Enterobacterial Common Antigen Biosynthesis in *Shigella flexneri*

Nicholas Tadeusz Maczuga

Bachelor of Science (Biomedical Science)

Bachelor of Science (Hons)

THE UNIVERSITY of ADELAIDE

Submitted for the Degree of Doctor of Philosophy

Research Centre for Infectious Diseases

School of Biological Sciences

The University of Adelaide

Adelaide, South Australia, Australia

January 2022

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Nicholas Tadeusz Maczuga

January 2022

Abstract

The complex outer membrane (OM) polysaccharides of bacteria play crucial roles in cellular homeostasis, fitness and provide resistances to extracellular pressures. *Enterobacteriaceae* possess two major OM polysaccharides; lipopolysaccharide (LPS), which is characteristic of Gram negative bacteria and Enterobacterial Common Antigen (ECA), which is ubiquitously expressed on the OM of all *Enterobacteriaceae*. Both of these polysaccharides are in-part biosynthesised by separate homologs of the Wzy-dependent pathway, the most common bacterial polysaccharide biosynthetic pathway. Due to the importance of LPS in virulence, the majority of the studies directed towards the key proteins of the Wzy-dependent pathway have been directed towards the LPS homologs, with little research being directed towards the ECA specific homologs. The objective of this thesis was to investigate ECA biosynthesis in *Shigella flexneri*, with particular focus on WzyE, the Wzy protein from ECA biosynthesis. For the first time, a WzyE protein was directly investigated and the data showed that WzyE is uncharacteristic of other Wzy proteins; showing a high amount of sequence conservation amongst Enterobacteriales. Furthermore, through experimental topology mapping and site-directed mutagenesis, the data showed a plausible central cavity which may be involved in the polymerisation mechanism of WzyE. Through the necessity of the project, for the first time, two different *wzyE* mutants were generated through lambda Red mutagenesis which showed alternative sensitivities to Colicin E2 and deoxycholate. I ultimately determined that this was due to the regulatory disruption of the adjacent gene, *wecG*, in one of the mutants and highlighted WecG's importance in the biosynthetic pathway. Subsequently I investigated WecG which, like WzyE, had not been directly investigated. The data revolutionised the understanding of WecG revealing that WecG is a protein which is peripherally associated with the inner membrane (IM) via its three, C-terminal helices. Further, critical residues along the second helix were shown to be important for both WecG's membrane association as well as its function and, demonstrated that WecG is likely maintained to the IM via interactions with ECA lipid-I. Ultimately this allowed me to place WecG as the second protein in the novel glycosyltransferase fold family, GT-E. Through investigating WzyE, I noticed an interdependence between ECA and LPS O antigen (Oag) biosynthesis. I subsequently demonstrated that the two OM polysaccharide pathways are fundamentally linked due to their reliance on undecaprenyl phosphate (Und-P) where, *wzy* mutations in one of the pathways caused a reduction in the OM polysaccharide of the un-mutated pathway. Overall, the work presented here provides new insights and revolutionises our understanding of the ECA biosynthetic pathway. I demonstrated indirect cross-talk between the two OM polysaccharide pathways of *S. flexneri* for the first time and provide a platform for future studies to investigate Wzy proteins and other key proteins from the ECA biosynthetic pathway.

Acknowledgements

First and foremost, I would like to thank my primary supervisor and mentor Assoc Prof Renato Morona. Thank you for giving me the opportunity to commence a PhD in your lab and for all the supervision, ideas and enthusiasm you've given me throughout my time in your lab both in honours and PhD. Countless times things didn't go to plan, results didn't make sense and we kept on finding novel, exciting discoveries until the very end (including when we really didn't want to) and in each of those time, your guidance and enthusiasm always propelled me forward making me want to chase every discovery until the end! Thank you for enabling me to be enthusiastic about the science I did and for allowing me, sometimes unwillingly, to chase problems and ideas! I hope you thoroughly enjoy your well-deserved retirement but that we stay in touch.

Next, I would like to thank Dr. Elizabeth Tran (Liz) my teacher and friend. Thank you Liz for always being there for us until the end. You were always there to look through data, read drafts and still somehow keep the lab afloat and under control. You really helped me focus on what needed to be done and empowered me through chocolate and encouragement to get things done. I will miss our Yum Cha lunches, the impromptu bubble tea runs and our meetings in your office where I always left with lollies and more chocolate! Thank you for being there for me, and us, and for always being so willing to help us with everything.

To my fellow Moronians both past and present, thank you! Thank you Dr. Min Teh (Min) for all the help you've given me over the time we spent together in the lab. I wish you all best in the future. To my best japan buddie (Japan bothers) Dr. Jilong Qin (Brad). Thanks for all the times we spent together both in and outside of the lab. Your constant hunger for science and want to get things done was the foundation of how I learnt to do my science and I learnt so much from you over the years. Keep being a great scientist so I have someone to keep on chasing after! Also, don't swap the door it's the same chance %, let us know BEFORE you engage and Cooper's mud is #1! Thanks for being a great friend and role model, stay in touch and all the best to you, Yuna and little Coady!

To Dr. Vincenzo Leo (Vince and/or Vance/Finch) what a time its been. You've always been around as a source of fun in the lab and have made the lab space, for all of us, such a funny and great place to be. You've become one of my best friends and I thank you for all the times we spent 'working' and not designing board games/playing geography games/doing quizes at the back! Thank you for being such an honest person and great friend. I'll never forget the early morning champions league matches, the constant laughs and all the good times we spent together. Of note, those weren't noodles, it became a mouse accuracy game, you're right about the league

and petrol prices but the doors are the same chance so don't swap and no, your arm doesn't grow. Keeping being who you are and I look forward to all the fun we'll have in the future.

To the morning coffee crew Ali, Kate and Tony. You've been a source of daily joy for the past 5 years! I've always enjoyed the news updates, the political insights and Tony, all of your stories over the years. T-bone, keep perfecting your coffee and pizza and don't stress too much about Kokkinakis. Go on more bike rides, keep on being 'polite' during your cycling incidents and I hope you can soon enjoy your retirement and learn some proper Italian!

To my friends outside of the science, thank you all for being so supportive and encouraging over the years. Special mention to the volleyball crew and Oliver, thanks for always being there for a chat and understanding over my absences, I hope we hangout more soon. And, a special thank you to my good friend Dr to be Mr Eden Blazejak. You've always been there over the past years to discuss life and ideas, our PhD journeys, Uni management and teaching woes and everything in between. I've always enjoyed our time together and I hope we'll have many more catch ups to come. Good luck with completing your PhD! Also thank you to Yasir for the fun we've spent together, life has moved us around but I hope we can still stay in touch.

To my family, both immediate and extended, thank you so much. Thank you Ma for always asking how things were going and for always making the best lunches. Your encouragement throughout my studies has always pushed me to be better and I can't thank you enough for what you've done for me. Thank you Pa for NOT always asking how things were going, now that things are done, I will make the time for us to do the things we said we wanted to do over the years. Thank you for always supporting me through my education and I look forward to our long overdue trips together! To ate Kasia, thank you for always being supportive of me and my studies, I know I'm not always the easiest brother but I appreciate the support and all that you've done for me. To Babcia and Dziadzio, I hope I've made you proud in Heaven and I hope that we will meet again. I love you all and again, thank you for all the support for my studies over the years. And to Lorella and Paolo, thank you for always making me feel comfortable and for making me feel a part of your family.

Lastly, to Alice. Thank you for all the time we've spent together and for always being there for me. PhD has been a long, stressful journey but, thanks to you, I've made it out somehow on the other side. Thank you for being my partner and for sticking with me through all stressful PhD times. I know that you will be a fantastic scientist and that you will excel in whatever you choose to do after. I hope we can travel again soon as I love our adventures together. Thank you for everything you've done for me and I look forward to our years together! *Ti amo tanto, baci.*

Abbreviations

AA	Amino acid	nm	Nanometres
AGRF	Australian Genome Research Facility	Oag	O Antigen
Amp	Ampicillin	OD_{600}	Optical density at wavelength of 600 nm
CDS	Coding sequence	OM	Outer membrane
Cml	Chloramphenicol	ON	Overnight
Cyc	Cyclic	PAGE	Polyacrylamide gel electrophoresis
DDT	dichloro-diphenyl-trichloroethane	pg	Phosphatidyl glycerol
ECA	Enterobacterial common antigen	PEtN	Phosphoethanolamine
Fuc*p*4NAc	4-acetamido-2,4-dideoxygalactose	Rha	L-rhamnose
g	Grams	R-LPS	Rough lipopolysaccharide
GlcNAc	*N*-acetylglucosamine	RO	Reverse osmosis
hrs	Hours	RT	Room temperature
IL	Interleukin	Sec	Seconds
IM	Inner membrane	SDS	Sodium dodecyl sulphate
IPTG	isopropyl-β-D-thiogalactopyranoside	S-LPS	Smooth O antigen
Kan	3-deoxy-D-manno-oct-2-ulosonic acid	SN	Supernantant
LB	Lysogeny broth medium	S-Oag	Short O antigen
LBA	Lysogeny broth agar	SR-LPS	Semi-rough lipopolysaccharide
LPS	Lipopolysaccharide	Tet	Tetracycline
mA	Milliamps	TBS	Tris-buffered saline
mg	Micro grams	TTBS	Tris-buffered saline tween
Man*p*NAcA	N-acetyl-mannosaminuronic acid	V	Volts
min	Minutes	v/v	Volume per volume
MOP	Multidrug/Oligosaccharidyl-lipid/Polysaccharide	VL-Oag	Very long O antigen
mPEG	Polyetheylene glycol maleimide	WC	Whole cell
MQ	Milli-Q	w/v	Weight per volume
MSA	Multiple sequence alignment	WM	Whole membrane
ng	Nana Grams	VC	Vector control

Thesis Style and Layout

This thesis is submitted in the style of a “Combination of conventional thesis and thesis by publication formats”. As such, the result chapters are three research publications (Chapter 3, 4 & 5) and a traditional chapter (Chapter 6). In Chapter 1, the studies and research in the field of this work is reviewed and research gaps identified as aims for investigation. In Chapter 2, the Materials and Methods used in this thesis are outlined in detail.

As for publication purposes, Chapters 3, 4 and 5 are presented as per the requirement of the journal, which includes all the information that will be submitted for publication. Author contributions for each publication are stated in the Statement of Authorship section. For Chapter 6 materials and methods used are outlined in Chapter 2 and are cross-referenced. Each result chapter has a discussion section to explore the underlying meaning of the work. Chapter 7 draws conclusions of the outcomes and the significance of this study, and points out the future research directions.

Declaration……iii

Abstract……iv

Acknowledgements……...v

Abbreviations……vii

Thesis Style and Layout ……..…viii

Contents……xi

Chapter 1: Introduction **2**

1.1 ***Shigella*** **2**

1.2 **Outer membrane glycolipids of *Shigella flexneri*** **2**

1.3 **Wzy-dependent Pathway** **5**

1.3.1 Wzx flippase 5
1.3.2 Wzy Polymerase 8
1.3.3 Wzz Polysaccharide Co-Polymerase 9

1.4 **Enterobacterial Common Antigen** **14**

1.4.1 Phosphatidyl linked ECA, ECA_{pg} 14
1.4.2 Lipid A linked ECA, ECA_{lps} 14
1.4.3 Cyclic ECA, ECA_{cyc} 16
1.4.4 ECA biosynthesis 16
1.4.5 *wec* operon mutants and the Rcs Pathway 18
1.4.6 *wec* operon cross-complementation 21
1.4.7 Bacteriophage N4 and ECA 21
1.4.8 Biotechnological applications of *wec* operon mutants 22
1.4.9 Roles of ECA - Pathogenesis 23
1.4.10 Roles of ECA – Self regulation 23
1.4.11 Roles of ECA – Homeostasis 24
1.4.12 Roles of ECA - Biosynthetic intermediates 25
1.4.13 ECA based vaccine 25

1.5 **Lipopolysaccharide** **26**

1.5.1 Lipid A 26
1.5.2 Core sugars 29
1.5.3 O antigen 31
1.5.4 O antigen function 31
1.5.5 Lpt System 34

1.6 **Peptidoglycan** **36**

1.6.1 Undecaprenyl phosphate 38

1.7 **Research aims and hypothesis** **40**

Chapter 2: Materials and Methods **42**

2.1 **Bacterial strains and growth conditions** **42**

2.1.1 Strains and plasmids 42
2.1.2 Growth media and conditions 42
2.1.3 Antibiotics and additives 42
2.1.4 Growth and maintenance of bacterial strains 42
2.2 Antibodies 42
2.3 DNA techniques 43
2.3.1 Plasmid isolation 43
2.3.2 Whole chromosome extraction 43
2.3.3 DNA quantification 43
2.3.4 Restriction digests 43
2.3.5 Oligonucleotides for PCR 43
2.3.6 PCR 43
2.3.7 Preparation of boiled lysate for PCR DNA template 44
2.3.8 DNA NaCl precipitation 44
2.3.9 PCR product DNA purification kit 44
2.3.10 Agarose gel electrophoresis 44
2.3.11 DNA gel extraction 45
2.3.12 DNA phosphorylation 45
2.3.13 DNA ligation 45
2.3.14 Site-directed mutagenesis 45
2.3.15 Inverse PCR to generate point substitutions, insertions and deletions 45
2.3.16 Chromosomal deletions via λ Red mutagenesis 45
2.3.17 Generation of nested DNA deletions 46
2.3.18 DNA sequencing 46
2.4 Bacterial Transformation and fixing 47
2.4.1 Preparation of chemically competent cells 47
2.4.2 Preparation of ultra competent cells 47
2.4.3 Preparation of electro-competent cells 47
2.4.4 Heat-shock transformation of chemically competent cells 47
2.4.5 Electroporation transformation of electro-competent cells 48
2.5 Protein techniques 48
2.5.1 Generation of whole cell lysate protein samples 48
2.5.2 Generation of whole membrane samples 48
2.5.3 *In vivo* protein crosslinking 48
2.5.4 Protein chaotropic disassociation 49
2.5.5 Whole membrane protein PEGylation 49
2.5.6 *In vivo* whole cell protein PEGylation 49
2.5.7 Protein SDS-PAGE 50
2.5.8 SDS-PAGE Western transfer 50
2.5.9 Protein detection by immunoblotting 50
2.6 Polysaccharide techniques 50
2.6.1 Preparation of polysaccharide samples 50
2.6.2 Detection of LPS by silver-staining 51
2.6.3 Detection of ECA by SDS-PAGE and Western immunoblotting 51
2.7 Bacterial sensitivity and growth analysis 51
2.7.1 Deoxycholate resistance assays 51
2.7.2 Colicin sensitivity assays 51
2.7.3 Analysis of growth kinetics by growth curves 52
2.7.4 Analysis of growth kinetics by colony forming unit (CFU) counting 52

2.8 Microscopy 52
2.8.1 Preparation of phase contrast slides 52
2.8.2 Visualization of cells by phase contrast microscopy 52
2.8.3 Cell measurements 52
2.9 Enzymatic assays 52
2.9.1 Alkaline phosphatase (PhoA) assay 52
2.9.2 β-galactosidase (LacZ) assay 53
2.9.3 Generation of Normalized Activity Ratio (NAR) 53
2.10 Bioinformatic analysis 53
2.10.1 DNA and peptide sequences 53
2.10.2 Multiple Sequence Alignments (MSA) 53
2.10.3 Secondary structure prediction 53
2.10.4 Protein topology prediction 53
2.10.5 Phylogenetic and Principal Component Analysis (PCA) 54
2.10.6 Isoelectric point analysis 54
2.10.7 Tertiary structure obtainment and prediction 54
2.10.8 Protein structure modelling and manipulation 54
2.11 Statistical analysis 54
2.11.1 Student T-tests 54
2.11.2 One-way ANOVA 54

Chapter 3: Topology of the Enterobacterial Common Antigen polymerase, WzyE 56
3.1 Statement of Authorship 56
3.2 Article Abstract 57
3.3 Article Introduction 58
3.4 Article Methods 60
3.4.1 Ethics statement 60
3.4.2 Bacterial strains, growth media and growth conditions 60
3.4.3 DNA methods 60
3.4.4 Construct generation 60
3.4.5 Generation of nested DNA deletions 60
3.4.6 Generation of targeted DNA deletions 61
3.4.7 Generation of arginine to glycine substitutions 61
3.4.8 Whole cell protein sample preparation 61
3.4.9 ECA and LPS sample preparation 61
3.4.10 Western immunoblotting 62
3.4.11 Bioinformatic analysis 62
3.4.12 Alkaline phosphatase assays 62
3.4.13 β-galactosidase assays 62
3.4.14 Normalized Enzymatic Activity Ratio (NAR) calculation 63
3.5 Article Results 65
3.5.1 *In silico* analysis of WzyE 65
3.5.2 Topology mapping of WzyE 69
3.5.3 Characterization of WzyE conserved arginine residues 72
3.6 Article Discussion 74

3.7 Article References 77

3.8 Article Supporting Information 81

3.8.1 Supporting Figures 81
3.8.2 Supporting Tables 85
3.8.3 Supporting Dataset. 89

3.9 Article Acknowledgements 96

3.10 Article Conflicts of interest 96

3.11 Additional Results 97

3.11.1 Analysis of WzyE cystine residues 97
3.11.2 Probing of $WzyE_{SF}$ cysteine residues with mPEG 99
3.11.3 Summary of Additional Results 102

Chapter 4: Interdependence of *Shigella flexneri* surface polysaccharides ECA and LPS. 104

4.1 Statement of Authorship 104

4.2 Article Abstract 106

4.3 Article Importance 106

4.4 Article Introduction 107

4.5 Article Methods 109

4.5.1 Ethics statement 109
4.5.2 Bacterial strains, growth media and growth conditions 109
4.5.3 DNA methods 109
4.5.4 Chromosomal mutagenesis 109
4.5.5 Generation of complementing constructs 109
4.5.6 LPS/ECA sample preparation 110
4.5.7 LPS SDS-PAGE and silver staining 110
4.5.8 ECA PAGE and Western immunoblotting 110
4.5.9 Measuring OM polysaccharide abundance 110
4.5.10 Growth curves 111
4.5.11 Colony Forming Units (CFU) counting 111
4.5.12 Microscopy 111
4.5.13 Statistical analysis 111

4.6 Article Results 114

4.6.1 Investigating the biosynthetic effects of the *ΔwzyE* mutation 114
4.6.2 Investigating the biosynthetic effects of the *ΔwzyB* mutation 120
4.6.3 Improving cellular undecaprenyl pool rescues mutant *wzy* phenotypes 123
4.6.4 The sequestration of Und-P and not the loss of OM polysaccharides is correlated to the *wzy* mutant phenotypes 127

4.7 Article Discussion 129

4.8 Article Reference 133

4.9 Article Supporting Information 138

4.9.1 Supporting Figures 138

4.10 Article Acknowledgements 139

4.11 Article Conflicts of interest 139

Chapter 5: Subcellular localization of WecG in *Shigella flexneri.* 141

5.1 Statement of Authorship 141

5.2 Article Abstract 142

5.3 Article Importance 142

5.4 Article Introduction 143

5.5 Article Methods 145

5.5.1 Ethics statement 145
5.5.2 Bacterial strains, growth media and growth conditions 145
5.5.3 DNA methods 145
5.5.4 Chromosomal mutagenesis 145
5.5.5 Construct generation/DNA sequencing 145
5.5.6 Whole cell protein sample preparation 146
5.5.7 ECA sample preparation 146
5.5.8 Membrane fractionation 146
5.5.9 Dissociation assays 146
5.5.10 *In vivo* DSP Crosslinking 147
5.5.11 Western immunoblotting 147
5.5.12 Bioinformatic analysis 147

5.6 Article Results 150

5.6.1 Bioinformatic analysis of WecG 150
5.6.2 WecG associates with the membrane 150
5.6.3 Peripheral association of WecG with the membrane is facilitated via CTD helices. 153
5.6.4 Key hydrophobic residues in WecG CTD helix II facilitate membrane association. 154
5.6.5 WzyE polymerase impacts WecG's membrane association. 159
5.6.6 Lipid interactions are crucial for WecG's association with the membrane. 159

5.7 Article Discussion 163

5.8 Article References 167

5.9 Article Supporting Information 171

5.9.1 Supporting Figures 171

5.10 Article graphical abstract 173

5.11 Article Acknowledgments 173

5.12 Article Conflicts of interest 173

Chapter 6: Additional phenotypes of *wzyE* mutant. 175

6.1 Introduction. 175

6.2 Generation of *ΔwzyE* and *ΔwzyE*$_{10\text{-}440}$ mutants by λ Red mutagenesis. 175

6.3 Characterization of the outer membrane polysaccharides of *wzyE* mutants. 176

6.4 Cell sensitivities of *wzyE* mutants to deoxycholate (DOC) and Colicin E2. .176

Chapter 7: Conclusion.....184

7.1 Novel Wzy topology characteristics are found in WzyE.....184

7.2 ECA and LPS Oag pathways are indirectly interdependent on one another in *Shigella flexneri.*.....185

7.3 WecG is a GT-E fold protein and is maintained to the membrane by ECA lipid-I.....186

7.4 The accumulation of biosynthetic intermediates provides strong resistances in *wzyE* mutants. 187

7.5 Summary 188

Appendix A: Bacterial Laboratory Strains……...191

Appendix B: Bacterial Strains Generated…….....192

Appendix C: Plasmids…….204

Appendix D: Oligonucleotides……206

Appendix E: Thesis Bibliography……..213

Chapter One

INTRODUCTION

Chapter 1: Introduction

1.1 *Shigella*

Shigella belong to the *Enterobacteriaceae* bacterial family and are non-motile intracellular bacterial pathogens specific to humans. Infection by *Shigella* causes shigellosis, a clinical syndrome which is an acute infection of the epithelial lining of the terminal lumen, colon and rectum and accounts for 164,000 annual deaths (Kotloff et al. 2018). *Shigella* comprises of four species: *S. dysentery*, *S. boydii*, *S. flexneri* and *S. sonnei* and their global disease burden is geographically distinct with different *Shigella* species more prevalent across first and third world countries. *Shigella* burden is most prevalent in developing countries in low and middle income settings where infections in children aged 1 to 4 years old contribute to the majority of the disease burden (Kotloff et al. 2018). In these settings, *S. flexneri* is associated with significantly higher mortality rate compared to other *Shigella* species (Gentle et al. 2016), whereas in higher income countries, *S. sonnei* is the most prevalent species (Ram et al. 2008) as illustrated in Figure 1.1.

