-antigen lipopolysaccharide (LPS) chain length regulators. *J Biol Chem* **287**: 15696-15705.
- Kato, A., Chen, H. D., Latifi, T., and Groisman, E. A. (2012) Reciprocal control between a bacterium's regulatory system and the modification status of its lipopolysaccharide. *Mol Cell* **47**: 897-908.
- Keenleyside, W. J., and Whitfield, C. (1995) Lateral transfer of *rfb* genes: a mobilizable ColE1-type plasmid carries the *rfbO*:54 (O:54 antigen biosynthesis) gene cluster from *Salmonella enterica* serovar Borreze. *J Bacteriol* **177**: 5247-5253.
- Keenleyside, W. J., and Whitfield, C. (1996) A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. *J Biol Chem* **271**: 28581-28592.
- Kelly, T. M., Stachula, S. A., Raetz, C. R., and Anderson, M. S. (1993) The *firA* gene of *Escherichia coli* encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase. The third step of endotoxin biosynthesis. *J Biol Chem* **268**: 19866-19874.
- Kido, N., Torgov, V. I., Sugiyama, T., Uchiya, K., Sugihara, H., Komatsu, T., Kato, N., and Jann, K. (1995) Expression of the O9 polysaccharide of *Escherichia coli*: sequencing of the *E. coli* O9 *rfb* gene cluster, characterization of mannosyl

- transferases, and evidence for an ATP-binding cassette transport system. *J Bacteriol* **177**: 2178-2187.
- Klena, J. D., and Schnaitman, C. A. (1993) Function of the *rfb* gene cluster and the *rfe* gene in the synthesis of O antigen by *Shigella dysenteriae* 1. *Mol Microbiol* **9**: 393-402.
- Kuhn, H. M., Neter, E., and Mayer, H. (1983) Modification of the lipid moiety of the enterobacterial common antigen by the "*Pseudomonas factor*". *Infect Immun* **40**: 696-700.
- Larue, K., Kimber, M. S., Ford, R., and Whitfield, C. (2009) Biochemical and structural analysis of bacterial O-antigen chain length regulator proteins reveals a conserved quaternary structure. *J Biol Chem* **284**: 7395-7403.
- Lehrer, J., Vigeant, K. A., Tatar, L. D., and Valvano, M. A. (2007) Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189**: 2618-2628.
- Lehrman, M. A. (1991) Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. *Glycobiology* **1**: 553-562.
- Liu, D., Cole, R. A., and Reeves, P. R. (1996) An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. *J Bacteriol* **178**: 2102-2107.
- Liu, D., Haase, A. M., Lindqvist, L., Lindberg, A. A., and Reeves, P. R. (1993) Glycosyl transferases of O-antigen biosynthesis in *Salmonella enterica*: identification and characterization of transferase genes of groups B, C2, and E1. *J Bacteriol* **175**: 3408-3413.
- Lloyd, A. J., Brandish, P. E., Gilbey, A. M., and Bugg, T. D. (2004) Phospho-N-acetylmuramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues. *J Bacteriol* **186**: 1747-1757.
- Lovering, A. L., Safadi, S. S., and Strynadka, N. C. (2012) Structural perspective of peptidoglycan biosynthesis and assembly. *Annu Rev Biochem* **81**: 451-478.
- Lugowski, C., and Romanowska, E. (1983) Identification of a trisaccharide repeating-unit in the enterobacterial common-antigen. *Carbohydr Res* **118**: 173-181.
- MacLachlan, P. R., Keenleyside, W. J., Dodgson, C., and Whitfield, C. (1993) Formation of the K30 (group I) capsule in *Escherichia coli* O9:K30 does not require attachment to lipopolysaccharide lipid A-core. *J Bacteriol* **175**: 7515-7522.
- Mamat, U., Meredith, T. C., Aggarwal, P., Kuhl, A., Kirchhoff, P., Lindner, B., Hanuszkiewicz, A., Sun, J., Holst, O., and Woodard, R. W. (2008) Single amino

- acid substitutions in either YhjD or MsbA confer viability to 3-deoxy-D-manno-oct-2-ulosonic acid-depleted *Escherichia coli*. *Mol Microbiol* **67**: 633-648.
- Mao, Y., Doyle, M. P., and Chen, J. (2001) Insertion mutagenesis of *wca* reduces acid and heat tolerance of enterohemorrhagic *Escherichia coli* O157:H7. *J Bacteriol* **183**: 3811-3815.
- Marolda, C. L., Li, B., Lung, M., Yang, M., Hanuszkiewicz, A., Rosales, A. R., and Valvano, M. A. (2010) Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. *J Bacteriol* **192**: 6160-6171.
- Marolda, C. L., Vicarioli, J., and Valvano, M. A. (2004) Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* **150**: 4095-4105.
- Mazur, A., Marczak, M., Krol, J. E., and Skorupska, A. (2005) Topological and transcriptional analysis of *pssL* gene product: a putative Wzx-like exopolysaccharide translocase in *Rhizobium leguminosarum* bv. *trifolii* TA1. *Arch Microbiol* **184**: 1-10.
- McGrath, B. C., and Osborn, M. J. (1991) Localization of the terminal steps of O-antigen synthesis in *Salmonella typhimurium*. *J Bacteriol* **173**: 649-654.
- Meier, U., and Mayer, H. (1985) Genetic location of genes encoding enterobacterial common antigen. *J Bacteriol* **163**: 756-762.
- Meier-Dieter, U., Starman, R., Barr, K., Mayer, H., and Rick, P. D. (1990) Biosynthesis of enterobacterial common antigen in *Escherichia coli*. Biochemical characterization of Tn10 insertion mutants defective in enterobacterial common antigen synthesis. *J Biol Chem* **265**: 13490-13497.
- Mengin-Lecreulx, D., Texier, L., Rousseau, M., and van Heijenoort, J. (1991) The *murG* gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J Bacteriol* **173**: 4625-4636.
- Meredith, T. C., Aggarwal, P., Mamat, U., Lindner, B., and Woodard, R. W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS Chem Biol* **1**: 33-42.
- Meredith, T. C., Mamat, U., Kaczynski, Z., Lindner, B., Holst, O., and Woodard, R. W. (2007) Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in *Escherichia coli*. *J Biol Chem* **282**: 7790-7798.

- Merino, S., Jimenez, N., Molero, R., Bouamama, L., Regue, M., and Tomas, J. M. (2011) A UDP-HexNAc:polyprenol-P GalNAc-1-P transferase (WecP) representing a new subgroup of the enzyme family. *J Bacteriol* **193**: 1943-1952.
- Metzger, L. E. t., and Raetz, C. R. (2009) Purification and characterization of the lipid A disaccharide synthase (LpxB) from *Escherichia coli*, a peripheral membrane protein. *Biochemistry* **48**: 11559-11571.
- Metzger, L. E. t., and Raetz, C. R. (2010) An alternative route for UDP-diacylglucosamine hydrolysis in bacterial lipid A biosynthesis. *Biochemistry* **49**: 6715-6726.
- Morrison, J. P., and Tanner, M. E. (2007) A two-base mechanism for *Escherichia coli* ADP-L-glycero-D-manno-heptose 6-epimerase. *Biochemistry* **46**: 3916-3924.
- Mulford, C. A., and Osborn, M. J. (1983) An intermediate step in translocation of lipopolysaccharide to the outer membrane of *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* **80**: 1159-1163.
- Murata, T., Tseng, W., Guina, T., Miller, S. I., and Nikaido, H. (2007) PhoPQ-mediated regulation produces a more robust permeability barrier in the outer membrane of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **189**: 7213-7222.
- Murray, G. L., Attridge, S. R., and Morona, R. (2006) Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. *J Bacteriol* **188**: 2735-2739.
- Narita, S., and Tokuda, H. (2009) Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. *FEBS Lett* **583**: 2160-2164.
- Nikaido, H., and Vaara, M. (1985) Molecular basis of bacterial outer membrane permeability. *Microbiol Rev* **49**: 1-32.
- Nsahlai, C. J., and Silver, R. P. (2003) Purification and characterization of KpsT, the ATP-binding component of the ABC-capsule exporter of *Escherichia coli* K1. *FEMS Microbiol Lett* **224**: 113-118.
- Patel, K. B., Ciepichal, E., Swiezewska, E., and Valvano, M. A. (2012a) The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* **22**: 116-122.
- Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology* **20**: 1389-1401.

- Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012b) Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol* **194**: 2646-2657.
- Pavelka, M. S., Jr., Wright, L. F., and Silver, R. P. (1991) Identification of two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. *J Bacteriol* **173**: 4603-4610.
- Perez, J. M., McGarry, M. A., Marolda, C. L., and Valvano, M. A. (2008) Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. *Mol Microbiol* **70**: 1424-1440.
- Popham, D. L., and Young, K. D. (2003) Role of penicillin-binding proteins in bacterial cell morphogenesis. *Curr Opin Microbiol* **6**: 594-599.
- Price, N. P., and Momany, F. A. (2005) Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. *Glycobiology* **15**: 29R-42R.
- Raetz, C. R., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* **76**: 295-329.
- Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**: 635-700.
- Rahman, A., Barr, K., and Rick, P. D. (2001) Identification of the structural gene for the TDP-Fuc4NAc:lipid II Fuc4NAc transferase involved in synthesis of enterobacterial common antigen in *Escherichia coli* K-12. *J Bacteriol* **183**: 6509-6516.
- Rahn, A., Beis, K., Naismith, J. H., and Whitfield, C. (2003) A novel outer membrane protein, Wzi, is involved in surface assembly of the *Escherichia coli* K30 group 1 capsule. *J Bacteriol* **185**: 5882-5890.
- Rick, P. D., Hubbard, G. L., and Barr, K. (1994) Role of the *rfe* gene in the synthesis of the O8 antigen in *Escherichia coli* K-12. *J Bacteriol* **176**: 2877-2884.
- Rick, P. D., Hubbard, G. L., Kitaoka, M., Nagaki, H., Kinoshita, T., Dowd, S., Simplaceanu, V., and Ho, C. (1998) Characterization of the lipid-carrier involved in the synthesis of enterobacterial common antigen (ECA) and identification of a novel phosphoglyceride in a mutant of *Salmonella typhimurium* defective in ECA synthesis. *Glycobiology* **8**: 557-567.
- Rick, P. D., Mayer, H., Neumeyer, B. A., Wolski, S., and Bitter-Suermann, D. (1985) Biosynthesis of enterobacterial common antigen. *J Bacteriol* **162**: 494-503.

- Rinno, J., Golecki, J. R., and Mayer, H. (1980) Localization of enterobacterial common antigen: immunogenic and nonimmunogenic enterobacterial common antigen-containing *Escherichia coli*. *J Bacteriol* **141**: 814-821.
- Ruan, X., Loyola, D. E., Marolda, C. L., Perez-Donoso, J. M., and Valvano, M. A. (2012) The WaaL O-antigen lipopolysaccharide ligase has features in common with metal ion-independent inverting glycosyltransferases. *Glycobiology* **22**: 288-299.
- Ruiz, N. (2008) Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *Proc Natl Acad Sci U S A* **105**: 15553-15557.
- Saldías, M. S., Ortega, X., and Valvano, M. A. (2009) *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *J Med Microbiol* **58**: 1542-1548.
- Saldías, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**: 440-453.
- Samuel, G., and Reeves, P. (2003) Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338**: 2503-2519.
- Sekine, S., Shimada, A., Nureki, O., Cavarelli, J., Moras, D., Vassylyev, D. G., and Yokoyama, S. (2001) Crucial role of the high-loop lysine for the catalytic activity of arginyl-tRNA synthetase. *J Biol Chem* **276**: 3723-3726.
- Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**: a000414.
- Sleytr, U. B., and Beveridge, T. J. (1999) Bacterial S-layers. *Trends Microbiol* **7**: 253-260.
- Sleytr, U. B., and Messner, P. (1983) Crystalline surface layers on bacteria. *Annu Rev Microbiol* **37**: 311-339.
- Sperandeo, P., Lau, F. K., Carpentieri, A., De Castro, C., Molinaro, A., Deho, G., Silhavy, T. J., and Polissi, A. (2008) Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *J Bacteriol* **190**: 4460-4469.
- Stead, C. M., Pride, A. C., and Trent, M. S. (2011) Genetics and Biosynthesis of Lipid A. In: *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis, and Interaction with Host Cells*. Valvano, Y. A. K. a. M. A. (ed). New York: SpringerWien, pp. 163-194.

- Steiner, K., Novotny, R., Patel, K., Vinogradov, E., Whitfield, C., Valvano, M. A., Messner, P., and Schaffer, C. (2007) Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *J Bacteriol* **189**: 2590-2598.
- Steiner, K., Novotny, R., Werz, D. B., Zarschler, K., Seeberger, P. H., Hofinger, A., Kosma, P., Schaffer, C., and Messner, P. (2008) Molecular basis of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus*. *J Biol Chem* **283**: 21120-21133.
- Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996) Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**: 4885-4893.
- Stout, V., Torres-Cabassa, A., Maurizi, M. R., Gutnick, D., and Gottesman, S. (1991) RcsA, an unstable positive regulator of capsular polysaccharide synthesis. *J Bacteriol* **173**: 1738-1747.
- Sutherland, I. W. (1969) Structural studies on colanic acid, the common exopolysaccharide found in the enterobacteriaceae, by partial acid hydrolysis. Oligosaccharides from colanic acid. *Biochem J* **115**: 935-945.
- Tarshis, L. C., Yan, M., Poulter, C. D., and Sacchettini, J. C. (1994) Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry* **33**: 10871-10877.
- Tatar, L. D., Marolda, C. L., Polischuk, A. N., van Leeuwen, D., and Valvano, M. A. (2007) An *Escherichia coli* undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling. *Microbiology* **153**: 2518-2529.
- Tocij, A., Munger, C., Proteau, A., Morona, R., Purins, L., Ajamian, E., Wagner, J., Papadopoulos, M., Van Den Bosch, L., Rubinstein, J. L., Fethiere, J., Matte, A., and Cygler, M. (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* **15**: 130-138.
- Touze, T., Tran, A. X., Hankins, J. V., Mengin-Lecreulx, D., and Trent, M. S. (2008) Periplasmic phosphorylation of lipid A is linked to the synthesis of undecaprenyl phosphate. *Mol Microbiol* **67**: 264-277.
- Tran, A. X., Trent, M. S., and Whitfield, C. (2008) The LptA protein of *Escherichia coli* is a periplasmic lipid A-binding protein involved in the lipopolysaccharide export pathway. *J Biol Chem* **283**: 20342-20349.
- Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**: s452-471.
- Valvano, M. A. (2008) Undecaprenyl phosphate recycling comes out of age. *Mol Microbiol* **67**: 232-235.

- Valvano, M. A., Furlong, S. E., and Patel, K. B. (2011) Genetics, Biosynthesis and Assembly of O-Antigen. In: *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis, and Interaction with Host Cells*. Valvano, Y. A. K. a. M. A. (ed). New York: SpringerWien, pp. 275-310.
- Venkatachalam, K. V., Fuda, H., Koonin, E. V., and Strott, C. A. (1999) Site-selected mutagenesis of a conserved nucleotide binding HXGH motif located in the ATP sulfurylase domain of human bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase. *J Biol Chem* **274**: 2601-2604.
- Videira, P. A., Garcia, A. P., and Sa-Correia, I. (2005) Functional and topological analysis of the *Burkholderia cenocepacia* priming glucosyltransferase BceB, involved in the biosynthesis of the cepacian exopolysaccharide. *J Bacteriol* **187**: 5013-5018.
- Villa, R., Martorana, A. M., Okuda, S., Gourlay, L. J., Nardini, M., Sperandio, P., Deho, G., Bolognesi, M., Kahne, D., and Polissi, A. (2013) The *Escherichia coli* Lpt transenvelope protein complex for lipopolysaccharide export is assembled via conserved structurally homologous domains. *J Bacteriol* **195**: 1100-1108.
- Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598-2604.
- Wang, L., and Reeves, P. R. (1994) Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen subunit processing. *J Bacteriol* **176**: 4348-4356.
- Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**: 39-68.
- Whitfield, C., Kaniuk, N., and Fridrich, 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.
- Whitfield, C., and Roberts, I. S. (1999) Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* **31**: 1307-1319.
- Woodward, R., Yi, W., Li, L., Zhao, G., Eguchi, H., Sridhar, P. R., Guo, H., Song, J. K., Motari, E., Cai, L., Kelleher, P., Liu, X., Han, W., Zhang, W., Ding, Y., Li, M., and Wang, P. G. (2010) *In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat Chem Biol* **6**: 418-423.
- Yethon, J. A., Vinogradov, E., Perry, M. B., and Whitfield, C. (2000) Mutation of the lipopolysaccharide core glycosyltransferase encoded by *waaG* destabilizes the outer membrane of *Escherichia coli* by interfering with core phosphorylation. *J Bacteriol* **182**: 5620-5623.

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

Chapter 2

Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase

This chapter has been published:

Sarah E. Furlong and Miguel A. Valvano (2012). Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase. *Protein Science*, 21, 1366-1375.

2.1 Introduction

The synthesis of lipid-linked glycan precursors is essential for the production of glycoconjugates in prokaryotes and eukaryotes such as polysaccharides, glycopeptides, and a wide range of glycosylated proteins (Valvano, 2003, Valvano, 2011). The enzymes that catalyze the initiation of the synthesis of these lipid-linked glycans fall into two broad families of membrane proteins: (i) the polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferases (PNPT), and (ii) the polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPT). These enzymes catalyze the transfer of a sugar-1-phosphate from a nucleoside sugar diphosphate to a phosphoisoprenyl lipid carrier resulting in the formation of a phosphoanhydride bond (Valvano, 2003, Valvano, 2011). Enzymes of the PHPT family occur only in bacteria and they catalyze the synthesis of O antigen (Patel *et al.*, 2012a, Patel *et al.*, 2012b, Patel *et al.*, 2010, Saldías *et al.*, 2008, Wang *et al.*, 1996, Merino *et al.*, 2011), various exopolysaccharides (Patel *et al.*, 2012b, Xayarath and Yother, 2007) and glycans precursors for protein glycosylation (Glover *et al.*, 2006, Chamot-Rooke *et al.*, 2007, Steiner *et al.*, 2007, Power *et al.*, 2000).

PNPT proteins occur in prokaryotes and eukaryotes. The eukaryotic members are UDP-N-acetylglucosamine (UDP-GlcNAc):dolichyl-phosphate (Dol-P) GlcNAc-1-P transferases and reside in the rough endoplasmic reticulum membrane (Lehrman, 1991). In contrast to the eukaryotic members, bacterial PNPTs use several different diphospho N-acetylaminosugar nucleoside substrates (Anderson *et al.*, 2000). Also, whereas eukaryotic PNPTs are specific for Dol-P (Mankowski *et al.*, 1975), the bacterial enzymes function only with undecaprenyl-phosphate (Und-P) (Rush *et al.*, 1997). WecA and MraY are prototypic bacterial PNPTs; the former is an UDP-GlcNAc:Und-P GlcNAc-1-P

transferase (Alexander and Valvano, 1994) that initiates the synthesis of O antigen and enterobacterial common antigen, and the latter is a UDP-N-acetylmuramyl (MurNAc)-pentapeptide:Und-P MurNAc-pentapeptide-1-P transferase that initiates the synthesis of cell wall peptidoglycan (Al-Dabbagh *et al.*, 2008).

These enzymes have multiple transmembrane (TM) helices, and several studies suggest that all the cytosolic loops contribute residues to form a putative catalytic site (Al-Dabbagh *et al.*, 2008, Lehrer *et al.*, 2007). Enzyme kinetics studies on MraY have suggested that the reaction involves the formation of an enzyme- phospho-MurNAc-pentapeptide intermediate prior to its transfer to Und-P (Heydanek *et al.*, 1969). However, recent experiments using purified MraY proteins suggests a one-step catalytic mechanism in which a base contributed by the side chain of a conserved aspartic acid residue would permit the deprotonation of a hydroxyl of the Und-P terminal phosphate. This deprotonated hydroxyl group would be involved in the nucleophilic attack on the β -phosphate of the UDP-MurNAc-pentapeptide (Al-Dabbagh *et al.*, 2008).

Despite the differences in the PNPT family regarding N-acetylaminosugar and lipid specificities, these proteins share highly conserved amino acid motifs, suggesting a common enzymatic mechanism (Price and Momany, 2005). To better understand the catalytic mechanism of PNPT enzymes, several of these conserved regions have been characterized in MraY and WecA. For example, the DDxxD motif in the cytoplasmic loop II of WecA has been proposed to bind the cofactor Mg^{2+} (Fig. 2.1A) (Lehrer *et al.*, 2007). The highly conserved D156 residue has been postulated to play a critical role in the catalytic mechanism as mutations to this residue causes the protein to lose all activity *in vivo* and *in vitro* (Lehrer *et al.*, 2007, Amer and Valvano, 2002). A conserved HIIHH

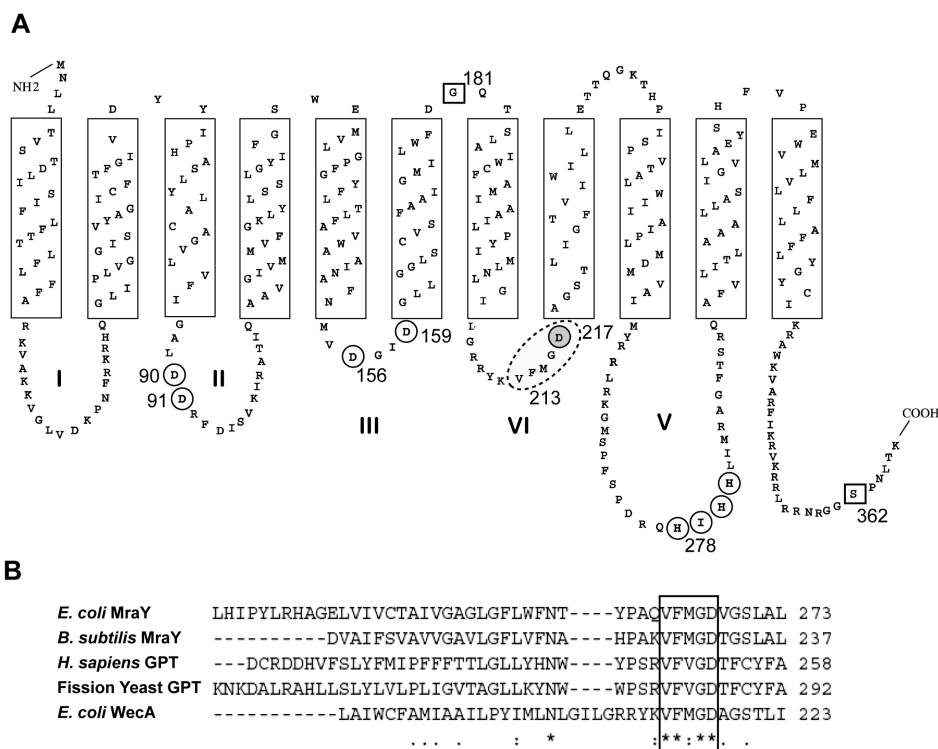


Figure 2.1. Topological model of the *E. coli* WecA. A. Refined topological model of the *E. coli* WecA. This model was previously established by a combination of bioinformatics, substituted cysteine accessibility experiments, and reporter gene fusions (Lehrer *et al.*, 2007). A dotted ellipse indicates the highly conserved V₂₁₃FMGD₂₁₇ motif and its topology has been determined in this work (see text in Results). The highly conserved residues that are important for function based on previous research (Amer and Valvano, 2001, Amer and Valvano, 2002, Lehrer *et al.*, 2007) are circled. Shading indicates D217. Squares show the control residues for cysteine scanning accessibility experiments used in this work, G181 and S362. B. ClustalW alignment of PNPT members shows the highly conserved VFMGD motif. Alignment shows protein sequences from *E. coli* MraY, *B. subtilis* MraY, *Homo sapiens* GPT, Fission Yeast GPT, and *E. coli* WecA. The completely conserved amino acids are marked with an asterisk (*) and partially conserved amino acids are marked with a colon (:).

motif present in loop V is predicted to be part of a carbohydrate recognition domain (Amer and Valvano, 2001, Anderson *et al.*, 2000). Another highly conserved region shared among eukaryotic and prokaryotic PNPT family members is the VFMGD motif (Fig. 2.1B) (Price and Momany, 2005). Lloyd *et al* proposed that the highly conserved aspartic acid residue 267 (D267) in the VFMGD motif of MraY serves as a catalytic nucleophile in a proposed double displacement catalytic mechanism that involves the cleavage of the pyrophosphate bond of the nucleoside (Heydanek *et al.*, 1969, Lloyd *et al.*, 2004). The comparable aspartic acid residue in the VFMGD motif of WecA is D217, which was predicted to be located in a transmembrane region (Lehrer *et al.*, 2007). In this study, we investigated the topology and function of the WecA V₂₁₃FMGD₂₁₇ motif. Our results demonstrate that D217 and the rest of the residues in the motif are located facing the cytosol, but D217 is not a catalytic nucleophile. However, the polarity and size of amino acid side chains at the D217 position are important for enzymatic activity, as were the replacements of the V213, F214, G216, and D217 residues with alanine. We propose that the highly conserved VFMGD motif defines a region in PNPT proteins that contributes to the active site, likely involved in the binding and/or recognition of the nucleotide moiety of the nucleoside phosphate precursor.

2.2 Materials and Methods

2.2.1 Bacterial strains, plasmids, media, and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 2.1. Bacteria were grown aerobically at 37°C in Luria-Bertani (LB) medium (Difco Laboratories, Sparks, MD, USA) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). Growth medium

Table 2.1. Characteristics of the bacterial strains and plasmids used in this study.

| Strain or Plasmid | Relevant Properties | Source or Reference |
|-------------------|---|-------------------------------|
| <i>Strains</i> | | |
| DH5 α | <i>E. coli</i> K-12 F ⁻ Φ 80 <i>lacZ</i> Δ M15 <i>endA recA</i> | Laboratory stock |
| MV501 | <i>hsdR</i> (r _K ⁻ m _K ⁻) <i>supE thi gyrA relA</i> <i>E. coli</i> VW187; <i>wecA</i> :Tn10 Tc ^R | (Alexander and Valvano, 1994) |
| VW187 | O7:K1; clinical isolate | (Valvano and Crosa, 1989) |
| <i>Plasmids</i> | | |
| pBAD24 | Cloning vector, inducible by arabinose; Amp ^R | (Guzman <i>et al.</i> , 1995) |
| pBAD-His | His ₆ inserted into pBAD24; Amp ^R | (Lehrer <i>et al.</i> , 2007) |
| pKV1 | Expresses <i>WecA</i> -FLAG-5xHis from <i>wecA</i> -FLAG cloned into pBAD-His | (Lehrer <i>et al.</i> , 2007) |
| pJL7 | pKV1 expressing cysteine-less <i>WecA</i> -FLAG-7xHis | (Lehrer <i>et al.</i> , 2007) |
| pJL7-G181C | pJL7 expressing cysteine-less <i>WecA</i> -G181C-FLAG-7xHis | (Lehrer <i>et al.</i> , 2007) |
| pJL7-S362C | pJL7 expressing cysteine-less <i>WecA</i> -S362C-FLAG-7xHis | (Lehrer <i>et al.</i> , 2007) |
| pSEF8 | pKV1 expressing <i>WecA</i> -D217N-FLAG-5xHis | This study |
| pSEF9 | pKV1 expressing <i>WecA</i> -D217E-FLAG-5xHis | This study |
| pSEF17 | pKV1 expressing <i>WecA</i> -D217A-FLAG-5xHis | This study |
| pSEF18 | pKV1 expressing <i>WecA</i> -D217K-FLAG-5xHis | This study |
| pSEF36 | pJL7 expressing cysteine-less <i>WecA</i> -D217C-FLAG-7xHis | This study |
| pSEF37 | pJL7 expressing cysteine-less <i>WecA</i> -S220C-FLAG-7xHis | This study |
| pSEF43 | pKV1 expressing <i>WecA</i> -V213A-FLAG-5xHis | This study |
| pSEF44 | pKV1 expressing <i>WecA</i> -F214A-FLAG-5xHis | This study |
| pSEF45 | pKV1 expressing <i>WecA</i> -M215A-FLAG-5xHis | This study |
| pSEF46 | pKV1 expressing <i>WecA</i> -G216A-FLAG-5xHis | This study |
| pSEF48 | pKV1 expressing <i>WecA</i> -D217S-FLAG-5xHis | This study |

were supplemented with 100 $\mu\text{g/ml}$ ampicillin and/or 20 $\mu\text{g/ml}$ tetracycline. DH5 α cells were transformed with plasmids by calcium chloride method (Cohen *et al.*, 1972) and MV501 cells were transformed with plasmids by electroporation (Dower *et al.*, 1988).