The increase occurrence of multi-drug resistant *Shigella* prompted the Centre for Disease Control (CDC) to denote multi-drug resistant *Shigella* as a 'serious threat' (CDC 2019). The growing resistance to ampicillin and trimethoprim-sulfamethoxazoles (Sivapalasingam et al. 2006) requires treatment of shigellosis to reply on drugs such as ciprofloxacin, azithromycin and ceftriaxone, however resistance to these drugs too is emerging (Boumghar-Bourtchai et al. 2008; Chung The et al. 2016). Bioconjugate vaccines to help combat *Shigellae* have been developed, one such vaccine comprises of *S. flexneri* 2a O-antigen (Oag) conjugated to *Pseudomonas aeruginosa* endotoxin A which has been shown to protect against shigellosis in phase 2 trials (Talaat et al. 2021).

1.2 Outer membrane glycolipids of *Shigella flexneri*.

The cell wall of Gram negative bacteria consists of three distinct domains: the inner membrane (IM), the outer membrane (OM) and the separating space between them, the periplasm. The cell wall functions as a physical barrier separating and protecting the bacterium from the extracellular environment. The OM of *S. flexneri* comprises of an inner and outer leaflet of phospholipids and lipopolysaccharide (LPS), respectively. In addition to LPS, *S. flexneri* also possess the glycolipid Enterobacterial Common Antigen (ECA) and the two glycolipids make up the lipid composition of the outer leaflet of the OM. Both OM polysaccharides are biosynthesised by independent homologs of the Wzy-dependent pathway however, in other bacteria, the ABC transporter and synthase pathways have also been described for the biosynthesis of LPS O antigen (Bi et al. 2018). The general construction of the cell wall is illustrated in Figure 1.2.

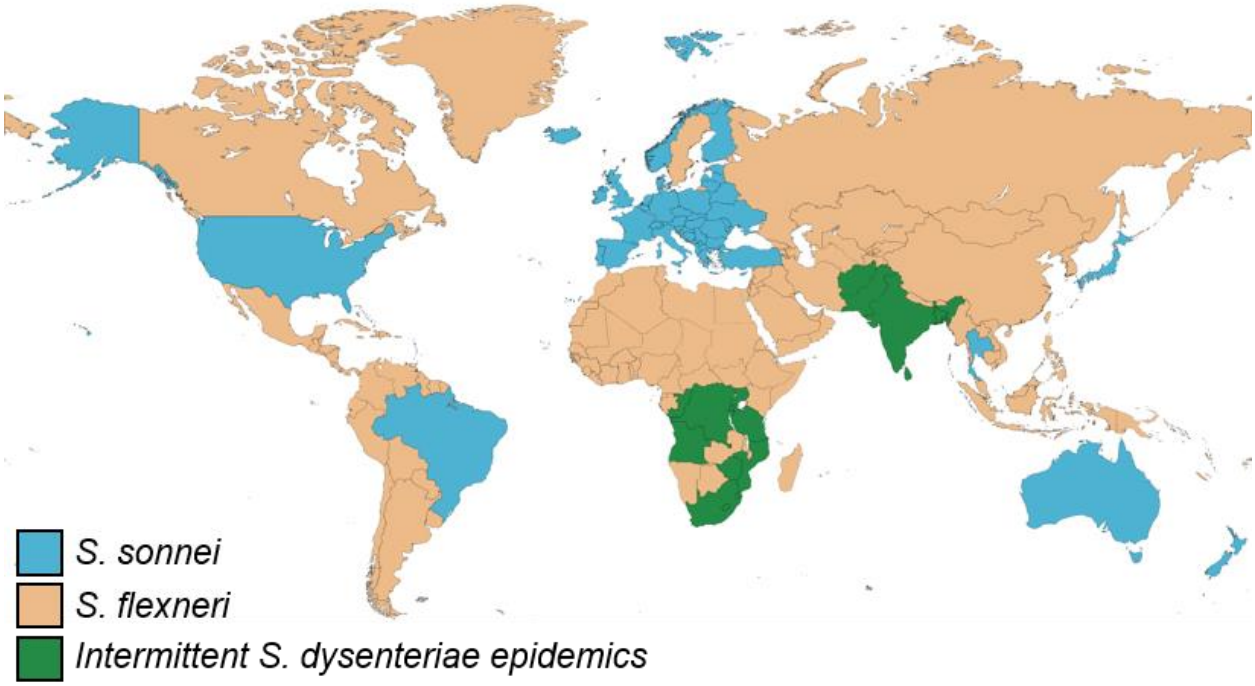


Figure 1.1: Global distribution of *Shigella* disease burden.

Shigella disease morbidity is of global concern with different *Shigella* species causing the majority of disease burden in different countries. Figure adapted from (Bennish 2012)

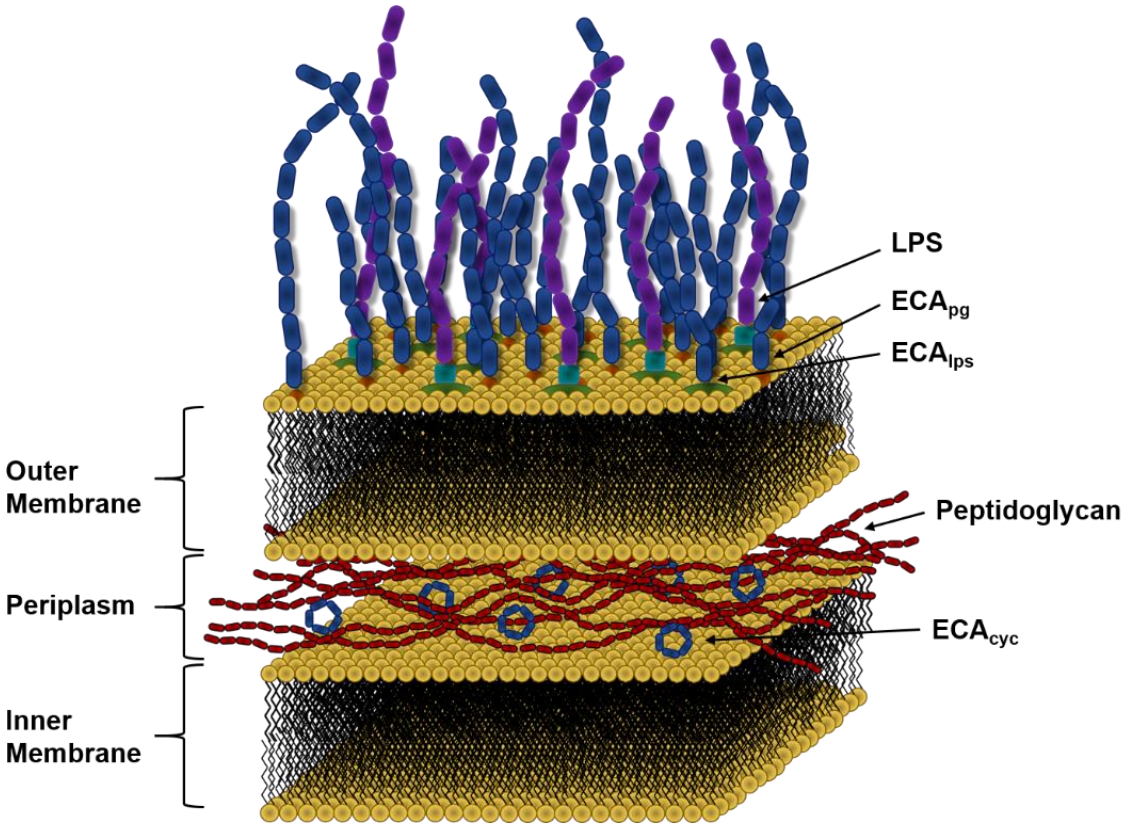


Figure 1.2: Cell wall features of *Enterobacteriaceae.*

The cell wall of *Enterobacteriaceae* comprises of three distinct layers; the outer membrane, periplasm and inner membrane. On the exterior leaflet of the outer membrane comprises of LPS whose O antigen protrudes from the surface. Along with LPS O antigen, ECA also protrudes from the surface where it is found commonly as ECA_{pg}. ECA_{lps} is found in strains lacking O antigen whereas ECA_{cyc} is restricted to the periplasm. ECA_{pg}=phosphatidyl-linked ECA, ECA_{lps}=Lipid A core-linked ECA, ECA_{cyc}=cyclic ECA.

The OM polysaccharides of *S. flexneri* are crucial in its pathogenesis and as such a large portion of research, not only in *S. flexneri* but in enterobacterial pathogens, has been directed to the LPS O antigen. Subsequently ECA, which is poorly studied, and its roles in pathogenesis have either been overlooked or not properly investigated due to the difficulties in isolating non-pleotropic mutants.

1.3 Wzy-dependent Pathway

The Wzy-dependent pathway the most common bacterial polysaccharide biosynthetic pathway which consists of three integral membrane proteins Wzx, Wzy and Wzz. Wzx, the flippase, is responsible for the translocation of lipid linked RUs (O antigen or ECA) from the cytoplasmic to the periplasmic leaflet of the IM. Here Wzy, the polymerase, polymerizes RUs into linear polysaccharide chain and Wzz, the polysaccharide co-polymerase (PCP) and modal length regulator specifies the degree of polymerisation performed by Wzy (Islam & Lam 2014; Kalynych, Morona & Cygler 2014). The Wzy-dependent pathway is described in Figure 1.3.

A number of speculative models of Wzy-Wzz mediated polymerisation have been proposed by various research groups. Bastin et.al. (1993) presented the 'clock model'; Morona et.al. (1995) proposed the 'chaperone model'; Tocilj et al., (2008) presented the 'organizing scaffold model'; Kintz and Goldberg (2011) proposed the 'Ruler model' while Kalynych et. al (2012) presented the 'Chain-feedback model'. The two most recent models include the 'Hybrid (chain-feedback-ruler)' model proposed by Islam and Lam (2014) and the hybrid (ruler-stopwatch) model proposed by Collins et. al (2017). The Hybrid (Ruler-Stopwatch) proposed by (Collins et al. 2017), incorporates ideas from both the Hybrid (Chain-Ruler-Feedback model) and the (stopwatch model) proposed by Lam et.al (2014) and Bastin et.al (1993) respectively. The model is illustrated in Figure 1.4.

1.3.1 Wzx flippase

Wzx proteins belong to the polysaccharide transport (PST) family, which itself belongs to the multidrug-oligosaccharide lipid-polysaccharide exporter (MOP) super family and are responsible for the translocation of lipid-linked biosynthetic intermediates from the inner to the outer leaflet of the IM (Hvorup et al. 2003).

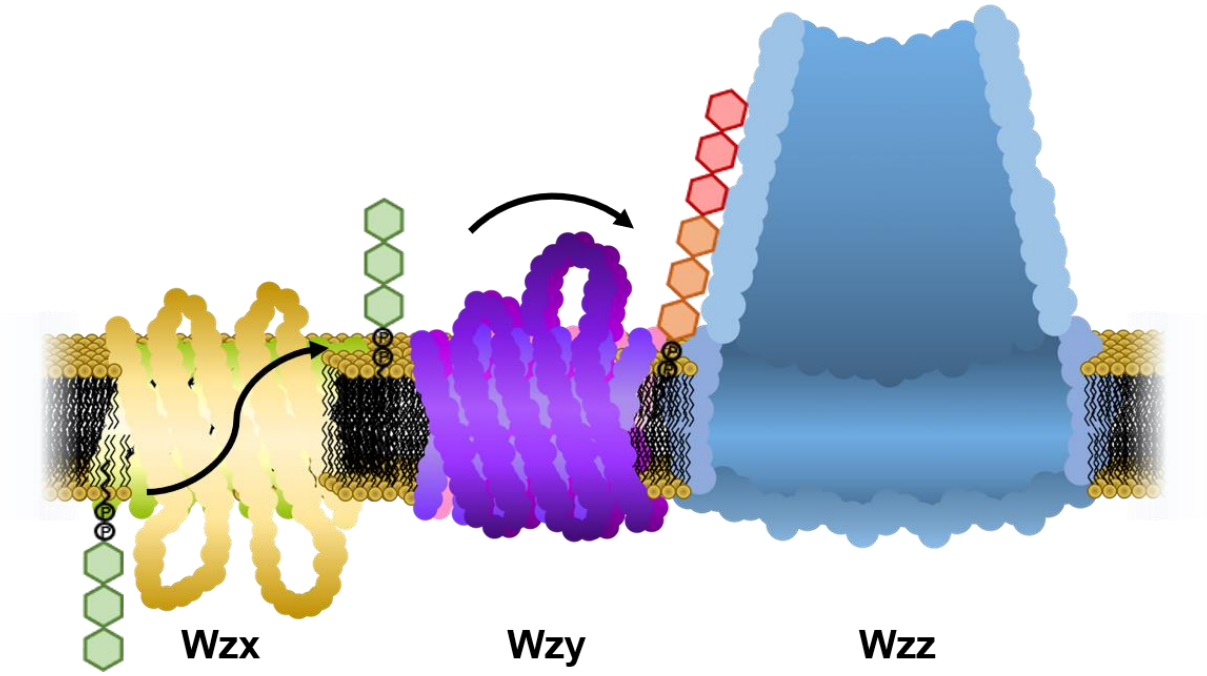


Figure 1.3: Wzy-dependent pathway.

The Wzy-dependent pathway consists of three proteins; Wzx the flippase which translocates lipid linked intermediates across the IM, Wzy the polymerase which polymerizes repeat units into linear chains and Wzz polysaccharide co-polymerase which controls the length of the polysaccharide assembled by Wzy. The LPS O antigen homologs of the Wzy-dependent pathway are WzxB, WzyB and WzzB, whereas the ECA homologs are WzxE, WzyE and WzzE.

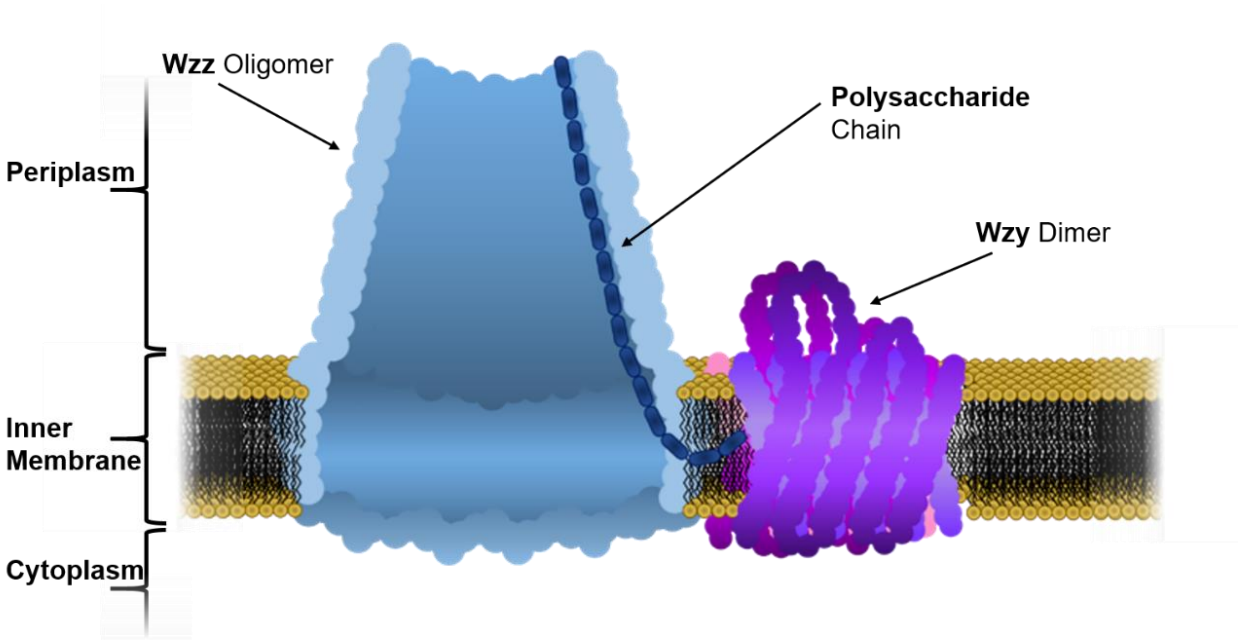


Figure 1.4: The hybrid Ruler-stopwatch model of Wzy Wzz polymerization.

Wzy dimers engage the transmembrane domains of a Wzz promoter within the conserved bell-shaped oligomeric structure. After a few rounds of polymerization, the growing chain is inserted within the IM through periplasmic 'gaps' within the transmembrane regions of the Wzz homooligomer into the central cavity where it's kept in a polymerisation permissive state, allowing for the further addition of repeat units. As the chain grows progressively longer and rigid, higher order structures form and cause the destabilization of the linear chain from the bell; leading to mechanical feedback being transmitted down the chain to Wzy where the chain is disengaged or alternatively until, Wzy disassociates from the Wzz oligomer (Collins et al. 2017).

Currently there are no tertiary structures of Wzx flippase, however structures do exist for MurJ and PglK which translocate PG lipid-II and lipid linked oligosaccharides (LLOs) which are required in the N-linked protein glycosylation pathway, respectively (Perez et al. 2017; Sham et al. 2014). The structures of MurJ and PglK have clarified, to a degree, the mechanism of action for PST flippases which can be divided into two separate steps; substrate binding followed by substrate translocation.

The structure of MurJ resembled two lobes with a cation lumen between them, which could be accessed using a lateral gate and was the suggested site for substrate binding. This was shown to be correct as the binding the substrate in multiple confirmations was observed in this domain. This included the substrate binding to the gate followed by a confirmation where the substrate was captured by the lumen in an inward-facing confirmation. The proposed mechanism of substrate translocation is called the "rocket switch" mechanism where simply the two lobes squeeze the lumen which forces the substrate out (Kuk et al. 2019; Kuk, Mashalidis & Lee 2017; Kumar et al. 2019). Due to the fact that Wzx and MurJ both belong to the PST family, it is likely that Wzx functions to translocate RUs in a similar way to MurJ.

Mutations in Wzx proteins are known to accumulate lipid-linked intermediates presumably on the cytoplasmic leaflet of the IM as was shown for WzxB in *S.enterica* (Liu, Cole & Reeves 1996). Additionally, the expression of WecA in an *Escherichia coli* K12 *wzxE* mutant led to the lysis of cells (Rick et al. 2003) and growth curves of wzx_{016} and *wzxE* mutants expressing WecA showed a decrease in cellular growth kinetics (Marolda et al. 2006).

While most cell wall biosynthetic pathways have pathway specific homologs of Wzx, cross complementation and promiscuity of Wzx proteins have been observed. The overexpression of Wzx has been shown to lead to substrate promiscuity revealing the ability of Wzx to transport non-intended substrates (Liu, Morris & Reeves 2019). Due to this, overexpression of Wzx was shown to complement mutations in MurJ, restoring PG synthesis (Sham et al. 2018).

1.3.2 Wzy Polymerase

Wzy proteins have been extensively studied from several species of bacteria including *E. coli, S. flexneri, Pseudomonas aeruginosa, Francisella tularensis* and *Rhizobium leguminosarum* where, the majority of studies have been directed towards WzyB, the Wzy functioning in the O antigen biosynthetic pathway.

Wzy proteins belong to the Shape, Elongation, Division and Sporulation (SEDS) protein family which are all characterized as being polytopic inner membrane proteins involved in cell

wall biosynthetic processes (Meeske et al. 2016). They themselves consist of 10-14 transmembrane segments (TMs) in the form of α helices and two large periplasmic loops (PL), PL3 and PL5, which are known to contain catalytic domains, and functional motifs (Islam et al. 2011; Nath & Morona 2015b). The topology map of *S. flexneri* WzyB is shown in Figure 1.5 (Nath & Morona 2015b). As Wzy is an integral membrane protein which are notoriously hard to crystalize, there are no available crystal structures to date; the precise tertiary structure fold adopted by the polypeptide chain is unknown (Islam & Lam 2014). Originally, topology mapping of Wzy was conducted through the use of PhoA and LacZ reporter fusion proteins which revealed the first recorded topology map of Wzy_{SF} (Daniels et al. 1998). This was then followed by studies investigating the topology of Wzy from *P. aeruginosa* and *R. leguminosarum* also using PhoA::LacZ fusion proteins to determine the topology (Islam et al. 2010; Mazur et al. 2003).

The majority of analysis performed on Wzy has been conducted in *P. aeruginosa* with the discovery of key motifs in PL3 and PL5 (Islam et al. 2011). They were both shown to contain a Rx10G motif which proved to be important in the proposed model for Wzy function in O antigen polymerization, named 'Catch and Release' described in Figure 1.6. Nath et.al (2015b) studied $WzyB_{SF}$ from *S. flexneri* and showed a similar Rx15G motif in both PL3 and PL5, starting from R164 in PL3 and R289 in PL5, which supports the importance of conserved arginine groups from within the loops (Nath & Morona 2015b; Nath, Tran & Morona 2015a).

Wzy exhibits low sequence conservation between different bacterial species and different serotypes; for instance among the different serotypes of *P. aeruginosa*, there is considerably low sequence conservation between Wzy homologs (Islam & Lam 2014). *Islam et. al* (2013) performed extensive work on Wzy_{PA} and conducted a '*jackhammer'* search to find the homologs of Wzy_{PA}; however their results showed that Wzy_{PA} is not related to the Wzy of *Enterobacteriaceae* (Nath & Morona 2015b). The 'Catch and Release' model proposed for Wzy function, would not be supported by Wzy_{SF} due to the similar pKa charges observed for PL3 and PL5 (Nath & Morona 2015b).

1.3.3 Wzz Polysaccharide Co-Polymerase

Wzz is known to be the modal chain length regulator of Wzy-dependent surface polysaccharides (Morona et al. 2009; Nath et al. 2015b). Wzz proteins are divided into three groups based on the chemical nature of their polysaccharide, their association with the Wzy-dependent or ABC cassette transporter dependent pathway and the presence or absence of an additional cytoplasmic domain where those involved with LPS O antigen and ECA biosynthesis are denoted as PCP-1 (Morona et al. 2009).

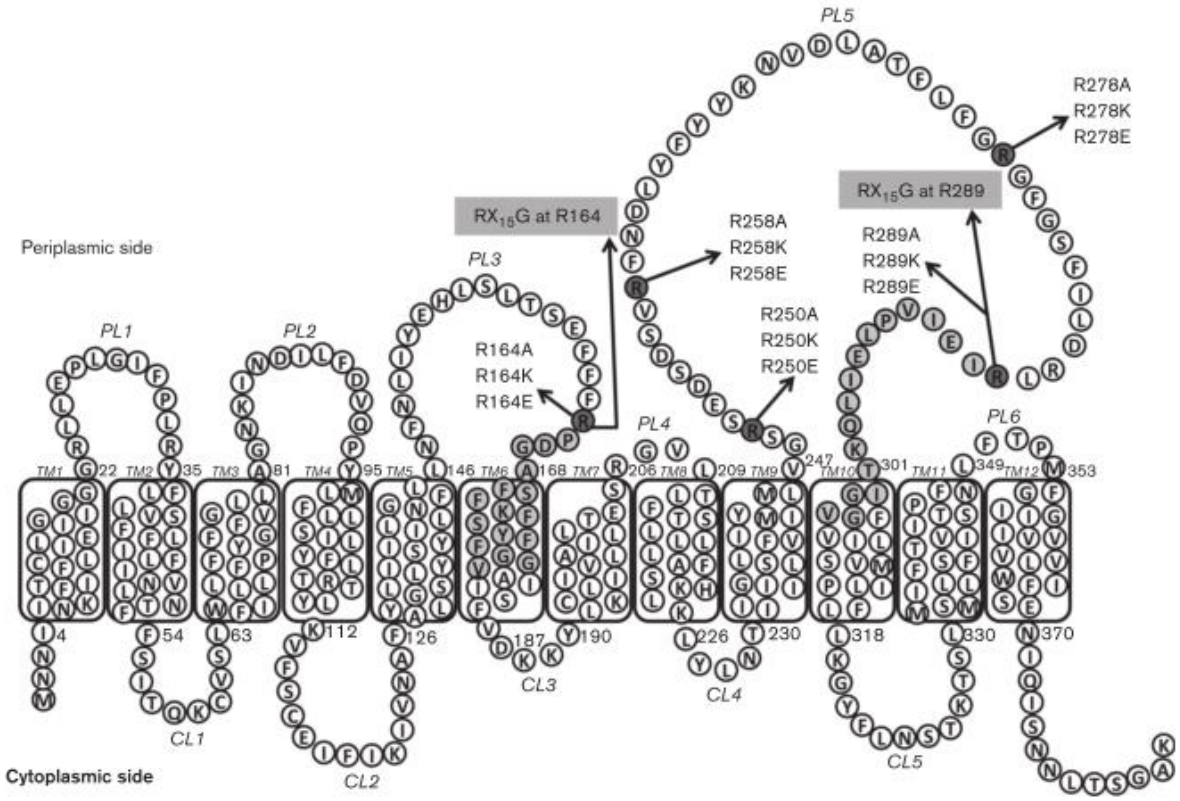


Figure 1.5: Topology map of $WzyB_{SF}$.

Wzy_{SF} follows a similar topology pattern as all other Wzy proteins, consisting of 12 TMS and two large periplasmic loops; PL3 and PL5. Nath et. al (2015b) performed site-directed mutagenesis on arginine groups present in PL3 and PL5 and showed that they were important for polymerisation activity. Figure adapted from Nath & Morona (2015b).

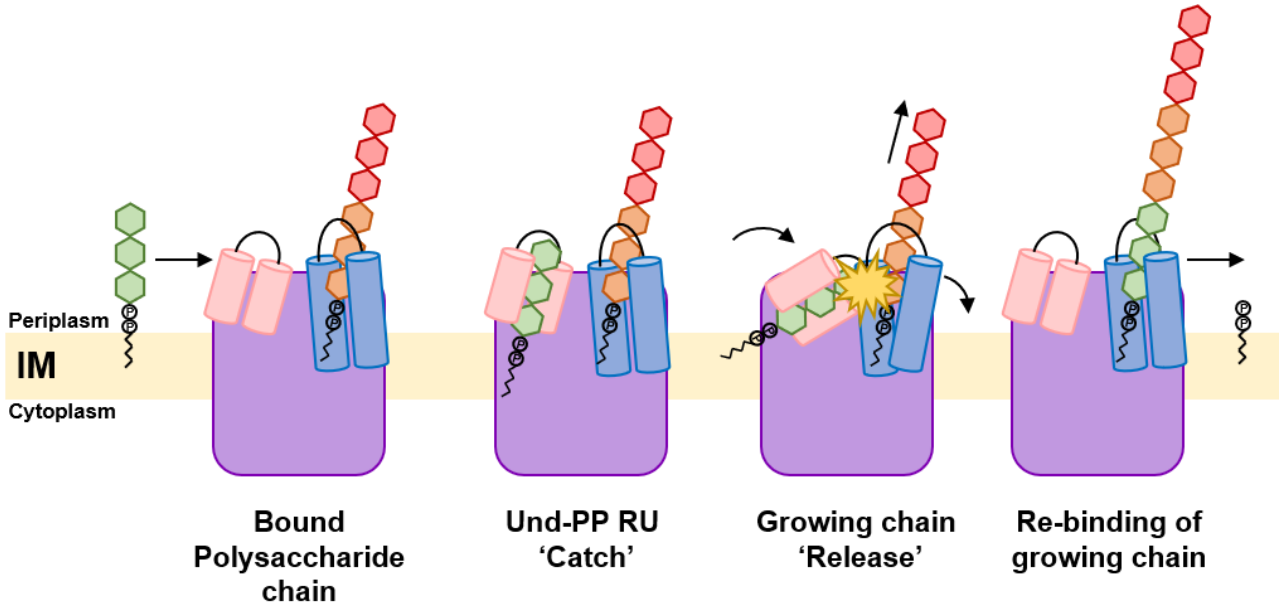


Figure 1.6: 'Catch and Release' model of Wzy function.

Opposingly charged periplasmic loops, PL3 and PL5, are thought to facilitate the capture or 'catch' of lipid bound repeat units and to facilitate their subsequent transfer or 'release' to the reducing end of the growing polysaccharide chain (Islam et al. 2011; Islam et al. 2010).

PCP-1 Wzzs are further categorised into groups where the main PCP from OM polysaccharide pathways are denoted as PCP-1a; WzzB and WzzE and secondary PCPs, which control an additional modal length, are denoted as PCP-1b which includes Wzz_{pSH2} and FepE from *S. flexneri* and *Salmonella enterica,* respectively (Morona et al. 2009).

Despite the categorisation, structural analysis of full length Wzz proteins revealed that they adopt a similar tertiary and quaternary folding pattern where the Wzz monomers form a bell-like structure (Wiseman et al. 2021); the structure of Wzz is illustrated in Figure 1.7. The structure of Wzz can be sub-divided into two major segments, the transmembrane region which comprise of two separate α-helixes and the periplasmic regions which protrude into the periplasm and each play specific roles. The two transmembrane regions TM1 have been shown to be important for the interaction between Wzz monomers where they are believed to directly interact with the TM1 regions of adjacent monomers. Whereas TM2 has been shown to be important in the interaction between Wzy and Wzz where TM2 substitutions between WzzE and WzzB chimera proteins showed that substitution of TM2 was sufficient to allow for the cross complementation of the two Wzy pathway homologs (Leo et al. 2020). Furthermore, a functionally important GXXXG motif adjected to proline-rich segment in known to exist in TM2; mutagenesis of residues within these motifs causes unregulated polysaccharide lengths (Papadopoulos et al. 2016; Wiseman et al. 2021).

The periplasmic region of Wzz is believed to facilitate the binding of the growing polysaccharide chain however whether the polysaccharide is polymerized within or on the exterior of the Wzz bell is controversial. It was long accepted that the polymerisation of polysaccharide chains occurred on the exterior of the bell, however mutations from within the bell were shown to affect LPS modal length and control (Papadopoulos, Magdalene & Morona 2010). This subsequently led to the current models of Wzy-dependent polymerisation as discussed in 1.3.

The oligomeric state of the Wzz complex is also controversial where it has been shown experimentally to form pentamers, hexamers and octamers by various research groups (Kalynych et al. 2015; Tocilj et al. 2008; Wiseman et al. 2021). Most recently, researchers were able to show Wzz in an octameric complex through the use of Ni+ resin purified $WzzB_{EC}$, cryoelectron microscopy and n-Dodecyl-B-D-Maltoside (DDM) detergent (Wiseman et al. 2021). The group revealed that the transmembrane segments of each protomer do not physically interact and that the presence of 20-30 Å ‘gaps’ between each protomer may allow for polysaccharide insertion by Wzy (Collins et al. 2017; Wiseman et al. 2021). The literature supports that these oligomeric complex

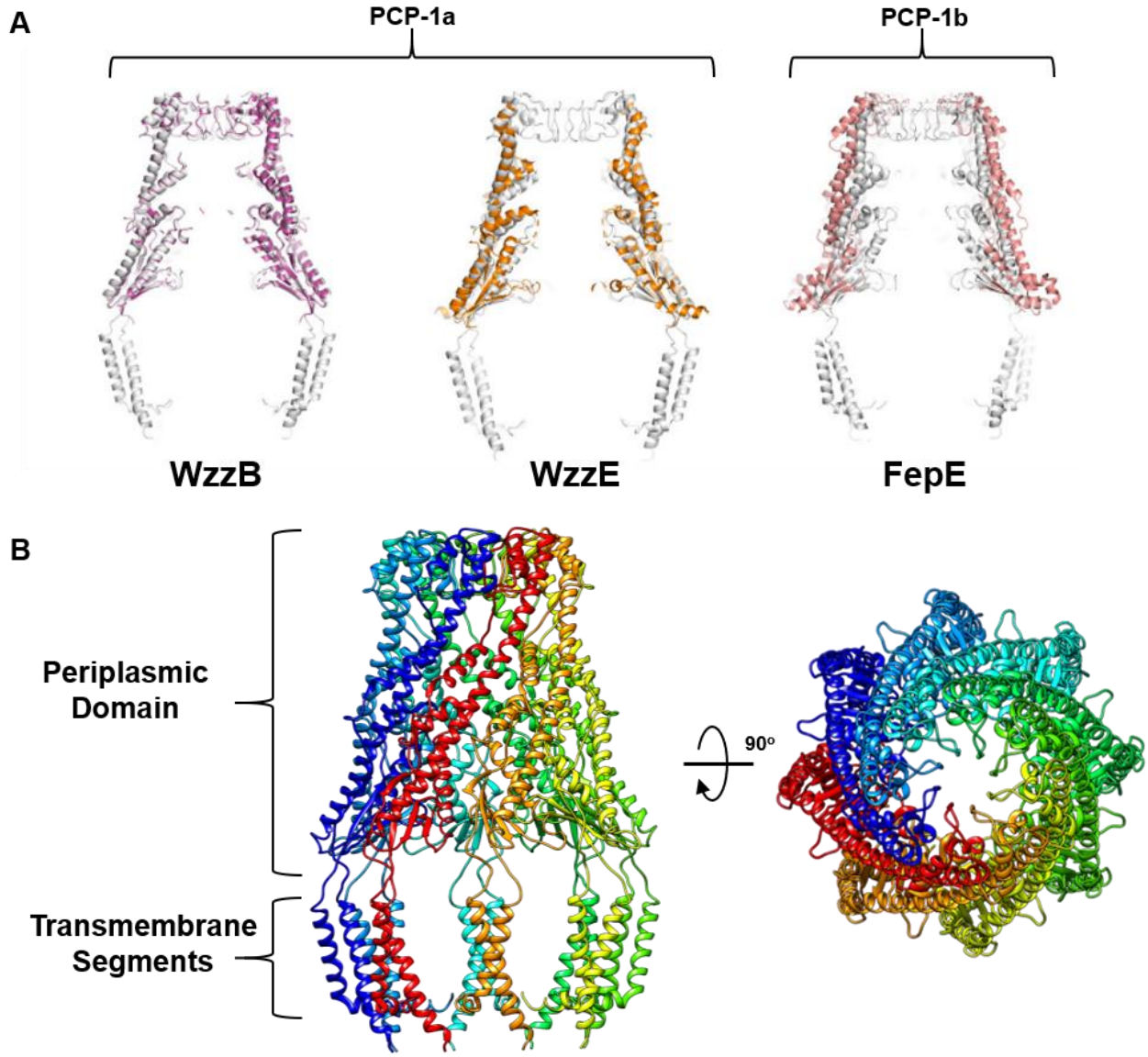


Figure 1.7: Wzz protomer and oligomeric structure.

A) Structures of the PCP-1a and PCP-1b Wzz protomers displayed front on showing the degree of similarity between the three Wzz proteins; WzzE, WzzB and FepE. B) Structure of full length WzzB oligomer from PDB. The structure displays the two separate domains of Wzz bells, the transmembrane domain and periplasmic domain. Structure was deposited by Wiseman et al. (2021) PDB accession ID = 6RBG. Different monomers of Wzz are displayed in different colours.

compete for available Wzy with an oligomeric state as an octamer showing the most intimate promoter-promoter contacts, suggesting this to be the most stable oligomer (Carter et al. 2009; Collins et al. 2017; Kalynych et al. 2015; Wiseman et al. 2021).

1.4 Enterobacterial Common Antigen

The Enterobacterial Common Antigen (ECA) is a bacterial heteropolysaccharide which is composed of a trisaccharide repeat unit: - >3)-a-D-Fuc*p*4NAc-(1->4)-B-D-Man*p*NAcA-(1->4)-a-D-Glc*p*NAc-(1-> (Gozdziewicz, Lugowski & Lukasiewicz 2014), where Fuc*p*4NAc refers to 4-acetamido-2,4-dideoxygalactose, Man*p*NAcA to N-acetyl-mannosaminuronic acid and Glc*p*NAc to N-acetylglucosamine. Predominantly linked to the OM as a linear chain, ECA exists in three arrangements; a phosphatidylglycerol (PG) linked form (ECA_{pg}), which is ubiquitously expressed on the OM of all *Enterobacteriaceae*, a core-oligosaccharide, LPS associated form (ECA_{lps}) and a cyclic form which contains no lipid anchor and is restricted to the periplasmic space (ECA_{cyc}) (Gozdziewicz, Lugowski & Lukasiewicz 2014; Kajimura, Rahman & Rick 2005; Rai & Mitchell 2020). These arrangements are illustrated in Figure 1.8.

1.4.1 Phosphatidyl linked ECA, ECA_{pg}

ECA_{pg} is the most abundant form of ECA found on *Enterobacteriaceae* where it constitutes roughly 0.2% of the cellular dry weight of *E. coli* K12 (Hella-Monika Kuhn 1988). In this form, ECA repeat units are directly attached onto a phosphatidylglycerol backbone in a WaaL ligase independent manner (Barr et al. 1999). The attachment is mediated via a glyosidic linkage between the phosphate moiety of the phosphatidylglycerol residue and the reducing end of the GlcNAc ECA repeat trisaccharide as it was shown to be sensitive to cleavage mediated by phospholipase D (Kajimura et al. 2006; Kuhn et al. 1987).

1.4.2 Lipid A linked ECA, ECA_{lps}

ECA_{lps} is the only naturally immunogenic form of ECA and is only present on strains lacking O antigen which contain either a R1, R2, R4 or K-12 core structure (Barr, Klena & Rick 1999; Gozdziewicz, Lugowski & Lukasiewicz 2014; Hella-Monika Kuhn 1988; Maciejewska et al. 2020). Whilst this very strongly supported, it has been suggested that in in *Y. enterocolitica* O:3 that strains expressing O antigen and ECA_{lps} can co-exist on the OM (Muszynski et al. 2013). In the ECA_{lps} arrangement, ECA trisaccharide repeat units are directly linked to the LPS core-sugars, where they are thought to constitute <5% of the total membrane bound ECA population and ECA_{pg} remains the dominant arrangement.

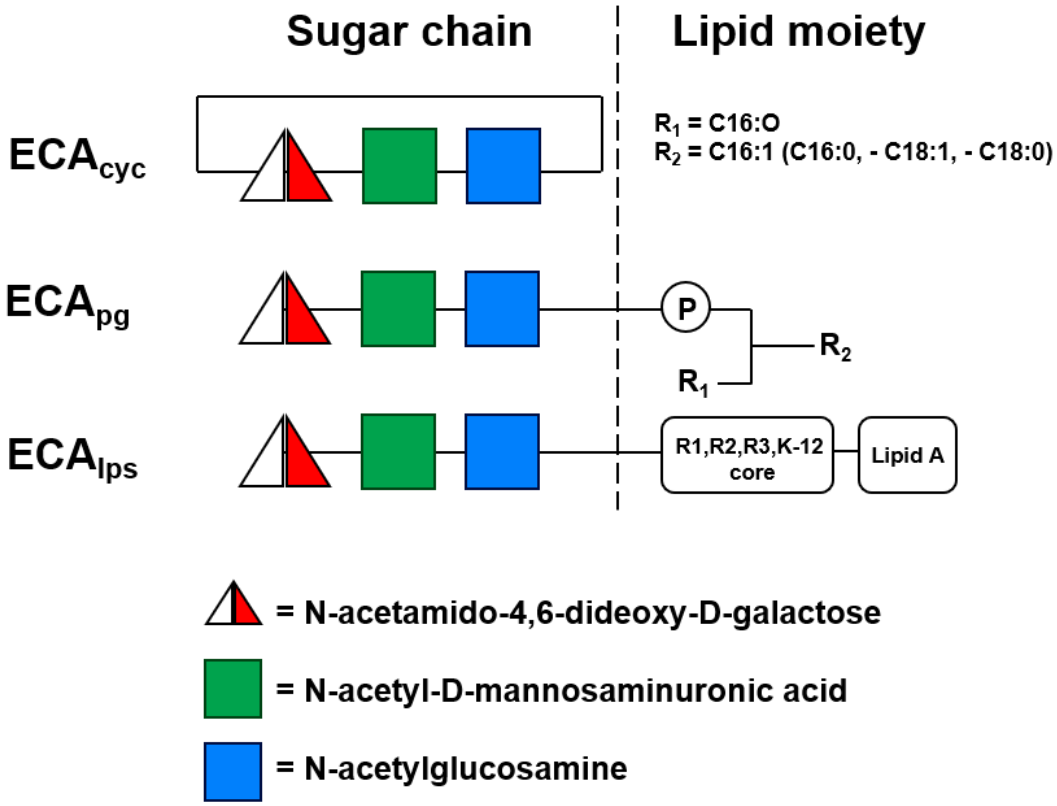


Figure 1.8: Arrangements of ECA.