2.2.2 Site-directed mutagenesis

Site directed mutations were made in pKV1, an arabinose inducible pBAD vector expressing the WecA-FLAG-Hisx5 protein. For sulfhydryl labeling experiments, novel cysteines were introduced in pJL7, similar to pKV1 but expressing the cysteine-less WecA-FLAG-Hisx7 protein. Replacements were introduced by the polymerase chain reaction (PCR) using primers containing the desired mutations and *Pfu* AD polymerase (Stratagene, Santa Clara, CA, USA). *Dpn*I was added to PCR reactions for overnight digestion of parental plasmid DNA at 37°C. The resulting DNA was introduced into *E. coli* DH5 α by transformation, and transformants were selected on LB-agar containing 100 $\mu\text{g/ml}$ of ampicillin.

2.2.3 Growth conditions of cells for protein preparation

Briefly, MV501 cells containing the appropriate arabinose-inducible plasmids were grown overnight in LB containing 100 $\mu\text{g/ml}$ of ampicillin and 20 $\mu\text{g/ml}$ of tetracycline. From overnight cultures, cells were diluted to an optical density measured at 600 nm (OD_{600}) of 0.1 in LB broth containing the respective antibiotics. Cells were grown at 37°C until an OD_{600} of 0.5, and cells were induced with 0.2% arabinose for 3 h at 37°C. Cells

were harvested by centrifugation 8 000 χ g and cell pellets were frozen at -20°C until needed.

2.2.4 Total membrane preparation and immunoblotting

Cells were resuspended in 50 mM TAE pH 8.5 with protease inhibitors and lysed by French press (Thermo Scientific, Rockville, MD, USA). Cell debris and unlysed cells were pelleted at 27, 216 χ g. Total membranes were isolated by centrifugation in microfuge tubes at 39, 191 χ g and resuspended in 1X TAE, unless otherwise stated. Protein concentrations were determined by Bradford protein assay (Bio-Rad, Hercules, CA, USA). Total membrane fractions were used for *in vitro* transferase assays and immunoblotting. Immunoblotting was performed by separating total membrane fractions by 14% SDS-PAGE and transferring to a nitrocellulose membrane. Membranes were blocked overnight in 5% Western blocking reagent (Roche Diagnostics Canada, Laval, QC, Canada) and TBS. The primary antibody, 4.6 mg/ml anti-FLAG M2 monoclonal antibody (Sigma, Saint Louis, MO, USA), was diluted to 1: 10, 000 and applied for 1.5 h, and the secondary antibody, 2 mg/ml goat anti-mouse Alexa fluor 680 IgG antibodies (Invitrogen Molecular Probes, Eugene, OR, USA) was diluted to 1: 20, 000 and applied for 20 min. Western blots were developed using LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE, USA). Bio-Rad Precision Plus Protein Standards were used for all Western blots.

2.2.5 Sulfhydryl labeling of cysteine residues using biotin maleimide and protein purification

MV501 cells expressing cysteine-less WecA mutant proteins were grown as described in *Growth conditions of cells for protein preparation* section. Cells were centrifuged at 8 000 \times g, washed twice with 0.1 M sodium phosphate buffer (pH 7.2) and resuspended in this buffer. Whole cells were either pre-treated with 0.1 M MTSET (2-(trimethylammonium) ethyl] methanethiosulfonate bromide) (Toronto Research Chemicals Inc., Toronto, ON, Canada), a membrane impermeable blocking reagent or directly labeled with the membrane permeable 0.5 mM N^α-(3-maleimidylpropionyl)biocytin (biotin maleimide) (Invitrogen Molecular Probes, Eugene, OR, USA). The reaction was terminated by adding 2% 2-mercaptoethanol. Cells were washed three times with 0.1 M sodium phosphate buffer (pH 7.2) and resuspended in sodium phosphate buffer prior to lysing by French press.

Total membranes were prepared and quantified as described above and solubilized for 16 h in 350 μ l of 0.1 M sodium phosphate buffer (pH 7.8), 8 M urea, and 0.5% Triton X-100 at 4°C. Solubilized samples were centrifuged at 39, 000 \times g for 35 min. Supernatants, containing soluble WecA-FLAG-Hisx7 protein, were applied to Co²⁺ charged sepharose beads for 2 h at 4°C. Beads were washed twice with wash buffer [0.1 M sodium phosphate buffer (pH 7.8), 8 M urea, 0.5% Triton X-100, 300 mM NaCl, 50 mM imidazole (pH 8)]. Samples were eluted with elution buffer [0.1 M sodium phosphate buffer (pH 7.8), 8 M urea, 0.5% Triton X-100, 300 mM NaCl, 400 mM imidazole (pH 8)]. Purified samples were separated by 14% SDS-PAGE and transferred to nitrocellulose membrane. Western blotting using anti-FLAG antibodies was performed

as described above, and a 1: 25, 000 dilution of streptavidin-conjugated infrared 800 dye (Rockland Immunochemicals Inc., Gilbertsville, PA, USA) was used to detect biotinylated proteins.

2.2.6 LPS analysis

In vivo transferase activity was assessed by observing the ability of mutant WecA proteins to complement O7 Ag synthesis in the *wecA*-defective strain MV501 (*wecA*::Tn10). MV501 cells harboring the respective plasmids were grown overnight in 5 ml of LB broth and induced with 0.002% arabinose. The OD₆₀₀ of overnight cultures were adjusted to a turbidity of 2 and cells were lysed by boiling and treated with proteinase-K. LPS samples were extracted using a hot phenol method as described by (Marolda *et al.*, 2006). LPS samples were separated by 14% acrylamide gels using a Tricine/SDS buffer system and stained with silver nitrate (Marolda *et al.*, 2006).

2.2.7 *In vitro* transferase assay

Total membranes were prepared from MV501 cells containing plasmids expressing wild-type and mutant WecA proteins. 40 µg or 60 µg of membranes, containing endogenous Und-P, were incubated for 30 min at 37 °C with 78.8 pmol of ¹⁴C-labeled UDP-GlcNAc in a 250 µl reaction buffer containing 50 mM Tris-HCl (pH 8), 40 mM MgCl₂, 0.5 mM EDTA, 0.5% CHAPS, 5 mM 2-mercaptoethanol, and 50 mM sucrose. For kinetic curves, 7.88-709.2 pmol of radiolabeled sugar was used. The lipid fraction was extracted twice with butanol, washed with water, and the butanol phase (containing the lipid fraction) was added to scintillation cocktail (Ecolume, MP Biomedical, Solon, OH, USA) and

measured by scintillation counter (Beckman Coulter Canada, Inc., Mississauga, Ontario, Canada) to determine the radioactive counts per minute (CPM).

2.2.8 Thin layer chromatography

To determine the radioactive lipid products formed in the *in vitro* transferase assay, lipid extractions from these assays (described above) were dried overnight and resuspended in 20 μ l of 2:1 chloroform-methanol solution. The extractions were spotted on Whatman TLC plates (silica gel, type PE SIL G) and placed in a saturated TLC tank with diisobutylketone, acetic acid, and water (80:50:10) solvent. The TLC plate was removed, dried, and exposed to a PhosphorImager storage screen overnight. The screen was imaged using PhosphorImager (Storm 840; Amersham Biosciences) and pixel densities were measured by ImageJ computer software.

2.2.9 Tunicamycin binding activity of WecA

The *in vitro* binding activity of proteins was assessed using a tunicamycin competition assay (Amer and Valvano, 2001). This assay is based on the ability of the WecA proteins to bind and be irreversibly inhibited by tunicamycin, a substrate-product transition-state analogue (Heifetz *et al.*, 1979). 40 μ g of total membranes containing the mutant proteins were incubated for 10 min with a concentration of tunicamycin that inhibits 100% activity of 20 μ g of total membranes containing wild-type WecA-FLAG enzyme. After which, membranes were centrifuged at 39,191 x g and the supernatants were added to 20 μ g of total membranes containing wild-type WecA. The *in vitro* transferase assay was performed (as described above). Any residual tunicamycin that was not bound by the

mutant WecA protein will inhibit the wild-type enzyme activity and can be described as a percentage of tunicamycin bound by wild-type WecA.

2.3 Results

2.3.1 The substituted cysteine accessibility method reveals that aspartic acid residue D217 is exposed to the cytosol

In silico analyses using topology prediction programs for membrane proteins placed the aspartic acid residue D217 in the eighth predicted transmembrane (TM VIII) segment of WecA (Lehrer *et al.*, 2007). However, the predicted location of this residue did not agree with its proposed function as a nucleophile (Lloyd *et al.*, 2004, Price and Momany, 2005). D217C and S220C derivatives of WecA were constructed in a cysteine-less WecA-FLAG-Hisx7 protein (Lehrer *et al.*, 2007). Bacterial cells expressing the WecA proteins with the Cys replacements were preincubated with MTSET, a membrane impermeable thiol-blocking reagent, and subsequently treated with the membrane permeable biotin maleimide. Biotinylated proteins were visualized as described in Material and Methods. As controls, aliquots of the same cells were treated with biotin maleimide only. We also used G181C (periplasmic) and S362C (cytoplasmic) replacement mutants as topology controls (Lehrer *et al.*, 2007). The results demonstrated that the Cys residue at position D217, similar to the residue at position S362 (cytoplasmic control), was labeled with biotin maleimide irrespective of pre-treatment with MTSET. In contrast, the periplasmic control Cys residue at position G181 was labeled with biotin maleimide but as expected due to its periplasmic location, labeling was blocked by pretreatment with the membrane-impermeable agent MTSET (Fig. 2.2) (Lehrer *et al.*,

2007). Therefore, we conclude that D217 is exposed to the cytoplasmic space. Interestingly, the replacement of S220 with Cys could not be labeled with biotin maleimide, suggesting that this residue may reside at the membrane boundary or within TM VIII (Fig. 2.1A).

2.3.2 Substitution of aspartic acid residue D217 by non acidic but polar residues results in a more active enzyme

To investigate the role of D217, we constructed $\text{WecA}_{\text{D217N}}$ and assessed the function of this protein *in vivo* by determining its ability to restore O7 Ag synthesis in the *E. coli* *wecA*-defective strain MV501 (Alexander and Valvano, 1994). The Asn replacement did not abolish the synthesis of polymeric O antigen (Fig. 2.3A), suggesting that this aspartic acid residue is not involved in catalysis. We also tested the enzymatic activity of membrane extracts containing $\text{WecA}_{\text{D217N}}$ by measuring the incorporation of ^{14}C -GlcNAc into a lipid fraction (Fig. 2.3B). Surprisingly, $\text{WecA}_{\text{D217N}}$ had increased *in vitro* transferase activity relative to the parental enzyme. This result was confirmed by densitometry quantitation of the reaction product after separating the lipid fractions by thin layer chromatography (TLC) (Fig. 2.4A). The results were normalized by the amount of WecA protein present in the membrane extract. The densitometric ratios (pixel density of Und-P-P-GlcNAc product formed per amount of WecA protein) were 1.75 for the parental WecA and 3.38 for the D217N mutant enzyme (Fig. 2.4B, C). Together, these experiments demonstrate that D217 does not function as a catalytic nucleophile.

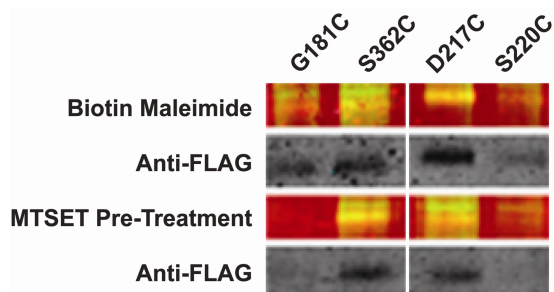


Figure 2.2. Sulfhydryl labeling accessibility of the cysteine replacement in WecA_{D217C}. The top panel shows the results obtained with biotin maleimide labeling. The second and fourth panels show western blots with anti-FLAG antibodies. The third panel shows the results of biotin maleimide labeling with MTSET-pretreated samples. WecA was purified by Co²⁺-affinity chromatography and separated by 14% SDS-PAGE. After transfer to a nitrocellulose membrane biotin maleimide labeling was detected by a streptavidin-conjugated fluorophore (green signal) and WecA-FLAG proteins were detected using anti-FLAG antibodies (red signal). Biotinylated WecA-FLAG proteins are observed by the merged images, which appear yellow using Odyssey LI-COR Scanner. The cysteine replacement of D217 in the cysteine-less WecA-FLAG protein is labeled with biotin maleimide irrespective of MTSET pre-treatment, suggesting that this residue faces the cytosol.

To further explore the role of D217, we replaced this residue with amino acids Glu, Ala, Lys, and Ser. Mutant proteins were tested for expression in total membrane fractions (Fig. 2.3C) and for functional activity *in vivo*. WecA_{D217E} and WecA_{D217S} supported polymeric O antigen synthesis, although the amount of O antigen produced was reduced relative to that obtained with parental WecA (Fig. 2.3A). MV501 expressing WecA_{D217A} could only produce one O antigen unit, while WecA_{D217K} was not functional (Fig. 2.3A). The *in vitro* transferase activity in membrane fractions showed that the replacement of D217 by glutamic acid (WecA_{D217E}) did not affect the *in vitro* enzyme activity. As for D217N, the D217S replacement resulted in a WecA protein with higher transferase activity than the parental enzyme, while replacements with Ala and Lys, resulted in decreased activity and loss of activity, respectively (Fig. 2.3B), in agreement with the *in vivo* functional data. None of these differences could be attributed to differential protein expression, as the parental and mutated forms of WecA were expressed at similar levels in the membrane protein extracts used for the enzymatic analysis (Fig. 2.3C).

2.3.3 The D217N WecA mutant has a higher V_{\max} but no significant change in K_m

The differences observed above prompted us to investigate whether WecA_{D217N} produces more Und-P-P-GlcNAc than the parental WecA due to increased binding affinity for UDP-GlcNAc. Enzyme kinetics was determined using increasing amounts of ¹⁴C-labeled UDP-GlcNAc (Fig. 2.5), and data were analyzed by nonlinear regression the Michaelis-Menten equation to calculate V_{\max} and K_m values (Table 2.2). The parental WecA and WecA_{D217N} have V_{\max} values of 2.6 ± 0.3 pmol/min/mg and 6.7 ± 0.7 pmol/min/mg,

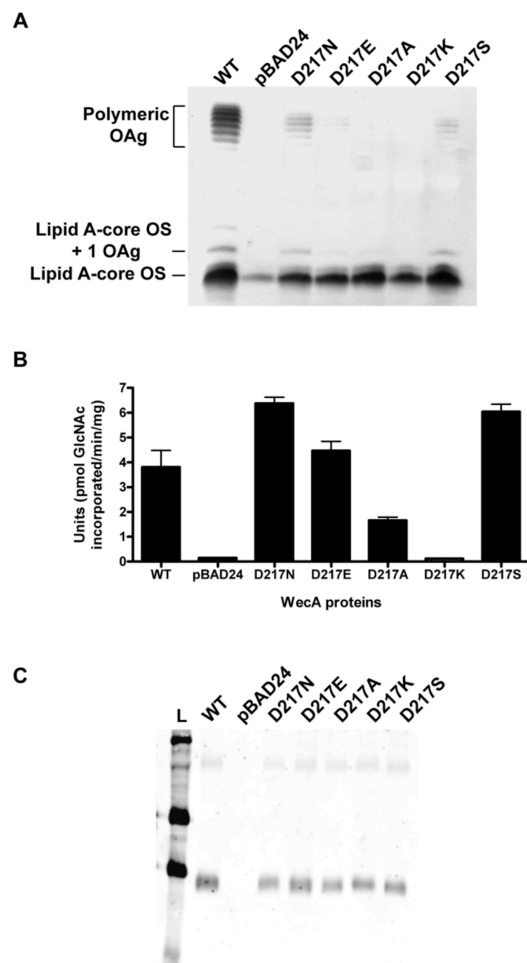


Figure 2.3. *In vivo* and *in vitro* complementation of the D217 replacement mutants.

A. Complementation of O7 LPS expression in *E. coli* MV501 by plasmids encoding parental WecA (WT) and various D217 replacement mutants, as determined with silver-stained polyacrylamide gels. The pBAD24 cloning vector was used as a negative control.

B. GlcNAc-1-P transferase activity assays performed with total membrane extracts prepared from *E. coli* MV501 cells carrying pBAD24 or plasmids encoding parental WecA (WT) and D217 replacement mutants. Bars represent means and standard errors of triplicate experiments.

C. Expression of the WecA proteins in the membrane extracts used for enzymatic assays, as detected by Western blotting using anti-FLAG antibodies.

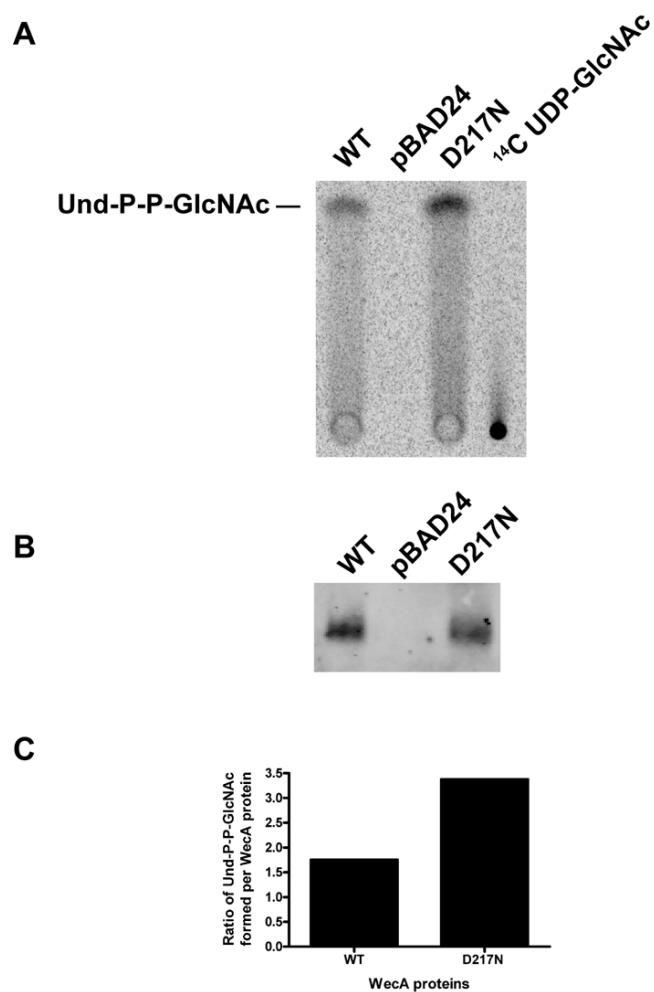


Figure 2.4. The WecA_{D217N} protein mediates increased Und-P-P-GlcNAc production compared to parental WecA. A. TLC of lipid extractions from MV501 total membranes containing parental WecA (WT) and the D217N mutant WecA total proteins that were incubated with ¹⁴C-UDP-GlcNAc. B. Western blot of parental WecA (WT) and D217N mutant proteins using anti-FLAG antibodies. C. Graph depicting the ratio of Und-P-P-GlcNAc product formed per WecA-FLAG protein.

respectively, while the K_m values of the two proteins did not change significantly (Table 2.2). These results demonstrate that the Asp to Asn replacement results in ~3-fold increased catalytic efficiency compared to the parental WecA, in agreement with the observed increased product conversion (Fig. 2.4A). Tunicamycin is an inhibitor of the transfer reaction that acts as a putative transition state analogue inhibitor of WecA and PNPT enzymes in general (Heifetz *et al.*, 1979). Therefore, to gain more information on the function of the residue at position 217, we assessed the ability of WecA mutant proteins D217N, D217E, D217A, and D217K to bind tunicamycin at subinhibitory concentrations, an assay previously used to estimate UDP-GlcNAc binding in WecA mutant proteins (Amer and Valvano, 2001). In this assay, the level of transferase activity of the parental WecA is a function of the residual tunicamycin concentration in the supernatant after exposure to membranes containing the appropriate WecA mutant protein. The results show no differences in the binding capacity of the mutants relative to the parental enzyme (Fig. 2.6), suggesting that the change in enzymatic activity is likely not attributed to a change in the ability of the enzyme to bind a putative transition state intermediate.

2.3.4 Alanine replacements of other residues of the VFMGD motif result in diminished enzymatic activity

To assess the *in vivo* activity of alanine replacements in other VFMGD residues, mutant proteins were expressed in MV501 and tested for membrane expression and their ability to restore O7 Ag synthesis. All mutant proteins were detected in the total membrane fraction by anti-FLAG Western blotting indicating they are correctly expressed at similar

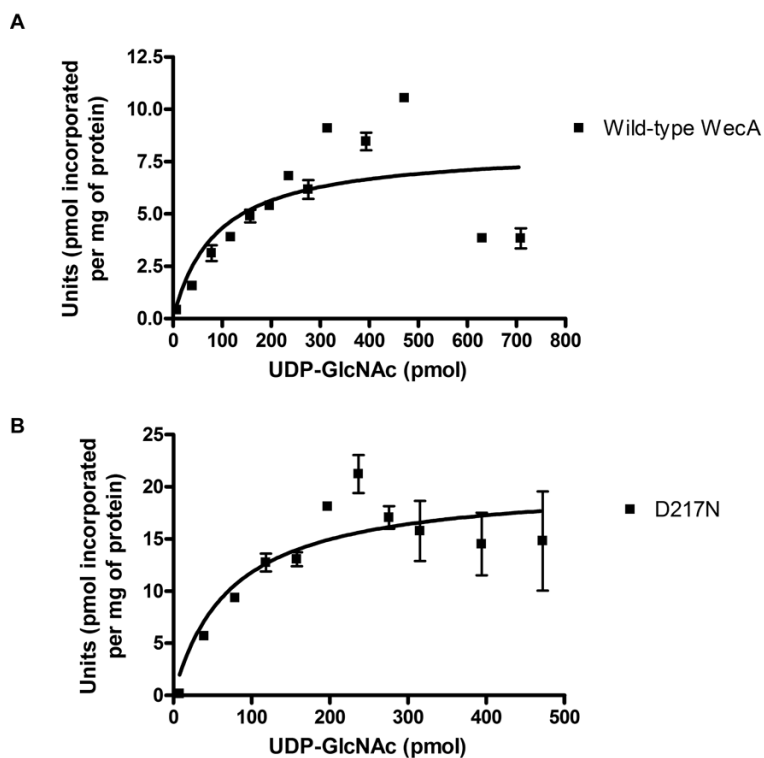


Figure 2.5. GlcNAc-1-P transferase activity of parental WecA (A) the D217N mutant (B). 40 μ g of total membranes were incubated with increasing amounts of 14 C-UDP-GlcNAc for 30 min at 37°C. The lipid fraction was extracted and counts were quantified in a scintillation counter. The enzyme units are expressed as pmol of UDP-GlcNAc incorporated per mg of protein. The experiment was performed in triplicate.

Table 2.2. Kinetic parameters of UDP-GlcNAc for parental WecA and D217N mutant enzymes.

| Protein | K_m (μ M) | V_{max} (pmol/min/mg) | V_{max}/K_m |
|---------------|------------------|----------------------------|---------------|
| Parental WecA | 0.08 ± 0.04 | 2.6 ± 0.3 | 32.5 |
| D217N | 0.07 ± 0.03 | 6.7 ± 0.7 | 95.7 |

Transfer assays were performed with total membrane extracts prepared from MV501 cells carrying the plasmid expressing parental or mutant WecA as described in Material and Methods.

levels and exported to the bacterial membrane (Fig. 2.3C, 2.7C). The LPS was extracted from cells expressing these mutant enzymes, separated by SDS-PAGE and stained with silver nitrate (Fig. 2.7A). The *WecA*_{V213A} and *WecA*_{M215A} complemented similar to the parental enzyme, however the *WecA*_{D217A} could only produce one O antigen unit and *WecA*_{G216A} produced a reduced amount of polymeric O antigen. Although F214A mutant enzyme could only produce one O antigen unit, the F214Y can complement O antigen synthesis similar to the wild-type *WecA* (data not shown) suggesting that the hydrophobicity of this residue is important for function. The alanine replacement of the completely conserved residues, V213, F214, G216, and D217 also had dramatic effects on the *in vitro* transferase activity (Fig. 2.7B), while the replacement of the M215, the less conserved residue of the motif, did not have any effect on *in vitro* transferase activity. The G216A mutation also revealed a severe effect on activity, suggesting this residue plays a structural role. Together, these data indicate that the highly conserved residues of the VFMGD motif may delineate a structural scaffold of the *WecA*'s substrate active site rather than being directly involved in the catalytic process.

2.4 Discussion

The VFMGD motif is a highly conserved region among eukaryotic and prokaryotic PNPT family members (Fig. 2.1B). Our initial hypothesis was that this region and in particular the conserved acidic residue D217 of the *E. coli* *WecA*, is critical to the enzyme's catalytic mechanism, as it was suggested for the analogous residue D267 in the *E. coli* *MraY* (Lloyd *et al.*, 2004). However, the *WecA* VFMGD motif was predicted to be part of TM VIII, which is inconsistent with a putative catalytic role of D217. In this

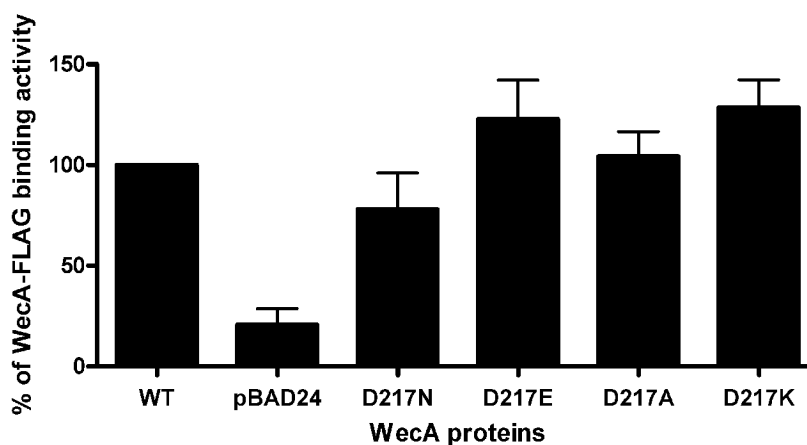


Figure 2.6. Tunicamycin binding assay using parental WecA and D217 replacement mutants. 40 μg of total membranes containing the various D217 mutant proteins were incubated with a concentration of tunicamycin that inhibits 100% activity of 20 μg of total membranes containing wild-type WecA. The residual tunicamycin, unbound by the mutant proteins, was determined by adding the supernatants to 20 μg of total membranes containing wild-type WecA. The data was represented as a percentage of WecA-FLAG binding tunicamycin activity. Data represent duplicates from two independent experiments.