ECA is arranged in two membrane associated forms; ECA_{pg} where the ECA polysaccharide is linked to phosphatidylglycerol and ECA_{lps} where ECA is linked to the core-sugars of lipid A. A non-membrane associated from, ECA_{cyc}, exists as well but is localized to the periplasm. Figure adapted form Kuhn et al, 1988. Sugar moiety symbols used in accordance to SNFG (Neelamegham et al. 2019).

The presence of ECA_{lps} is O antigen ligase (WaaL) dependent with *waaL* mutants being unable to substitute ECA repeat units onto the Lipid A core-sugars (Barr, Klena & Rick 1999). Recently, it was reported for the first time that in *Y. enterocolitica* O:3, ECA trisaccharide repeat units were found to be directly linked to the LPS core-sugar Kdo (Noszczynska et al. 2015). However, whether this translates to other *Enterobacteriaceae* remains to be explored.

1.4.3 Cyclic ECA, ECA_{cyc}

ECA_{cyc} is localized to the periplasm in a β-glycan cyclic form, where unlike other arrangements of ECA, it does not contain a lipid anchor. Contrary to other forms of ECA, the presence of ECA_{cyc} is dependent on functional WzzE, and, as a protein which would facilitate its circulation has yet to be determined, it has therefore been suggested that its presence is an artefact of membrane bound ECA biosynthesis (Kajimura et al. 2005). The functions of ECA_{cyc} have been heavily investigated recently and it's possible biological functions include a role in the regulation of osmotic pressure (due to similarities in chemical structure to other common osmoregulatory periplasmic glycans (Lee, Cho & Jung 2009), as a homeostasis marker for the induction of the Regulator of Capsule Synthesis (Rcs) stress response (Castelli & Vescovi 2011; Majdalani & Gottesman 2005), and as a microbe-associated molecular pattern (Paunova-Krasteva et al. 2014).

1.4.4 ECA biosynthesis

The process of ECA biosynthesis is illustrated in Figure 1.9. ECA biosynthesis begins on the cytosolic side of the inner membrane (IM) with the transfer of GlcNAc 1-P from the UDP-linked nucleotide sugar UDP-GlcNAc to yield Und-PP-GlcNAc (lipid I) by WecA (Erbel et al. 2003). This is followed by the sequential transfer of ManNAcA and Fuc4NAc from donor UDP-linked nucleotide sugars UDP-ManNAcA and TDP-Fuc4NAc substrates, catalysed by WecG and WecF to yield lipid II and lipid III, respectively. The full-length lipid-linked ECA trisaccharide, lipid III, is then translocated across the IM to the periplasm via WzxE flippase (Kajimura et al. 2006). ECA repeats are then sequentially polymerized together by WzyE polymerase to form the full length ECA chain where the modal, or average chain length, is determined by WzzE co-polymerase. Lastly, mature ECA polysaccharides are most commonly ligated to a phosphatidylglycerol lipid carrier and then translocated to the OM via an unknown mechanism. The 6 position of GlcNAc residues within the repeat trisaccharide are nonstoichimetrially substituted with *O*-acetyl groups by WecH (Kajimura et al. 2006). The process of ECA biosynthesis is illustrated in Figure 1.9.

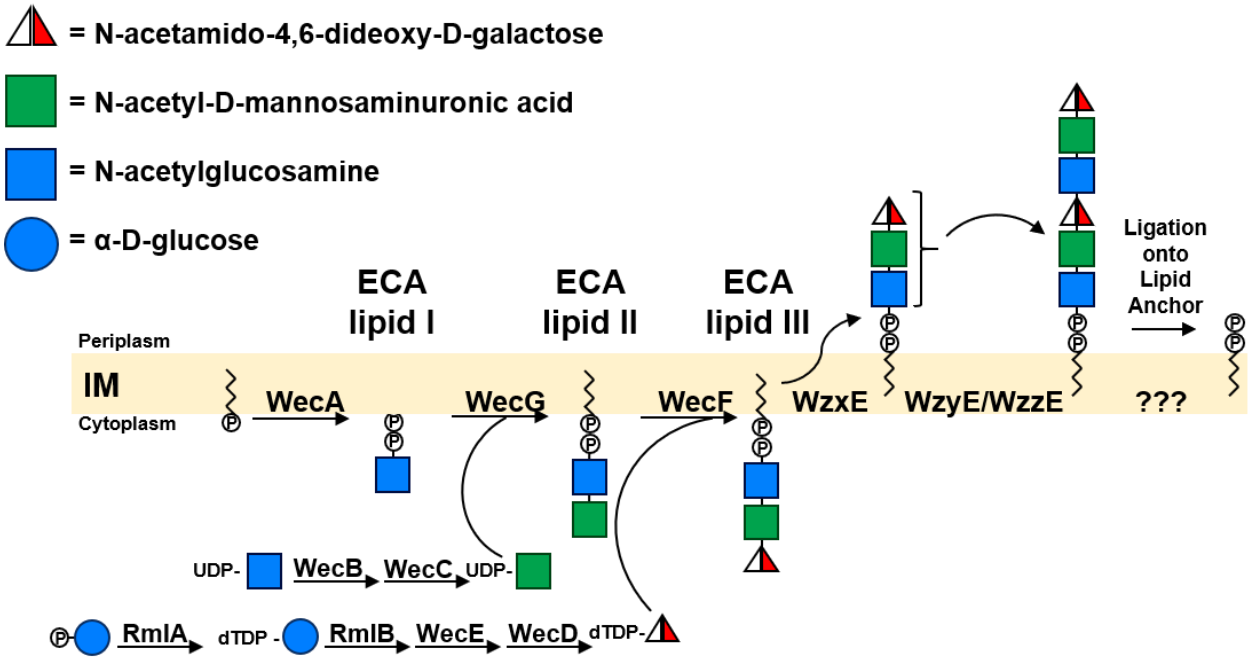


Figure 1.9: ECA biosynthesis in *Enterobacteriaceae.*

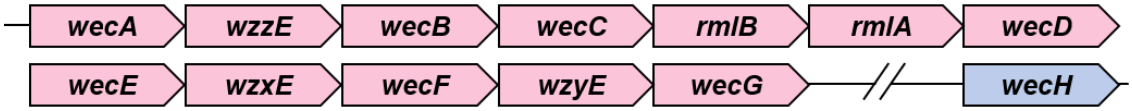
ECA biosynthesis begins on the inner leaflet of the IM where the glycosyltransferase WecA, WecG and WecF sequentially add sugar moieties to Und-P. The complete ECA repeat units is then translocated across the IM by WzxE where it is polymerized into controlled lengths by WzyE-WzzE prior to ligation onto its membrane anchor prior to export to the OM (Eade et al. 2021). ECA_{cyc} does not contain a membrane anchor. Sugar moiety symbols used in accordance to SNFG (Neelamegham et al. 2019).

1.4.5 *wec* operon mutants and the Rcs Pathway

The biosynthetic genes of ECA lie within the *wec* operon, listed in Figure 1.10. The integrity of these genes has been shown to be important in the viability of *Enterobacteriaceae*. In *E. coli* K12, mutations within the genes involved in nucleotide sugar precursor biosynthesis; *wecB*, *wecD*, *wecE*, *rmlB* and RU assembly; *wecG*, *wecF*, *wzxE* lead to the accumulation of undecaprenyl linked ECA-biosynthetic intermediates which lead swelling, morphological abnormalities and to the activation of the Rcs stress response pathway (Jorgenson et al. 2016; Ramos-Morales et al. 2003). Mutations in *E. coli* K12 *wzyE* have been observed to be lethal and *wzzE* mutants in *Serratia marcescens* have also been reported to induce the Rcs pathway (Baba et al. 2006; Castelli et al. 2008; Castelli & Vescovi 2011).

The Rcs pathway is a cellular stress response pathway which detects and responds to envelope stresses which include OM damage, defects in LPS/ECA biosynthesis and peptidoglycan (PG) stresses, and is illustrated in Figure 1.11 (Meng et al. 2020; Tao, Narita & Tokuda 2012). The pathway itself is composed of five proteins which include RcsF, a lipoprotein associated with OMPs in the OM, RcsC and RcsD which act as transmembrane phosphorelay proteins, and RcsA and RcsB which homo- or heterodimerise and act as transcription factors (Meng, Young & Chen 2021). For example, homodimerization of RcsB leads to the activation of *rprA* and *gadA* leading to biofilm and glutamate decarboxylase gene expression (Huesa et al. 2021) whereas, heterodimerization with RcsA leads to *cps* and *flhDC* activation leading to K12 capsule synthesis and motility associated gene expression (Ebel & Trempy 1999).

Many cell wall mutant phenotypes have been associated with pleiotropic effects due to the activation of the Rcs pathway, most commonly in the analysis of ECA biosynthetic mutants as described above. Additionally, due to these pleiotropic effects, non-related pathways have been associated with one-another where, it has been later clarified that the activation of the Rcs pathway was the reason why the two pathways seemingly associated with one another. This is best exemplified with ECA and flagellar synthesis in *S. marcescens* where, mutations in *wecD*, *wzxE* and *wzyE* seemed to prevent flagellar biogenesis however, it was later discovered that the Rcs pathway, which the mutants induced, led to this phenotype (Castelli et al. 2008; Castelli & Vescovi 2011). This was also seen in *rffG* (*rmlB*) mutants preventing the swarmer cell phenotype in *Proteus mirabilis* which likewise was due to the induction of the Rcs pathway (Little, Tipping & Gibbs 2018).



Gene	Activity
wecA	Undecaprenyl-phosphate α-N-acetylglucosaminyl transferase
wzzE	ECA polysaccharide chain length regulator
wecB	UDP-N-acetylglucosamine 2-epimerase
wecC	UDP-N-acetyl-D-mannosamine dehydrogenase
rmlD	dTDP-glucose 4,6-dehydratase
rmlA	dTDP-glucose pyrophosphorylase
wecD	dTDP-4-amino-4,6-dideoxy-D-galactose acyltrasferase
wecE	dTDP-4-dehydro-6-deoxy-D-glucose transaminase
wzxE	ECA lipid-III flippase
wecF	4-acetamido-4,6-dideoxy-D-galactose transferase
wzyE	Wzy protein involved in ECA polysaccharide elongation
wecG	UDP-N-acetyl-D-mannosaminuronic acid transferase
wecH	O-acetyltransferase

Figure 1.10: Genes of the *wec* operon.

Genes of the *wec* operon listed in chronological order as they are arranged on the chromosome. Genes are not drawn to scale.

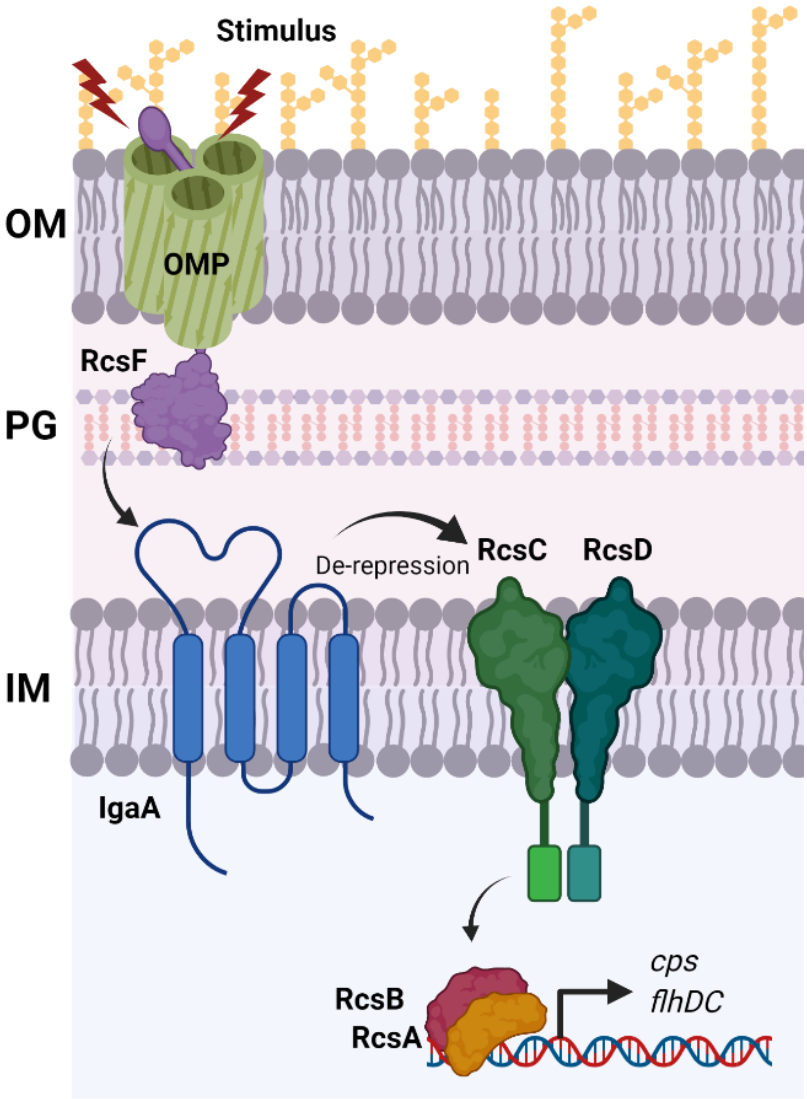


Figure 1.11: Rcs stress response pathway in *E. coli* K12.

The Rcs pathway, under homeostasis, monitors the integrity of the cell wall. RcsF, resent in OMPs, can translocate to IgaA when activated which prevents IgaA from repressing RcsC. RcsC and RcsD activation leads to the recruitment of RcsB which can either homo- or heterodimerise with RcsA or other transcription factors. Once in a dimer, RcsB can induce gene expression in response to the stimulus of RcsF (Meng, Young & Chen 2021).

1.4.6 *wec* operon cross-complementation

Of the genes in the *wec* operon, two genes from the nucleotide sugar biosynthetic portion of the pathway (*wecB, wecC, wecD, wecE, rmlA, rmlB*) in addition to *wecA* are involved in or have homologs in O antigen biosynthesis in *E. coli* and *S. flexneri*, they being *rmlA* and *rmlB.* Both *rmlA* and *rmlB* have functional homologs involved in the biosynthesis of dTDP-L-Rhamnose in O antigen biosynthesis (Marolda & Valvano 1995). It was shown that in *E. coli* K12 *rmlA* and *rmlB* mutants from the *rfb* cluster involved in dTDP-L-Rhamnose biosynthesis, were able to produce LPS, observed via LPS gels, where this was attributed to the unmodified homologs present within the *wec* operon (Phan et al. 2013). WecA is the first glycosyltransferase involved in the ECA and O antigen biosynthetic pathway through the synthesis of Und-PP-GlcNAc which commits Und-P towards OM polysaccharide biosynthesis.

Cross complementation is also observed with the key proteins of the Wzy-dependent pathway due to structural and substrate similarities. It was shown by Marolda et al. (2006) that a *wzxB* mutant could be complemented by WzxE however only in a *wzzE wzyE* double mutant background (Marolda et al. 2006). Likewise this was observed for *wzzE/wzzB* mutants and WzzE/WzzB proteins by Leo et al. (2020) who found that WzzB could complement a *wzzE* mutant, and *vice versa* either as the restoration of modal length control or an increase in abundance of OM polysaccharides.

1.4.7 Bacteriophage N4 and ECA

In the past, ECA has been linked to be the glycan receptor for the N4 bacteriophage belonging to the *Schitoviridae* family which is known to infect *E. coli* species (Kiino et al. 1993). Recently a publication by Sellner et al. (2021) clarified this and showed that it was not ECA which was the receptor but a novel glycan named NGR (N4 Glycan Receptor) which was partially biosynthesised by WecB. It was shown that the protein PdeL, a phosphodiesterase, protected *E. coli* cells from N4 mediated killing by binding to the *wec* operon, specifically *wecB* where it acted as a repressor. WecB had previously been implemented in affecting the progression of P22 bacteriophage infection in *S. enterica* (Bohm et al. 2018). To determine whether or not the lack of ECA due to the *wecB* suppression caused the protection, they examined N4 phage mediated killing on cells lacking ECA and found that they were more resistant to killing. However, when only removing *wecA*, despite lacking ECA the strains were sensitive to N4 killing which was likewise observed in an ECA KO strain supplemented with ectopic *wecB* expression (Junkermeier & Hengge 2021; Sellner et al. 2021). Ultimately, Sellner et al suggested that WecB is involved in the biosynthesis of a N-acetylmannosamine based glycan receptor, NGR, which is common receptor

for other related *E. coli* targeting phages, *Vequintavirinae* and *Enquatrovirus*, which showed reduced killing in *wecB* mutants (Maffei et al. 2021; Sellner et al. 2021). This represent the first *wec* operon gene, excluding *wecA*, which is involved in an alternative biosynthetic process other than ECA biosynthesis under homeostatic conditions.

1.4.8 Biotechnological applications of *wec* operon mutants

Besides investigating the roles of ECA and pleotropic phenotypes of *wec* mutants, some mutants of the *wec* operon have been used in biotechnological applications. This is well exemplified for *wecA*, which has been used in the vaccine development against *S. enterica* due to the fact that *wecA* mutants are attenuated upon oral challenge, presumably due to the lack of ECA (Bridge et al. 2015; Gilbreath et al. 2012). It is important to note that in *S. enterica*, WbaP biosynthesises the O antigen lipid-I, not WecA like in *E. coli* strains or *S. flexneri* species subsequently, hence *wecA* mutants do not disrupt O antigen biosynthesis in *S. enterica*. This has been taken further with a study investigating if *wec* mutants could be used as a heterologous antigen vehicle to deliver antigen. In the study by Liu et al. (2020), it was shown that *S. enterica wecA* mutant strains expressing *Streptococcus pneumoniae* antigen PspA did indeed lead to the generation of anti-PspA IgG in mice (Liu et al. 2020). This was additionally attempted with *Pseudomonas aeruginosa* O11 O antigen expressed from a *S. enterica wecA* mutant which likewise produced IgG antibodies which were shown to protect against lethal challenge (Bridge et al. 2016). This highlights the possibilities of using *wec* mutations to at attenuate strains of Enterobacterial pathogens if the first committed step to O antigen biosynthesis is not shared with ECA biosynthesis.

In addition to vaccine development, *wec* operon mutants have been investigated for potential roles in industrial processes. A common practice in these studies is the deletion of the whole *wec* operon and sometimes additional biosynthetic pathways to improve cell growth or to replace with other biosynthetic processes. In one such study, the removal of the *wec* operon and genes associated with flagellar biosynthesis in *E. coli* MG1655 demonstrated that the productivity and efficiency of the cells could be improved leading to potential industrial benefits (Qiao et al. 2021). Furthermore the deletion of the *wec* operon, O antigen and colonic acid biosynthetic genes in *E. coli* MG1655 enabled the researchers to replace the O antigen biosynthetic genes with genes with encoded O antigen from *Bordetella pertussis* (Wang et al. 2021). Lastly, the *wec* operon was removed in *E. coli* MG1655 and replaced with a ~9-10 kb segment of DNA encoding *Campylobacter jejuni* genes required for asparagine-linked protein glycosylation (Yates et al. 2019). The researchers showed that this resulted in increased N-glycosylation of proteins efficiency, higher glycoprotein titres and better growth phenotypes (Yates et al. 2019).

1.4.9 Roles of ECA - Pathogenesis

The role of ECA in pathogenesis is difficult to clearly define due to; 1) the nature of how ECA is investigated in these systems, and 2) the pleiotropic phenotypes presented by ECA biosynthetic mutants due to induction of stress pathways. In the majority of the publications which link ECA to pathogenesis, ECA biosynthetic mutants are used throughout the publication. An example of this is Ramos-Morales et al. (2003) who linked ECA's importance in *S. enterica* serovar Typhimurium mouse infection through the use of *wecA* and *wecD* mutants. The study showed that *wecD* mutants were sensitive to deoxycholate (DOC), where this was partially reduced in *wecD wecA* double mutants, linking this interaction with bile salt resistance suggesting that ECA played a role in DOC resistance but that the *wecD* mutant was affected by pleiotropic phenotypes (Figure 1.9), (Ramos-Morales et al. 2003). ECA's role in bile resistance was similarity demonstrated again with *S. enterica wecD* and *wecA* mutants where it was shown that they mutants were 40x more sensitive to bile than wildtype cells (Maire et al. 2004) and in *S. enterica* with *wecB* mutants (May & Groisman 2013). Thus, ECA's role in providing resistance to bile salts is well documented in the literature.