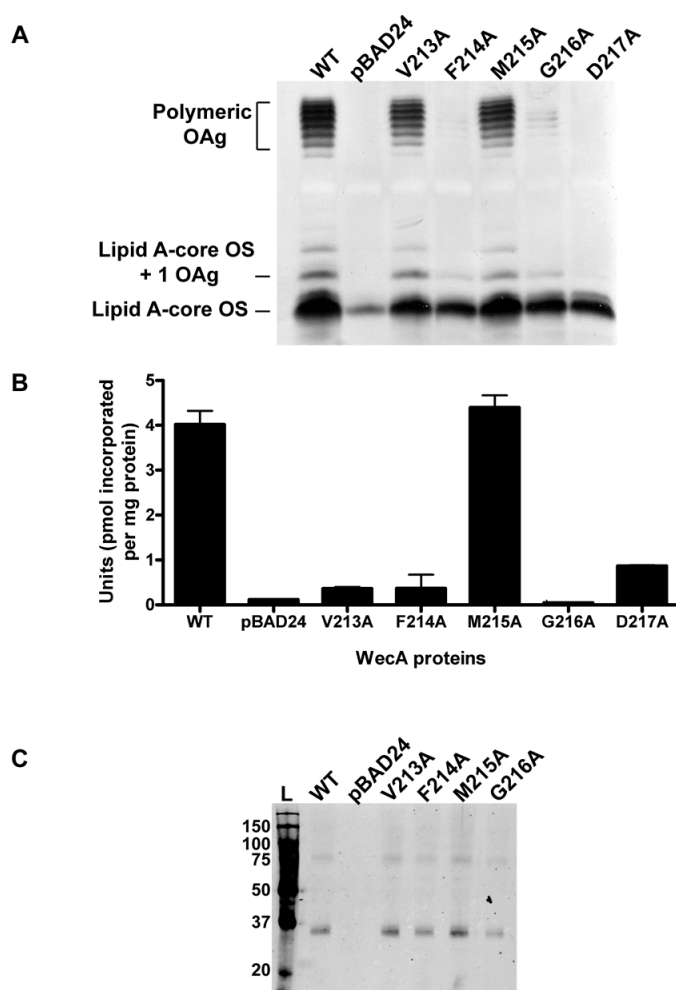


Figure 2.7. *In vivo* and *in vitro* complementation of alanine replacement mutants in the VFMGD motif. A. Complementation of O7 LPS in *E. coli* MV501. B. GlcNAc-1-P transferase activity of parental WecA and alanine replacement mutant proteins. Enzyme units are expressed as pmol of UDP-GlcNAc incorporated per mg of protein. Experiments were performed in triplicate. C. Expression of the mutant proteins in total membranes as detected by Western blotting using anti-FLAG antibodies.

work, we provide experimental evidence based on substituted cysteine mutagenesis that D217 resides in the cytoplasmic space. Confirming this conclusion, the S220C replacement could not be labeled with biotin maleimide (Fig. 2.2), indicating that S220 resides within TM VIII, likely at the boundary between membrane and cytosolic space. Therefore, we have updated the topology model for WecA (Lehrer *et al.*, 2007) by shifting D217 and the rest of the VFMGD motif to the cytoplasmic space in loop IV and placing S220 at the border of TM VIII (Fig. 2.1A). This topology refinement is consistent with the cytoplasmic location of the VFMGD motif in *MraY* (Bouhss *et al.*, 1999).

Further experiments to assess the functional role of D217 demonstrated that, unlike the observations with *MraY* D267 (Lloyd *et al.*, 2004), D217N and D217A replacements did not abolish the *in vitro* enzymatic activity of the mutant WecA proteins (Fig. 2.3B) or their ability to support O antigen synthesis *in vivo* (Fig. 2.3A). In particular, Asn and Ser replacements at the 217 position resulted in enhanced enzymatic activity, as determined by an ~2-fold increase in the reaction product (Fig. 2.4C). These differences could not be attributed to changes in the ability of the mutant proteins to bind the UDP-GlcNAc substrate as compared to parental WecA (Fig. 2.5). Indeed, kinetic experiments revealed that parental and D217N WecA proteins have no significant difference in their K_m values, which is also consistent with the tunicamycin binding experiment results (Fig. 2.6).

Despite no differences in K_m values, kinetic data showed that V_{max} of WecA_{D217N} was 2.5-fold greater than that of the parental WecA (Table 2.2). Changes to the velocity of the enzyme without alteration in the affinity for the substrate may suggest a faster release of the reaction product (Dealwis *et al.*, 1995, Chaidaroglou and Kantrowitz, 1989, Matten and Maness, 1987, Gavilanes *et al.*, 1983). However, we have no direct proof that

this is the mechanism behind the differences between parental WecA and the D217N mutant. Additional mutagenesis of the other residues in the VFMGD motif supports a structural role for this motif instead of a direct role in enzyme catalysis. This stems from the following observations: (i) the replacements showing a severe defect in the ability of WecA to produce O antigen and to perform catalysis *in vitro* did not affect the localization of the mutated protein to the membrane; (ii) the lack of function of the F214A replacement was rescued by a replacement with tyrosine, suggesting a bulky hydrophobic residue is required at this position; (iii) the inability of G216A WecA mutant to remain functional, despite a relatively conservative substitution from Gly to Ala. Furthermore, our results do not support the notion that the VFMGD motif contains a catalytic nucleophile (Lloyd *et al.*, 2004). Previous work in our laboratory demonstrated that the residue D156, located in cytosolic loop III is a catalytic nucleophile, as this is the only aspartic residue we have investigated to date that is absolutely required for WecA activity *in vivo* and *in vitro* when replaced with Asn (Lehrer *et al.*, 2007).

In conclusion, we propose that the conserved VFMGD motif provides a structural scaffold to a putative catalytic site. This model is consistent with our results showing that subtle side chain changes at position D217 and neighboring residues have drastic effects on enzyme activity without changing the binding affinity for UDP-GlcNAc, which suggests that the VFMGD motif may contribute to the structure of the enzyme's active site.

2.5 Chapter two references

- Al-Dabbagh, B., Henry, X., El Ghachi, M., Auger, G., Blanot, D., Parquet, C., Mengin-Lecreulx, D., and Bouhss, A. (2008) Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. *Biochemistry* **47**: 8919-8928.
- Alexander, D. C., and Valvano, M. A. (1994) Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing N-acetylglucosamine. *J Bacteriol* **176**: 7079-7084.
- Amer, A. O., and Valvano, M. A. (2001) Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine. *Microbiology* **147**: 3015-3025.
- Amer, A. O., and Valvano, M. A. (2002) Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in *Escherichia coli*. *Microbiology* **148**: 571-582.
- Anderson, M. S., Eveland, S. S., and Price, N. P. (2000) Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiol Lett* **191**: 169-175.
- Bouhss, A., Mengin-Lecreulx, D., Le Beller, D., and Van Heijenoort, J. (1999) Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol* **34**: 576-585.
- Chaidaroglou, A., and Kantrowitz, E. R. (1989) Alteration of aspartate 101 in the active site of *Escherichia coli* alkaline phosphatase enhances the catalytic activity. *Protein Eng* **3**: 127-132.
- Chamot-Rooke, J., Rousseau, B., Lanternier, F., Mikaty, G., Mairey, E., Malosse, C., Bouchoux, G., Pelicic, V., Camoin, L., Nassif, X., and Dumenil, G. (2007) Alternative *Neisseria* spp. type IV pilin glycosylation with a glyceramido acetamido trideoxyhexose residue. *Proc Natl Acad Sci U S A* **104**: 14783-14788.
- Cohen, S. N., Chang, A. C., and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A* **69**: 2110-2114.
- Dealwis, C. G., Chen, L., Brennan, C., Mandrecki, W., and Abad-Zapatero, C. (1995) 3-D structure of the D153G mutant of *Escherichia coli* alkaline phosphatase: an enzyme with weaker magnesium binding and increased catalytic activity. *Protein Eng* **8**: 865-871.

- Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**: 6127-6145.
- Gavilanes, F., Peterson, D., Bullis, B., and Schirch, L. (1983) Structure and reactivity of cysteine residues in mitochondrial serine hydroxymethyltransferase. *J Biol Chem* **258**: 13155-13159.
- Glover, K. J., Weerapana, E., Chen, M. M., and Imperiali, B. (2006) Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **45**: 5343-5350.
- Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Heifetz, A., Keenan, R. W., and Elbein, A. D. (1979) Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase. *Biochemistry* **18**: 2186-2192.
- Heydanek, M. G., Jr., Struve, W. G., and Neuhaus, F. C. (1969) On the initial stage in peptidoglycan synthesis. 3. Kinetics and uncoupling of phospho-N-acetylmuramyl-pentapeptide translocase (uridine 5'-phosphate). *Biochemistry* **8**: 1214-1221.
- Lehrer, J., Vigeant, K. A., Tatar, L. D., and Valvano, M. A. (2007) Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189**: 2618-2628.
- Lehrman, M. A. (1991) Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. *Glycobiology* **1**: 553-562.
- Lloyd, A. J., Brandish, P. E., Gilbey, A. M., and Bugg, T. D. (2004) Phospho-N-acetylmuramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues. *J Bacteriol* **186**: 1747-1757.
- Mankowski, T., Sasak, W., and Chojnacki, T. (1975) Hydrogenated polyprenol phosphates - exogenous lipid acceptors of glucose from UDP glucose in rat liver microsomes. *Biochem Biophys Res Commun* **65**: 1292-1297.
- Marolda, C. L., Lahiry, P., Vines, E., Saldias, S., and Valvano, M. A. (2006) Micromethods for the characterization of lipid A-core and O-antigen lipopolysaccharide. *Methods Mol Biol* **347**: 237-252.
- Matten, W. T., and Maness, P. F. (1987) V_{max} activation of pp60^{c-src} tyrosine kinase from neuroblastoma Neuro-2A. *Biochem J* **248**: 691-696.

- Merino, S., Jimenez, N., Molero, R., Bouamama, L., Regue, M., and Tomas, J. M. (2011) A UDP-HexNAc:polyprenol-P GalNAc-1-P transferase (WecP) representing a new subgroup of the enzyme family. *J Bacteriol* **193**: 1943-1952.
- Patel, K. B., Ciepichal, E., Swiezewska, E., and Valvano, M. A. (2012a) The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* **22**: 116-122.
- Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology*.
- Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012b) Functional characterization of UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol*.
- Power, P. M., Roddam, L. F., Dieckelmann, M., Srikhanta, Y. N., Tan, Y. C., Berrington, A. W., and Jennings, M. P. (2000) Genetic characterization of pilin glycosylation in *Neisseria meningitidis*. *Microbiology* **146 (Pt 4)**: 967-979.
- Price, N. P., and Momany, F. A. (2005) Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. *Glycobiology* **15**: 29R-42R.
- Rush, J. S., Rick, P. D., and Waechter, C. J. (1997) Polyisoprenyl phosphate specificity of UDP-GlcNAc:undecaprenyl phosphate N-acetylglucosaminyl 1-P transferase from *E. coli*. *Glycobiology* **7**: 315-322.
- Saldías, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**: 440-453.
- Steiner, K., Novotny, R., Patel, K., Vinogradov, E., Whitfield, C., Valvano, M. A., Messner, P., and Schaffer, C. (2007) Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *J Bacteriol* **189**: 2590-2598.
- Valvano, M., Furlong, SE., Patel, KB. (2011) Genetics, Biosynthesis and Assembly of O-Antigen. In: Bacterial Lipopolysaccharides Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells. Yuriy A. Knirel, M. A. V. (ed). SpringerWienNewYork, pp. 275-310.
- Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**: s452-471.

- Valvano, M. A., and Crosa, J. H. (1989) Molecular cloning and expression in *Escherichia coli* K-12 of chromosomal genes determining the O7 lipopolysaccharide antigen of a human invasive strain of *E. coli* O7:K1. *Infect Immun* **57**: 937-943.
- Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598-2604.
- Xayarath, B., and Yother, J. (2007) Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J Bacteriol* **189**: 3369-3381.

Chapter 3

Topological analysis of the *Escherichia coli* WcaJ reveals a new conserved configuration for the polyisoprenyl-phosphate hexose-1-phosphate transferase family of initiating enzymes

Sarah E. Furlong, Lorena Albarnez Rodriguez and Miguel A. Valvano

3.1 Introduction

WcaJ is an integral membrane protein found in *Escherichia coli* K-12 and other *Enterobacteriaceae*, which is involved in the synthesis of colanic acid (CA) capsule (or M-antigen). The protein catalyzes the transfer of glucose-1-P from UDP-Glc to the lipid carrier undecaprenyl-phosphate (Und-P) resulting in the formation of Und-P-P-Glc, which is the first step in CA synthesis (Stevenson *et al.*, 1996, Patel *et al.*, 2012b). CA is a high molecular weight extracellular polysaccharide (EPS) that is produced by many bacteria and is critical for the formation of biofilms, resistance to desiccation, and cell stress responses (Whitfield, 2006, Danese *et al.*, 2000). Unlike other capsular polysaccharides, CA is loosely associated with the outer membrane, but under specific conditions units of CA can be attached to the lipopolysaccharide (LPS) core, resulting in M_{LPS} (Meredith *et al.*, 2007). The production of CA is regulated by the *rcs* (regulator of capsule synthesis) genes *rcsA*, *rcsB*, and *rcsC*. RcsA and RcsB proteins are positive transcriptional regulators and are required for maximal CA production (Gottesman and Stout, 1991).

WcaJ is a member of a bacterial family of initiating glycosyltransferases known as the polyisoprenyl-phosphate hexose-1-phosphate transferase family of initiating enzymes (PHPTs). Unlike polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferases (PNPTs), PHPTs have no known eukaryotic homologs (Valvano, 2003). PHPT members can initiate the synthesis of EPSs in various bacteria (Valvano, 2003) and virulence factors such as O antigen (O Ag) from *Salmonella enterica*, which is initiated by WbaP (Saldias *et al.*, 2008) and S-layers in *Geobacillus stearothermophilus* (Steiner *et al.*, 2007). A majority of these family members also share a similar predicted topology to

the prototypic PHPT member, the *S. enterica* WbaP, with five predicted transmembrane domains, a large periplasmic loop region and a long C-terminal tail region (Fig. 3.1) (Saldias *et al.*, 2008, Patel *et al.*, 2010, Wang *et al.*, 1996). Even amongst a broad spectrum of Gram-negative, Gram-positive, and archaeal PHPTs, the C-terminus is highly conserved consistent with the notion that the last 219 amino acids in C-terminus of WbaP, termed WbaP_{CT}, is sufficient for enzymatic activity (Saldias *et al.*, 2008, Patel *et al.*, 2010, Patel *et al.*, 2012a). Patel *et al.* suggested that the large soluble loop domain may reside in the cytoplasm since protease cleavage experiments in spheroplasts demonstrated that the N-terminal His₆-TrxA fusion partner faces the cytosol. However, it was unclear if this was due to the rapid folding of the TrxA fusion partner, or due to the C-terminal WbaP protein itself (Patel *et al.*, 2010). Therefore, it is important to address the topology of WbaP without the N-terminal TrxA fusion, but we have been unable to characterize the topology of the protein due to our inability to clone the full-length gene and truncations of WbaP are relatively unstable and/or poorly expressed (Saldias *et al.*, 2008). We opted instead to study the topology of similar PHPT member, WcaJ.

To date there is only one topological analysis of a PHPT family member, BceB, an enzyme that initiates the synthesis of Cepacian EPS in *Burkholderia cenocepacia* (Videira *et al.*, 2005). BceB was proposed to have seven transmembrane domains and lacks a distinct large periplasmic loop region, which is inconsistent with previous topological predictions. However, a detailed topological analysis was not performed, as only five C-terminal LacZ and PhoA fusions were studied (Videira *et al.*, 2005). The *E. coli* WcaJ shares a similar *in silico* predicted topology of WbaP (Fig. 3.1) (Wang *et al.*, 1996, Saldias *et al.*, 2008, Patel *et al.*, 2010); therefore, we hypothesized that the last 212

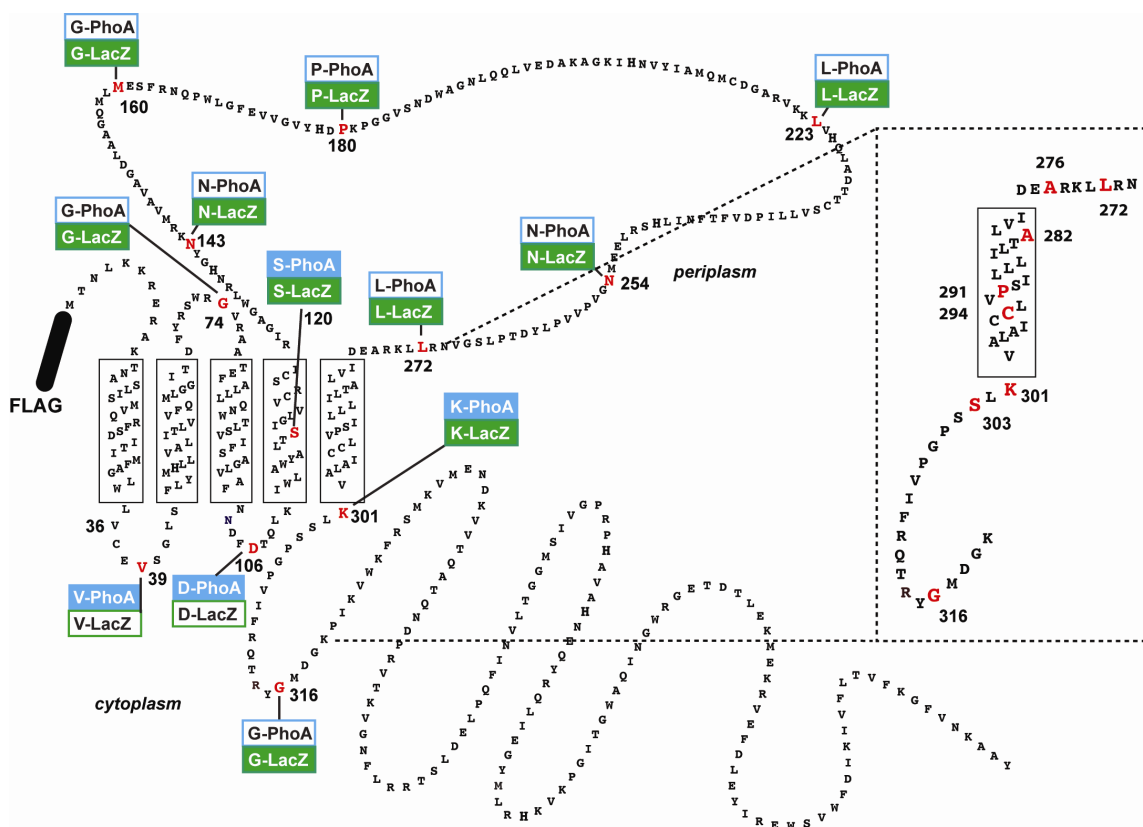


Figure 3.1. Predicted topology of *E. coli* W3110 WcaJ using TMHMM results from PhoA/LacZ reporter fusions. The prediction program suggests a periplasmic N-terminus, a large periplasmic loop, and a long C-terminal cytosolic tail region. The fusion endpoints are indicated by PhoA/LacZ boxes. Positive results are coloured boxes and negative results are white boxes. The last transmembrane domain is shown in the dotted insert and is of particular interest because of conflicting reporter fusion data with positive LacZ results on either side of the membrane domain. The red residues in the box were analyzed using the cysteine replacement sulfhydryl labeling technique.

amino acids of WcaJ (analogous to the last 219 amino acids of WbaP) are also sufficient for enzymatic activity and that they share a similar topology.

Here, we report a detailed topological analysis of a PHPT member using two methods: LacZ and PhoA reporter fusions and sulfhydryl labeling of novel cysteines that are introduced into a Cysteine-less (Cys-Less) WcaJ protein by PEGylation, a technique described by Koch *et al.* 2012. Our results reveal that the topology of Flag-WcaJ is inverted compared to the *in silico* topological predictions, where the N-terminus, C-terminal tail and the large soluble loop reside in the cytoplasm. Notably, we also find that the last membrane domain does not fully span the membrane and is likely ‘pinched in’. Together these data refine the topological model of the PHPT family of enzymes and may provide insight into the role of the large soluble loop domain.

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids, media, and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 3.1. Bacteria were grown aerobically at 37°C in Luria-Bertani (LB) medium (Difco Laboratories, Sparks, MD, USA) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). Growth media were supplemented with 100 µg/ml ampicillin, 20 µg/ml tetracycline, 40 µg/ml kanamycin or 30 µg/ml chloramphenicol. DH5α cells were transformed with plasmids by calcium chloride method (Cohen *et al.*, 1972), CC118 and XBF1 cells were transformed with plasmids by electroporation (Dower *et al.*, 1988).

Table 3.1. List of strains and plasmids

| Strain or Plasmid | Relevant Properties | Source or Reference |
|-------------------|---|--------------------------------|
| Strains | | |
| CC118 | $\Delta(ara\ leu)\ \Delta lac\ phoA\ galE\ galK\ thi\ rpsL\ rpsB\ argE\ recA$ | Laboratory stock |
| DH5 α | F ⁻ $\Phi 80 lacZ\ \Delta M15\ endA\ recA\ hsdR(r_K^- m_K^-)\ nupG\ thi\ glnV\ deoR\ gyrA\ relA1\ \Delta(lacZYA-argF)U169$ | Laboratory stock |
| W3110 | <i>rph-1</i> IN(<i>rrnD-rrnE</i>)1, <i>wbbL::IS5</i> | Laboratory stock |
| XBF1 | W3110, $\Delta wcaJ::aph$, KnR | (Patel <i>et al.</i> , 2012b) |
| Plasmids | | |
| pAH01 | pBAD vector expressing Flag-Wzx-K367-PhoA | (Marolda <i>et al.</i> , 2010) |
| pAH18 | pBAD vector expressing Flag-PhoA | (Marolda <i>et al.</i> , 2010) |
| pAH1809 | pBAD vector expressing Flag-Wzx-T242-PhoA | (Marolda <i>et al.</i> , 2010) |
| pBAD24 | Cloning vector, inducible by arabinose, AmpR | (Guzman <i>et al.</i> , 1995) |
| pBADNTF | pBAD24 vector for N-terminal Flag fusions, AmpR | (Marolda <i>et al.</i> , 2004) |
| pHASoxYZ | <i>E. coli</i> <i>tatA</i> promoter controlling expression of <i>P. panthotrophus</i> HA- <i>soxY</i> and <i>soxZ</i> in pSU20, CmR | (Bauer <i>et al.</i> , 2011) |
| pLA1 | pBAD expressing His5-WcaJ (full length) | This study |
| pLA3 | pBAD expressing Flag-WcaJ (full length) | This study |
| pLA5 | pBAD expressing Flag-WcaJ _{E252-Y464} (C-terminal) | This study |
| pLA6 | pBAD expressing Flag-WcaJ- A134-PhoA | This study |
| pLA7 | pBAD expressing Flag-WcaJ-D106-PhoA | This study |
| pLA8 | pBAD expressing Flag-WcaJ-V39-PhoA | This study |
| pLA9 | pBAD expressing Flag-WcaJ-G74-PhoA | This study |
| pLA11 | pBAD expressing Flag-WcaJ-S360-PhoA | This study |
| pLA18 | pBAD expressing Flag-WcaJ-L223-PhoA | This study |
| pLA19 | pBAD expressing Flag-WcaJ-N254-PhoA | This study |
| pLA20 | pBAD expressing Flag-WcaJ-L272-PhoA | This study |
| pLA23 | pBAD expressing Flag-WcaJ-G316-LacZ | This study |
| pLA24 | pBAD expressing Flag-WcaJ-G316-PhoA | This study |
| pLA30 | pBAD expressing Flag-WcaJ-P180-LacZ | This study |
| pLA31 | pBAD expressing Flag-WcaJ-A134-LacZ | This study |
| pLA32 | pBAD expressing Flag-WcaJ-D106-LacZ | This study |
| pLA33 | pBAD expressing Flag-WcaJ-V39-LacZ | This study |
| pMLTraP-BCAS0627 | promoter of <i>bcas0627</i> cloned in front of full <i>lacZ</i> | Valvano Lab (Unpublished) |
| pSEF62 | pBAD expressing Flag-WcaJ-G74-LacZ | This study |
| pSEF63 | pBAD expressing Flag-WcaJ-L272-LacZ | This study |
| pSEF64 | pBAD expressing Flag-WcaJ-N254-LacZ | This study |
| pSEF65 | pBAD expressing Flag-WcaJ-L223-LacZ | This study |
| pSEF66 | pBAD expressing Flag-WcaJ-S120-LacZ | This study |

| | | |
|---------|--|-------------------------------|
| pSEF67 | pBAD expressing Flag-WcaJ-N143-LacZ | This study |
| pSEF68 | pBAD expressing Flag-WcaJ-M160-LacZ | This study |
| pSEF69 | pBAD expressing Flag-WcaJ-K301-LacZ | This study |
| pSEF71 | pBAD expressing Flag-WcaJ-S120-PhoA | This study |
| pSEF72 | pBAD expressing Flag-WcaJ-N143-PhoA | This study |
| pSEF73 | pBAD expressing Flag-WcaJ-M160-PhoA | This study |
| pSEF74 | pBAD expressing Flag-WcaJ-K301-PhoA | This study |
| pSEF75 | pBAD expressing Flag-WcaJ-P180-PhoA | This study |
| pSEF102 | pBAD expressing Flag-WcaJ-Cys-Less | This study |
| pSEF104 | pBAD expressing Flag-WcaJ-Cys-Less-A282C | This study |
| pSEF105 | pBAD expressing Flag-WcaJ-Cys-Less-A276C | This study |
| pSEF106 | pBAD expressing Flag-WcaJ-Cys-Less-A294C | This study |
| pSEF107 | pBAD expressing Flag-WcaJ-Cys-Less-G316C | This study |
| pSEF108 | pBAD expressing Flag-WcaJ-Cys-Less-L272C | This study |
| pSEF109 | pBAD expressing Flag-WcaJ-Cys-Less-S303C | This study |
| pSEF110 | pBAD expressing Flag-WcaJ-Cys-Less-K301C | This study |
| pWQ499 | pKV102 containing <i>rcsAK30</i> ; TcR | C. Whitfield |
| pXF1 | pBAD24 encoding W3110 WcaJ | (Patel <i>et al.</i> , 2012a) |

3.2.2 Cloning

3.2.2.1 Construction of pLA1

The plasmid pLA1 was constructed by introducing a His₅ tag onto the N-terminus of WcaJ using a site directed mutagenesis strategy. Primers 4779 and 4780 containing novel histidines were used to introduce histidines into the coding region of the N-terminal *wcaJ* gene in pXF1 (using Pfu AD polymerase (Stratagene, Santa Clara, CA, USA) (Table 3.2). *DpnI* was added to PCR reactions for overnight digestion of parental plasmid DNA at 37°C. The resulting DNA was introduced into *E. coli* DH5α by transformation, and transformants were selected on LB-agar containing 100 µg/ml of ampicillin. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.2.2 Cloning of pLA3

To create an N-terminal Flag fusion to WcaJ, the gene encoding this protein was PCR amplified from pLA1 using primers 4911, containing a *SmaI* restriction site, and 4912, containing a *HindIII* restriction site (Table 3.2). *DpnI* was added to the PCR product to digest the parental plasmid template. The PCR product was then cleaned using Qiagen PCR purification kit (Toronto, Ontario, Canada). The pBADNTF vector and PCR product were digested with *SmaI* and *HindIII*. After which the digested DNA was cleaned again with Qiagen PCR purification kit. The digested pBADNTF vector was treated with alkaline phosphatase and then ligated with *wcaJ* PCR product using T4 DNA ligase. The resulting DNA was introduced into *E. coli* DH5α by transformation, and transformants were selected on LB-agar containing 100 µg/ml of ampicillin. Colony PCR was performed using pBADNTF external primers 252 and 258 (Table 3.2). DNA sequencing was performed at York University in Toronto, Ontario, Canada.

Table 3.2. List of primers used for cloning.

| Primer Name | Primer Number | Sequence (5'-3') |
|----------------|---------------|---------------------------------------|
| pBAD forward | 252 | GATTAGCGGATCCTACCTGA |
| pBAD reverse | 258 | GACCGCTTCTGCGTTCTGAT |
| R-His-WcaJ | 4779 | ATGGTGGTGATGGTGCATCCTTGTTCCCTCCATGGTG |
| F-His-WcaJ | 4780 | CACCATCATCACCATACAAATCTAAAAAAGCGC |
| F-SmaI-WcaJ | 4911 | CTAGCCCGGGACAAATCTAAAAAAGCGC |
| R-HindIII-WcaJ | 4912 | CTAGAAGCTTTCAATATGCCGCTTTGTTA |
| F-PstI-WcaJCT | 5094 | CCTGCTGCAGGAGATGAACGGCGTACCG |
| R-PstI-WcaJCT | 5095 | TCGACTGCAGCTTGTCATCGTCATCCTTG |
| F-XbaI-PhoA | 5257 | CCCGTCTAGAGTCGACCTGCAGCCTGTTCTG |
| R-XbaI-V39 | 5364 | CGCATCTAGAGACTTCGCAAACCAGCCATAG |
| R-XbaI-G74 | 5363 | CGCATCTAGAACCGCGCCATGAGCGATAAAA |
| R-XbaI-D106 | 5259 | CGCATCTAGAGTCGAAATCATTGTTGAACGC |
| R-XbaI-S120 | 5841 | CGCATCTAGAGCTGGTCAGCGCATACCACGCCA |
| R-XbaI-A134 | 5258 | CGCATCTAGACGCCCAATGCGAATACACGA |
| R-XbaI-N143 | 5840 | CGCATCTAGAGTTATAGCCATGATTACGCAGCC |
| R-XbaI-M160 | 5839 | CGCATCTAGACATCAGCATTTGCCCGGCGGCTA |
| R-XbaI-P180 | 5093 | CGCATCTAGACGGGTCGTGGTAAACGCCAC |
| R-XbaI-L223 | 5394 | CGCATCTAGAGACCAGTTTTTTTCACTCGCGCGC |
| R-XbaI-N254 | 5393 | CGCATCTAGAGCCGTTTCATCTCTTCGAGGCGTGA |
| R-XbaI-L272 | 5392 | CGCATCTAGACAGGCGGTTAACCCCGGAAAG |
| R-XbaI-K301 | 5867 | CGCATCTAGATTTACCACCAGCGCAATACAGC |
| R-XbaI-G316 | 5580 | CGCATCTAGAGCCGTAGCGAGTCTGGCGG |

3.2.2.3 Construction of pLA5

The pLA5 vector was made by PCR amplifying the C-terminus of *wcaJ*_(E252-Y464) using the forward primer 5094 and the reverse primer 5095 from pLA3 using Pfu AD polymerase (Table 3.2). *DpnI* was added to each PCR product, cleaned with PCR purification kit and digested with *PstI*. The digest was cleaned, and ligated into digested pBADNTF with T4 ligase, transformed into *E. coli* DH5 α cells and selected on LB-agar supplemented with 100 μ g/ml ampicillin. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.2.4 Construction of Flag-WcaJ-A134-PhoA

The Flag-WcaJ-A134-PhoA fusion was made by amplifying *wcaJ* using *Pwo* polymerase and primers 4911 and 5258 (Table 3.2). Add *DpnI* to digest parental plasmid, clean PCR product using PCR purification kit and digest with the PCR product and pSEF74 with *SmaI* and *XbaI*. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.2.5 Construction of Flag-WcaJ-PhoA at the following positions: V39, G74, D106, L223, N254, L272

The WcaJ V39 and G74 PhoA constructs were made by using the direct primer 5257 that amplifies from the *phoA* gene in pLA6 onwards and reverse primers 5364, 5363 that amplify from various positions in the *wcaJ* gene. The D106 fusion was made using the direct primer 5257 that amplifies the *phoA* gene in pLA7 onwards and the reverse primer 5259 that amplifies from D106 position in the *wcaJ* gene. The L223, N254, L272 fusions were made using the direct primer 5257 that amplifies the *phoA* gene in pLA11 and

reverse primers 5394, 5393, and 5392 (Table 3.2). Each primer pair contains an *Xba*I site. The *Pfu* AD polymerase extends around the entire plasmid. *Dpn*I was added to each PCR product, cleaned with PCR purification kit and digested with *Xba*I. Digested plasmids were ligated with T4 DNA ligase, transformed into *E. coli* DH5 α cells and selected on LB-agar supplemented with 100 μ g/ml ampicillin. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.2.6 Construction of Flag-WcaJ-PhoA S120, N143, M160, P180, K301

The S120, N143, M160, P180, and K301 fusions were made by amplifying the *wcaJ* gene from pLA3 using direct primer 4911 and reverse primers 5841, 5840, 5839, 5093, 5867 using *PWO* polymerase (Table 3.2). *Dpn*I was added to each PCR product to digest the parental plasmid. The PCR product was cleaned using PCR purification kit. The PCR product and pAH18 plasmid was digested with *Sma*I and *Xba*I, cleaned and ligated together using T4 ligase, transformed into *E. coli* DH5 α cells and selected on LB-agar supplemented with 100 μ g/ml ampicillin. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.2.7 Construction Flag-WcaJ-LacZ at the following positions: V39, G74, D106, S120, A134, N143, M160, P180, L223, N254, L272, K301, G316

These constructs were made by using the direct primer 5640 that amplifies from the *lacZ* gene in pLA23 onwards and reverse primers, 5364, 5363, 5259, 5841, 5258, 5840, 5839,

5093, 5394, 5393, 5392, 5867, and 5580 that amplify from various positions in the *wcaJ* gene (Table 3.2). Each primer pair contains an *XbaI* site. The *Pfu* AD polymerase extends around the entire plasmid. *DpnI* was added to each PCR product, cleaned with PCR purification kit and digested with *XbaI*. Digested plasmids were ligated with T4 ligase, transformed into *E. coli* DH5 α cells and selected on LB-agar supplemented with 100 μ g/ml ampicillin. DNA sequencing was performed at York University in Toronto, Ontario, Canada.

3.2.3 Creation of cysteine-less WcaJ and introduction of novel cysteines using site directed mutagenesis

The cysteine-less Flag-WcaJ was constructed by replacement of cysteines with alanines using site directed mutagenesis. Replacements were introduced by PCR using primers containing the desired mutations and *Pfu* Turbo AD polymerase (Stratagene, Santa Clara, CA, USA). *DpnI* was added to PCR reactions for overnight digestion of parental plasmid DNA at 37°C. The resulting DNA was introduced into *E. coli* DH5 α by transformation, and transformants were selected on LB-agar containing 100 μ g/ml of ampicillin.

Novel cysteines were introduced in pSEF102, similar to pLA3 but expressing the Flag-WcaJ_{Cys-Less} protein. Replacements were introduced by PCR using primers containing the desired mutations and *Pfu* Turbo AD polymerase (Stratagene, Santa Clara, CA, USA). *DpnI* was added to PCR reactions for overnight digestion of parental plasmid DNA at 37°C. The resulting DNA was introduced into *E. coli* DH5 α by transformation, and transformants were selected on LB-agar containing 100 μ g/ml of ampicillin. DNA

sequencing was performed at York University in Toronto, Ontario, Canada or Eurofins MWG Operon in Huntsville, Alabama, USA.

3.2.4 Complementation of colanic acid in *E. coli* $\Delta wcaJ$, XBF1

Plasmids expressing various Flag-WcaJ replacement proteins were transformed into XBF1/pWQ499 by electroporation and selected on LB-agar supplemented with 100 $\mu\text{g/ml}$ of ampicillin, 20 $\mu\text{g/ml}$ of tetracycline, and 40 $\mu\text{g/ml}$ of kanamycin. Cells were re-plated onto LB-agar with antibiotics with and without 0.2% arabinose. Plates were incubated at 37°C overnight and then incubated at room temperature for an additional 24-48 h to observe mucoidy because CA is optimally produced at 20°C (Gottesman and Stout, 1991).

3.2.5 Growth conditions of cells for protein preparation

Briefly, bacterial cells containing the appropriate arabinose-inducible plasmids were grown overnight in LB containing the appropriate antibiotics (100 $\mu\text{g/ml}$ of ampicillin, 40 $\mu\text{g/ml}$ kanamycin, 20 $\mu\text{g/ml}$ of tetracycline, 30 $\mu\text{g/ml}$ chloramphenicol). From overnight cultures, cells were diluted to an optical density measured at 600 nm (OD_{600}) of 0.2 in LB broth containing the respective antibiotics. Cells were grown at 37°C until an OD_{600} of 0.5-0.7, and cells were induced with 0.2% arabinose for 3 h at 37°C. Cells were harvested by centrifugation 8 000 \times g and cell pellets were frozen at -20°C until needed.

3.2.6 Total membrane preparation and immunoblotting

Cells were resuspended in 50 mM Tris-HCl pH 8 with protease inhibitors and lysed at 10,000 PSI with a cell disruptor (Constant Systems Ltd, Kennesaw, GA). Cell debris was pelleted at 27,216 χ g. Total membranes were isolated by centrifugation in microfuge tubes at 39,191 χ g and resuspended in 50 mM Tris-HCl pH 8. Protein concentrations were determined by Bradford protein assay (Bio-Rad, Hercules, CA, USA). Total membrane preparations were used for immunoblotting and PEGylation labeling. Immunoblotting was performed by separating total membrane preparations by 14% SDS-PAGE and transferring to nitrocellulose membrane. Membranes were blocked overnight in 5% Western blocking reagent (Roche Diagnostics Canada, Laval, QC, Canada) and TBS. The primary antibody, 4.6 mg/ml anti-FLAG M2 monoclonal antibody (Sigma, Saint Louis, MO, USA), was diluted to 1: 10,000 and applied for 1.5 h, and the secondary antibody, 2 mg/ml goat anti-mouse Alexa fluor 680 IgG antibodies (Invitrogen Molecular Probes, Eugene, OR, USA) was diluted to 1: 20,000 and applied for 20 min. Immunoblots were developed using LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE, USA). Bio-Rad Precision Plus Protein Standards were used for all immunoblots.

3.2.7 *In vivo* alkaline phosphatase assay

CC118 cells expressing the WcaJ-PhoA fusion proteins were plated on LB-agar plates supplemented with 100 μ g/ml of ampicillin, 0.2% arabinose to induce the expression of

proteins, and 60 µg/ml of 5-Bromo-4-chloro-3-indolylphosphate *p*-toluidine salt (X-P) (Sigma, Markham, Ontario Canada).

3.2.8 *In vivo* β-galactosidase assay

DH5α cells expressing the WcaJ-LacZ fusion proteins were plated on LB-agar plates supplemented with 100 µg/ml of ampicillin, 0.2% arabinose to induce the expression of proteins, and 40µg/ml of 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Boehringer Mannheim, Germany).

3.2.9 *In vitro* alkaline phosphatase assay

Protocol was previously described by (Manoil, 1991). Briefly, overnight cultures of CC118 cells were diluted and adjusted to an OD₆₀₀ of 0.2 in 5 ml of fresh medium supplemented with 100 µg/ml of ampicillin. Cells were grown for 1.5 h at 37°C. After cells reach an OD₆₀₀ of 0.5, cells were induced with 0.2% arabinose and grown for an additional 3 h at 37°C. All cells were centrifuged in a bench top centrifuge at 16, 100 χ g. Cells were washed once with 1 ml of Buffer 1 (10 mM Tris-HCl pH 8, 10 mM MgSO₄, 1 mM iodoacetamide) and centrifuged again in a bench top centrifuge at 16, 100 χ g. Cells were resuspended in 1 ml of cold Buffer 2 (1 M Tris-HCl pH 8, 1 mM iodoacetamide). The OD₆₀₀ was measured and then 0.1 mL of each undiluted sample was added to 0.9 mL of Buffer 3 (1 M Tris-HCl pH 8, 0.1 mM ZnCl₂, 1 mM iodoacetamide). To permeabilize cells, 50 µl of 0.1% SDS and 50 µl chloroform were added to each sample, vortexed and incubated at 37°C for 5 min. Cells were placed on ice for 5 min. After 0.1 ml of 0.4% *p*-nitrophenyl phosphate (PNPP) (Boehringer Mannheim, Germany) (in 1M Tris-HCl pH 8)

was added, samples were mixed and incubated at 37°C and the time was noted. Samples were incubated until yellow colour was observed and then 120 µl 1M KH₂PO₄ was added to each sample and tubes were placed tubes in an ice water bath to stop the reaction. The time was taken. The OD₅₅₀ and OD₄₂₀ were measured. The Miller units of activity were calculated using the following formula:

$$\text{Units activity} = \frac{[\text{OD}_{420} - (1.75 \times \text{OD}_{550})] 1000}{\text{Time (min)} \times \text{OD}_{600} \times \text{volume cells}}$$

3.2.10 *In vitro* β-galactosidase assay

Overnight cultures of cells were diluted and adjusted to an OD₆₀₀ to 0.2 in 5 ml of fresh medium supplemented with 100 µg/ml of ampicillin. Cells were grown for 1.5 h at 37°C. After cells reach an OD₆₀₀ of 0.5, cells were induced with 0.2% arabinose and grown for an additional 3 h at 37°C. All cells were centrifuged in a bench top centrifuge at 16, 100 χ g. Cells were washed once with 1 ml of 1X M9 media, and centrifuged again and resuspended in a final volume of 1 ml in 1X M9 media. At this point, the OD₆₀₀ was measured. Then 0.1 ml of each undiluted sample was added to 0.9 mL of 1X M9 media 30 µg/ml of chloramphenicol. Incubate at 30°C for 10 minutes. 30 µl of lysis solution was added and samples were shaken vigorously by hand for approximately 20 min or until the solution became clear. After this, 200 µl of 4 mg/ml 2-Nitrophenyl-β-D-galactopyranoside (ONPG) (Fluka Biochemika, Germany) was added and the time was recorded. Reactions were stopped by adding 500 µl of 20% sodium carbonate and the time was recorded again. The OD₅₅₀ and OD₄₂₀ were measured. The Miller units of activity were determined using the following formula:

Units activity = $[\text{OD}_{420} - (1.75 \times \text{OD}_{550})] \times 1000$

Time (min) x OD₆₀₀ x volume cells

3.2.11 Labeling of novel cysteines in the Flag-WcaJ_{Cys-Less} protein using Methoxypolyethylene glycol maleimide (PEG-Mal) in total membrane preparations

DH5 α cells were transformed with plasmids expressing Flag-WcaJ mutant proteins and pHASoxYZ plasmid, constitutively expressing the cytoplasmic control protein SoxY that contains an N-terminal hemagglutinin tag (Bauer *et al.*, 2011). The pHASoxYZ plasmid was provided as a generous gift from Tracey Palmer. Cells were cultured in 50 mls of LB broth at an OD₆₀₀ of 0.2 at 37°C. Cells were grown to an OD₆₀₀ of 0.5-0.7 and added 0.2% arabinose to induce expression of Flag-WcaJ proteins. Cells were grown for an additional 3 h, centrifuged at 12,096 χ g for 10 min at 10°C. Cell pellets were frozen at -20°C. Cell pellets were resuspended in 1 ml of 50 mM HEPES pH 6.8 with 5 mM MgCl₂ and lysed at 10,000 PSI using the Cell Disruptor constant system. Total membranes were isolated by centrifugation and quantified using Bradford assay. In a final volume of 10 μ l, 10 μ g (unless otherwise stated) of total membranes were either left untreated, incubated with 1 mM of PEG-Mal (PEG ~ 5,000 Da; Sigma, Markham, Ontario, Canada) or 1% SDS and 1 mM of PEG-Mal for 1 h at room temperature. Reaction was quenched with 45 mM DTT for 10 min at room temperature. Each sample was separated by 12% SDS-PAGE and transferred to nitrocellulose membrane. Immunoblotting for the membrane fraction was performed using 1: 10,000 dilution of mouse anti-Flag antibodies. Membranes were

probed with the 1: 20, 000 secondary antibody anti-mouse Alexa-fluor 680 and developed using Licor Scanner.

3.2.12 Labeling of novel cysteines in the Flag-WcaJ_{Cys-Less} protein using PEG-Mal in EDTA permeabilized whole cells

DH5 α cells were transformed with plasmids expressing Flag-WcaJ mutant proteins and pHASoxYZ plasmid, constitutively expressing the cytoplasmic control protein SoxY that contains an N-terminal hemagglutinin tag (Bauer *et al.*, 2011). Cells were cultured in 50 mls of LB broth at an OD₆₀₀ of 0.2 at 37°C. Cells were grown to an OD₆₀₀ of 0.5-0.7 and added 0.2% arabinose to induce expression of _{Flag-}WcaJ proteins. Cells were grown for an additional 3 h, centrifuged at 12, 096 χ g for 10 min at 10°C. Wash cells in 5 mls of 50 mM HEPES pH 6.8 with 5 mM MgCl₂. Cells were centrifuged at 12, 096 χ g for 2 min at 10°C. Cell pellets were resuspended in 1 ml of 50 mM HEPES pH 6.8 with 5 mM MgCl₂. The OD₆₀₀ was measured. For each sample, three volumes were aliquoted with a final OD₆₀₀ of 5. One volume was for the whole cell PEG-Mal treatment, one volume was for the non-treatment condition, and one volume was for the lysed cell PEG-Mal treatment.

For the whole cell treatment, cells were permeabilized with 5 mM EDTA for 10 min at room temperature. Then 1 mM of PEG-Mal was added and incubated at room temperature for 1 h. Reaction was quenched with 45 mM DTT for 10 min at room temperature. Cells were centrifuged for 1 min at 16, 100 χ g in a benchtop centrifuge. Supernatants were removed and cell pellets were frozen at -20°C.

For the non-treatment condition, cells were incubated at room temperature for 1 h, were centrifuged for 1 min at 16, 100 χ g in a benchtop centrifuge and supernatants were removed and cell pellets were frozen at -20°C.

For the lysed cell PEG-Mal treatment, cells were incubated at room temperature for 1 h, were centrifuged for 1 min at 16, 100 χ g in a benchtop centrifuge and supernatants were removed and cell pellets were frozen at -20°C. Cell pellets were resuspended in 2 mL of cold 50mM HEPES pH 6.8 with 5 mM MgCl₂ and protease inhibitors. Cells were lysed at 10, 000 PSI using the Cell Disruptor constant system. Then 1 mM of PEG-Mal was added and incubated at room temperature for 1 h. Reaction was quenched with 45 mM DTT for 10 min at room temperature. The rest of the sample cell pellets were resuspended in 1 mL of cold 50mM HEPES pH 6.8 with 5 mM MgCl₂ and protease inhibitors. Cells were lysed at 10, 000 PSI using the Cell Disruptor constant system. After which, the membrane and cytoplasmic fractions were isolated by centrifugation and quantified using Bradford assay. The cytoplasmic fraction was separated using 18% SDS-PAGE and the membrane fraction was separated using 12% SDS-PAGE. Immunoblotting for cytoplasmic fraction performed using 1: 10, 000 dilution of mouse anti-HA antibodies and immunoblotting for the membrane fraction was performed using 1: 10, 000 dilution of mouse anti-Flag antibodies. Both membranes were probed with the 1:20, 000 dilution of the secondary antibody anti-mouse Alexa-fluor 680. Blots were developed using Licor Scanner.

3.3 Results

3.3.1 The last 212 amino acids of the C-terminus of the *E. coli* WcaJ is sufficient for enzymatic activity *in vivo*

Using commonly employed algorithms for the topological prediction of transmembrane helices (TMHs) members of the PHPT family were proposed to have five predicted TMHs that are separated into three distinct domains: an N-terminal domain encompassing TMH-I to TMH-IV, a large periplasmic domain between TMH-IV and TMH-V and a large C-terminal cytoplasmic domain (Fig. 3.1) (Saldias *et al.*, 2008, Valvano *et al.*, 2011). However, the topology of PHPT proteins has not been experimentally validated. While the predicted N-terminal and periplasmic domains are not always present in PHPT proteins, the TMH-V and the C-terminal domain are highly conserved features in all family members, and we have demonstrated that this region in WbaP and PssY is sufficient for enzymatic activity both *in vivo* and *in vitro* (Wang *et al.*, 1996, Saldias *et al.*, 2008, Patel *et al.*, 2010, Patel *et al.*, 2012a). Therefore, prior to using WcaJ as a prototypic PHPT member for topological analysis, it was critical to determine if the C-terminal domain of WcaJ was functionally comparable to the C-terminus of WbaP and PssY. Full-length Flag-WcaJ as well as Flag-WcaJ_{CT}, a truncated version carrying residues E252-Y464, which are analogous to the WbaP_{CT} derivative (F258-Y476) used before (Saldias *et al.*, 2008, Patel *et al.*, 2010), were constructed and expressed from an arabinose inducible promoter (Fig. 3.2A). The plasmids encoding these proteins (pLA3 and pLA5, Table 3.1) were introduced by transformation into *E. coli* $\Delta wcaJ$ containing a plasmid encoding RcsA (pWQ499), the positive transcriptional regulator of CA synthesis (Bauer *et al.*, 2011). In the presence of excess RcsA,

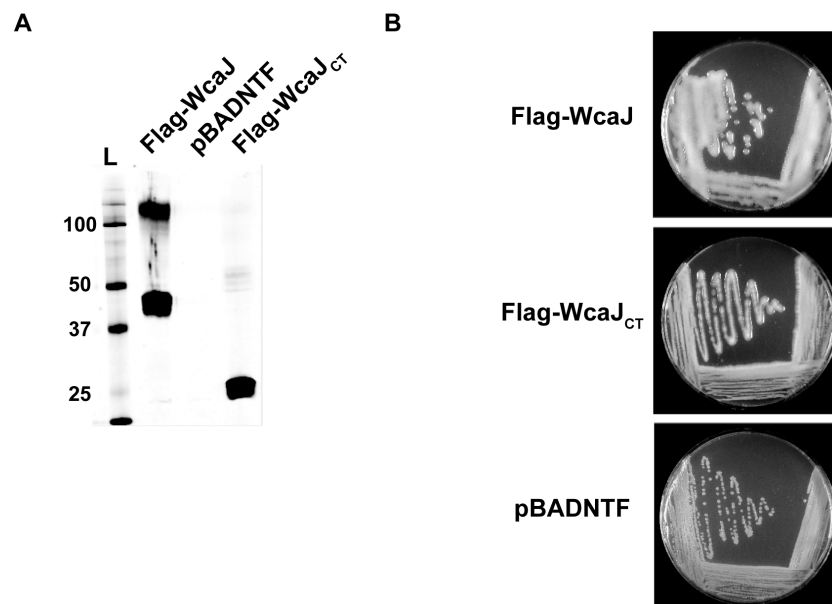


Figure 3.2. Membrane protein expression of Flag-WcaJ and Flag-WcaJ_{CT} constructs and complementation of CA in XBF1/pWQ499. Flag-WcaJ and Flag-WcaJ_{CT} were expressed from the arabinose-inducible vector pBADNTF. A) Total membranes were isolated from DH5α cells expressing constructs and immunoblotting was performed using anti-Flag antibodies. B) XBF1/pWQ499 cells expressing constructs were plated on LB-agar supplemented with appropriate antibiotics and 0.2% arabinose. The Flag-WcaJ and Flag-WcaJ_{CT} constructs produce CA similarly and the vector control cannot complement this phenotype.

production of CA generates a highly mucoid colony phenotype (Stout *et al.*, 1991). The results showed that expression of Flag-WcaJ_{CT} and Flag-WcaJ restored mucoidy, while no mucoidy was observed in $\Delta wcaJ$ cells transformed with the pBADNTF vector control (Fig. 3.2B). These data support the notion that the WcaJ C-terminal domain is sufficient for enzymatic activity, as in the other members of the PHPT family.

3.3.2 Topological analysis of WcaJ using LacZ and PhoA reporter fusions reveals an unexpected membrane topology

To validate the *in silico* predicted membrane topology of WcaJ (Fig. 3.1) we employed a dual reporter fusion strategy based on LacZ and PhoA reporters, which were C-terminally fused to Flag-WcaJ truncations and expressed in *E. coli* strains CC118 ($\Delta phoA$) and DH5 α (*lacZ* Δ M15), respectively. Colonies were examined for blue or white phenotypes on LB-agar plates that were supplemented with chromogenic substrates X-P and X-Gal, as appropriate. PhoA is only active in the periplasm, as it requires a reducing environment to form disulphide bonds (Derman and Beckwith, 1991, Derman and Beckwith, 1995). Conversely, LacZ functions in the cytosol. Also, the *in vitro* alkaline phosphatase and β -galactosidase activities were measured in Miller Units (MU), as described in Material and Methods (Manoil, 1991) (Table 3.3). The results revealed an inverted membrane topology compared to the *in silico* predicted topology (Fig. 3.1). Flag-WcaJ-LacZ fusions at residues valine-39 and aspartic acid-106, which are located in predicted cytoloops I and II, did not exhibit β -galactosidase activity *in vivo* or *in vitro* (Table 3.3). On the contrary, the corresponding Flag-WcaJ-PhoA fusions at the same positions yielded strong blue colonies in X-P and a positive alkaline phosphatase *in vitro*

(2112 ± 95 MU, 938 ± 70 MU) (Table 3.3), suggesting that these residues reside in the periplasmic space. Despite relatively low *in vitro* activity, the LacZ fusion to glycine-74, a residue predicted to be in periplasmic loop I, yielded blue colonies on X-gal plates, and the corresponding PhoA fusion gave no color on X-P plates, suggesting this residue is exposed to the cytosolic space. Furthermore, LacZ fusions at residues in the predicted periplasmic loop II (asparagine-143, methionine-160, proline-180, leucine-223, asparagine-254, and leucine-272) gave blue colonies on X-Gal plates and most showed strong *in vitro* β -galactosidase activity (Fig. 3.1 and Table 3.3). Conversely, none of the corresponding PhoA fusions demonstrated blue color in X-P plates and had very low or no detectable alkaline phosphatase activity *in vitro* (Table 3.3). Therefore, these results demonstrate that the large soluble domain of WcaJ resides in the cytoplasm, in contrast to the *in silico* prediction (Fig. 3.1).

LacZ and PhoA fusions at serine-120 afforded inconclusive results as both gave relatively low *in vitro* enzymatic activity for both enzymes and colonies with both constructs produced blue colonies when plated on media containing X-gal or X-P (Table 3.3). This is commonly seen in fusions to amino acids located in a TMH, so we interpret these results as an indication that serine-120 likely resides within TMH-IV (Fig. 3.1). A similar conclusion was drawn for the fusions at position lysine-301, which is predicted *in silico* to be located at the border of TMH-V (Fig. 3.1). In contrast, the fusions at glycine-316, which is expected to be part of the functional C-terminal domain of WcaJ, gave results consistent with its predicted cytosolic location (Fig. 3.1 and Table 3.3).

Immunoblotting was performed on total membrane preparations of CC118 cells expressing Flag-WcaJ-PhoA or DH5 α cells expressing Flag-WcaJ-LacZ fusion proteins.

Table 3.3. Alkaline phosphatase and β -galactosidase results of Flag-WcaJ-LacZ/PhoA fusion proteins expressed in DH5 α or CC118 cells. Plate assays are observed by cleavage of X-gal or X-P substrate on plates and *in vitro* alkaline phosphatase and β -galactosidase assay were observed by the hydrolysis of PNPP and ONPG.

| Amino Acid Fusion | Predicted Location | Plate assay | Miller Units | Location of Residue |
|---|--------------------|-------------|-------------------|---------------------|
| V39-PhoA | CL1 | + | 2111.8 \pm 94.6 | PL1 |
| V39-LacZ | CL1 | - | N.D. | PL1 |
| G74-PhoA | PL1 | - | 4.4 \pm 2.2 | CL1 |
| G74-LacZ | PL1 | + | 0.2 \pm 0.1 | CL1 |
| D106-PhoA | CL2 | + | 938.0 \pm 70.5 | PL2 |
| D106-LacZ | CL2 | - | N.D. | PL2 |
| S120-PhoA | TMIV | + | 97.0 \pm 2.5 | TMIV |
| S120-LacZ | TMIV | + | N.D. | TMIV |
| A134-PhoA | PL2 | - | 15.4 \pm 1.3 | CL2 |
| A134-LacZ | PL2 | - | N.D. | CL2 |
| N143-PhoA | PL2 | - | 15.4 \pm 3.9 | CL2 |
| N143-LacZ | PL2 | + | 1.3 \pm 0.2 | CL2 |
| M160-PhoA | PL2 | - | 13.4 \pm 3.5 | CL2 |
| M160-LacZ | PL2 | + | 1.2 \pm 0.2 | CL2 |
| P180-PhoA | PL2 | - | N.D. | CL2 |
| P180-LacZ | PL2 | - | 0.3 \pm 0.1 | CL2 |
| L223-PhoA | PL2 | - | N.D. | CL2 |
| L223-LacZ | PL2 | + | 65.8 \pm 3.2 | CL2 |
| N254-PhoA | PL2 | - | N.D. | CL2 |
| N254-LacZ | PL2 | + | 15.5 \pm 0.6 | CL2 |
| L272-PhoA | PL2 | - | N.D. | CL2 |
| L272-LacZ | PL2 | + | 21.9 \pm 0.4 | CL2 |
| K301-PhoA | CT | + | 611.1 \pm 43.9 | TMV/CT? |
| K301-LacZ | CT | + | 4.8 \pm 0.5 | TMV/CT? |
| G316-PhoA | CT | - | N.D. | CT |
| G316-LacZ | CT | + | 12.9 \pm 1.2 | CT |
| S360-PhoA | CT | - | N.D. | CT |
| S360-LacZ | CT | N/A | N/A | CT |
| Flag-Wzx-K367-PhoA (+ve control) | P | + | 992.5 \pm 53.9 | N/A |
| Flag-Wzx-T242-PhoA (-ve control) | C | - | N.D. | N/A |
| Vector control | N/A | - | N.D. | N/A |
| Constitutive promoter of <i>bcas0627</i> cloned in front of <i>lacZ</i> | N/A | + | 1405.5 \pm 14.6 | N/A |
| Vector control | N/A | - | N.D. | N/A |

Note: All samples were tested at the same time, in triplicate. Miller units represent mean \pm standard deviation. Activity that was not detectable is marked N.D.