ECA was also implicated in promoting the survival of *Yersina pestis* cells during murine macrophage challenge. The genes *wecB* and *wecC* were identified via transposon insertion mutagenesis under negative selection and were found to decrease the viability of cells engulfed by macrophages (Klein et al. 2012). In addition to these examples, ECA has also been partially tied in part to the pathogenesis of *S. marcescens* and *P. mirabilis* however, upon further investigation it was found that the Rcs pathway induction was the cause of the linkage between ECA and pathogenesis as mention above.

Other potential roles of ECA providing resistance to cationic anti-microbial peptide (cAMP) have not been explored however, due to the negative charge nature of the polysaccharide it is unlikely that ECA can perform a protective role.

1.4.10 Roles of ECA – Self regulation

Cyclic ECA has also been implemented in the regulation of ECA_{pg} biosynthesis through unknown interactions with the protein ElyC (Rai et al. 2021). This was displayed in *E. coli* MG1655 through using complicated ECA mutants *ΔwzzE ΔwaaL* to isolate strains producing only ECA_{pg}. It was shown that the deletion of *elyC* seemed to make more surface ECA in the *ΔwzzE* background observed through anti-ECA dot blots and anti-ECA Western immunoblotting (Rai et al. 2021). Subsequently deletion of *elyC* also seemed to increase the amount of ECA_{lps} however the ectopic expression of WaaL did not produce an increase. Due to these observations it seemed

that due to lack of ECA_{cyc}, which is WzzE dependent, that ECA_{pg} biosynthesis was unregulated in the *elyC* mutant background, linking that ECA_{cyc} regulated ECA_{pg} biosynthesis mediated through ElyC.

However, the study did not correctly address the potential pleotropic effects of *elyC* mutants which are involved in PG lipid-II metabolism (Paradis-Bleau et al. 2014) and are hence very likely to lead to pleotropic mutant phenotypes. Multiple observations throughout the study attributing ECA_{pg} regulation to ECA_{cyc} and ElyC can be equally explained through the potential impacts of pleotropic phenotypes from the *elyC* mutant itself. There may further be a link between ElyC and ECA biosynthesis however, the current work performed is insufficient to make the claims it has without properly investigating the roles of known phenotypes in cell wall biosynthesis mutants using appropriate methods.

1.4.11 Roles of ECA – Homeostasis

Although the integrity of the *wec* operon is essential for proper cellular homeostasis, the full biological significance of ECA is not fully understood. The most well-known function of ECA is its ability to confer resistance to bile salts which was shown in *S. enterica* (Ramos-Morales et al. 2003). This is presumably through repulsive electrostatic interactions between the negatively charged polysaccharide chain of the membrane associated forms of ECA, ECA_{pg} and $ECA_{lps,}$ and the amphipathic bile salt molecules, as *wecD* and *wecA* mutants are sensitive to deoxycholate, one of the most common bile salts (Ramos-Morales et al. 2003). ECA has also been shown to play roles in the resistance to SDS and EDTA where *wecA* mutants in *E. coli K12* were shown to be sensitive (Mitchell, Srikumar & Silhavy 2018). Additionally, as all *wec* mutants induce the Rcs pathway, including *wecA* mutants, it suggests that the presence of ECA itself is important to the cell (Castelli & Vescovi 2011; Jorgenson et al. 2016).

The role of cyclic ECA, ECA_{cyc}, perhaps is the best investigated ECA form in relation to it's role in homeostasis. Recently, ECA_{cyc} been implicated in maintaining the permeability barrier of the OM in a YhdP dependent manner (Mitchell, Srikumar & Silhavy 2018). It was shown that *yhdP* mutants, which are sensitive to vancomycin and SDS which suggests OM permeability defects, could be made resistant with the deletion of *wecA* in *E. coli* K12. Upon further investigating, including the role of Und-P usage and PG, it was found that deletion of *wzzE*, the protein critical of ECA_{cyc} production albeit for an unknown reason, restored the OM permeability barrier of the *ydhP* mutant (Mitchell, Srikumar & Silhavy 2018). Subsequently, ECA_{cyc} was identified to be the cause of the sensitivity in the *ydhP* mutant and it was suggested that YdhP interacts with ECA_{cyc} to maintain the OM barrier. The rationale behind this was that ECA_{cyc} may

assist in the shuttling of specific molecules to and from the OM to the periplasm in a similar way as other cyclodextrins do (Mitchell, Srikumar & Silhavy 2018).

1.4.12 Roles of ECA - Biosynthetic intermediates

The role of biosynthetic intermediates in the phenotypes of ECA mutants have often been directed in their negative roles of impacting other cell wall pathways due to the sequestering of Und-P (Jorgenson et al. 2016; Jorgenson & Young 2016). This, while true, does not highlight the potential unforeseen benefits of the accumulation of lipid intermediates in the cell.

The accumulation biosynthetic intermediates of ECA have been show to restore the barrier function of the IM in *E. coli* K12 *tol-pal* mutants. It was shown in *wecC*, *wecG* and *wecE* mutants, which accumulate ECA lipid-I and ECA lipid-II respectively (Figure 1.9), that the accumulation of the ECA biosynthetic intermediates was able to partially restore vancomycin resistance in *tolA* mutants (Jiang et al. 2020). This was also reported, but not discussed, when identifying suppressor mutations in the *ydhP* mutant background to restore OM permeability to vancomycin and SDS. In addition to mutations in *wecA*, mutations in *wecC* and *wecF*, which accumulate ECA lipid-I and ECA lipid-II respectively, also restored the OM barrier (Mitchell, Srikumar & Silhavy 2018).

This has also been observed in *wecE* and *wzxE* mutants which gained resistance to gentamycin, and nalidixic and amikacin, respectively, and this was contributed to the fact that they accumulated ECA lipid-II and ECA lipid-III (Girgis, Hottes & Tavazoie 2009; Tamae et al. 2008). The mechanisms as to why the accumulation of ECA biosynthetic intermediates provide OM barrier resistance remains unknown and requires further investigation.

1.4.13 ECA based vaccine

Due to its abundance amongst all *Enterobacteriaceae*, ECA has also been investigated for its potential as a glyco-conjugate vaccine. Liu, et al. (2015) successfully conjugated synthetic tri- and hexassacharide derivatives of ECA onto bovine serum albumin and showed via ELISA that the conjugates could be detected by anti-ECA monoclonal antibodies (Liu et al. 2015). However, it was shown that passive immunization with human monoclonal ECA antibodies produced no protective effect during sepsis caused by *Enterobacteriales* (Albertson et al. 2003). Additionally, only passive protection was observed in mice when passively immunised with rabbit serum from rabbits inoculated with *E. coli* O14 (Valtonen et al. 1976). This suggests that the effectivity of an ECA vaccine may be short lived and non-protective long term.

1.5 Lipopolysaccharide

Lipopolysaccharide (LPS) is an amphiphilic glycolipid which is found on the OM of Gram negative bacteria, where it contributes to approximately ~75% of the total OM lipid population (Le Brun et al. 2013). LPS comprises of three domains; 1) lipid A, the proximal, hydrophobic anchor (also known as endotoxin), 2) non-repeating, core-oligosaccharides which can be further broken down into the inner and outer core sugars, and 3) O antigen chains, the distal, highly variable oligosaccharide repeat domain (Kalynych, Morona & Cygler 2014). The O antigen (Oag) is composed of repeating sequences of three to six sugar residues; O antigen RUs are linked to the lipid A via the core (Raetz & Whitfield 2002). The complete LPS structure with full length Oag chains is termed 'smooth' (S-LPS), while LPS devoid of Oag is termed 'rough' (R-LPS), and LPS with a single O unit is termed 'semi-rough' (SR-LPS) as illustrated in Figure 1.12 (Kalynych, Morona & Cygler 2014; Nath & Morona 2015b).

LPS is acknowledged as a major virulence factor for Gram-negative bacteria as its domains are known to play separate key roles in the resistance of host immune mechanisms, toxins as well as allowing the cells to persist in harsh conditions. These roles will be discussed below.

1.5.1 Lipid A

Lipid A is the proximal hydrophobic segment of LPS which anchors the glycolipid to the OM. The glucosamine-based lipid itself makes up the outer leaflet of the OM with ~10^6 present with a cell (Raetz & Whitfield 2002) and provides a negative charge to the OM. In *S. flexneri* fatty acetylation occurs by four (R)-3-hydroxy fatty acids at positions O-2, O-3, O-2' and O-3 (Steimleet al. 2016). Lipid A is biosynthesised by the nine-step Raetz pathway which is highly conserved and uses of the Lpx proteins.

Lipid A biosynthesis begins on the cytoplasmic side of the IM where briefly, LpxA/LpxC which adds a β-hydroxymyristoyl chain to UDP-GlcNAc (Anderson & Raetz 1987). The addition of a second β-hydroxymyristoyl chain by LpxD forms UDP-2,3-diacyl-GlcN which is then cleaved by LpxH, to form 2,3-diacyl-GlcN-1-phosphate also known as lipid X (Babinski, Ribeiro & Raetz 2002). LpxB then joins UDP-2,3-diacyl-GlcN with lipid X. Following this, LpxK phosphorylation leads to the formation of lipid IV_A where post addition of the Kdo sugars, LpxL and LpxM complete lipid A biosynthesis (Brozek & Raetz 1990). This process is illustrated in Figure 1.13.

Lipid A structures differ among bacteria with the substitution of different sugar moieties as well as the number, length and location of the acyl chains (Steimle et al. 2016). Of note are modifications to which alter the charge of lipid A such as the addition of 4-amino-4-deoxy-L-arabinose by PrmA which neutralizes the negative charge of the 4' phosphate

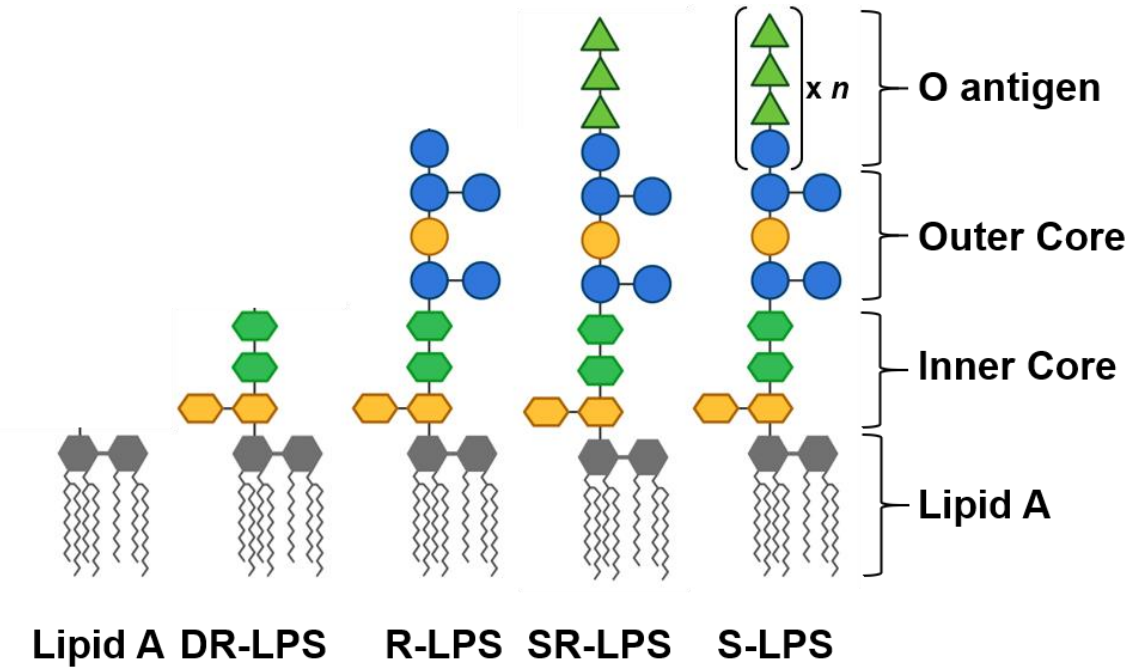


Figure 1.12: Lipopolysaccharide structures in *S. flexneri.*

Lipopolysaccharide is composed of three separate domains; Lipid A core, core sugars (which are divided into inner and outer sugars) and O antigen. Full length LPS with multiple RUs of O antigen is called 'smooth LPS' (S-LPS), LPS with a single O antigen RU is called 'semi rough LPS' (SR-LPS), LPS devoid of O antigen is called 'rough LPS' (R-LPS), LPS devoid of outer core sugars is called 'deep rough LPS' (DR-LPS) and LPS devoid of core sugars is referred to as Lipid A (Kalynych, Morona & Cygler 2014; Nath & Morona 2015b). Sugar moiety symbols used in accordance to SNFG (Neelamegham et al. 2019).

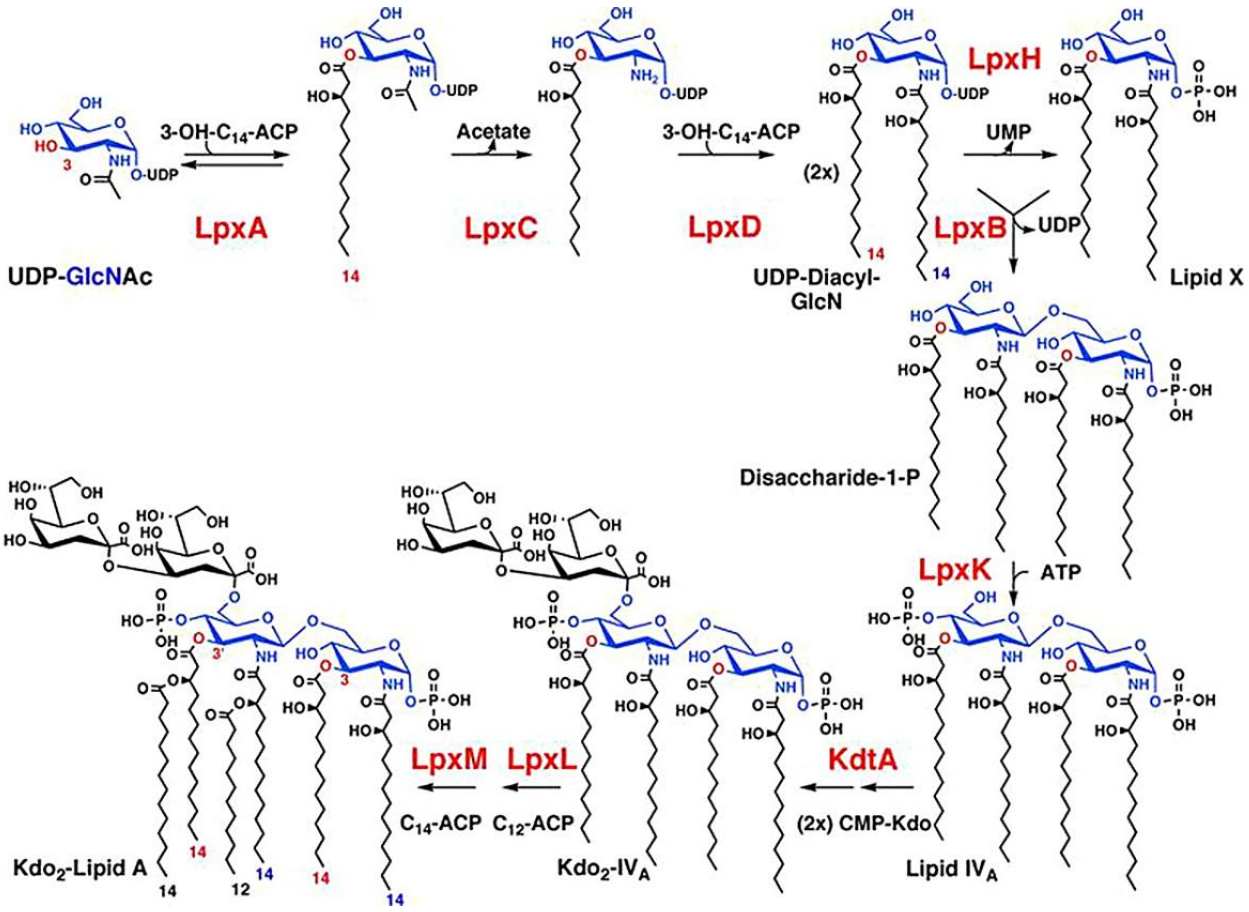


Figure 1.13: Lipid A biosynthesis by the Raetz pathway.

The Raetz pathway is a conserved bacterial biosynthetic pathway which biosynthesises Lipid A which makes up the majority of the extior leaflet of the outer membrane. The nine enzymes are shown in red and their products are labelled beneath their structures. Abbreviations: CMP, cytidine monophosphate; UMP, uridine monophosphate; UDP-GlcNAc, uridine diphosphate N-acetylglucosamine. Figure adapted from Whitfield & Trent, (2014).

groups, reducing the susceptibility of *S. enterica* to cationic antimicrobial peptides (Trent et al. 2001). Lipid A, also known as endotoxin, is pathogen associated molecular pattern (PAMP) and is a highly immunostimulatory molecule recognised by host pattern recognition receptors (PPRs). Lipid A is recognised by host TLR4 receptors (Chilton, Embry & Mitchell 2012) which leads to the activation type 1 interferons and the transcription factor NF_KB which subsequently leads to the activation of the immune system (Heumann & Roger 2002).

1.5.2 Core sugars

The LPS core sugars are the sugars which link O antigen to the Lipid A and can be sub-divided into the inner and outer core sugars. Different core structures have been described from across Gram negatives with five major cores structures present in *E. coli*: K12, R1, R2, R3, R4 of which R1, R3 and R4 are present in *Shigella* (Knirel et al. 2011).

The core sugars are assembled on top of lipid A while the molecule remains on the cytoplasmic leaflet of the IM where core sugar biosynthesis begins with the addition of two Kdo residues by WaaA onto the glucosamines of Lipid A. WaaC and WaaF then extend the inner core with the sequential addition of two heptose residues (Gronow, Brabetz & Brade 2000). Lastly WaaP phosphorylates the first heptose moiety followed by WaaQ, which transfers an additional heptose moiety to the second heptose of the inner core which is then phosphorylated by WaaY and completes the inner core (Yethon et al. 1998). Outer core synthesis begins with the addition of glucose to the second heptose by WaaG, which is then acted upon by WaaI and WaaJ which subsequently add glucose and galactose groups. The final steps of outer core synthesis involves the addition of a glucose by WaaJ, to which O antigen is ligated upon, followed by the another glucose group is added by WaaD (Whitfield, Kaniuk & Frirdich 2003). This process is illustrated in Figure 1.14. At this stage the complete lipid A and core-sugars structure is translocated across the inner membrane by MsbB, a flippase belonging to the ABC transporter family, where the macromolecule can accept O antigen and prior to export to the OM (Zhou et al. 1998).

The core-sugars can also be modified where, these modifications often occur on the outer core sugars. Some of these modifications include the addition phosphoethanolamine (PEtN), or rhamnose groups, or an additional Kdo moieties, to the second Kdo sugar by EptB, WaaS or WaaZ, respectively where, for example, the addition of PEtN provides great resistance to polymyxin B in *E. coli* (Klein et al. 2015). Inner core sugar modifications include the addition of PEtN to the first heptose sugar of the inner core by EptC (Klein, Gracjana et al. 2013) or the addition of different glucose derivatives to the third heptose residue of the inner core by WaaH.