All of the fusion proteins were detectable by immunoblot, confirming that the reduced enzymatic activity *in vivo* and *in vitro* was likely not due to membrane protein expression (Fig. 3.3 and 3.4).

Together, the reporter fusion data suggest that the topology of WcaJ is different from that predicted by *in silico* methods. The C-terminal domain and the large soluble loop between TMH-IV and -V are cytoplasmic, and the other loop regions have an inverted topology compared to the prediction. However, the most remarkable consequence from the experimental results is that TMH-V separates two cytosolic soluble domains, suggesting this TMH has an unusual structure.

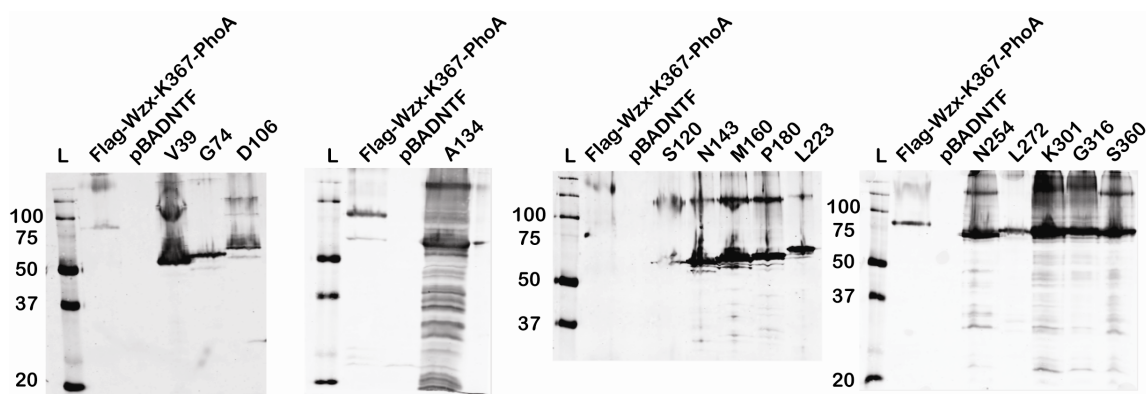


Figure 3.3. Membrane expression of Flag-WcaJ-PhoA fusion proteins in CC118 cells. 30 μ g of total membrane preparations from CC118 expressing Flag-WcaJ-PhoA fusions were separated by 12% SDS-PAGE and immunoblotting was performed using anti-Flag antibodies. Flag-Wzx-K367-PhoA (Marolda *et al.*, 2010) was used as positive control for PhoA western blots and experiments.

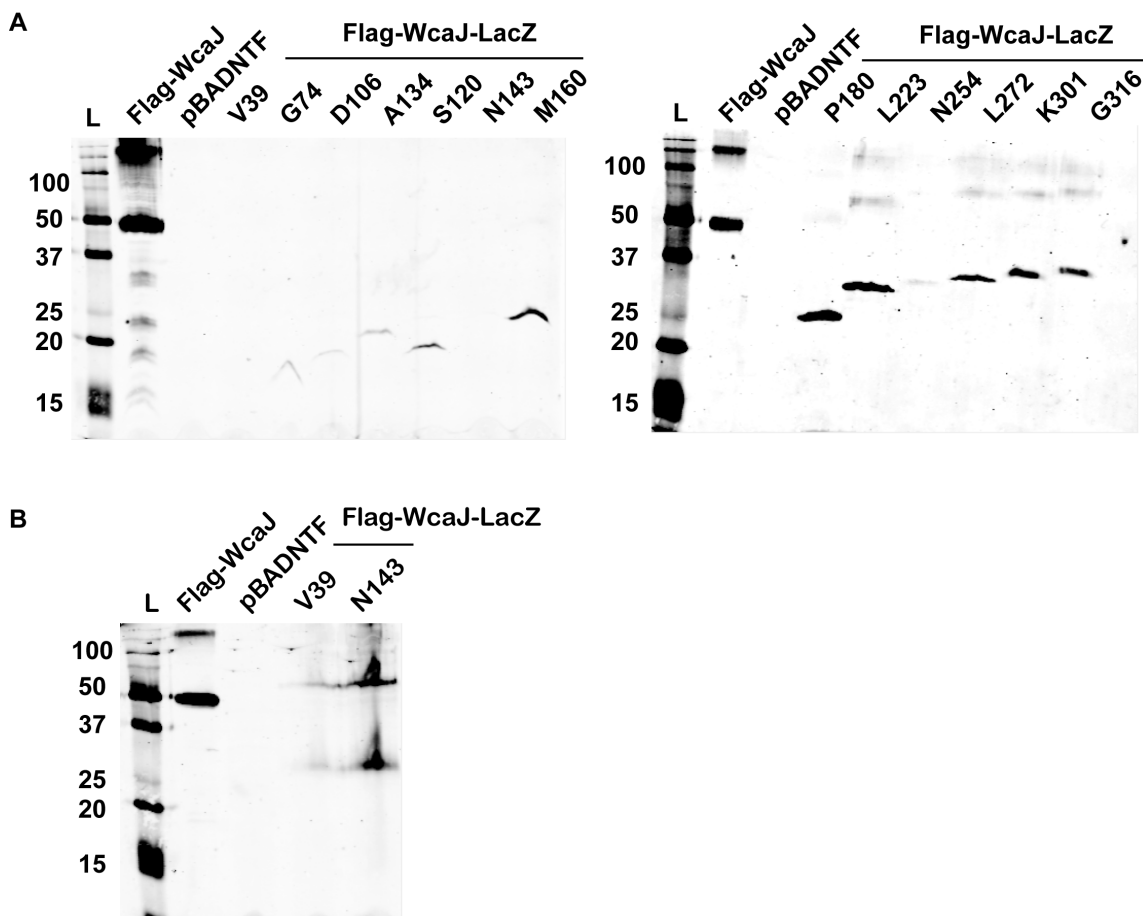


Figure 3.4. Membrane expression of Flag-WcaJ-LacZ fusion proteins in DH5α cells.

Total membrane preparations were separated by 18% SDS-PAGE and immunoblotting was performed using anti-Flag antibodies. A) 10 μ g of total membranes from DH5α expressing Flag-WcaJ were used as a positive control. 30 μ g of total membranes from DH5α expressing Flag-WcaJ-LacZ fusions were separated by SDS-PAGE. B) 10 μ g of total membranes from DH5α expressing Flag-WcaJ were used as a positive control. 50 μ g of total membranes from DH5α expressing Flag-WcaJ-LacZ fusions were separated by SDS-PAGE.

3.3.3 Analysis of the borders of TMH-V suggests a helix-break-helix structure

To examine in more detail the borders and the characteristics of TMH-V, we turned to a substituted cysteine labeling approach using sulfhydryl reactive reagents. Using this strategy the native structure of the protein is maintained with minimal disruption. The seven native cysteines in Flag-WcaJ were replaced with alanine residues to create a cysteine-less protein, WcsJ_{Cys-less}. Despite slight reduction in protein expression/stability of WcaJ_{Cys-less} observed by immunoblotting (Fig. 3.5A), the protein complements CA synthesis to similar levels as those mediated by the parental enzyme, as observed by the restoration of mucoidy in $\Delta wcaJ$ /(pWQ499) (Fig. 3.5B). Novel cysteines were introduced by site-directed mutagenesis at various positions around and within TMH-V including residues leucine-272, alanine-276, alanine-282, proline-291, cysteine-294 (by re-introducing the native cysteine), lysine-301, serine-303, and glycine-316. The cysteine replacement proteins were tested for their ability to complement CA synthesis by transforming plasmids into $\Delta wcaJ$ /(pWQ499). All proteins, except P291C and G316C, complement like the wild-type enzyme (Fig. 3.6). The lack of complementation with P291C and G316C was attributed to poor membrane protein expression and/or stability (Fig. 3.7).

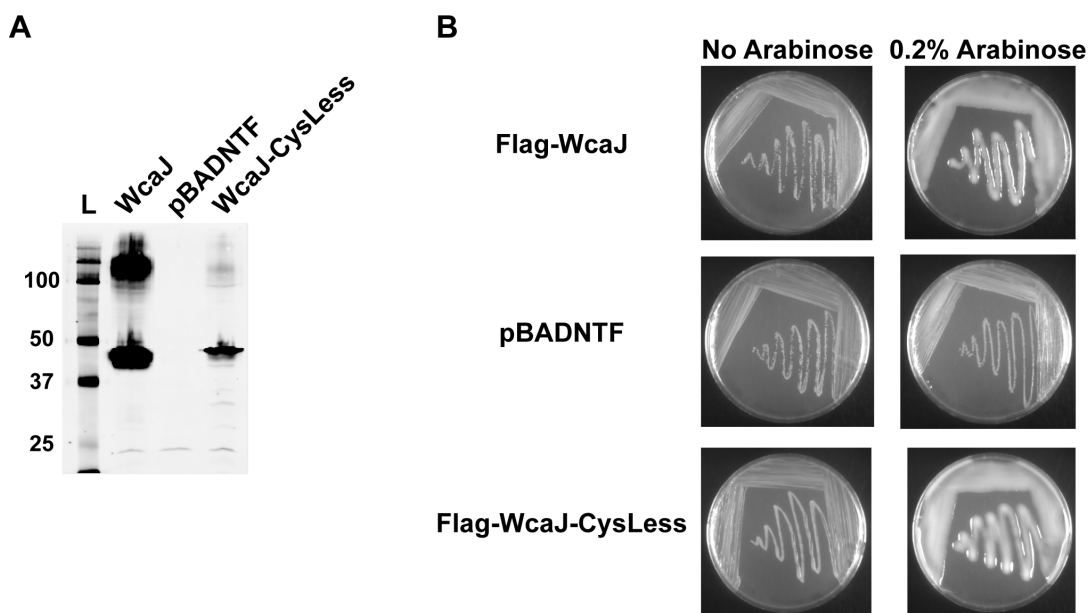


Figure 3.5. Membrane protein expression of Flag-WcaJ_{Cys-Less} and complementation of CA. A) 10 μ g of total membrane preparations from DH5 α cells expressing Flag-WcaJ and Flag-WcaJ_{Cys-Less} were separated by 14% SDS-PAGE and immunoblotting was performed using anti-Flag antibodies. B) Flag-WcaJ and Flag-WcaJ_{Cys-Less} proteins were expressed in XBF1/pWQ499 by plating cells on LB-agar supplemented with appropriate antibiotics and 0.2% arabinose.

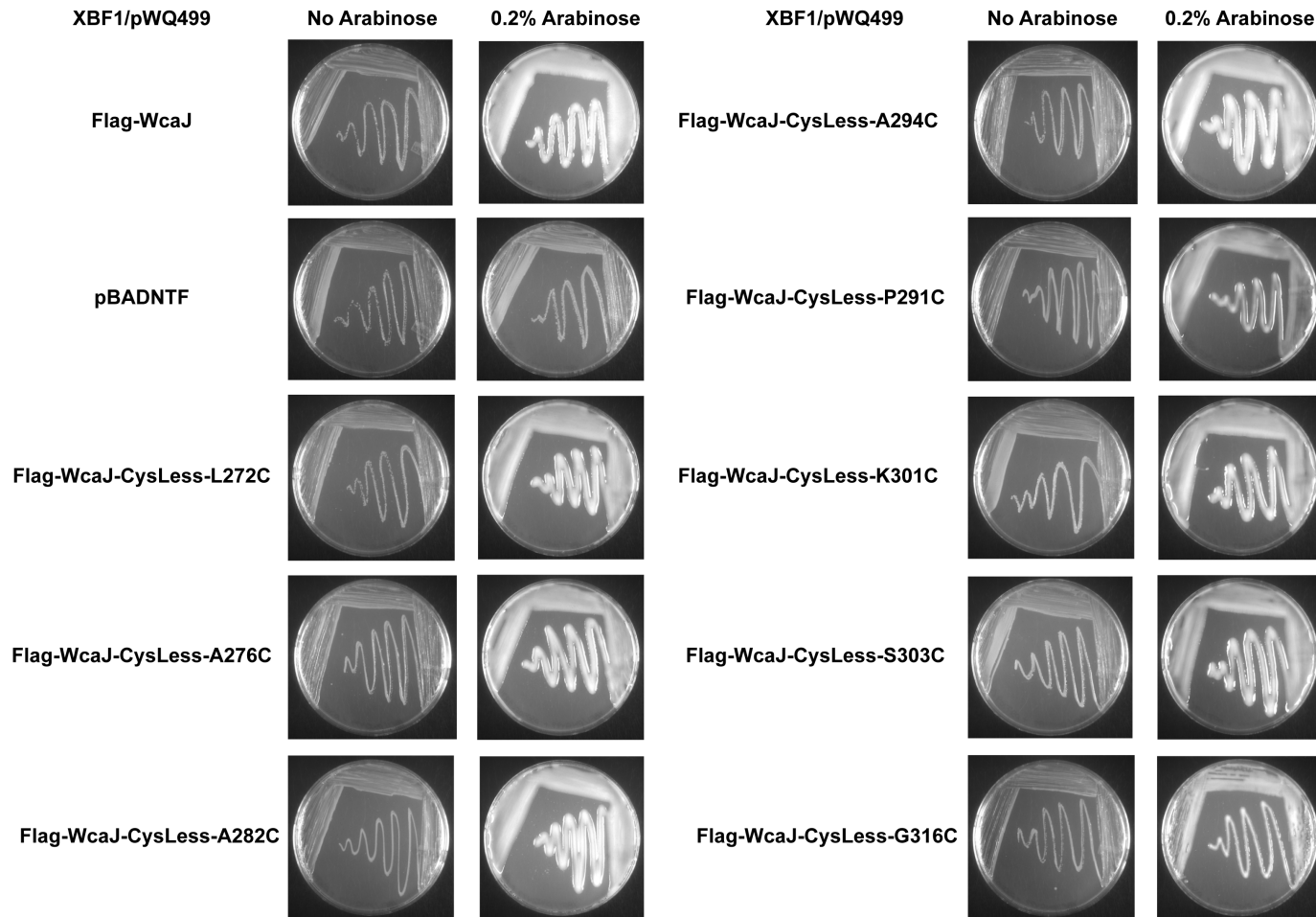


Figure 3.6. Complementation of CA in XBF1/pWQ499 by cysteine replacement proteins. Proteins were expressed in XBF1/pWQ499 by plating cells on LB-agar supplemented with appropriate antibiotics and 0.2% arabinose.

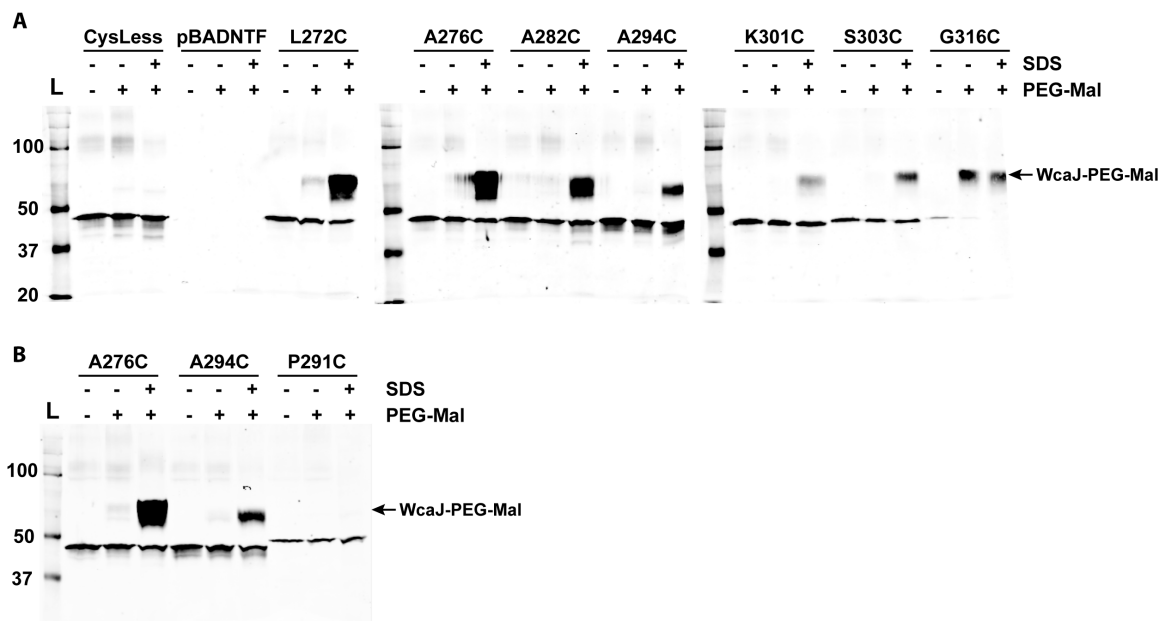


Figure 3.7. Substituted cysteine labeling using PEG-Mal in total membrane preparations expressing Flag-WcaJ_{Cys-Less} cysteine replacement proteins. Total membranes were untreated, incubated with 1 mM PEG-Mal to reveal exposed residues, or 1 mM PEG-Mal with 1% SDS to reveal membrane domain residues. Band shifting of proteins indicates PEG-Mal labeled proteins. 10 μ g of each membrane fraction was labeled, except for P291C, which had 30 μ g.

To determine the topology of the novel cysteines in WcaJ, we performed sulfhydryl-labeling experiments with PEG-Mal as described in Materials and Methods. As an internal control and also a control for cell lysis, the cysteine replacement proteins were co-expressed with HA-SoxY, a cytoplasmic protein with an N-terminal hemagglutinin tag and a single available cysteine on a flexible arm (Sauve *et al.*, 2007, Koch *et al.*, 2012). Labeling was performed in total membrane preparations with and without treatment with 1% SDS to differentiate between solvent exposed and lipid bilayer buried residues. Labeling of intact membranes with PEG-Mal detected only L272C and G316C. In contrast, the other residues, except P291C, were PEGylated only in the membranes treated with SDS indicating that these residues are buried in the membrane (Fig. 3.7). To assess whether L272C and G316C, were cytoplasmic, we labeled EDTA permeabilized whole cells and found that only the periplasmic control, the wild-type protein with intact native periplasmic cysteine C37 could be labeled in whole cells. In the whole cell labeling conditions, cell membrane integrity was not compromised, as the HA-SoxY protein was not PEGylated. Only upon lysing the cells prior to PEG-Mal labeling, could we observe PEGylation of L272C and G316C concomitantly with PEGylation of HA-SoxY (Fig. 3.8). These results provided strong evidence that amino acids at position L272 and G316 are exposed to the cytosol, while the remaining amino acids are located in TMH-V. P291C could not be labeled with PEG-Mal under any treatment condition, including SDS, despite that the nearby residue A294C could be readily labeled after the addition of PEG-Mal. This suggests that P291C may adopt a conformation in the membrane that prevents labeling. Also, the WcaJ_{P291C} was poorly expressed and migrated

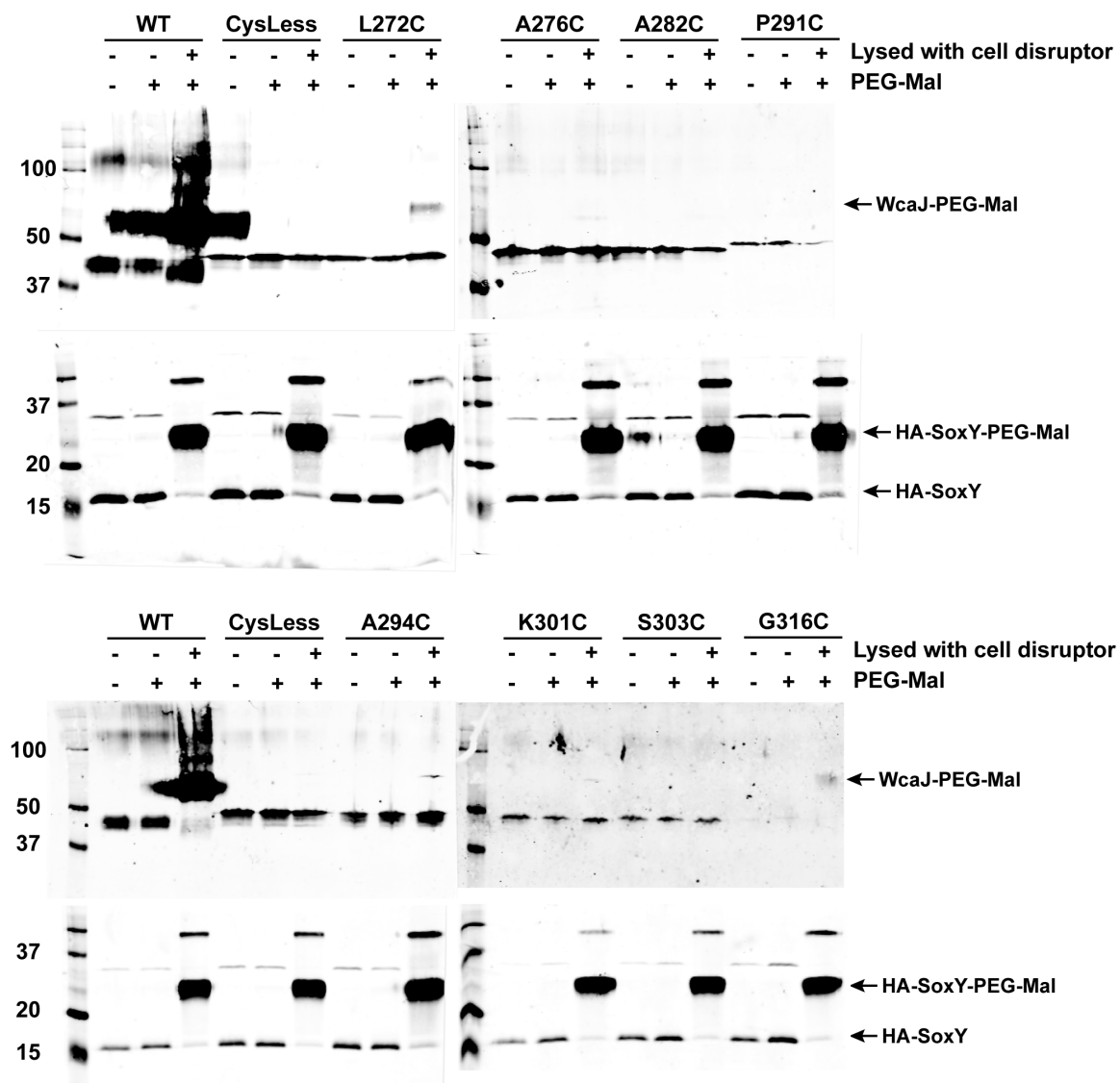


Figure 3.8. Substituted cysteine labeling using PEG-Mal in EDTA-permeabilized whole cells and lysed cells. EDTA-permeabilized whole cells were treated with PEG-Mal to assess if residues were periplasmic exposed. Cells were lysed with cell disruptor prior to PEG-Mal treatment to identify cytoplasmic exposed residues. Wild-type (WT) WcaJ was used as a positive periplasmic control (C37) and HA-SoxY was used as a lysis control. 10 μ g of each membrane fraction was separated by SDS-PAGE, except for P291C, which had 30 μ g.

abnormally compared to the other proteins, supporting the idea of an altered secondary structure in this mutant. To gain additional insight on the potential structure of this TMH, we analyzed its secondary structure with PSIPRED, which predicted a helical conformation between leucine-273 and leucine-302 (Fig. 3.9A). This prediction was in agreement with the experimental boundaries determined for TMH-V. The region R₂₇₀LLK...G₃₀₆ was modeled with TMKink, a program that identifies kinks within helices. Meruelo *et al* demonstrate that there are residue preferences that contribute to kinks and that a score above the designated threshold has a kink in the helix (Meruelo *et al.*, 2011). Using TMKink, we found that this region has a predicted kink (residues are red) in TMH-V (Fig. 3.9B). We also modeled this region with PEP-FOLD, which revealed a helix-break-helix predicted structure (Fig. 3.9C). The break region corresponded to serine-290 and proline-291. Further analysis using of WcaJ residues E252-Y464 using ConSeq revealed that this proline is buried and highly conserved in PHPT homologs (Fig. 3.10), suggesting the structural features of TMH-V are also conserved in the family members (Berezin *et al.*, 2004). Together, our results strongly suggest that TMH-V has a helix-break-helix structure and we have therefore modified the protein topology of WcaJ (Fig. 3.11).

3.1 Discussion

This study provides experimental elucidation of the topology of WcaJ, demonstrating the existence of a single transmembrane helix domain (TMH-V) flanked by two large cytosolic regions, one of which corresponds to the C-terminal domain of the protein. Protease accessibility experiments previously performed in a truncated version of WbaP,

**Legend:****The conservation scale:**

1 2 3 4 5 6 7 8 9

Variable Average Conserved

e - An exposed residue according to the neural-network algorithm.

b - A buried residue according to the neural-network algorithm.

f - A predicted functional residue (highly conserved and exposed).

s - A predicted structural residue (highly conserved and buried).

X - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.

Figure 3.10. ConSeq results of WcaJ_{CT}. Using multiple programs, ConSeq identifies multiple homologs, compiles them into a multiple sequence alignment, and calculates conservation scores for each residue (pink most conserved), and predicts if they are buried (b) or exposed (e).

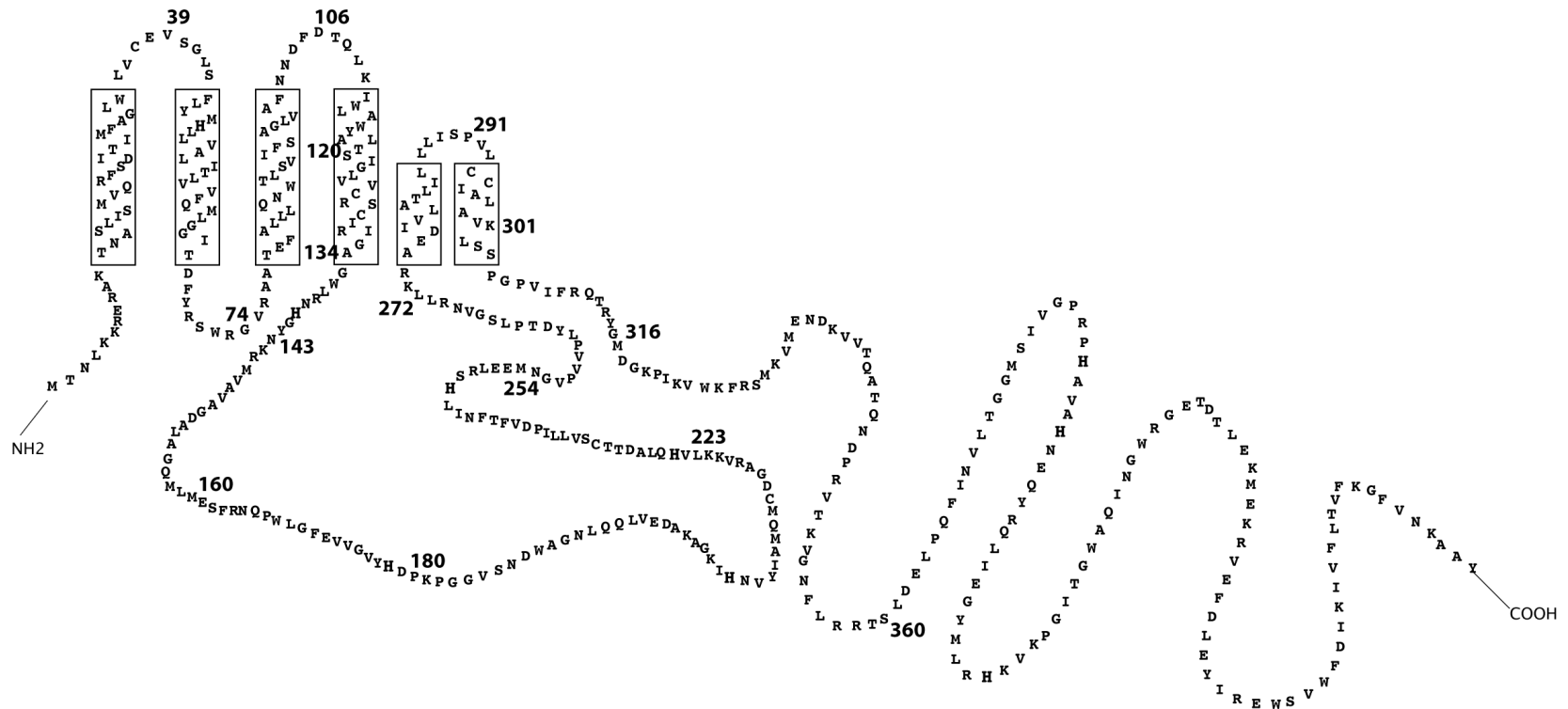


Figure 3.11. Refined topological model of the *E. coli* WcaJ after using reporter fusions and substituted cysteine labeling with PEG-Mal. The refined topology has four transmembrane domains, one ‘pinched in’ domain, two periplasmic loops, one large cytoplasmic loop, and a long C-terminal cytosolic tail.

another PHPT family member, suggested that TMH-V does not fully span the membrane (Patel *et al.*, 2010). The WbaP construct investigated had a His₆-N-terminal TrxA protein fused to part of the predicted periplasmic loop also including TMV and the entire C-terminal tail. Protease cleavage experiments in spheroplasts and lysed cells demonstrated that the His₆-N-terminal TrxA fusion partner localized to the cytoplasmic space. However, it was unclear if the resulting topology was an artifact attributable to the naturally cytoplasmic TrxA protein partner or due to the native structure of TMV not fully spanning the membrane (Patel *et al.*, 2010).

The availability of a cysteine-less WcaJ provided us with the opportunity to investigate the topology of TMV and its neighboring soluble regions utilizing a substituted cysteine replacement sulfhydryl labeling technique with PEG-Mal. These experiments demonstrate that residues flanking the N- and C-terminal end of TMH-V (leucine-272 and glycine-316, respectively) face the cytosol (Fig. 3.7 and 3.8). In contrast, the other residues at positions alanine-276, alanine-282, cysteine-294 (native), lysine-301, and serine-303 are buried in the lipid bilayer (Fig. 3.7). Alignment of the region corresponding to TMH-V in WcaJ and other PHPT members reveals a conserved proline (at position 291 in WcaJ) (Fig. 3.10). The substitution of this proline with cysteine resulted in a WcaJ derivative that was poorly expressed, could not be labeled with PEG-Mal under any condition including treatment with SDS, and migrated abnormally in the gel. Similar observations were made when P291 was replaced with alanine (data not shown). These results suggest the proline-291 is a critical site to maintain the secondary structure of TMV. Investigation of TMH-V sequences with the program TMKink revealed they contain 8 amino acids, always including the conserved

proline, which has a high probability score for a helix kink, supporting the notion that the TMH-V does not fully span the membrane bilayer.

Studies on the membrane protein caveolin-1 have identified a helix-break-helix structure causing the N- and C-terminal domains of this helix to reside on the same side of the membrane forming an “intra-membrane horseshoe conformation” that does not fully span the membrane. A proline and an isoleucine were among the key residues contributing to the “break” in helix, and their replacement by alanine dramatically disrupted the helix-break-helix structure (Lee and Glover, 2012). Interestingly, the TMH-V is of similar size to the helix-break-helix structure (~27 residues) and has a highly conserved proline and isoleucine at positions 291 and 289, respectively. Therefore, like in the reduced CA complementation in $\Delta wcaJ$, the altered protein migration on SDS-PAGE and reduced membrane expression or stability of $WcaJ_{P291C}$, provide strong support to the notion that TMH-V contains a helix-break-helix structure and cannot span the membrane bilayer, thus explaining why a single predicted TM is found between two regions of the proteins that are facing the cytosol.

In summary, this study reveals an unexpected membrane topology for $WcaJ$, which is proposed as a signature for all members of the PHPT family. Further investigations are underway in our laboratory to elucidate the function of the central cytosolic domain, previously thought to be located on the periplasmic space.

3.5 Chapter three references

- Bauer, J., Fritsch, M. J., Palmer, T., and Uden, G. (2011) Topology and accessibility of the transmembrane helices and the sensory site in the bifunctional transporter DcuB of *Escherichia coli*. *Biochemistry* **50**: 5925-5938.
- Berezin, C., Glaser, F., Rosenberg, J., Paz, I., Pupko, T., Fariselli, P., Casadio, R., and Ben-Tal, N. (2004) ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics* **20**: 1322-1324.
- Cohen, S. N., Chang, A. C., and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A* **69**: 2110-2114.
- Danese, P. N., Pratt, L. A., and Kolter, R. (2000) Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J Bacteriol* **182**: 3593-3596.
- Derman, A. I., and Beckwith, J. (1991) *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. *J Bacteriol* **173**: 7719-7722.
- Derman, A. I., and Beckwith, J. (1995) *Escherichia coli* alkaline phosphatase localized to the cytoplasm slowly acquires enzymatic activity in cells whose growth has been suspended: a caution for gene fusion studies. *J Bacteriol* **177**: 3764-3770.
- Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**: 6127-6145.
- Gottesman, S., and Stout, V. (1991) Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. *Mol Microbiol* **5**: 1599-1606.
- Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Koch, S., Fritsch, M. J., Buchanan, G., and Palmer, T. (2012) *Escherichia coli* TatA and TatB proteins have N-out, C-in topology in intact cells. *J Biol Chem* **287**: 14420-14431.
- Lee, J., and Glover, K. J. (2012) The transmembrane domain of caveolin-1 exhibits a helix-break-helix structure. *Biochim Biophys Acta* **1818**: 1158-1164.
- Manoil, C. (1991) Analysis of membrane protein topology using alkaline phosphatase and beta-galactosidase gene fusions. *Methods Cell Biol* **34**: 61-75.

- Marolda, C. L., Li, B., Lung, M., Yang, M., Hanuszkiewicz, A., Rosales, A. R., and Valvano, M. A. (2010) Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. *J Bacteriol* **192**: 6160-6171.
- Marolda, C. L., Vicarioli, J., and Valvano, M. A. (2004) Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* **150**: 4095-4105.
- Meredith, T. C., Mamat, U., Kaczynski, Z., Lindner, B., Holst, O., and Woodard, R. W. (2007) Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in *Escherichia coli*. *J Biol Chem* **282**: 7790-7798.
- Meruelo, A. D., Samish, I., and Bowie, J. U. (2011) TMKink: a method to predict transmembrane helix kinks. *Protein Sci* **20**: 1256-1264.
- Patel, K. B., Ciepichal, E., Swiezewska, E., and Valvano, M. A. (2012a) The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* **22**: 116-122.
- Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology* **20**: 1389-1401.
- Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012b) Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol* **194**: 2646-2657.
- Saldias, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**: 440-453.
- Sauve, V., Bruno, S., Berks, B. C., and Hemmings, A. M. (2007) The SoxYZ complex carries sulfur cycle intermediates on a peptide swinging arm. *J Biol Chem* **282**: 23194-23204.
- Steiner, K., Novotny, R., Patel, K., Vinogradov, E., Whitfield, C., Valvano, M. A., Messner, P., and Schaffer, C. (2007) Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *J Bacteriol* **189**: 2590-2598.
- Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996) Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**: 4885-4893.

- Stout, V., Torres-Cabassa, A., Maurizi, M. R., Gutnick, D., and Gottesman, S. (1991) RcsA, an unstable positive regulator of capsular polysaccharide synthesis. *J Bacteriol* **173**: 1738-1747.
- Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**: s452-471.
- Valvano, M. A., Furlong, S. E., and Patel, K. B. (2011) Genetics, Biosynthesis and Assembly of O-Antigen. In: *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis, and Interaction with Host Cells*. Valvano, Y. A. K. a. M. A. (ed). New York: SpringerWien, pp. 275-310.
- Videira, P. A., Garcia, A. P., and Sa-Correia, I. (2005) Functional and topological analysis of the *Burkholderia cenocepacia* priming glucosyltransferase BceB, involved in the biosynthesis of the cepacian exopolysaccharide. *J Bacteriol* **187**: 5013-5018.
- Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598-2604.
- Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**: 39-68.

Chapter 4

Investigating the role of the N-terminal domain of the *Escherichia coli* UDP-glucose:undecaprenyl phosphate glucose-1-phosphate transferase WcaJ

Sarah E. Furlong, Mohamad A. Hamad, and Miguel A. Valvano

4.1 Introduction

Two major enzyme families catalyze the initiation of lipid-linked polysaccharide synthesis: the polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPTs) and the polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferases (PNPTs) (Valvano, 2003). Enzymes in both families are membrane proteins that transfer a sugar-1-phosphate donated by a nucleotide diphospho-sugar to a lipid carrier. Unlike PNPTs, PHPTs have no known eukaryotic homologs and therefore only utilize the lipid acceptor, bacterial undecaprenyl phosphate (Und-P) (Valvano, 2003). PHPTs initiate the synthesis of various polysaccharides such as O antigen (O Ag) in *Salmonella enterica* (WbaP) (Wang *et al.*, 1996, Saldias *et al.*, 2008), S-layers in *Geobacillus stearothermophilus* (WsaP) (Steiner *et al.*, 2007), capsules in *Streptococcus pneumoniae* (CpsE) (Cartee *et al.*, 2005), and colanic acid (CA) in *E. coli* and other *Enterobacteriaceae* members (WcaJ) (Stevenson *et al.*, 1996, Patel *et al.*, 2012). The topology of WcaJ has been refined (Chapter 3), and using *in silico* prediction programs, many of these members share similar predicted topology with five transmembrane domains, a large cytoplasmic loop domain (previously known as “extracytoplasmic” loop), and a long C-terminal tail domain that has the glycosyltransferase activity (Wang *et al.*, 1996, Saldias *et al.*, 2008, Patel *et al.*, 2010, Patel *et al.*, 2012).

All PHPTs have a C-terminal domain (CT), which encompasses part of the large loop, the last membrane domain, and the cytoplasmic tail and it has been well established that this domain is sufficient for enzymatic activity (Patel *et al.*, 2010, Wang *et al.*, 1996, Saldias *et al.*, 2008) (Fig. 4.1). However, not all PHPTs contain an N-terminal domain, comprising the first four transmembrane domains and the large loop domain. Examples

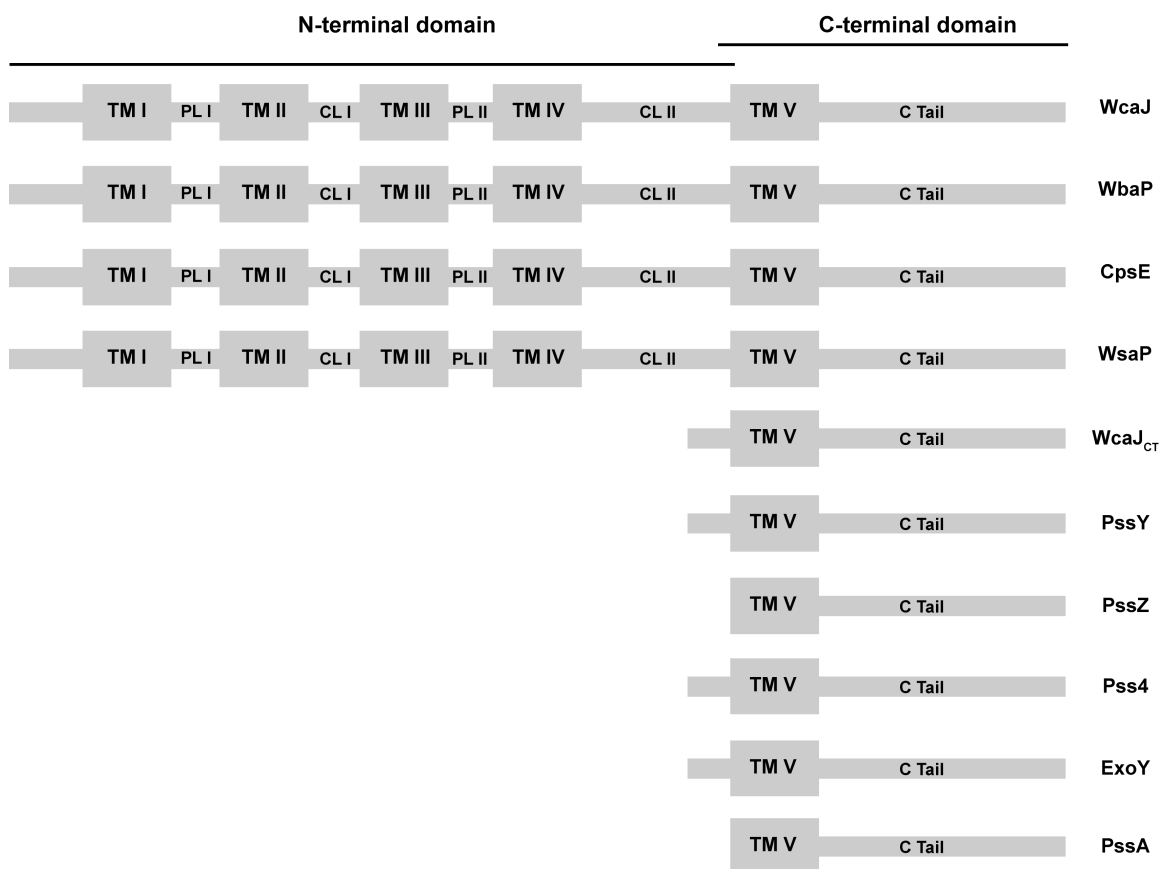


Figure 4.1. Topological alignments of various PHPTs. The N-terminal domain comprises of the first four transmembrane domains and the large cytosolic loop domain. The C-terminal domain comprises of the last membrane domain V and the C-terminal tail (C Tail). Alignments are of the *E. coli* WcaJ (Patel *et al.*, 2012), *S. enterica* WbaP (Wang *et al.*, 1996, Saldias *et al.*, 2008), *S. pneumoniae* CpsE (Xayarath and Yother, 2007), *G. stearothermophilus* WsaP (Steiner *et al.*, 2007), *C. crescentus* PssY and PssZ (Patel *et al.*, 2012), *Rhizobium leguminosarum* PssA and Pss4, and *Rhizobium meliloti* ExoY (Wang *et al.*, 1996). Topology alignments are based off of various prediction programs and refined topological analysis of WcaJ (Chapter 3).

of these enzymes include PssY and PssZ, two enzymes that initiate holdfast synthesis in *Caulobacter crescentus* (Patel *et al.*, 2012), potential sugar transferases PssA and Pss4 of two *Rhizobium leguminosarum* strains and ExoY of *Rhizobium meliloti* (Fig. 4.1). The role of the N-terminal domain has not yet been established experimentally. Wang *et al.* suggested the N-terminal domain contributes to the release of the Und-P linked sugar end product (Wang *et al.*, 1996). Xayarath and Yother proposed that the N-terminal domain might be important for interactions with other assembly proteins, as suppressor mutations accumulated in the CT and “extracytoplasmic” loop domain that relieve lethality of *wzy* and *wzx* gene deletions. These authors suggested that the loop domain must have some important function if suppressor mutations occurred in this region (Xayarath and Yother, 2007). Work by Saldias *et al* suggested that the N-terminal domain of PHPTs contributes to the proper folding and/or stability of WbaP and that the large loop domain may interact with Wzz, a protein that regulates O Ag chain length (Saldias *et al.*, 2008).

In this study, we attempt to elucidate the function of the N-terminal domain, including the large cytoplasmic loop domain, of the *E. coli* UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferase WcaJ.

4.2 Materials and Methods

4.2.1 Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 4.1. Bacteria were grown aerobically at 37°C in Luria-Bertani (LB) medium (Difco Laboratories, Sparks, MD, USA) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). Growth medium were supplemented with 100 µg/ml ampicillin, 20 µg/ml tetracycline, 40 µg/ml

Table 4.1. Characteristics of the bacterial strains and plasmids used in this study.

| Strain or Plasmid | Relevant Properties | Source or Reference |
|--------------------------|--|---------------------------------|
| <i>Strains</i> | | |
| <i>E. coli</i> | | |
| C43 (DE3) | F ⁻ <i>ompT hsdSB</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i> (DE3) | (Arechaga <i>et al.</i> , 2000) |
| DH5 α | F ⁺ ϕ 80lacZ Δ M15 <i>endA recA hsdR</i> (r _K ⁻ m _K ⁻) <i>supE thi gyrA relA</i> | Laboratory stock |
| MH110 | W3110 (chromosomal Flag-tagged <i>wcaJ</i>) | This study |
| MH111 | W3110 (chromosomal Flag-tagged <i>wcaJ</i> _{CT} E252-Y464) | This study |
| W3110 | <i>rph-1</i> IN(<i>rrnD-rrnE</i>)1, <i>wbbL</i> ::IS5 | Laboratory stock |
| XBF1 | W3110, Δ <i>wcaJ</i> :: <i>aph</i> , KnR | (Patel <i>et al.</i> , 2012) |
| <i>Salmonella</i> | | |
| LT2 | <i>Salmonella enterica</i> serovar Typhimurium, wild-type | S. Maloy |
| MSS2 | LT2, Δ <i>wbaP</i> :: <i>cat</i> , CmR | (Saldias <i>et al.</i> , 2008) |
| <i>Plasmids</i> | | |
| pBAD24 | Cloning vector, inducible by arabinose, AmpR | (Guzman <i>et al.</i> , 1995) |
| pBADNTF | pBAD24 vector for N-terminal Flag fusions, AmpR | (Marolda <i>et al.</i> , 2004) |
| pBADFLAG | pBAD24 vector for C-terminal Flag fusions, AmpR | (Saldias <i>et al.</i> , 2008) |
| pDI-SceI-SacB | ori _{pBBR1} , mob ⁺ , P _{dhfr} , <i>sacB</i> , encodes I-SceI endonuclease TcR | (Hamad <i>et al.</i> , 2010) |
| pGP-SceI-2 | ori _{R6K} , mob ⁺ , carries I-SceI cut site, TpR | (Hamad <i>et al.</i> , 2012) |
| pLA3 | pBAD expressing Flag-WcaJ (full length) | Chapter 3 |
| pLA5 | pBAD expressing Flag-WcaJ _{CT} (E252-Y464) | Chapter 3 |
| pSEF86 | pBAD expressing Flag-WcaJ- Δ N143-L272 | This study |
| pSEF94 | pBAD expressing Flag-WcaJ- Δ L152-N164 (Δ α -helix 1) | This study |
| pSEF95 | pBAD expressing Flag-WcaJ- Δ N192-A202 (Δ α -helix 2) | This study |
| pSEF97 | pBAD expressing Flag-WcaJ- Δ D216-D229 (Δ α -helix 3) | This study |
| pSEF99 | pBADFLAG expressing WcaJ-R131-D278-Flag | This study |
| pSEF111 | pBAD expressing Flag-WcaJ-T2-I130 | This study |
| pSEF112 | pBAD expressing Flag-WcaJ-R131-Y464 | This study |
| pSEF113 | pBAD expressing Flag-WcaJ-T2-D278 | This study |
| pWQ499 | pKV102 containing <i>rcsAK30</i> ; TcR | C. Whitfield |

kanamycin or 30 µg/ml chloramphenicol. DH5α cells were transformed with plasmids by calcium chloride method (Cohen *et al.*, 1972). C43 and XBF1 cells were transformed with plasmids by electroporation (Dower *et al.*, 1988).

4.2.2 Construction of WcaJ cytoplasmic loop domain and α-helix deletions in pLA3

Deletion of the large cytoplasmic loop region, and individual α-helices 1 (L152-N164), 2 (N192-A202), and 3 (D216-D229) of the loop region in WcaJ was performed by PCR amplification around each deletion endpoint using template plasmid, pLA3, *Pfu* Turbo AD polymerase (Stratagene, Santa Clara, CA, USA) and primer pairs containing an *Xho*I or *Xba*I restriction enzyme site (Table 4.2). *Dpn*I was added to PCR reactions for overnight digestion of parental plasmid DNA at 37°C. Each *Dpn*I digest was cleaned with a PCR purification kit, and subsequently digested with either *Xho*I or *Xba*I. Each digest was cleaned, and ligated using T4 DNA ligase. The resulting DNA was introduced into *E. coli* DH5α by transformation, and transformants were selected on LB-agar containing 100 µg/ml of ampicillin. Colony PCR was performed using pBAD external primers 252 and 258. DNA sequencing was performed at Eurofins MWG Operon in Huntsville, Alabama, USA.

4.2.3 Cloning various WcaJ constructs with N- or C-terminal Flag fusions

The WcaJ constructs were made by amplifying WcaJ from the template pLA3 using *Pwo* polymerase and the following primer pairs: 6632 and 6633 for WcaJ-R131-D278;

Table 4.2. Primers used in this study.

| Primer Name | Primer Number | Sequence (5'-3') |
|-------------------|---------------|--|
| pBAD Forward | 252 | GATTAGCGGATCCTACCTGA |
| pBAD Reverse | 258 | GACCGCTTCTGCGTTCTGAT |
| pGPI-SceI Forward | 3957 | CAACGAACGATTCATGACCGTGC |
| pGPI-SceI Reverse | 4021 | GCTCAATCAATCACCGGATCCC |
| F-SmaI-WcaI | 4911 | CTAG <u>CCCCGGG</u> ACAAATCTAAAAAAGCGC |
| R-HindIII-WcaI | 4912 | CTAGA <u>AAGCTTT</u> CAATATGCCGCTTTGTTA |
| F-XhoI-ΔN143-R271 | 6306 | ACGACCTCGAGCCCTGCTCAAACGTGCGGA A |
| R-XhoI-ΔN143-R271 | 6308 | ACGACCTCGAGGTTATAGCCATGATTACG |
| F-XbaI-ΔL152-N164 | 6509 | TACATCTAGACAGCCGTGGTTAGGG |
| R-XbaI-ΔL152-N164 | 6510 | TACATCTAGAATCCCCGCCACCGC |
| F-XbaI-ΔD216-D229 | 6511 | TACATCTAGAACACCTGTTCGGTG |
| R-XbaI-ΔD216-D229 | 6512 | TACATCTAGAGCACATTTGCATCGC |
| F-XbaI-ΔN192-A202 | 6513 | TACATCTAGAGGCAAGATTCATAAC |
| R-XbaI-ΔN192-A202 | 6514 | TACATCTAGAACCCGCCAGTCGTT |
| F-SmaI-R131-D278 | 6632 | CT <u>CCCCGGG</u> CGCATTGGGGCGGGC |
| R-SalI-R131-D278 | 6633 | GACTGTCGACGCGTCTTCCGCACGTTT |
| WcaI-FUS-NotI | 6644 | TAGCTAGGCGGCCGCTTTGACCGCTGTTTC CTGTTTGAC |
| WcaI-RUS-NdeI | 6645 | GATTTGTCATATGCGTTGTTCCCTGTTATTAG CCCCTTAC |
| Flag-FDS-NdeI | 6646 | TAGCTGACATATGGATTACAAGGATGACGA TGACAAG |
| pBAD-RDS-XbaI | 6647 | GTCGTGTCTAGAAATCTTCTCATCCGCCA AAACAGCC |
| F-WcaI screen | 6737 | CAGCGGGTCCGGTGGTGGACGCCATTG |

| | | |
|---------------|------|------------------------------|
| R-WcaJ screen | 6738 | CGATAATCACCAGCGACACGGTAAG |
| R-XbaI-D278 | 6798 | CAGCTCTAGATCAGTCTTCCGCACGTTT |
| R-XbaI-I130 | 6799 | CAGCTCTAGATCAAATACACGAACGGCA |
| F-SmaI-R131 | 6800 | CTAGCCCGGGCGCATTGGGGCGGGC |

4911 and 6798 for WcaJ-T2-D278; 4911 and 6799 for WcaJ-T2-I130; 6800 and 4912 for WcaJ-R131-Y464. Each primer pair contains various restriction sites. *DpnI* was added to the PCR product to digest the parental plasmid template. The PCR product was then cleaned using Qiagen PCR purification kit (Toronto, Ontario, Canada). The pBADNTF or pBADFLAG vectors and PCR products were digested with the appropriate restriction enzymes. Digested DNA was cleaned again with Qiagen PCR purification kit. The digested vectors were treated with alkaline phosphatase and then ligated with the *wcaJ* PCR product using T4 DNA ligase. The resulting DNA was introduced into *E. coli* DH5 α by transformation, and transformants were selected on LB-agar containing 100 μ g/ml of ampicillin. Colony PCR was performed using pBAD external primers 252 and 258. DNA sequencing was performed at Eurofins MWG Operon in Huntsville, Alabama, USA.

4.2.4 Cloning plasmids to make chromosomal Flag-tagged *wcaJ* and Flag-tagged *wcaJ*_{CT}

The upstream region of the *wcaJ* gene (~700 base pairs) was PCR-amplified using primers 6644 and 6645 and genomic DNA from *E. coli* W3110 as a template. Genes encoding the full-length Flag-WcaJ and Flag-WcaJ_{CT} (encoding residues E252-Y464) were amplified using primers 6646 and 6647 with pLA3 and pLA5 as templates, respectively. The upstream PCR product was digested with *NotI* and *NdeI* and the Flag-WcaJ PCR product was digested with *NdeI* and *XbaI*. Both fragments were ligated in the linearized pGP-SceI-2, which was digested with *NotI* and *XbaI* to create pGPI-SceI-2-Flag-WcaJ. To create a replacement vector for Flag-WcaJ_{CT}, the upstream fragment was digested with *NotI* and *NdeI* and the Flag-WcaJ_{CT} PCR product was digested with *NdeI* and *XbaI* and was ligated in pGP-SceI-2 linearized with *NotI* and *XbaI* to create pGPI-

SceI-2-Flag-WcaJ_{CT}. Successful cloning was screened by PCR using primers 3957 and 4021. DNA sequencing at Eurofins MWG Operon in Huntsville, Alabama, USA, confirmed positive clones.

4.2.5 Construction of Flag-tagged *wcaJ* and *wcaJ*_{CT} strains

To replace the native *wcaJ* gene with Flag-*wcaJ* and Flag-*wcaJ*_{CT} in *E. coli* W3110, the respective suicide vectors, pGPI-SceI-2-Flag-WcaJ and pGPI-SceI-2-Flag-WcaJ_{CT} were introduced into *E. coli* W3110 by electroporation and selection for vector integration was achieved by plating onto LB-agar containing 50 µg/ml of trimethoprim. pDI-SceI-SacB, a replicative vector expressing I-*SceI*, was then electroporated into the respective cointegrates and selected for using 20 µg/ml of tetracycline. Resulting colonies were first screened for trimethoprim sensitivity, which indicates the loss of the integrated suicide vector and completion of the second crossover event. Trimethoprim sensitive colonies were then screened for replacement of Flag-*wcaJ* or Flag-*wcaJ*_{CT} using primers 6646 and 6738. The strains with successful Flag-*wcaJ* and Flag-*wcaJ*_{CT} replacements were named MH110 and MH111 respectively. The replicative vector pDI-SceI-SacB was cured from MH110 and MH111 by plating cells on LB-agar containing 5% sucrose. The resulting colonies were assayed for vector loss by screening for tetracycline sensitivity.

4.2.6 Growth conditions of cells for protein preparation

Briefly, bacterial cells containing the appropriate arabinose-inducible plasmids were grown overnight in LB containing the appropriate antibiotics (100 µg/ml of ampicillin, 40 µg/ml kanamycin, and 20 µg/ml of tetracycline). From overnight cultures, cells were diluted to an optical density measured at 600 nm (OD₆₀₀) of 0.2 in LB broth containing

the respective antibiotics. Cells were grown at 37°C until an OD₆₀₀ of 0.5-0.7, and cells were induced with 0.2% arabinose for 3 h at 37°C. Cells were harvested by centrifugation 8 000 χ g and cell pellets were frozen at -20°C until needed.

4.2.7 Membrane preparation and immunoblotting

Cells were resuspended in 50 mM Tris-HCl pH 8 with protease inhibitors and lysed at 10,000 PSI by cell disruptor (Constant Systems Ltd, Kennesaw, GA). Cell debris was pelleted at 27,216 χ g. Total membranes were isolated by centrifugation in microfuge tubes at 39,191 χ g and resuspended in 50 mM Tris-HCl pH 8. Protein concentrations were determined by Bradford protein assay (Bio-Rad, Hercules, CA, USA). For immunoblotting, total membrane preparations were separated in 14% or 18% SDS-PAGE and transferred to nitrocellulose membranes. Membranes were blocked overnight in 5% Western blocking reagent (Roche Diagnostics Canada, Laval, QC, Canada) and TBS. The primary antibody, 4.6 mg/ml anti-FLAG M2 monoclonal antibody (Sigma, Saint Louis, MO, USA), was diluted to 1: 10,000 and applied for 1.5 h, and the secondary antibody, 2 mg/ml goat anti-mouse Alexa fluor 680 IgG antibodies (Invitrogen Molecular Probes, Eugene, OR, USA) was diluted to 1: 20,000 and applied for 20 min. Immunoblots were developed using LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE, USA). Bio-Rad Precision Plus Protein Standards were used for all immunoblots.