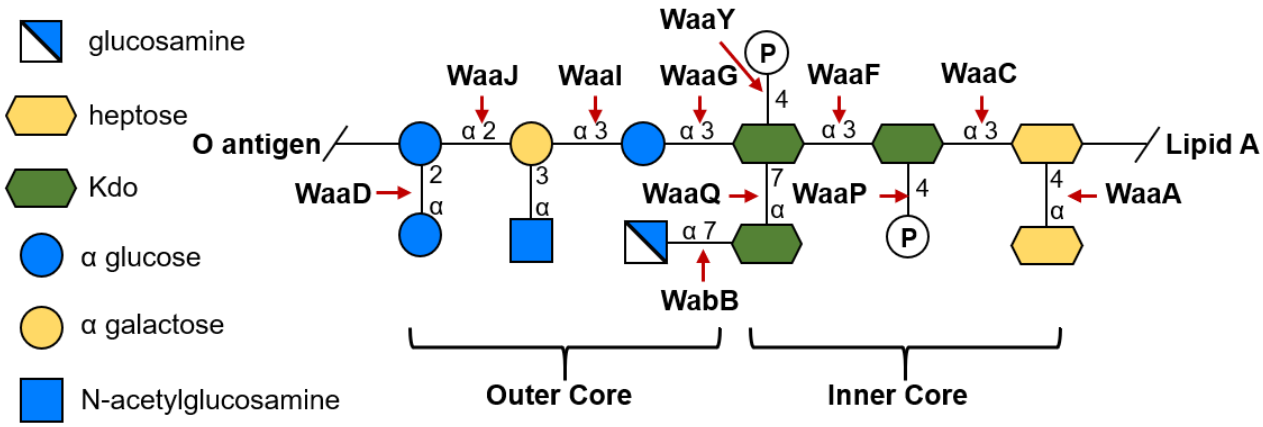


Figure 1.14: LPS R3 Core-sugar composition and biosynthetic proteins.

Schematic of the bonds, sugar moieties and the enzymes responsible for R3 LPS core biosynthesis in *S. flexneri*. Sugar moiety symbols used in accordance to SNFG (Neelamegham et al. 2019).

1.5.3 O antigen

O antigen is the distal polysaccharide portion of LPS and is composed entirely of sugar RUs. The composition of the polysaccharide differs between species as well as serotypes where 50 different O antigen serotypes have been described in *Shigella* alone with at least 19 serotypes in *S. flexneri* [1a, 1b, 1c/7a, 1d, 2a, 2b, 3a, 3b, 4a, 4av, 4b, 5a, 5b, 6, 7b, X, Xv, Y and Yv] (Liu et al. 2014; Muthuirulandi et al. 2017; Perepelov et al. 2012; Stenutz, Weintraub & Widmalm 2006).

In *S. flexneri*, O antigen biosynthesis is performed in part by the Wzy-dependent pathway where biosynthesis begins on the cytoplasmic leaflet of the IM where the glycosyltransferase WecA transferred UDP-GlcNAc onto Und-P forming Und-PP-GlcNAc, Oag lipid-I. the sequential additions of L-Rhamnose by RfbF and RfbG lead to the formation of Oag lipid-II/-III and -IV where, at which point, the Oag RU is complete (Macpherson, Manning & Morona 1994; Morona et al. 1994). The Oag RU is then translocated across the IM by WzxB where it is polymerized into a linear polysaccharide by WzyB whose activity is controlled by WzzB. The completed polysaccharide chain is then assembled on top of the distal heptose of the core-sugars by the glycosyltransferase WaaL and in doing so completes the mature LPS molecule which is then exported to the OM by the Lpt system (Islam & Lam 2014).c This biosynthetic process is illustrated in Figure 1.15.

In *S. flexneri* Y, Oag modifications are typically generated by prophage originating genetic sequences (Chung et al. 2016) which include the *gtr* locus and the genes *oacB* and *oacD* (Teh et al. 2020). As the structure of *S. flenxeri* Y Oag is GlcNAc-Rha^{I}-Rha^{II}-Rha^{III}, the proteins of the *gtr* locus are responsible for the addition of glucose where; GtrI and GtrIV glucosylate GlcNAc, GtrII glucosylates L-Rha^{I}, GtrV glucosylates L-Rha^{II} and GtrX glucosylates L-Rha^{III} (Mavris, Manning & Morona 1997; Teh et al. 2020). OacB is responsible for the O-acetylation of L-Rha^{III} where as OacD is responsible for the O-acetylation of GlcNAc (Sun et al. 2014). The full extent of known *S. flexneri* serotypes are displayed in Figure 1.16.

1.5.4 O antigen function

O antigen is involved in the pathogenesis, virulence and persistence of multiple Gram negative pathogens. The presence of O antigen itself provides the bacterium resistance to low pH and toxins (Martinić et al. 2011; Tran, Papadopoulos & Morona 2014), as well as the ability to resist attack from host complement proteins (Hong & Payne 1997; Pluschke et al. 1983). Serotype switching enables bacteria to evade host immune responses, rendering specific antibodies useless (Reeves 1995).

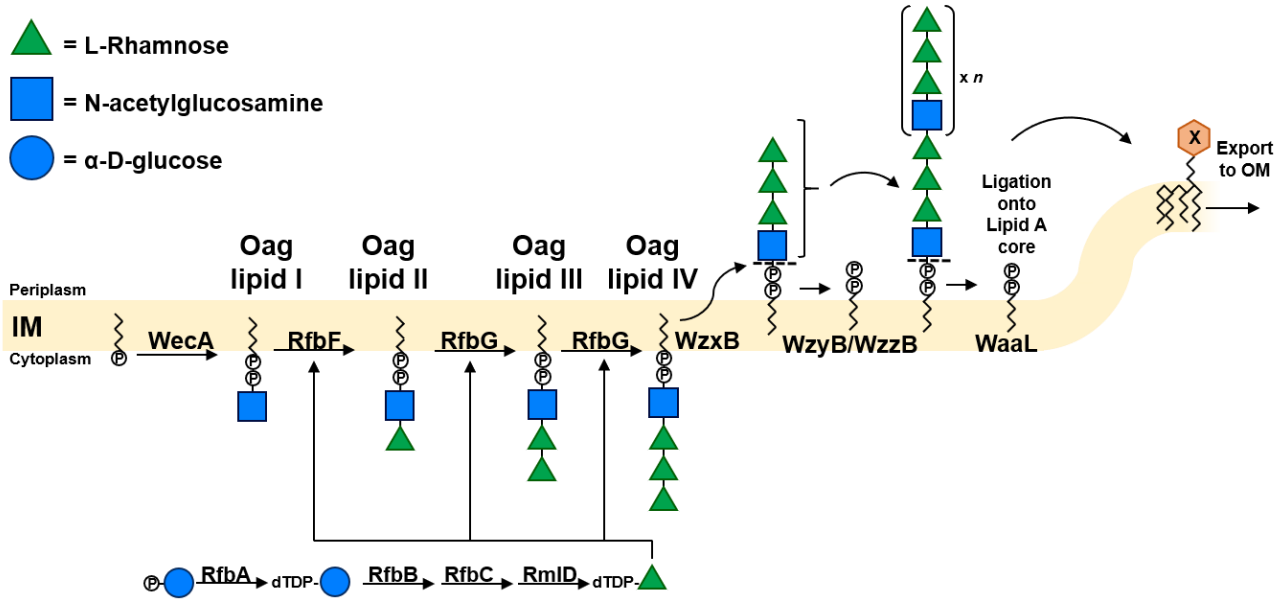


Figure 1.15: LPS O antigen biosynthesis in *S. flexneri* Y.

Biosynthesis of O antigen begins on the inner leaflet of the IM where the glycosltransferases WecA, RfbF and RffG add sugar moieties to undecaprenyl to complete repeat units consisting of GlcNAc-Rha^{I}-Rha^{II}-Rha^{III}. Once completed, the repeat unit is translocated across the IM by WzxB where it is acted upon by Wzy and Wzz to polymerize O antigen polysaccharide chains at specific lengths. Once the chain is completed it is transferred onto the outer core sugars of a complete Lipid A – core molecule by WaaL at which point the mature LPS molecule exported to the OM by the Lpt system. Sugar moiety symbols used in accordance to SNFG (Neelamegham et al. 2019).

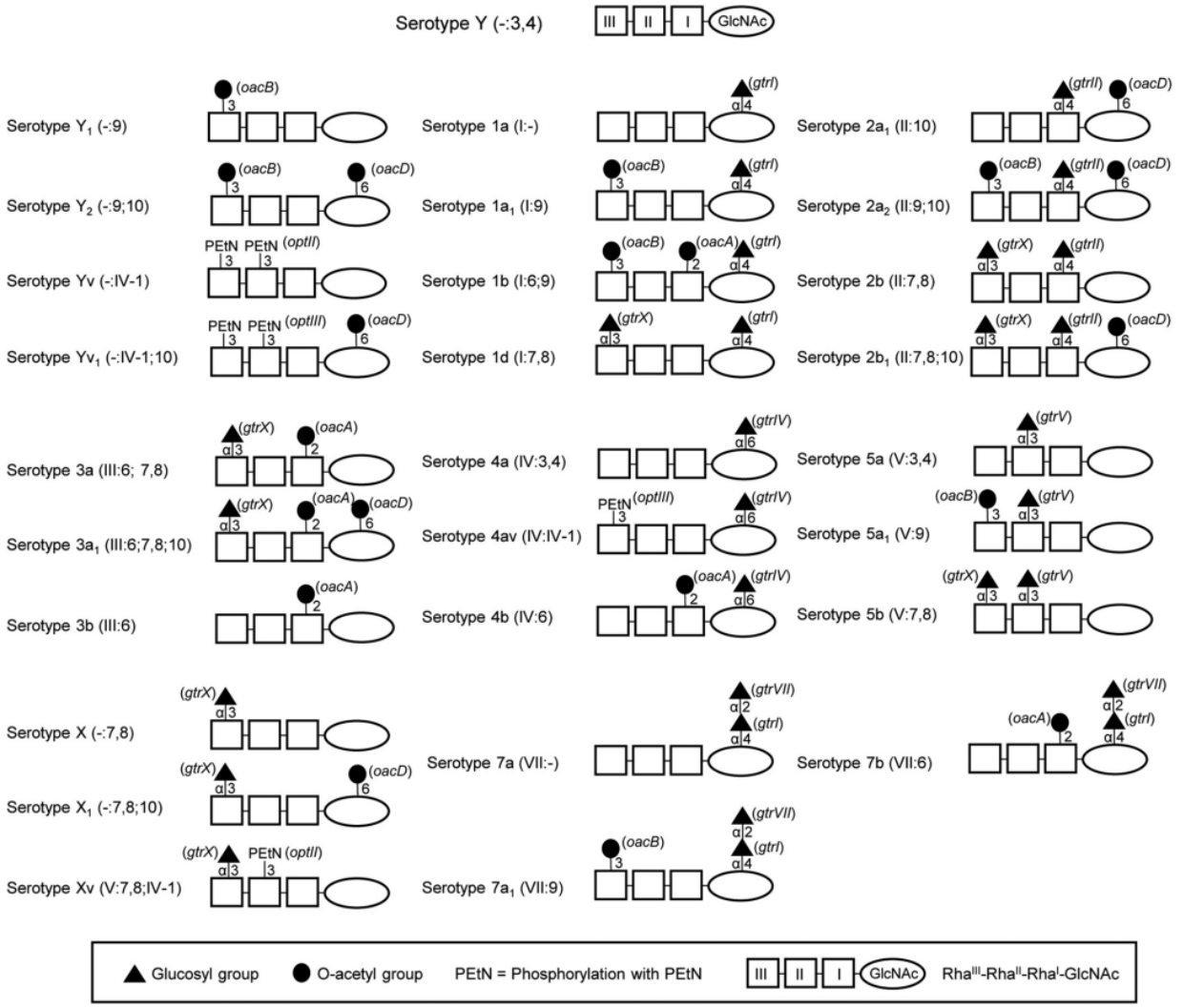


Figure 1.16: Serotypes of *S. flexneri*.

Graphical illustration of the structures of know S. flexneri O antigen serotypes. The basic, non-modified repeat unit of S. flexneri is GlcNAc-Rha^{I}-Rha^{II}-Rha^{III} which is referred to as serotype Y. Modifications by the addition of glycosyl groups, O-acetyl groups or phosphorylation with phosphoethanolamine (PEtN) provide the various serotypes observed. Additional complexity occurs from the use of different bonds used to modify the sugar residues as well as the number of modifications. Figure adopted from Teh, 2020.

O antigen is also used explicitly in key pathogenic niches. For example, in *S. flexneri* O antigen has been shown to act as an adhesin by interacting with host glycans, and in doing so, helps the bacterium invade host Goblet Cells during infection (Tran et al. 2020). O antigen is also known to play key roles in directing ABM as masking of IcsA by O antigen prevents the formation of F-actin comet tails (Morona & Van Den Bosch 2003).

1.5.5 Lpt System

The Lipopolysaccharide transport (Lpt) system is a system for the translocation of full length LPS molecules from the IM, across the periplasm to the OM. The system shields the acyl chains of the LPS molecule from the aqueous environment of the periplasm as the molecule is transported, and comprises of seven proteins which span the cell wall consisting of both aqueous and membrane bound proteins (Silhavy, Kahne & Walker 2010).

The transportation of the LPS molecule begins with the formation of the ABC transporter consisting of LptB dimers together with LptF and LptG named $LptB_2FG$, which utilizes cytoplasmic ATP to power the transport of LPS molecules (Sherman et al. 2014). The exact mechanism of how LPS molecules are extracted from the IM remains unknown. However, some study suggest that the transmembrane domains of LptF and LptG may facilitate this extraction from the IM as they are predicted to form a central cavity within $LptB_2FG$ (Luo et al. 2017). The LPS molecule is then bound by the periplasmic domain of LpxC, which consists of multiple antiparallel β-strands that possess a hydrophobic groove used to shield the acyl chains of the LPS molecule from the aqueous environment (Tran et al 2010). LpxC, LpxA and LpxD form the trans bilayer bridge where, it has been suggested that multiple LpxA oligomers may be incorporated into the bridge complex depending on the width of the periplasm (Suits et al. 2008).

LpxD and LpxE are the final proteins in the system and are involved with the insertion of the LPS molecule into the outer leaflet of the OM by forming the OM translocon (Qiao et al. 2014). LpxD is a large membrane imbedded β-barrel protein whereas, LpxE is a lipoprotein which is present within the central cavity of LpxD where its role in LPS translocation is unclear. It is proposed that LPS molecule enter the periplasmic domain of LpxD where the LPS acyl chains are deposited directly into the OM through gaps within the β-barrel structure and the hydrophilic portion of the LPS molecule passes through the central opening (Qiao et al. 2014). The complete process of LPS export via the Lpt system is described in Figure 1.17.

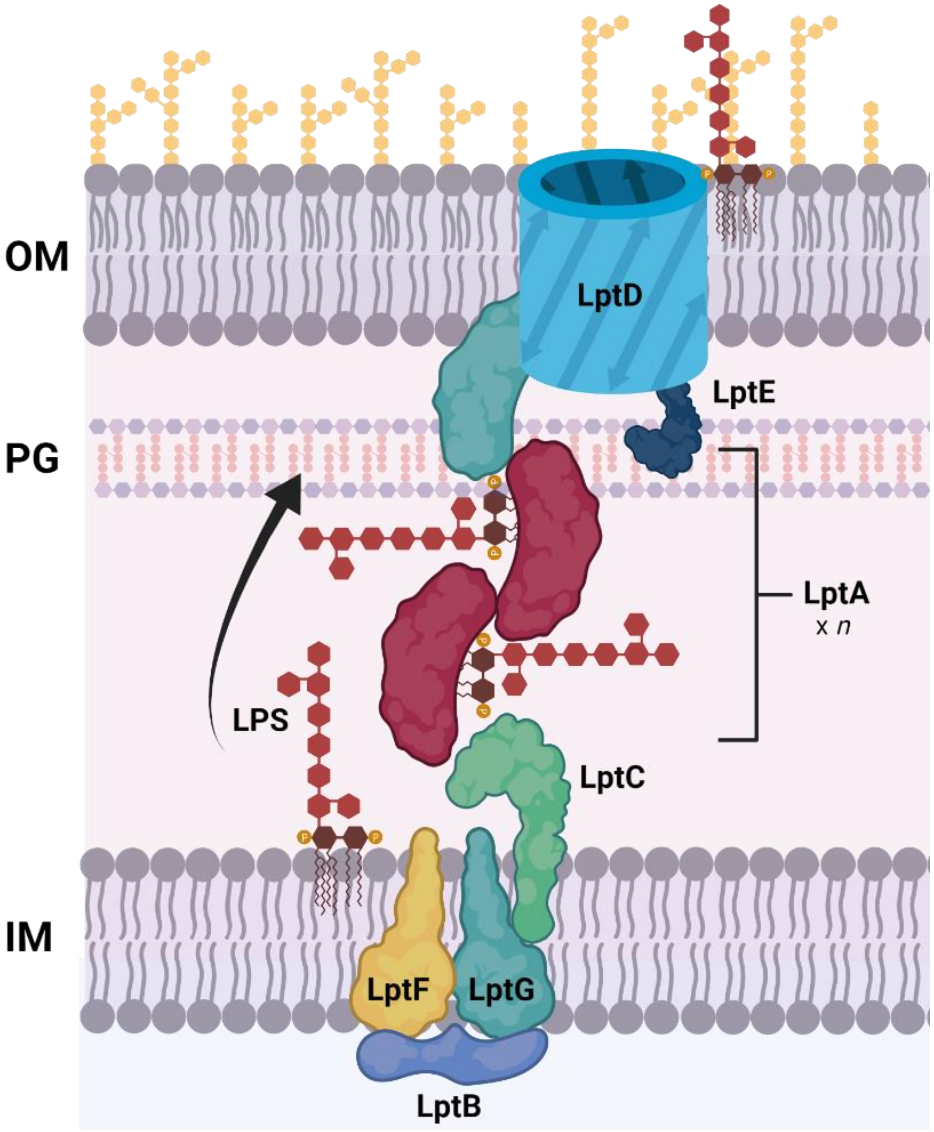


Figure 1.17: Lipopolysaccharide transport in *S. flexneri* by the Lpt system.

Upon ligation of O antigen on top of the core-sugars, the mature LPS molecule enters the $LptB_2FG$ translocon where it is bound by LptC. Using its hydrophobic groves, LptC is able to bind LPS and shield it's acyl chains from the periplasm. The LPS molecule passed through LptA which binds the molecule in a similar way prior to insertion into the OM by LptD where the acyl chains pass through gaps in LptD and the hydrophilic O antigen exits through the top. The process of LPS export is ATP dependent.

1.6 Peptidoglycan

Bacterial shape is mostly determined by peptidoglycan, a mesh of bacterial glycans which themselves are connected by short peptides forming the sacculus, the bacterial skeleton. The glycans consist of alternating residues of β-(1,4) *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) where the D-lactoyl group of each MurNAc residue is substituted by a short peptide sequence. This sequence is composed often as L-Ala-γ-D-Glu-L-Lys-D-Ala-D-Ala where the terminal D-Ala is removed during glycan crosslinking (Rogers 1980). Aside providing shape, peptidoglycan plays crucial roles in resisting osmotic and physical pressures as well as anchoring the OM to the cell through lipoproteins and OM protein (OMP) interactions (Park et al. 2012; Vollmer, Blanot & de Pedro 2008).

Peptidoglycan (PG) biosynthesis begins in the cytosol with the generation of UDP-MurNAc from UDP-GlcNAc by MurA and MurB. Peptide strand attachment onto UDP-MurNAc then occurs by the sequential addition of single peptides by MurC, MurD, MurE and MurF with MurI, DadX and DdlA involved in the biosynthesis of L-Glu and D-Ala respectively (Barreteau et al. 2008). The complete UDP-MurNAc unit including peptide stem is then ligated to undecaprenyl phosphate (Und-P) by MraY prior to the addition of UDP-GlcNAc at the non-reducing end leading to the formation of PG lipid-I and lipid-II, respectively. PG lipid-II is then translocated across the IM by MurJ where the PG disaccharides are attached to each other forming nascent PG chains by GTases such as penicillin binding protein 1 (PBP1)(Zheng et al. 2018). These chains are then attached to the nascent chains which are then attached to the sacculus, both facilitated by DD-transpeptidases. The GTases which from the nascent chain, PBP1a and PBP1b, RodA, form a large class of SEDS proteins, peptidoglycan glycosyltransferases (PGTs) where they are the first SEDS proteins to be identified (Meeske et al. 2016).

Peptidoglycan is constantly assembled and disassembled as the cell grows and divides as such, sacculus remodelling is performed by DD-/LD-/DL-carboxylpeptidases which trim peptides off and by DD-/LD- endopeptidases which cleave crosslinks (Egan, Errington & Vollmer 2020). LD-transpeptidaes and amidase are responsible for covalently and non-covalently attaching PG to lipoprotein Lpp and OmpA, respectively, which are embedded in the OM and facilitates anchoring of the OM to the cell and OM stability (Samsudin et al. 2016). This process is illustrated in Figure 1.18.