4.2.8 *In silico* secondary structure prediction tools

Secondary structure predictions of the large cytoplasmic loop domain of the *E. coli* WcaJ (amino acids R131-D278) were predicted using PSIPRED and HHpred programs. PSIPRED analysis provides amino acid information with α -helix and β -sheet predictions with confidence bars. HHpred provides a 3-D structural model of the input sequence based off of known protein structures.

4.2.9 Colanic acid complementation assay

Plasmids expressing various Flag-WcaJ proteins were transformed into XBF1/pWQ499 by electroporation and selected on LB-agar supplemented with 100 $\mu\text{g/ml}$ of ampicillin, 20 $\mu\text{g/ml}$ of tetracycline, and 40 $\mu\text{g/ml}$ of kanamycin. Cells were replated onto LB-agar with antibiotics with and without arabinose. Plates were incubated at 37°C overnight and then incubated at room temperature for an additional 24-48 h to observe mucoidy because CA is optimally produced at 20°C (Gottesman and Stout, 1991).

4.2.10 *In vitro* transferase assay

Total membranes were prepared from C43 cells containing plasmids expressing wild-type and WcaJ deletion proteins. Sixty μg of membranes, containing endogenous Und-P, were incubated for 30 min at 37 °C with 0.025 μCi of ^{14}C -labeled UDP-Glc (specific activity, 300 mCi/mmol) in a 200 μl reaction buffer containing 50 mM Tris-HCl (pH 8.5), 50 mM MgCl_2 , 6.25 mM EDTA, and 5 mM 2-mercaptoethanol. The lipid fraction was extracted twice with butanol, washed with water, and the butanol phase (containing the lipid fraction) was added to scintillation cocktail (Ecolume, MP Biomedical, Solon, OH, USA)

and measured by scintillation counter (Beckman Coulter Canada, Inc., Mississauga, Ontario, Canada) to determine the radioactive counts per minute (CPM).

4.3 Results

4.3.1 *In silico* structural prediction programs suggest that the large cytoplasmic loop domain of WcaJ comprises α -helices and β -sheets reminiscent of a Rossmann fold

The amino acid sequence of the large cytoplasmic loop domain of WcaJ, residues R131-D278, was entered in two *in silico* structural prediction programs: PSIPRED and HHpred. PSIPRED provides a graphical structural prediction indicating α -helices in cylinders and β -sheets in arrows, with confidence bars under each prediction. HHpred provides a 3-D structural prediction that is based on known protein structures. Both prediction programs suggest that the large cytoplasmic loop domain comprises a series of α -helices and β -sheets (Fig. 4.2 and 4.3) in an arrangement that is reminiscent of a Rossmann fold, which is known to participate in nucleotide binding (Rossmann *et al.*, 1975). Interestingly, the HHpred structure of the WcaJ loop resembles a structure observed in UDP-D-quinovosamine 4-dehydrogenase. Because this loop has a common structure with another UDP-sugar binding enzyme, it is tempting to speculate that this region may contribute to binding the nucleotide portion of UDP-Glc, or the release product, UMP, which may contribute to regulation of enzymatic activity.

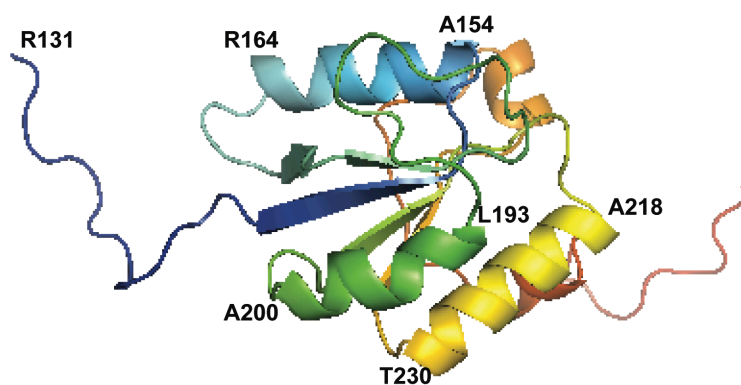


Figure 4.3. Structural prediction of the WcaJ large cytoplasmic loop domain using HHpred. The secondary structure depicts residues of the WcaJ loop domain R131-D278. The α -helices are indicated by residue numbers: α -helix 1 (R164-A154), α -helix 2 (L193-A200), and α -helix 3 (A218-T230).

4.3.2 The N-terminal domain of WcaJ may be important for protein folding and/or stability

The C-terminal domain of WcaJ (E252-Y464) was found to be sufficient for enzymatic activity *in vivo* when the protein was overexpressed in the *wcaJ* deletion strain XBF1 (see Chapter 3, Fig. 3.2) using medium supplemented with 0.2% arabinose to induce overexpression of the construct from the pBADNTF plasmid. However, we decided to investigate if the C-terminal domain could complement CA production under physiological expression conditions by creating chromosomally Flag-tagged *wcaJ* and *wcaJ_{CT}*, strains MH110 and MH111 respectively. Strains were transformed with pWQ499 to express RcsA, a positive transcriptional regulator of CA to observe mucoidy (Gottesman and Stout, 1991). Unfortunately we could not detect the chromosomally expressed Flag-tagged constructs by immunoblotting of total membranes (data not shown), likely due to very low basal levels of expression. However, the chromosomally expressed Flag-WcaJ construct still produced CA but chromosomally expressed Flag-WcaJ_{CT} could not (Fig. 4.4). This suggests that the N-terminal domain may be important for enzymatic function when expressed at wild-type levels.

Because we were unable to detect chromosomally expressed Flag-tagged constructs, we could not confirm if the N-terminal domain was affecting enzymatic function or protein stability. We therefore investigated if the C-terminal domain could complement CA production in lower induction conditions, 0.002% and 0.02% arabinose, which may cause higher expression than the physiological level of WcaJ in the wild-type *E. coli* W3110. XBF1 cells co-expressing RcsA and either Flag-WcaJ or Flag-WcaJ_{CT} were plated on medium containing increasing concentrations of arabinose (none, 0.002%,

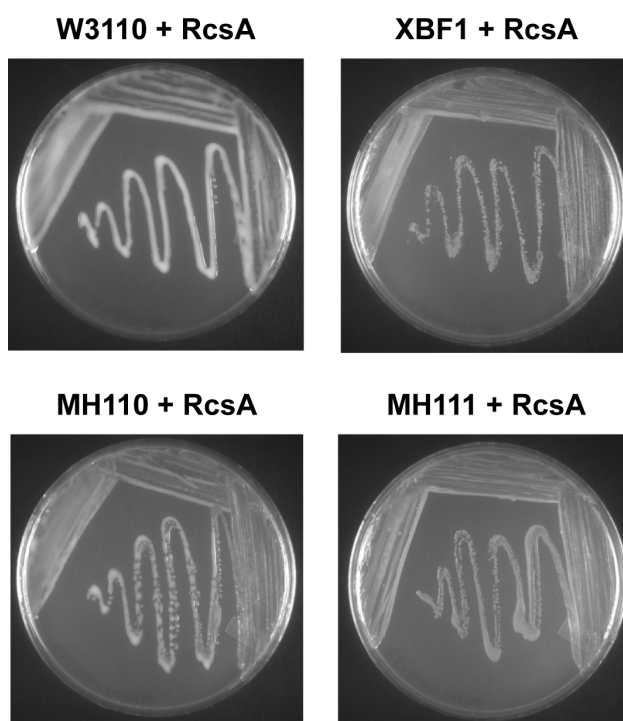


Figure 4.4. Colanic acid production in *E. coli* strains W3110, XBF1 (*wcaJ* deletion strain), MH110 and MH111 (chromosomally tagged Flag-WcaJ and Flag-WcaJ_{CT}) co-expressing RcsA.

0.02% and 0.2%). Although the full-length WcaJ could complement CA synthesis under all expression conditions, the C-terminal domain alone could not (Fig. 4.5); however, the expression of these constructs was affected under the lower induction conditions, 0.002% and 0.02% arabinose (Fig. 4.6). The Flag-WcaJ construct had a dose-dependent increase of protein expression with increasing arabinose concentrations. However, the Flag-WcaJ_{CT} construct only had detectable protein after 0.2% arabinose induction, which may explain why Flag-WcaJ_{CT} could only complement CA at the highest arabinose concentration. These experiments suggest that the N-terminal domain (M1-L251) may be important for enzymatic activity *in vivo* under physiological expression conditions by contributing to protein folding and/or protein stability.

4.3.3 Deletion of the large cytoplasmic loop domain affects protein expression and/or stability of WcaJ while deletion of individual α -helices in the loop domain negatively affect enzymatic activity *in vivo* and *in vitro* without compromising protein expression

We have hypothesized that if large cytoplasmic loop domain resembles a Rossmann fold and contributes to binding the nucleotide portion of UDP-Glc, or the release of the product, UMP, deletion of the loop domain entirely or portions of the loop domain may affect the enzymatic activity. Using a PCR deletion strategy, the large cytoplasmic loop domain was deleted, leaving the first four transmembrane domains, the fifth membrane domain and the C-terminal tail intact, resulting in Flag-WcaJ- Δ N143-L272, expressed from pSEF86. When this construct was co-expressed with RcsA in XBF1, these cells

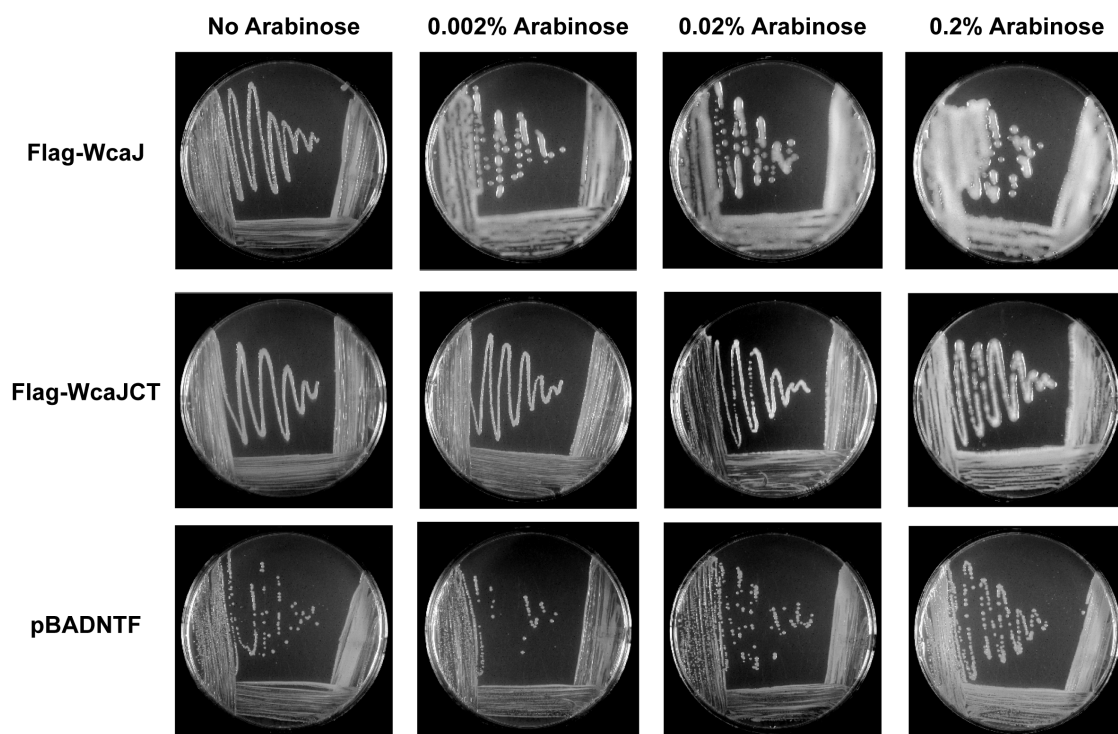


Figure 4.5. Complementation of colanic acid production in XBF1/pWQ499 by Flag-WcaJ and Flag-WcaJ_{CT} with increasing concentrations of arabinose. Medium was supplemented with 0.002%, 0.02%, and 0.2% arabinose to induce expression of Flag-WcaJ and Flag-WcaJ_{CT} proteins from pBADNTF plasmids.

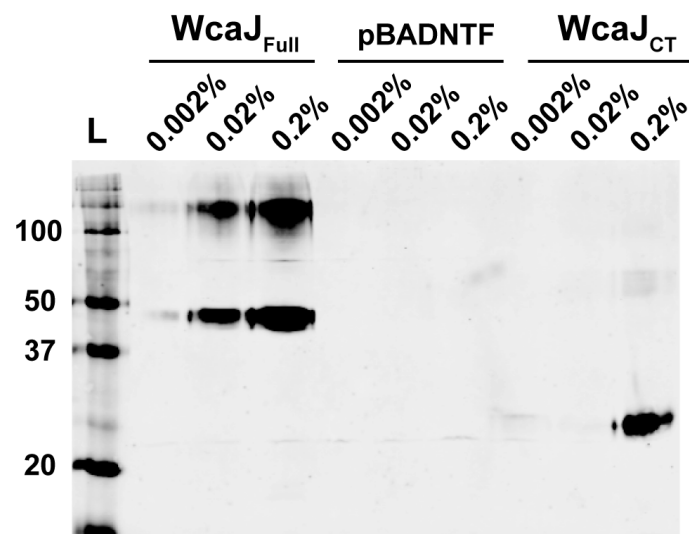


Figure 4.6. Immunoblotting of total membranes expressing Flag-WcaJ and Flag-WcaJ_{CT} with increasing concentrations of arabinose. 10 μ g of total membranes expressing each construct were separated by 14% SDS-PAGE, transferred to nitrocellulose membrane and probed with anti-Flag antibodies.

produce less CA when compared to the Flag-WcaJ (Fig. 4.7). The ability of Flag-WcaJ- Δ N143-L272 to transfer Glc-1-P to Und-P was assessed using the *in vitro* transferase assay. Total membranes of Flag-WcaJ proteins were incubated with ^{14}C -labeled UDP-Glc and the lipid fraction was extracted twice with butanol and measured by scintillation counting. The *in vitro* transferase assay revealed that the deletion of the WcaJ loop domain abolished the enzymatic activity (Fig. 4.8). However, the amount of protein that was detected in the membrane fraction was markedly reduced compared to the wild-type Flag-WcaJ (Fig. 4.8). This may suggest that removal of the large cytoplasmic loop region in the full-length protein may alter the expression, folding and/or stability of the protein.

The removal of the large cytoplasmic loop domain of WcaJ abolished activity completely, but it also affected membrane expression. Therefore, we opted to remove smaller regions of the loop domain. Using the predicted secondary structures of the loop domain (Fig. 4.2 and 4.3), three α -helices were deleted from Flag-WcaJ (Δ L152-N164, Δ N192-A202, Δ D216-D229) using a PCR deletion strategy. These constructs were co-expressed with RcsA in XBF1 and each deletion construct produced less CA compared to Flag-WcaJ (Fig. 4.9) and abolished the enzymatic activity *in vitro* despite that there was only a slight decrease in membrane protein expression (Fig. 4.10). These results suggest that α -helices in the large cytoplasmic loop domain may influence glycosyltransferase activity by maintaining the overall protein structure, nucleotide sugar binding via the Rossmann fold and/or interacting with other enzymes involved in CA synthesis. However, further experiments will need to be performed to elucidate if the loop domain contributes to nucleotide binding.

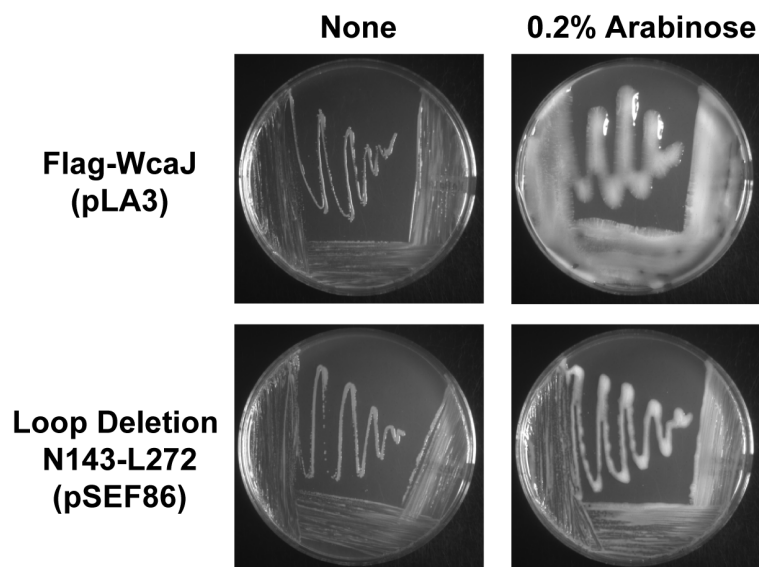


Figure 4.7. Complementation of colanic acid synthesis in XBF1/pWQ499 expressing the wild-type Flag-WcaJ and the loop deletion construct Flag-WcaJ Δ N143-L272. Medium was supplemented with 0.2% arabinose to induce expression of Flag-WcaJ proteins from pBADNTF plasmids.

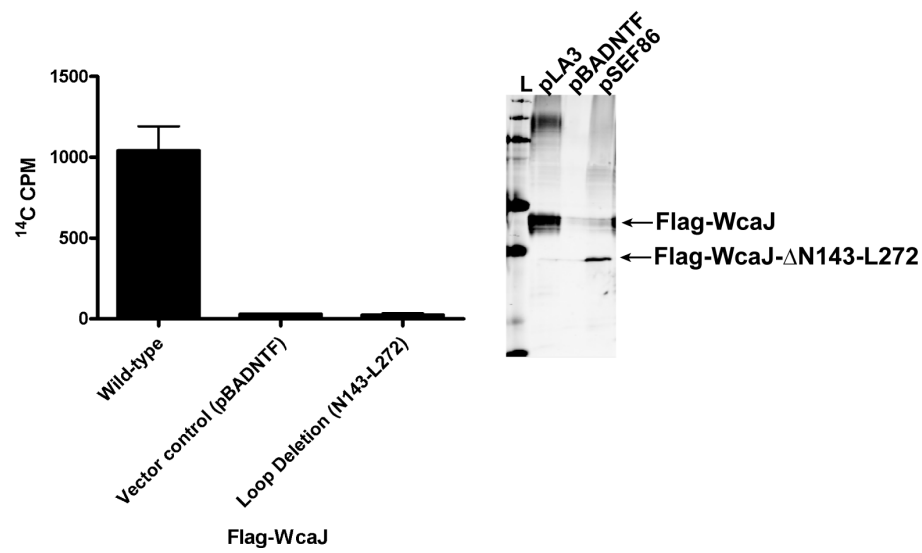


Figure 4.8. *In vitro* transferase assay of the wild-type Flag-WcaJ and the loop deletion construct Flag-WcaJ Δ N143-L272. 60 μg of total membranes prepared from C43 cells expressing the constructs were incubated with UDP- ^{14}C -Glc. The lipid fraction was extracted with butanol and the ^{14}C counts per minute were measured using scintillation counting. Immunoblotting was performed on 10 μg of total membranes expressing the constructs. Proteins were detected using anti-Flag antibodies.

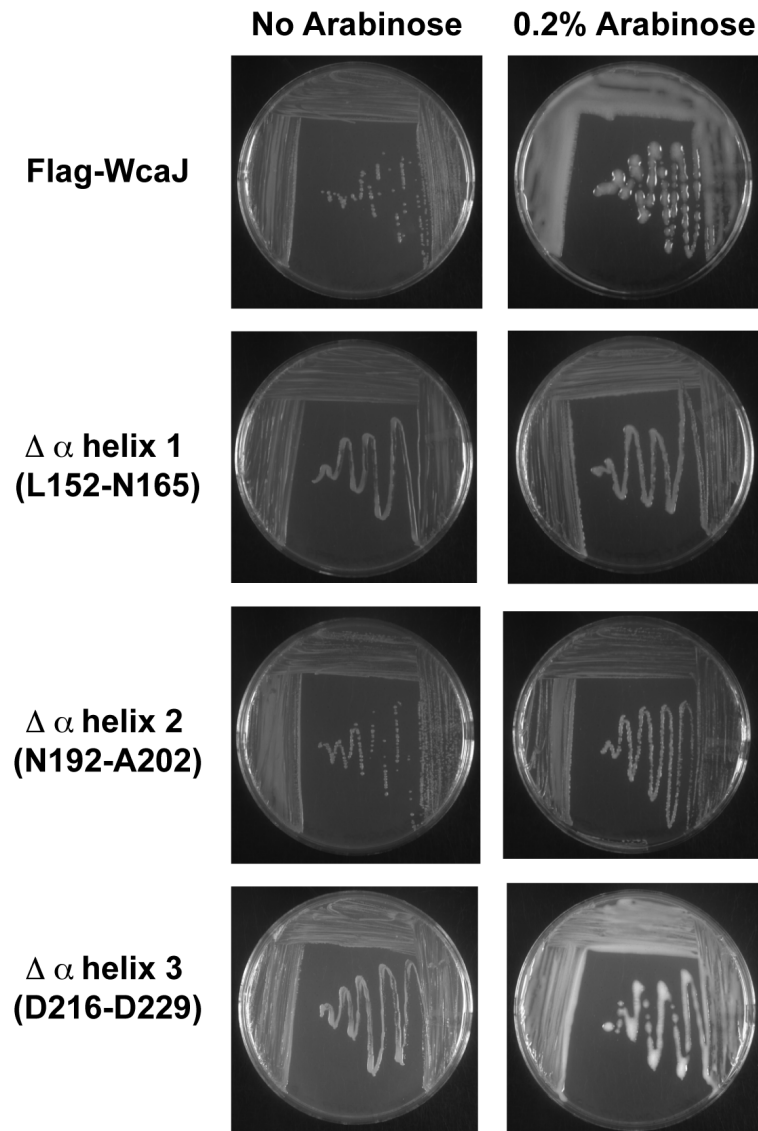


Figure 4.9. Complementation of colanic acid synthesis in XBF1/pWQ499 expressing the wild-type Flag-WcaJ and the α -helix deletion constructs. Medium was supplemented with 0.2% arabinose to induce expression of Flag-WcaJ proteins from pBADNTF plasmids.

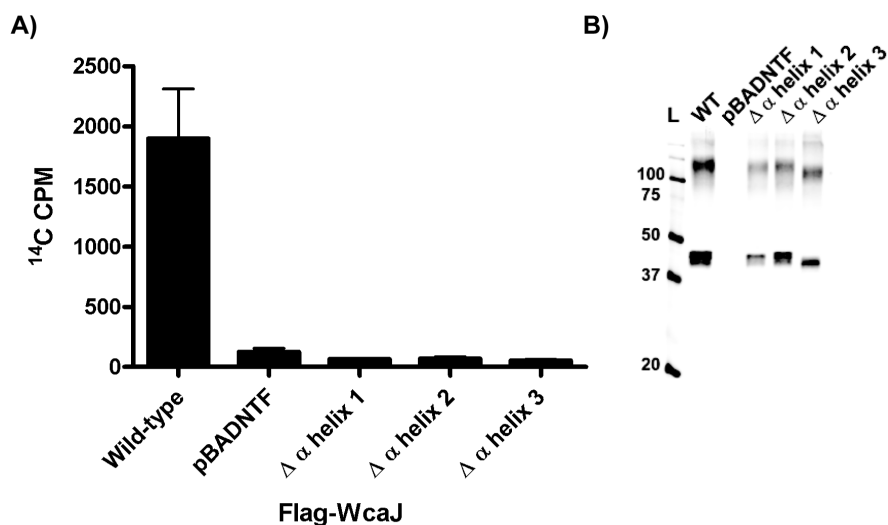


Figure 4.10. *In vitro* transferase assay of the wild-type Flag-WcaJ and the Flag-WcaJ α -helix deletion constructs (Δ L152-N164, Δ N192-A202, Δ D216-D229). 60 μg of total membranes prepared from C43 cells expressing the constructs were incubated with UDP- ^{14}C -Glc. The lipid fraction was extracted with butanol and the ^{14}C counts per minute were measured using scintillation counting. Immunoblotting was performed on 10 μg of total membranes expressing the constructs. Proteins were detected using anti-Flag antibodies.

4.3.4 Expression of various N-terminal domain constructs does not restore enzymatic function in MH111 (chromosomal Flag-WcaJ_{CT})

To assess if expressing the N-terminal domain in MH111, the chromosomally expressed Flag-WcaJ_{CT}, could restore CA synthesis, various WcaJ constructs were made (Fig. 4.11). Constructs were first expressed in XBF1 along with RcsA and we found that only the Flag-WcaJ_{R131-Y464} construct could complement CA production, which is consistent with the notion that the C-terminal domain is required for enzymatic function (Fig. 4.12). To assess if the N-terminal domain could restore CA in MH111, proteins were expressed in this strain. Unfortunately none of the N-terminal or loop domain constructs could restore the CA production in MH111 (Fig. 4.13) despite that all constructs were expressed (Fig. 4.14). These data suggest that the N-terminal domain cannot be expressed separately from the C-terminal domain to complement CA synthesis and that the N-terminal domain in the full-length WcaJ may be important for proper protein folding or stability.

4.4 Discussion

It is well established that the C-terminal domain of PHPTs is sufficient for glycosyltransferase activity (Wang *et al.*, 1996, Saldias *et al.*, 2008, Patel *et al.*, 2010). However, the function of the N-terminus, including the large loop domain in PHPTs has remained elusive. Other groups have hypothesized its function, suggesting that it may contribute to the release of the Und-P linked sugar end product (Wang *et al.*, 1996), protein interactions (Xayarath and Yother, 2007, Saldias *et al.*, 2008), or correct protein

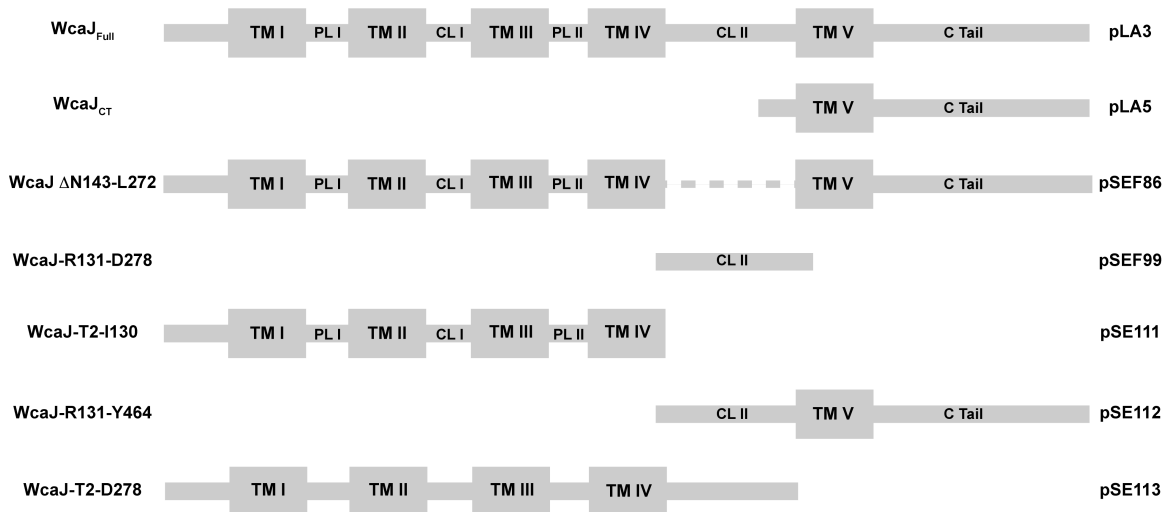


Figure 4.11. Diagram of the various WcaJ constructs cloned into pBAD vectors. All constructs, except for pSEF99, have an N-terminal Flag-tag. The construct expressed from pSEF99 has a C-terminal Flag-tag. The dashed line indicates a deletion in the protein.