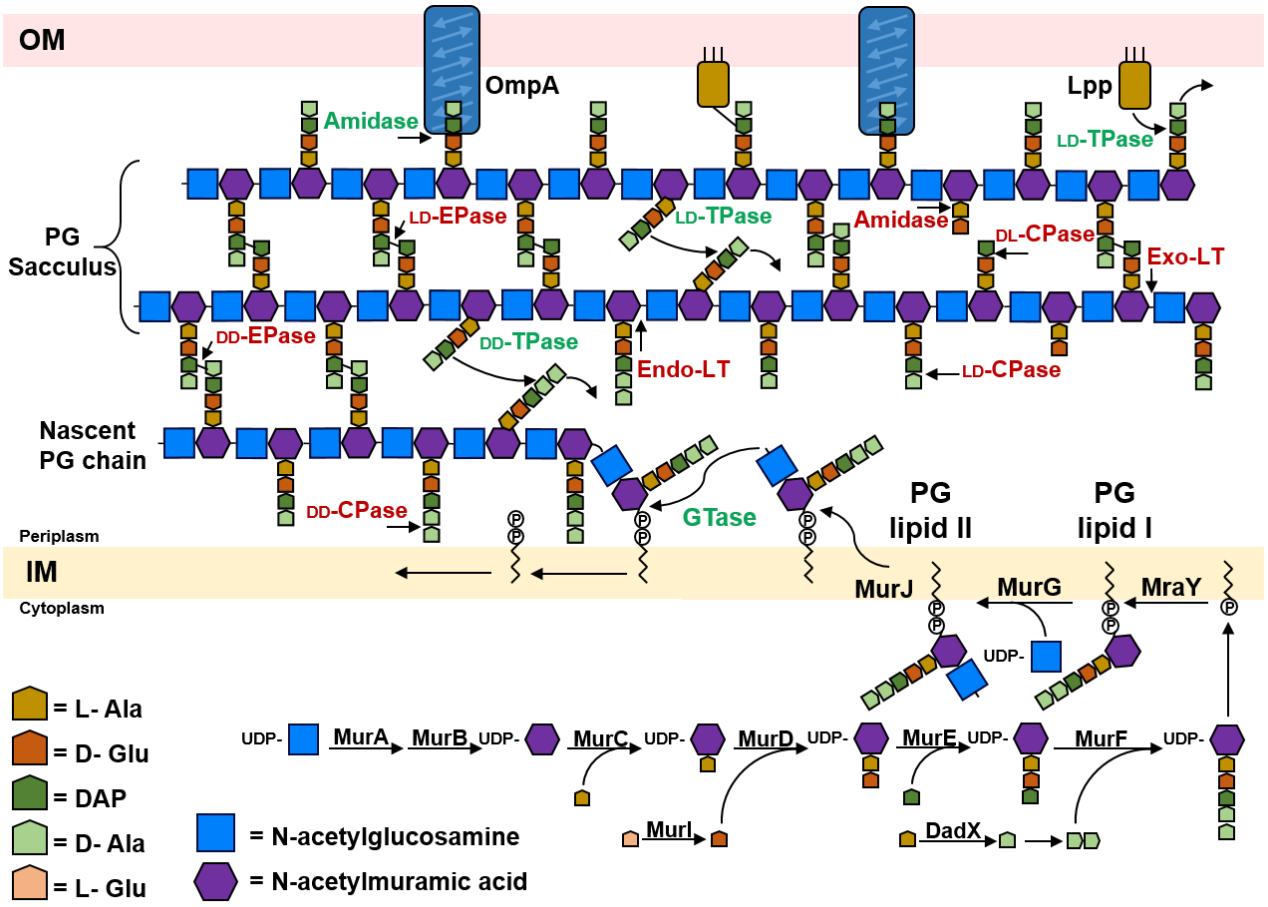


Figure 1.18 Peptidoglycan biosynthesis.

Peptidoglycan biosynthesis begins on the cytoplasmic leaflet of the inner membrane with the assembly of sugar repeat units to which the peptide sequence is added. Once complete, the repeat unit is attached to the Und-P forming PG lipid-I and post the addition of GlcNAc, PG lipid-II. PG lipid-II is then translocated across the IM by MurJ where it is assembled into nascent chains by GTases/SEDS PGTs, such as PBP1a and PBP1b, prior to attachment to the sacculus by transpeptidases. Remodeling of the sacculus is constant through the activity of DD-/LD-/DL-carboxylpeptidases (CPase) and DD-/LD- endopeptidases (EPase) and Exo/Endo lytic transglycoslases (Exo-LT/Endo-LT) (Egan, Errington & Vollmer 2020). The sacculus is attached to the OM to OmpA and Lpp which anchors the OM to the cell. Sugar moiety symbols used in accordance to symbol nomenclature for Glycans (SNFG) (Neelamegham et al. 2019).

Due to its importance in bacterial growth, peptidoglycan biosynthesis was a common target of antibiotics where for example, the first identified and mass-produced antibiotic, penicillin, targeted DD-transpeptidases preventing the addition of nascent chains to the sacculus (Kelly et al. 1982). Multiple classes of antibiotics have subsequently been used to target PG biosynthesis including β-lactams (of which include penicillin), cephalosporins, carbapenems and glycopeptides (Sarkar et al. 2017), however due to the immense selective pressure exerted by these drugs, multidrug resistance to these compounds is common (Nikolaidis, Favini-Stabile & Dessen 2014).

1.6.1 Undecaprenyl phosphate

Undecaprenyl phosphate is the universal lipid carrier molecule found in all bacteria which is involved in cell wall biosynthetic processes such as peptidoglycan, LPS O antigen and ECA biosynthesis. Comprising of a 55-carbon long chain isoprene lipids, Und-P exists as a finite resource within the cell where it is shared among the above pathways (Jorgenson et al. 2016; Jorgenson & Young 2016; Touze et al/ 2008) .

De novo synthesis of Und-P occurs on the cytoplasmic side of the IM and involves two major steps. Firstly the sequential condensation of trans,trans-C_{15}-PP with eight molecules of C_5-PP by UppS yielding C_{55}-PP (Touze et al. 2008) followed by the dephosphorylation of C_{55}-PP to yield C_{55}-P, Und-P, by UppP (El Ghachi et al. 2018). The processing by UppP translocates Und-P across the IM to the periplasmic leaflet (Tatar et al. 2007), there it is again translocated back across the IM by an unknown protein at which point it can be utilized by the cell wall biosynthetic pathways. The recycling of Und-PP occurs in a similar way: once released from cell wall biosynthetic pathways, Und-PP is acted upon by UppP yielding Und-P which is then translocated across the IM by an unknown process (Tatar et al. 2007). The process of Und-P biosynthesis, processing and recycling is illustrated in Figure 1.19.

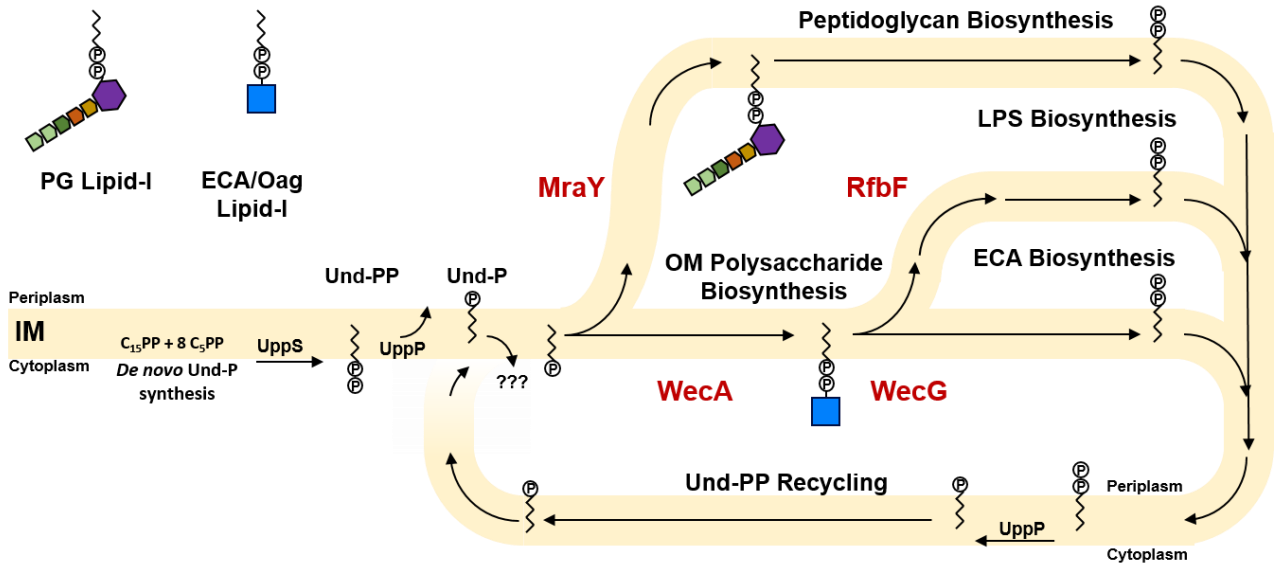


Figure 1.19: Undecaprenyl biosynthesis, usage and recycling in *S. flexneri.*

Undecaprenyl is first biosynthesised from *de novo* synthesis of $C_{15}PP$ and $8xC_5PP$ by UppS yielding Und-PP which is dephosphorylated by UppP exiting on the periplasmic leaflet of the IM (TouzÉ & Mengin-Lecreulx 2008). The process by which Und-P re-enters the cytoplasmic leaflet is unknown. Und-P is then acted upon by WecA or MraY, which commits that molecule of Und-P to PG or OM polysaccharide biosynthesis which is further divided into LPS O antigen or ECA biosynthesis when acted upon by RfbF or WecG respectively. Once the various repeat units are removed, Und-PP then is recycled by UppP yielding Und-P which can once again enter the cycle.

1.7 Research aims and hypothesis

Research into the Wzy-dependent pathway has been heavily orientated towards LPS O antigen biosynthesis with disregard towards the Wzy-dependent homologs from the ECA biosynthetic pathway. Likewise in the research of the roles of ECA, until recently, the majority of research conducted has been towards the understandings of why ECA biosynthetic mutants generate pleotropic phenotypes rather than focusing on the proteins themselves. The pleiotropic nature of the *wec* operon mutants as well as the similarities between ECA and LPS O antigen biosynthesis exemplifies the importance of investigating the potential indirect interactions between the two OM polysaccharide systems.

Subsequently these topics will be, in part, addressed in this thesis with the first study into investigating the characteristics of WzyE wherein WzyE was genetically, topologically investigated using experimental and *in silico* methods and mutated to discover functional residues. Another overlooked protein from the *wec* operon, WecG was also investigated where we showed that the historical subcellular localization of WecG was incorrect through mutagenesis and formed a hypothesis of how WecG is maintained to the membrane. Lastly, we investigated possible indirect linkages between ECA and LPS O antigen biosynthesis through the use of *wzyE* and *wzyB* mutants which we further linked with peptidoglycan biosynthesis, revealing how interdependent the three cell wall pathways are.

Aims:

1) To perform site-directed mutagenesis on $WzyE_{SF}$ conserved residues and to determine their role in the protein's function
2) To perform experimental topology mapping of $WzyE_{SF}$ through the use of PhoA::LacZα reporter fusions
3) To investigate the subcellular localization of WecG through mutagenesis and chaotropic reagent treatment
4) To investigate potential outer membrane polysaccharide biosynthetic cross talk between the two major pathways; LPS O antigen and ECA.

Chapter Two

MATERIALS AND METHODS

Chapter 2: Materials and Methods

2.1 Bacterial strains and growth conditions

2.1.1 Strains and plasmids

The bacterial strains using during this study are listed in Appendix A & B. Plasmids are listed in Appendix C.

2.1.2 Growth media and conditions

Bacterial strains were routinely grown in lysogeny broth (LB) (10 g tryptone L^{-1} (Difco), 5 g yeast extract L^{-1} (Difco), 5 g NaCl L^{-1}) or in LB agar (LBA) (10 g tryptone L^{-1} (Difco), 5 g yeast extract L^{-1} (Difco), 5 g NaCl L^{-1}, 15 g agar L^{-1}) containing the appropriate antibiotics.

2.1.3 Antibiotics and additives

For growth selection, the following antibiotics were used at the following concentrations; 100 µg ampicillin (Amp) ml^{-1}, 25 µg chloramphenicol (Cml) ml^{-1}, 50 µg kanamycin (Kan) ml^{-1}, 10 ng tunicamycin (Tnc) ml^{-1} with 3 µg polymyxin B nonapeptide ml^{-1} (PBMN; Sigma). Strains carrying pWSK29, pWKS30, pWALDO, pPLEO1, pQE30, pSUP203 or pBCKs $^{(+)}$ constructs requiring induction were grown in LB at 37 ºC with aeration for 16 hours, sub-cultured (1/20) into fresh broth and induced with 1 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG (Sigma)). Strains containing pBAD18 or pBAD33 constructs requiring induction were grown in LB at 37 ºC with aeration for 16 hours, sub-cultured (1/20) into fresh broth and induced with 0.2 % (w/v) L-arabinose. For topology mapping, LBA indicator plates were with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), 80 µg ml^{-1} 5-bromo-4-chloro-3-indolyl phosphate (BCIP (Sigma)), and 100 µg ml^{-1} 6-chloro-3-indolyl-β-D-galactoside (Red-Gal (Sigma)). For deoxycholate sensitivity assays, LBA plates were supplemented with 1 % (w/v) DOC (Sigma).

2.1.4 Growth and maintenance of bacterial strains

All bacterial strains were grown at 37 °C in liquid cultures using 20 mL McCartney bottles with aeration (200 rpm). Bacterial strains were stored at -80 °C. The bacterial strains were collected from LBA plates from which a bacterial lawn was generated from a single colony. Cultures were prepared from glycerol stocks by streaking out the bacterial strain onto selective agar and incubating at 37 °C overnight. A single colony was then used to inoculate a liquid broth as required.

2.2 Antibodies

The antibodies used in this study for Western immunoblotting are as follows: polyclonal rabbit anti-WzzE (Morona Lab), polyclonal rabbit anti-ECA (Morona Lab), mouse monoclonal

anti-FLAG (Sigma), monoclonal anti-His (GeneScript) HRP-conjugated goat anti-mouse IgG (KPL) and HRP-conjugated goat anti-Rabbit IgG (KPL).

2.3 DNA techniques

2.3.1 Plasmid isolation

Plasmid DNA was purified from overnight bacterial cultures (10 mL LB) and completed according to the QIAprep Spin Miniprep kit (Qiagen) protocol. MilliQ (MQ) H_2O (18 Ωcm) was used for elution and final preparations were stored at -20 °C in 1.5 mL reaction tubes.

2.3.2 Whole chromosome extraction

Cells from overnight cultures were harvested by centrifugation (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C) and the supernatant removed and the pellet resuspended in 3 ml of saline (0.85 %). 3 ml of phenol (pH 7.5) was then added and the mixture vortexed intermittently over a 2 min period. The mixture was then centrifuged as above and the aqueous phase collected; to which 3 ml of ice cold 100 % isopropanol was added. DNA was then precipitated by vigorous agitation of the mixture and incubated at -20 °C for 20 mins. Precipitated DNA was then washed in 70 % (v/v) ethanol, then transferred to a new 1.5 mL reaction tube and resuspended in 500 µL of MQ.

2.3.3 DNA quantification

The NanoDrop 2000c spectrophotometer (Thermo scientific) was used for quantification of all plasmid DNA samples using 1 µL at an absorbance of 260 nm.

2.3.4 Restriction digests

All restriction digests were performed using New England Biolabs (NEB) enzymes following the supplied protocols. Where possible, the enzymes were heat inactivated after use with heating at 65 °C if possible.

2.3.5 Oligonucleotides for PCR

The oligonucleotides for PCR used for this study were purchased from Integrated DNA Technologies (IDT) (Singapore) and are listed in Appendix C. Oligonucleotides were resuspended in MQ H_2O at a concentration of 100 µM for storage at -20 °C. Oligonucleotides was diluted to a working stock of 10 µM for use in PCR reactions.

2.3.6 PCR

PCR reactions were conducted according to the supplied protocols using either Taq DNA polymerase or Q5 DNA polymerase (NEB) using the Eppendorf Mastercycler Gradient PCR

thermocycler. Thermocycler settings were adjusted with each reaction with 34 amplification cycles used per reaction. For standard PCR cycle with Taq DNA polymerase, denaturation of the template was performed at 95 ºC for 30 sec, annealing of primers to the template DNA at a temperatures ranging from 55 ºC -70 ºC for 30 sec, and extension at 68 ºC for 1 min/kb of DNA. Deoxynucleic triphosphates (dNTPs) (Sigma) were used at a final concentration of 200 µM.

2.3.7 Preparation of boiled lysate for PCR DNA template

Single colonies of bacteria were resuspended in 100 µL of MQ H_2O in 1.5 mL reaction tubes, followed by incubation at 100 °C for 10 mins. Lysates were then centrifuged (16,250 × g, 1 min) to remove cell debris before the supernatant was used in PCR reactions as the DNA template.

2.3.8 DNA NaCl precipitation

Post PCR, 12.5 µl of 4 M NaCl and 500 µl of chilled 100% (v/v) ethanol was added to 250 µl of PCR reaction in a 1.5 mL reaction tube and was gently mixed and incubated on ice for 20 mins. 1.5 mL reaction tubes were then centrifuged at 16,250 × g, 10 min, 4 ºC, the supernatant discarded and the pellet gently washed with 500 µl of chilled 70 % (v/v) ethanol. 1.5 mL reaction tubes were then centrifuged again (16,250 × g, 1 min, 4 ºC), and the pellet washed once more in 70% (v/v) ethanol and pelleted as above. Tubes were then incubated at room temperature to evaporate off residual ethanol and the DNA pellet was resuspended in 50 µl of MQ H_20.

2.3.9 PCR product DNA purification kit

PCR product DNA was purified using the PCR DNA and Gel Band Purification Kit (GE Healthcare) according to the supplied protocols. DNA was eluted win 20 µL of MQ H_2O and stored at -20 °C.

2.3.10 Agarose gel electrophoresis

PCR amplicons were mixed with loading dye (0.1 % (w/v) bromophenol blue (Sigma-Aldrich), 20 % (v/v) glycerol, 0.1 mg/mL RNAase) 2:1 and separated on 1% (w/v) agarose gels with 1× Tris-borate-EDTA (TBE) buffer at 120 V for 30 min. *Bacillus subtilis* SPP1 bacteriophage DNA digested with *Eco*RI was used as a DNA maker (2.5 µL) (Morona laboratory) (sizes (kb): 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36, and 0.09). Gels were then visualised and photographed using the GelDoc XR system (BioRad).

2.3.11 DNA gel extraction

DNA bands were excised from 1% agarose (w/v) gels using a scalpel and DNA fragments were purified using the PCR DNA and Gel Band Purification Kit (GE Healthcare). DNA was eluted with 20 µL of MQ H_2O and stored at -20 °C.

2.3.12 DNA phosphorylation

Linear DNA fragments were subjected to 5' phosphorylation using T4 polynucleotide kinase (PNK) (NEB) according to the supplied protocols. Samples were then heat treated to 65 °C for 10 minutes to deactivate T4 polynucleotide kinase.

2.3.13 DNA ligation

Ligation of linear DNA fragments was performed using T4 DNA Ligase (NEB) according to the supplied protocols. Samples were incubated at 25 °C for 2 hrs or 4 °C overnight.

2.3.14 Site-directed mutagenesis

Generation of some point mutations were performed via site-directed mutagenesis using a QuikChange lightning site-directed mutagenesis kit (Catalog # 210518, Agilent) according to the manufacturer's instructions. Primers were designed on the Agilent website QuikChange Primer Design page: https://www.agilent.com/store/primerDesignProgram.jsp.

2.3.15 Inverse PCR to generate point substitutions, insertions and deletions

For mutagenesis through inverse PCR, DNA primers were designed such that the primer would anneal adjacent to the region interest. Primers were also designed such that when deletions or insertions occurred within the coding sequence, post DNA ligations the open reading frame (ORF) would remain intact. The entire plasmid backbone was PCR amplified (using Q5 DNA polymerase) with standard settings of denaturing at 98 °C for 10 secs, annealing from 55 °C – 70 °C for 30 secs and extension at 72 °C for 30 seconds/kb of DNA.

2.3.16 Chromosomal deletions via λ Red mutagenesis

Bacterial chromosomal gene deletion through the use of homology exchange with selective antibiotic resistance cassettes, λ Red mutagenesis, was performed as described in (Datsenko & Wanner 2000). In brief, primers were firstly designed to directly flank upstream and downstream (50 bps), respectively, of the target gene nucleotide sequence, as well as the FRT region from pKD3 or pKD4 (20 bps). Using the primers, the FRT-resistance cassette was amplified from pKD3 or pKD4 via PCR and the fragment size was checked by agarose gel electrophoresis.