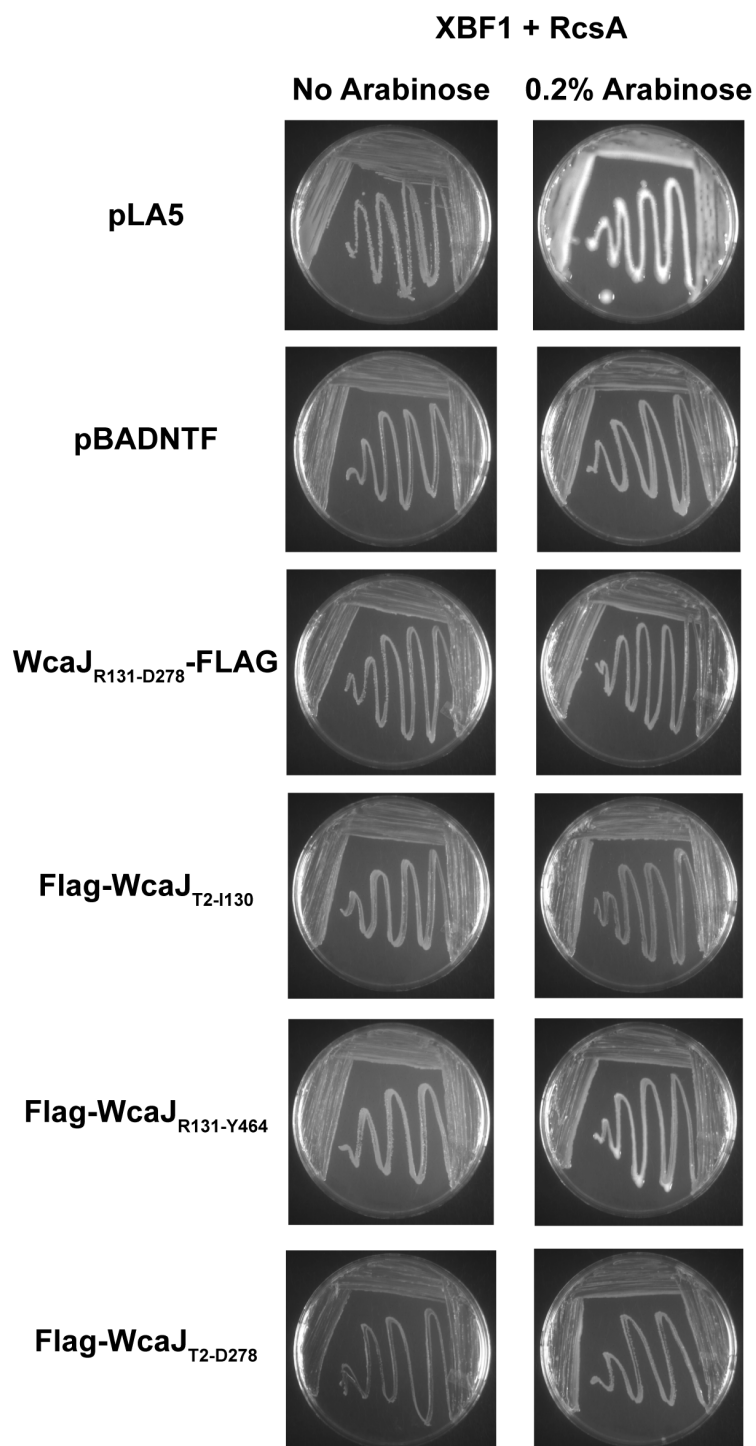


Figure 4.12. Expression of various WcaJ constructs in XBF1/pWQ499. 0.2% arabinose was supplemented in the medium to induce expression of Flag-WcaJ proteins from pBADNTF plasmids.

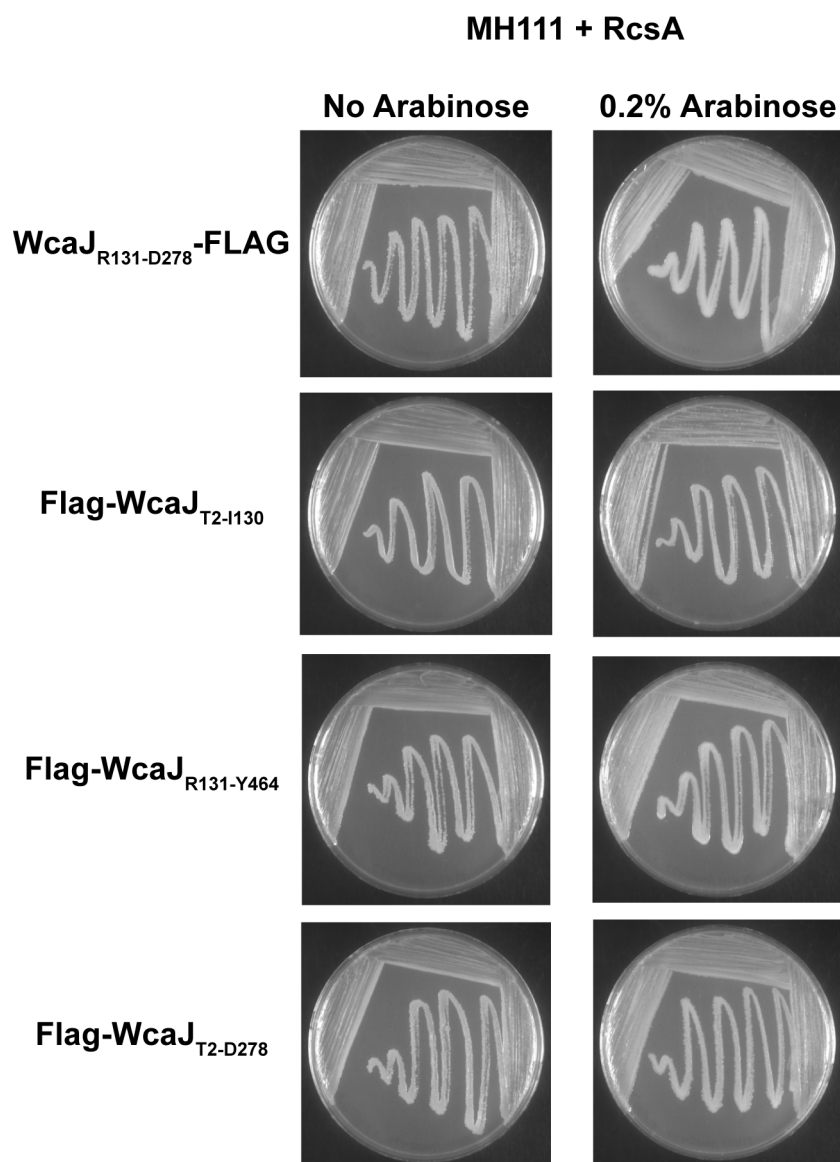


Figure 4.13. Expression of various WcaJ constructs in MH111 (Flag-*wcaJ*_{CT})/pWQ499. Medium was supplemented with 0.2% arabinose to induce expression of Flag-WcaJ proteins from pBADNTF plasmids.

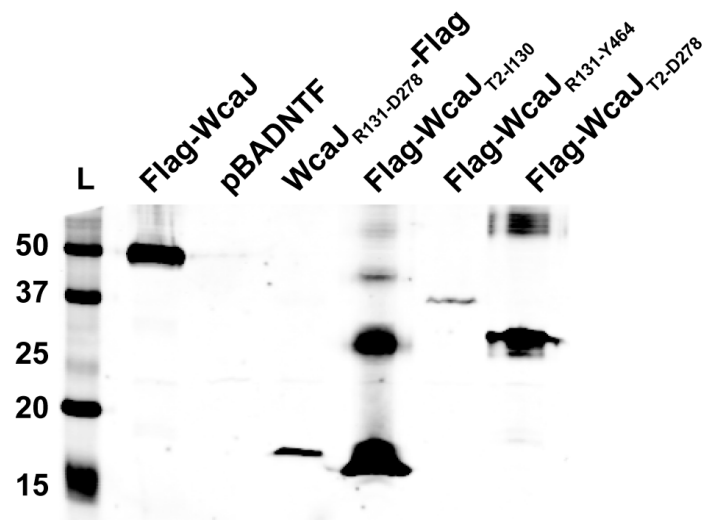


Figure 4.14. Immunoblotting of total membrane expressing various WcaJ constructs. 10 μ g of total membranes were separated by 18% SDS-PAGE. Proteins were detected by immunoblotting using anti-Flag antibodies.

folding and/or protein stability (Saldias *et al.*, 2008). What is perplexing about PHPTs is that some members, such as the *C. crescentus* UDP-Glc:Und-P Glc-1-P transferase PssY and PssZ, innately lack the N-terminal domain and resemble WcaJ_{CT} (Fig. 4.1), which supports the notion that the glycosyltransferase activity resides in the C-terminal domain but does not explain why other PHPTs have this N-terminal domain (Patel *et al.*, 2012). We have hypothesized that the large loop domain in the N-terminal domain may contribute to nucleotide-sugar substrate binding as this domain has a predicted secondary structure reminiscent of a Rossmann fold (Fig. 4.2 and 4.3).

Our experimental data suggest that the N-terminal domain of WcaJ (M1-D278) may play an important role for the stability of these glycosyltransferases, as complementation of CA in strains expressing Flag-WcaJ_{CT} constructs from the chromosome or in low induction conditions was severely affected (Fig. 4.4 and 4.5). Also Flag-WcaJ_{CT} expressed in low induction conditions, 0.002% and 0.02%, was not detectable despite that it was detectable during overexpression at 0.2% (Fig. 4.6). Furthermore, deletion of the loop alone, with the first four and fifth membrane domains intact had reduced protein expression (Fig. 4.8), reduced CA complementation in XBF1 (Fig. 4.7), and no *in vitro* transferase activity (Fig. 4.8). These results suggest that the N-terminus, including the large loop domain, may contribute to enzyme stability and/or protein expression and only under overexpression conditions these deletion constructs have enough membrane protein expression to initiate CA synthesis and be detected by immunoblotting. Interestingly, the deletion of individual α -helices of the large loop domain (Δ L152-N164, Δ N192-A202, Δ D216-D229), which could theoretically make up part of the Rossmann fold, had reduced CA complementation in XBF1 (Fig. 4.9) and no

enzymatic activity *in vitro* despite that these deletions did not dramatically alter the protein expression (Fig. 4.10). These data suggest that the large loop domain may contribute to nucleotide-sugar binding, although more experiments are needed to validate this notion.

The role of the N-terminal domain contributing to protein stability is supported by Saldias *et al* as these authors found that N-terminal domain of WbaP, another PHPT family member that initiates O Ag synthesis in *S. enterica*, was important for *in vitro* enzymatic activity as total membranes expressing this construct could not transfer radioactively labeled UDP-galactose to Und-P (Saldias *et al.*, 2008). This is also further supported by Patel *et al* as these authors found that WbaP_{CT}, analogous to WcaJ_{CT}, could only restore O Ag complementation like the wild-type strain when thioredoxin A (TrxA) was fused to the N-terminus. These authors also demonstrate that the WbaP_{CT} construct had enhanced protein stability after introducing this TrxA fusion partner (Patel *et al.*, 2010). Together this evidence strongly suggests that the role of the N-terminal domain, including the large cytoplasmic loop, may be important for either enzyme folding or stability in the membrane.

4.5 Chapter four references

- Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., de Kruijff, B., Runswick, M. J., and Walker, J. E. (2000) Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. *FEBS Lett* **482**: 215-219.
- Cartee, R. T., Forsee, W. T., Bender, M. H., Ambrose, K. D., and Yother, J. (2005) CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. *J Bacteriol* **187**: 7425-7433.
- Cohen, S. N., Chang, A. C., and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A* **69**: 2110-2114.
- Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**: 6127-6145.
- Gottesman, S., and Stout, V. (1991) Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. *Mol Microbiol* **5**: 1599-1606.
- Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Hamad, M. A., Di Lorenzo, F., Molinaro, A., and Valvano, M. A. (2012) Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. *Mol Microbiol* **85**: 962-974.
- Hamad, M. A., Skeldon, A. M., and Valvano, M. A. (2010) Construction of aminoglycoside-sensitive *Burkholderia cenocepacia* strains for use in studies of intracellular bacteria with the gentamicin protection assay. *Appl Environ Microbiol* **76**: 3170-3176.
- Marolda, C. L., Vicarioli, J., and Valvano, M. A. (2004) Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* **150**: 4095-4105.
- Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology* **20**: 1389-1401.
- Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012) Functional characterization of UDP-

- glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol* **194**: 2646-2657.
- Rossmann, M. G., Liljas, A., Branden, C. I., and Banaszak, L. J. (1975) Evolutionary and structural relationships among dehydrogenases. In: The enzymes. Boyer, P. D. (ed). Academic Press, New York, N. Y., pp. 61-102.
- Saldias, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**: 440-453.
- Steiner, K., Novotny, R., Patel, K., Vinogradov, E., Whitfield, C., Valvano, M. A., Messner, P., and Schaffer, C. (2007) Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *J Bacteriol* **189**: 2590-2598.
- Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996) Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**: 4885-4893.
- Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**: s452-471.
- Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598-2604.
- Xayarath, B., and Yother, J. (2007) Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J Bacteriol* **189**: 3369-3381.

Chapter 5

Discussion

5.1 PNPTs and PHPTs are two distinct families of enzymes that initiate lipid-linked glycan synthesis

This work contributes to our mechanistic and topological understanding of PNPTs and PHPTs. Although members of these two major enzyme families catalyze the transfer of a sugar-1-phosphate to an isoprenoid lipid carrier to initiate the synthesis of lipid-linked glycans, they have no amino acid sequence similarities and have distinct membrane topologies and catalytic domains. Here we discuss the refined topological model of the PNPT member WecA and the unique topology of the PHPT enzyme WcaJ as well as the differences in their catalytic domains.

5.1.1 PNPT and PHPT enzymes have distinct membrane protein topologies and catalytic domains

Using reporter fusions and sulfhydryl labeling experiments, the topology of various PNPT members has been analyzed and refined, including that of *E. coli* WecA and MrayY and the *S. aureus* MrayY. All these studies conclude that PNPTs have ten to eleven transmembrane (TM) domains with several cytosolic loop regions (Lehrer *et al.*, 2007, Bouhss *et al.*, 1999). Several studies suggest that most if not all the cytosolic loops contribute residues to form a putative catalytic site (Al-Dabbagh *et al.*, 2008, Lehrer *et al.*, 2007, Amer and Valvano, 2002, Amer and Valvano, 2001, Amer and Valvano, 2000). However, the topology prediction programs of PHPT members suggest a completely different topology compared to PNPTs, with only five TMs, a large periplasmic loop domain, and a long C-terminal tail (Saldias *et al.*, 2008, Wang *et al.*, 1996, Patel *et al.*, 2010). Interestingly, unlike PNPTs, PHPTs can be truncated and retain function, as the C-

terminal domain (including TM V) is sufficient for enzymatic activity (chapter 3) (Patel *et al.*, 2010, Patel *et al.*, 2012, Saldias *et al.*, 2008).

In this work (chapter 2) we were interested in refining the topology of the highly conserved VFMGD region of the *E. coli* WecA because of conflicting experimental evidence that the aspartic acid residue in this motif acts as a catalytic nucleophile in PNPTs. *In silico* prediction programs placed this region inside TM VIII, but using substituted cysteine accessibility labeling experiments, we found that this region faces the cytosol. However, amino acid replacements at D217 reveal that this conserved residue does not serve as a catalytic nucleophile, but the polarity and size of amino acid side chains at the D217 position are important for enzymatic activity, as were the replacements of the V213, F214, G216, and D217 residues with alanine. This work supports the notion that all cytosolic loop regions in PNPTs contribute to the transfer reaction and that the highly conserved VFMGD motif defines a region in PNPT that may be involved in the binding and/or recognition of the nucleotide moiety of the nucleoside phosphate precursor.

Unlike PNPTs where several cytosolic loops likely contribute to the enzyme active site, the glycosyltransferase activity of PHPTs resides in the C-terminal domain (chapter 3). PHPTs also have very different predicted topology compared to PNPT members; however, there was very little experimental validation of these predictions. Using a truncated version of a PHPT member with an N-terminal fusion partner, His₆-TrxA-WbaP_{CT}, and protease accessibility experiments in spheroplasts, Patel *et al.* first suggested that the periplasmic loop domain might face the cytosol. However, it was unclear if the resulting topology was an artifact attributable to the naturally cytoplasmic

TrxA protein partner or due to the native structure of TMV not fully spanning the membrane (Patel *et al.*, 2010). Therefore, it was our goal to elucidate the topology. Using both reporter fusion assays and substituted cysteine labeling using PEG-Mal, we provide the first detailed topological analysis of a PHPT family member, WcaJ, which has an unexpected membrane topology (chapter 3). We found that the topology was not only inverted compared to the *in silico* prediction tools, but the large “periplasmic” loop domain was found to face the cytosol. To have two cytosolic regions surrounding the last membrane domain V, we propose that it does not fully span the membrane. Analysis of TM V sequences with various prediction programs suggest that a conserved proline contributes to a helix kink. Interestingly, another membrane protein, caveolin-1 also has a helix-break-helix structure of similar size to the membrane domain of WcaJ with a highly conserved proline and isoleucine. We propose that this helix-break-helix structure in TM helix V is a signature for all members of the PHPT family.

5.1.2 The N-terminal domain of PHPTs may contribute to protein stability and binding of UDP

After revealing that the large loop domain in the PHPT family member, *E. coli* WcaJ, faces the cytosol (chapter 3), we investigated the role of the N-terminal domain, including the large cytoplasmic loop domain. This region was of particular interest to us because some members, such as the *C. crescentus* UDP-Glc:Und-P Glc-1-P transferase PssY and PssZ, innately lack the N-terminal domain. Our experimental data in chapter 4, suggest that the N-terminal domain of WcaJ (M1-D278) may play an important role for

the stability of these glycosyltransferases as removal of the N-terminus affects protein expression and *in vivo* complementation, which was also observed by Saldias *et al.* (Saldias *et al.*, 2008).

We have also hypothesized that the large cytoplasmic loop domain may contribute to nucleotide-sugar substrate binding as secondary structure prediction programs suggest that the loop domain resembles a Rossmann fold and deletions of individual α helices of the loop domain negatively affected the enzymatic activity without considerable defects in protein expression. However, further experiments will be required to determine the function of the N-terminal domain and the large cytoplasmic loop domain.

5.2 Future studies

5.2.1 Identification of the catalytic nucleophile of PNPT members

The catalytic mechanism for the transfer reaction of PNPTs has remained elusive. Two theories have been proposed: a two-step (or double-displacement) reaction mechanism and a one-step reaction mechanism. The two-step reaction mechanism proposes that there is a conserved aspartic acid residue that serves as a catalytic nucleophile to attack the phosphate of UDP-sugar residue to form a covalent intermediate with the enzyme. After that, the Und-P will complete the transfer reaction by attacking this enzyme intermediate forming the reaction product Und-P-P-sugar (Heydanek *et al.*, 1969, Lloyd *et al.*, 2004). The one-step reaction mechanism suggests that there is a single conserved aspartic acid residue in cytosolic loop II that acts as a base to deprotonate the hydroxyl of the terminal phosphate of Und-P, which would subsequently attack the UDP-GlcNAc substrate (Al-

Dabbagh *et al.*, 2008). However, mutation of this aspartic acid residue does not abolish enzymatic activity (Al-Dabbagh *et al.*, 2008, Lehrer *et al.*, 2007). If this residue is essential for enzymatic activity, one would expect that replacement of this amino acid should abolish activity.

Evidence supports the two-step reaction mechanism as the acyl-enzyme intermediate of the solubilized PNPT enzyme MraY and the radiolabeled UDP-MurNAc-pentapeptide substrate have been observed following gel filtration (Lloyd *et al.*, 2004). This supports the notion that there is catalytic nucleophile. Unfortunately, the authors did not demonstrate a loss of the enzyme-substrate intermediate with an aspartic acid mutant or specific amino acid linked to the sugar using mass spectrometry. To date there is only one residue in PNPTs that is completely conserved and replacement of this residue in WecA abolishes activity *in vivo* and *in vitro*. Therefore we propose that this residue, D156 in WecA, is the catalytic nucleophile in the two-step reaction mechanism.

It will be important to identify a catalytic nucleophile, however this has been hampered by our inability to overexpress, solubilize and purify WecA. Future studies in our laboratory would involve the trapping of a reaction intermediate using purified WecA and UDP-GlcNAc. After incubation of WecA with its sugar substrate, the enzyme can be digested with trypsin and fragments can be analyzed using reversed phase high performance liquid chromatography/electrospray ionization mass spectrometry to identify the specific residue that the covalent intermediate is bound to, a method that has been employed by Lairson *et al.*, to identify the catalytic nucleophile in *Neisseria meningitidis* protein LgtC, an α -1,4-galactosyltransferase (Lairson *et al.*, 2004). This should reveal which residue acts as catalytic nucleophile and reveal the reaction mechanism of PNPTs.

5.2.2 Structural analysis of the PHPT members

It will be critical to characterize structure of the TMH-V using other analysis, such as circular dichroism, NMR spectroscopy, and chemical shift indexing, which has been employed by Lee and Glover (Lee and Glover, 2012). Furthermore, it would be important to assess the role of the conserved proline and its contribution to the helix-break-helix structure. We have found that substitution of this proline with alanine is not well tolerated as indicated by reduced protein expression. Currently, work is being done in our laboratory to obtain the crystal structure of the C-terminal domain of WbaP. Future studies should also include the full-length and C-terminal domain of WcaJ as well as other PHPT family members. This should elucidate if there is a helix-break-helix protein structure in TMH V and if it is a conserved structure in all PHPT family members.

Structural analysis of the large cytoplasmic loop domain to date has been performed by *in silico* prediction tools; therefore it will be essential to address the protein structure using experimental methods, such as X-ray crystallography. Ideally, we should focus on purifying the full length WcaJ for structural studies and WcaJ should be crystallized alone and with UDP to assess if the large loop domain has Rossmann folds and interacts with UDP. This strategy has been used by Barreras *et al.* to identify the structure and nucleotide-binding domain in GumK, the membrane associated glucuronosyltransferase (Barreras *et al.*, 2008).

5.3 Concluding remarks and significance

This work contributes to cement the notion that the two major lipid-linked glycan initiating enzyme families, the PNPTs and the PHPTs, have distinct membrane topologies and catalytic domains, despite that they catalyze similar transferase reactions. These studies address major questions about these enzyme families, such as identifying a catalytic nucleophile in PNPTs and the topological location of the large loop domain of PHPTs. This work will guide future investigations taking advantage of new topological and mechanistic information that will hopefully pave the way for further detailed structural and mechanistic studies.

5.4 Chapter five references

- Al-Dabbagh, B., Henry, X., El Ghachi, M., Auger, G., Blanot, D., Parquet, C., Mengin-Lecreulx, D., and Bouhss, A. (2008) Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. *Biochemistry* **47**: 8919-8928.
- Amer, A. O., and Valvano, M. A. (2000) The N-terminal region of the *Escherichia coli* WecA (Rfe) protein, containing three predicted transmembrane helices, is required for function but not for membrane insertion. *J Bacteriol* **182**: 498-503.
- Amer, A. O., and Valvano, M. A. (2001) Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine. *Microbiology* **147**: 3015-3025.
- Amer, A. O., and Valvano, M. A. (2002) Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in *Escherichia coli*. *Microbiology* **148**: 571-582.
- Bouhss, A., Mengin-Lecreulx, D., Le Beller, D., and Van Heijenoort, J. (1999) Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol* **34**: 576-585.
- Heydanek, M. G., Jr., Struve, W. G., and Neuhaus, F. C. (1969) On the initial stage in peptidoglycan synthesis. 3. Kinetics and uncoupling of phospho-N-acetylmuramyl-pentapeptide translocase (uridine 5'-phosphate). *Biochemistry* **8**: 1214-1221.
- Lairson, L. L., Chiu, C. P., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C., and Withers, S. G. (2004) Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. *J Biol Chem* **279**: 28339-28344.
- Lee, J., and Glover, K. J. (2012) The transmembrane domain of caveolin-1 exhibits a helix-break-helix structure. *Biochim Biophys Acta* **1818**: 1158-1164.
- Lehrer, J., Vigeant, K. A., Tatar, L. D., and Valvano, M. A. (2007) Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189**: 2618-2628.
- Lloyd, A. J., Brandish, P. E., Gilbey, A. M., and Bugg, T. D. (2004) Phospho-N-acetylmuramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues. *J Bacteriol* **186**: 1747-1757.

- Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology* **20**: 1389-1401.
- Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012) Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol* **194**: 2646-2657.
- Saldias, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**: 440-453.
- Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598-2604.

Appendix- Copyright Release

Chapter 2: Sarah E. Furlong and Miguel A. Valvano (2012). Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase. *Protein Science*, 21, 1366-1375.

A License Agreement has been obtained between Sarah E Furlong and John Wiley and Sons provided by Copyright Clearance Center on June 12, 2013. As an author of this Wiley article, Sarah E Furlong has obtained a “limited license to reproduce the Materials for the purpose specified [Dissertaton/Thesis] in the licensing process”. License number: 3166680593301.

Curriculum Vitae- Sarah E Furlong

| | |
|---|---|
| Post Secondary Education and Degrees | <p>University of Western Ontario London, Ontario, Canada Honors Specialization in Microbiology and Immunology, Bachelor of Medical Science (BMSc) with distinction</p> <p>University of Western Ontario London, Ontario, Canada Microbiology and Immunology, PhD candidate 2008-present</p> |
| Honors and Awards | <p>Western Graduate Research Scholarship, UWO, 2008-2013</p> <p>Ontario Graduate Scholarship in Science and Technology (OGSST) Ontario Ministry of Training, Colleges and Universities, 2010-2011</p> <p>Cedarlane and ATCC Award for Best Graduate Poster Presentation Infection and Immunity Research Forum, London, Ontario, 2011</p> <p>Dean's Honor List, UWO, 2006, 2008</p> <p>Western Scholarship of Distinction, UWO, 2004</p> |
| Relevant Work Experience | <p>Teaching Assistant Microbiology and Immunology, UWO, 2009-2012</p> |
| Related Professional Experience | <p>Graduate Student Representative Undergraduate Education Committee for the Department of Microbiology and Immunology, UWO, 2008-2013</p> <p>Graduate Student Representative Outreach Committee for the Department of Microbiology and Immunology, UWO, 2011-2012</p> <p>Sponsor Liaison Infection and Immunity Research Forum Organizing Committee, UWO, 2011</p> |

Publications

Furlong SE, and Valvano MA. (2012) Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl- phosphate N-acetylaminosugar-1-phosphate transferase. *Protein Science*, 21(9): 1366-1375.

Valvano, MA, **Furlong, SE**, and Patel, KB. (2011) Genetics, Biosynthesis and Assembly of O-Antigen. In: *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis, and Interaction with Host Cells*. Valvano, Y. A. K. a. M. A. (ed). New York: SpringerWien, pp. 275-310.

Patel KB, **Furlong SE**, and Valvano MA. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology*. 20(11):1389-1401.

Poster Presentations

Sarah E. Furlong and Miguel A. Valvano. The highly conserved VFMGD motif of the sugar-phosphate transferase WecA may contribute to the enzyme's active site. FASEB Summer Research Conference: Microbial Polysaccharides of Medical, Agricultural, & Industrial Importance, Carefree, Arizona, USA. (June 2011)

Sarah E. Furlong and Miguel A. Valvano. Investigation of the highly conserved VFMGD motif in the *Escherichia coli* sugar-phosphate transferase WecA. Infection and Immunity Research Forum, London, Ontario. (November 2010)

Sarah E. Furlong and Miguel A. Valvano. Investigation of the catalytic mechanism of the *Escherichia coli* sugar-phosphate transferase WecA. 110th General Meeting American Society for Microbiology, San Diego, California, USA. (May 2010)

Kinnari B. Patel, **Sarah E. Furlong** and Miguel A. Valvano. Structure-function of the C-terminal domain of the *Salmonella enterica* undecaprenylphosphate: galactose-1-phosphate transferase WbaP. 110th General Meeting American Society for Microbiology, San Diego, California, USA. (May 2010)

Sarah E. Furlong and Miguel A. Valvano. Investigation of potentially catalytic aspartic acid residues in the *Escherichia coli* sugar-phosphate transferase WecA. Margaret Moffat Research Day, London, Ontario. (March 2009)

Sarah E. Furlong and Miguel A. Valvano. Investigation of potentially catalytic aspartic acid residues in the *Escherichia coli* sugar-phosphate transferase WecA. Infection and Immunity Research Forum, London, Ontario. (November 2008)

Kinnari B. Patel, **Sarah E. Furlong**, and Miguel A. Valvano. Identifying the regions of WbaP critical for function and O-antigen chain length distribution in *Salmonella enterica*. Microbial Polysaccharides of Medical, Agricultural, and Industrial Importance. Phoenix, Arizona. (2008)