Post confirmation, the PCR reaction was cleaned up 2.3.9 and then DpnI digested to remove any residual template plasmid DNA. Post DpnI digestion, the PCR product was once again cleaned and either stored at -20 °C or used immediately. Cells to be mutated carrying pKD46 were grown at 30 °C in the presence of 0.2 % (v/v) L-arabinose and were made electrocompetent. Following this, the concentrated PCR product was then electroporated into 100 µL of electrocompetent cells and transformants were selected for by overnight growth on selection LBA. Single colonies were then selected and patched onto individual Amp, Kan and LBA plates or Amp, Cml and LBA plates to confirm the loss of pKD46 and the chromosomal mutation was confirmed by PCR. In order to remove the resistance cassette, pCP20 was electroporated into each strain. Transformants were then inoculated into 10 ml LB and grown overnight where the following day, the overnight cultures were further incubated (2 hr, 42 °C). Cultures were then diluted and plated onto LB agar and further incubated (16 hr, 37°C). To confirm the loss of pCP20 and the resistance cassette, single colonies were selected and patched onto LB agar plates containing separately, Kan and Amp or Cml and Amp. Colonies which grew on neither antibiotic containing plate had their mutation confirmed via PCR with appropriate primers.

2.3.17 Generation of nested DNA deletions

Initially pPLEO1-WzyE^{IIa1350t} was digested with restriction enzymes *Pst*I and *Xba*I to generate 3' and 5' overhangs between *wzyE* and the *phoA::lacZα* fusion sequences. *Exo*III (NEB) was then used to digest *wzyE* 3'>5' prior to quenching by heating at 75 ºC for 20 minutes followed treatment with Mung bean nuclease (NEB) for 30 minutes at 30 ºC to remove 5' overhangs. DNA PolI (Klenow fragment) (NEB) was then used to remove remaining DNA overhangs prior to treatment with poly nuclease kinase and T4 DNA ligase. The resulting cocktail of truncated pPLEO1-WzyE^{IIa1350t} was then transformed into *E. coli* DH5α cells and plated onto LBA selection indicator plates. Pigmented colonies were then isolated and "truncated" pPLEO1-WzyE^{IIa1350t} plasmids extracted prior to DNA sequencing to identify the location of the truncation.

2.3.18 DNA sequencing

Oligonucleotide primer was added at a final concentration of 0.8 µM to 1000 – 1500 ng of purified DNA and adjusted to a final volume of 12 µL using MQ H_2O. The samples were submitted to and sequenced by Australian Genome Research Facility (AGRF).

2.4 Bacterial Transformation and fixing

2.4.1 Preparation of chemically competent cells

Mid-exponential phase (OD_{600} ~0.5) bacteria were grown in 10 mL LB broths and harvested via centrifugation (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C) and washed with 10 mL of chilled MQ H_2O. The bacterial pellet was then resuspended and washed in 5 mL of chilled 0.1 mM $MgCl_2$, centrifuged again and resuspended in 1 mL of chilled 100 mM $CaCl_2$.

2.4.2 Preparation of ultra competent cells

Bacteria were grown overnight at 37 °C for ~ 8 hrs and then were diluted 1:10 and 1:20 into 250 mL of SOB medium (20 g/L-^{1} tryptone [BD], 5 g/L^{-1} yeast extract [BD] and 0.5 g/L^{-1} NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, pH 7.0). Cultures were then incubated at 18 °C for 16 h with vigorous shaking (250-300 rpm). The bacterial culture which had closest OD_{600} ≤0.55 was then selected and continued to grow to OD_{600} reading of 0.55. Bacterial cells were then collected by centrifugation (Beckman Coulter Avanti J-26XPI; 10,000 x g, 10 min, 4 ^{o}C), and washed with 80 mL of precooled transformation buffer (55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl, 10 mM PIPES, pH 6.7), followed by centrifugation as above. The bacterial pellet was then resuspended in 20 mL chilled transformation buffer, and 1.5 mL of DMSO was added prior to aliquoting out the bacterial suspension at 200 µL aliquots and snap freezing the cells. The aliquots were stored at -80 °C (Inoue, Nojima & Okayama 1990).

2.4.3 Preparation of electro-competent cells

Mid-exponential phase (OD_{600} ~0.5) bacteria were grown in 10 mL LB broth and harvested via centrifugation (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C). The cells were washed 3x with 10 mL of chilled MQ H_2O, before being resuspended in 100 µL of 10 % (v/v) glycerol.

2.4.4 Heat-shock transformation of chemically competent cells

To pre-prepared chemically component cells, 4 µL plasmid DNA or the complete ligation mixture was added. The cell suspension was incubated on ice for 20 mins prior to incubation of the cells at 42 °C for 45 seconds. Following immediately, the cells were transferred to ice for 5 minutes, before being added to 900 µL of SOC recovery media. The cells were then incubated (1 hr, 37 °C) with aeration and then plated and grown at 37 °C overnight on selective LBA.

2.4.5 Electroporation transformation of electro-competent cells

To pre-prepared electro-component cells, 2 µL of purified DNA was added and the cell suspension was transferred to an electroporation cuvette (0.2 cm) (BioRad) following which it was incubated on ice for 20 mins. The cuvette was then transferred into the BioRad Gene Pulser and electroporated at 2.5 kV. Following the pulse, the cells were added to 900 µL of SOC recovery media and incubated (1 hr, 37 °C. The cells were then plated and grown 37 °C overnight on selective LBA.

2.5 Protein techniques

2.5.1 Generation of whole cell lysate protein samples

18 hour cultures were sub-cultured (1/20) into fresh LB broth, induced and grown as above and mid-exponential phase cells ($5x10^8$ cells) were collected by centrifugation (16,250 × g, 1 min, 4 ^oC). The cells were then resuspended in 100 µL of sample buffer (2 % (w/v) SDS, 10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol, 0.02 % (w/v) bromophenol blue, 62.5 mM Tris-HCl pH 7.0). The samples were then either heated to either 42 °C or 100 °C for 5 minutes, depending on whether membrane proteins were to be probed, before use in SDS-PAGE or stored at -20 °C.

2.5.2 Generation of whole membrane samples

18 hour cultures were sub-cultured (1/20) into 200 ml of fresh LB broth, induced and grown for 4 hours. Bacterial cells were then harvested by centrifugation (Beckman Coulter Avanti J-26XPI; 10,000 x g, 10 min, 4 ^oC) and the bacterial pellet resuspended in 10 mL sonication buffer (100 mM $NaCO_3$, 150 mM NaCl, pH 7.5). The cell suspension was then disrupted by sonication (Branson B15) and cell debris pelleted by centrifugation (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C). The supernatant was then transferred into ultracentrifugation tubes and the whole membrane was collected by ultra-centrifugation (Beckman Coulter Optima L-100 XP ultra-centrifuge; 250,000 x g, 1 hr, 4 ^oC).

2.5.3 *In vivo* protein crosslinking

18 hour cultures were sub-cultured (1/20) into 200 mL of fresh LB broth, induced and grown for 4 hours. Cells were then harvested by centrifugation (Beckman Coulter Avanti J-26XPI; 10,000 x g, 4 ^oC, 10 min) and the pellets were resuspended and washed in DSP crosslinking buffer (150 mM NaCl, 20 mM sodium phosphate buffer (Na_2PO_4/NaH_2PO_4), pH 7.2). Cells were then again centrifuged as above, resuspended in 5 mL of DSP crosslinking buffer, and treated with and without 1 mM DSP (Thermo Fischer Scientific). Cell suspensions were then incubated (37 ^oC, 20 min) and then excess DSP was quenched (20 mM Tris-HCl, pH 7.5) and left for 10 min at RT.

Following this, the cell suspension was centrifuged (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C) followed by washing with DSP crosslinking buffer as above.

2.5.4 Protein chaotropic disassociation

Whole membrane pellets were resuspended in 5 mL of chilled MQ H_2O and 500 µl of the whole membrane suspension was aliquoted into 1.5 ml reaction tubes to which 500 µl of 2x chaotropic buffers (3 M NaI, 200 mM $NaCO_3$, pH 7.0), (4 M NaCl, 200 mM $NaCO_3$, pH 7.0), (200 mM $NaCO_3$, pH 12.0, (200 mM, $NaCO_3$, pH 7.0) or PBS were added. The 1.5 mL reaction tubes were then incubated (RT, 1 hr) with agitation before the insoluble fraction membrane was collected via ultracentrifugation (Beckman Coulter Optima L-100 XP ultra-centrifuge; 250,000 x g, 1 hr, 4 ^oC).

2.5.5 Whole membrane protein PEGylation

Whole membranes were generated as above in 2.5.2. Whole membranes were then resuspended in 200 µL of chilled MQ H_2O and 200 µL of 2x PEGylation buffer (200 mM HEPES; 20 mM $MgCl_2$; pH 7.0). The membrane suspension was then split into two separate 1.5 mL reaction tubes to which 1 mM polyethylene glycol (PEG) maleimide (Sigma) was added to one tube and incubated (3 hr, 4 °C). Excess PEG malemide was then quenched by the addition of 45 mM dichloro-diphenyl-trichloroethane (DDT) (Sigma) and the tubes were incubated (20 min, 4°C).

2.5.6 *In vivo* whole cell protein PEGylation

Overnight cultures were sub-cultured (1/20) into 200 mL of fresh LB broth, induced and grown for 2 hours before cells were harvested by centrifugation (Beckman Coulter Avanti J-26XPI centrifuge; 10,000 x *g*, 10 min, 4 °C) and resuspended with 10 mL of chilled sonication buffer (100 mM $NaCO_3$, 150 mM NaCl, pH 7.5). 1 mL of cell suspension was then aliquoted into six 1.5 mL reaction tubes where, two tubes received 10 µL of 0.5 M EDTA and another two tubes received 200 µL of 10 % (w/v) SDS to make a set of 3 samples. All tubes then received 1 mM PEG malemide (Sigma) and were incubated at 4°C and 37°C for 1 hr before the addition of 45 mM DDT to quench excess PEG maleimide. Samples were then centrifuged (16,250 × g, 10 min, 4 ^oC) and the pellet was resuspended with 1 mL of chilled sonication buffer prior to the disruption of cells by sonication (Branson B15). The mixtures were then centrifuged (Thermo Scientific Labofuge 400 R centrifuge; 4,000 x *g*, 10 min, 4 °C) to pellet cell debris and the whole membrane sample collected by ultra-centrifugation (Beckman Coulter Optima L-100 XP ultra-centrifuge; 250,000 x g, 1 hr, 4 ^oC).

2.5.7 Protein SDS-PAGE

Protein samples were heated at 42 °C or 100 °C depending on if membrane proteins would be investigated, for 10 mins and separated on a 12% SDS (w/v) polyacrylamide gels in protein running buffer (200 mM glycine, 80 mM Tris-HCl, 0.1 % (w/v) SDS) at 200V (Biorad MiniProtean) for 1 hr. As a molecular size maker SeeBlue Plus2 molecular weight marker (Invitrogen) was used.

2.5.8 SDS-PAGE Western transfer

Post electrophoresis, SDS-PAGE gels were transferred onto a nitrocellulose membranes (NitroBind, Pure nitrocellulose, 0.45 µm, GE Water & Process Technologies) for 1 hr at 400 mA in transfer buffer (25 mM Tris-HCl, 200 mM glycine, 5 % (v/v) methanol).

2.5.9 Protein detection by immunoblotting

Following Western transfer, nitrocellulose membranes were incubated (1 hr, RT) with 100 mL blocking solution (5% (w/v) skim milk powder (Woolworths) in TTBS (16 mM Tris-HCl, 120 mM NaCl, 0.05 % (v/v) Tween 20, pH 7.4 (Sigma-Aldrich)) followed by incubation (16 hr, RT) with primary antibody in 2.5 % (w/v) skim milk powder in TTBS buffer with agitation. The following day, the membrane was washed 3x in TTBS (10 min), followed by incubation (2 hr, RT) with secondary antibody in TTBS buffer with agitation. The membrane was then washed 3x in TTBS buffer (5 min) and then 3x in TBS (16 mM Tris-HCl, 120 mM NaCl, pH 7.4) (5 min). For detection of bound secondary antibody, chemiluminescent peroxidase substrate-3 (Sigma) was used as described by the manufacturer's instructions. The fluorescent signals were detected by the BioRad ChemiDoc imaging system.

2.6 Polysaccharide techniques

2.6.1 Preparation of polysaccharide samples

18 hour cultures were sub-cultured (1/20) into 200 mL of fresh LB broth, induced and grown for 4 hours before mid-exponential phase cells ($1x10^9$ cells) were collected by centrifugation (16,250 × g, 1 min, 4 ºC), resuspended in 50 µL of 2x lysis buffer (2 % (w/v) SDS, 0.1 % (w/v) bromophenol blue (Sigma), 4% (v/v) β-mercaptoethanol (Sigma), 10% (v/v) glycerol, 660 mM Tris-HCl, pH 7.6) and heated (100 ºC, 10 min) before incubation (56 ºC, 2 hr) with 2.5 mg/ml proteinase K (Sigma-Aldrich). Samples were heated at 100 ºC for 1 minute prior to loading.

2.6.2 Detection of LPS by silver-staining

Polysaccharide samples were heated (100°C, 10 min) before loading onto 15% (w/v) SDS polyacrylamide gels as described by Macpherson et al. (1991). Gels were then electrophoresed (12 mA, 13 hr) in LPS running buffer (200 mM glycine, 80 mM Tris-HCl, 0.1 % (w/v) SDS). Silver-staining was performed as described by Tsai and Frasch (1982). Briefly, gels were incubated (2 hr) in fixing solution, oxidized (5 min) in oxidizing solution, washed in MQ H_2O (6×15 min), stained (10 min) in staining solution, and washed in MQ H_2O (5×10 min). Gels were then developed in pre-warmed (42 °C) developing solution, and the development reaction was stopped using stopping solution.

2.6.3 Detection of ECA by SDS-PAGE and Western immunoblotting

Polysaccharide samples were loaded onto a 15% (w/v) SDS polyacrylamide gel and electrophoresed (200 V,1 hr). SDS-PAGE gels were then transferred onto a nitrocellulose membrane (Bio-Rad) (400 mA, 1 hr) prior to membranes being incubated (1hr, RT) with blocking solution 5% (w/v) skim milk in TTBS buffer. Membranes were then incubated (16hr, RT) with polyclonal rabbit anti-ECA antibodies (1:500), diluted in 2.5% (w/v) skim milk in TTBS. The following day, the membrane was washed 3x in TTBS (10 min), followed by incubation (2 hr, RT) with secondary goat anti-rabbit horseradish peroxidase-conjugated antibodies (KPL) in TTBS buffer with agitation. The membrane was then washed 3x in TTBS buffer (5 min) and then 3x in TBS (16 mM Tris-HCl, 120 mM NaCl, pH 7.4) (5 min). For detection of bound secondary antibody, chemiluminescent peroxidase substrate-3 (Sigma) was used as described by the manufacturer's instructions. The chemiluminesent signals were detected by the BioRad ChemiDoc imaging system.

2.7 Bacterial sensitivity and growth analysis

2.7.1 Deoxycholate resistance assays

Overnight cultures were sub-cultured (1/20) into 10 mL of fresh LB broth, induced and grown for 2 hours. Cells ($1x10^8$ cells) were then harvested by centrifugation (16,250 × g, 1 min, 4 ^oC) and resuspended in 1 mL of LB broth. The cell suspension was then diluted 1:10 with LB prior to spotting 3 µL of cellular suspension onto LBA plates supplemented with 1% DOC (w/w). Plates were then incubated (16 hr, 37 °C).

2.7.2 Colicin sensitivity assays

Overnight cultures were sub-cultured (1/20) into 10 mL of fresh LB broth, induced and grown for 2 hours prior to spreading 100 µl of cell suspension onto a pre-dried 25 ml LBA plates.

Colicin E2 protein (7.5 mg/ml, Morona Laboratory) was serially diluted in MQ H_2O to 80, 60, 50, 40, 20, 15, 10, 5, 2, 1, 0.5, 0.25 µg/ml and 3 µL of Colicin E2 protein was spotted onto the bacterial spread plates. Plates were then incubated (16 hr, 37 °C).

2.7.3 Analysis of growth kinetics by growth curves

Bacteria from overnight cultures were collected ($1x10^7$ cells) via centrifugation (16,250 × g, 1 min, 4 ºC) and resuspended in 1 mL of LB broth. Cells were then sub-cultured (1/10) into 135 µL of fresh LB broth in a 96 well tray. The tray was then incubated (10 hr, 37 ºC) with aeration and agitation. OD_{600} absorbance readings were taken every 20 minutes (BioTek PowerWave XS2).

2.7.4 Analysis of growth kinetics by colony forming unit (CFU) counting

Bacteria from overnight cultures were normalized and sub-cultured into 10 mL of LB broth. 20 µL of cell culture was taken at time points 0, 2, 4, 6 and 8 hours and serially diluted 1:10 in LB broth prior to spotting 10 µL of the cell suspension from the range of 10^{-5} to 10^{-8} in triplicate onto LBA plates. Plates were then incubated (16 hr, 37 °C) and bacterial colonies counted.

2.8 Microscopy

2.8.1 Preparation of phase contrast slides

Overnight cultures were sub-cultured (1/20) into 10 mL of fresh LB broth, induced and grown for 2 hours prior to 10 µL of bacterial culture being spotted onto glass slides. The bacterial culture was then allowed to dry followed by mounting with 20% (v/v) Mowiol 4–88 (Calbiochem) and sealing with nail polish.

2.8.2 Visualization of cells by phase contrast microscopy

Bacterial cells were observed under a 100x oil objective lens (Olympus IX70) and further magnified by 1.5x body magnification to take higher magnification photos using Metamorph 7.5.6.

2.8.3 Cell measurements

The pixel length was normalized across all photos and cell lengths were then manually measured from pole to pole using Metamorph 7.5.6.

2.9 Enzymatic assays

2.9.1 Alkaline phosphatase (PhoA) assay

18 hour cultures were sub-cultured (1/20) into 10 mL of fresh LB broth, induced and grown for 2 hours. Cells ($1x10^8$) were harvested by centrifugation (16,250 × g, 1 min, 4 ºC) then washed in chilled wash buffer (10 mM Tris-HCl, pH 8.0, 10 mM $MgSO_4$) and resuspended in 1 mL of

chilled PM1 buffer (1 M Tris-HCl pH 8.0; 0.1 mM $ZnCl_2$; 1 mM Iodoacetamide). 100 µL of chloroform and 100 µL of 0.05 % (w/v) SDS was added and the mixture incubated (5 min, 37 ºC). 100 µL of the upper cell suspension was added to a 96 well tray followed by addition of 50 µl of pNPP solution (0.5 % (w/v) pNPP (Thermofisher); 1 M Tris-HCl pH 8.0). The tray was then incubated (2 hr, 37 ºC) with OD_{405} and OD_{595} readings taken every 2 minutes.

2.9.2 β-galactosidase (LacZ) assay

18 hour cultures were sub-cultured (1/20) into 10 mL of fresh LB broth, induced and grown for 2 hours. Cells ($1x10^8$) were harvested by centrifugation (16,250 × g, 1 min, 4 ºC) then washed and resuspended in chilled Z buffer (60 mM Na_2PO_4; 40 mM NaH_2PO_4; 10 mM KCl; 1 mM $MgSO_4$; 50 mM β-ME; pH 7.0). 100 µL of chloroform and 100 µL of 0.05 % (w/v) SDS was added and the mixture incubated (5 min, 37 ºC). 50 µL of the upper cell suspension was added to a 96 well tray followed by addition of 0.15 % (w/v) ONPG in Z buffer without β-ME. The tray was then incubated (2 hr, 37 ºC) with OD_{405} and OD_{595} readings taken every 2 minutes.

2.9.3 Generation of Normalized Activity Ratio (NAR)

The NAR value was determined as follows = (PhoA activity/Highest PhoA activity)/LacZ activity/Highest LacZ activity.

2.10 Bioinformatic analysis

2.10.1 DNA and peptide sequences

The nucleotide and peptide sequences used in this study were obtained NCBI.

2.10.2 Multiple Sequence Alignments (MSA)

Multiple sequence alignments were performed by the Clustal Omega servers where the peptide sequences were entered in FASTA format. The complete MSA was then inputted into Jalview for analysis and manipulation (Sievers et al. 2011; Waterhouse et al. 2009).

2.10.3 Secondary structure prediction

Secondary structure prediction was performed by Jpred. The peptide sequence was entered into the Jpred servers in FASTA format (Drozdetskiy et al. 2015).

2.10.4 Protein topology prediction

In silico topology prediction was performed using the CCTOP and TMHMM servers. The protein peptide sequence was entered into each server separately in FASTA format (Dobson, Reményi & Tusnády 2015; Krogh et al. 2001).

2.10.5 Phylogenetic and Principal Component Analysis (PCA)

Phylogenetic and PCA analysis was performed within Jalview using inbuilt tools. After the generation of the MSA, the MSA was analysed and a phylogenetic tree was drawn based on the similarities between sequences. PCA analysis was performed on top to visualize the degree of difference in three-dimensions (Waterhouse et al. 2009).

2.10.6 Isoelectric point analysis

Iso-electric point analysis was performed in the iso-electric point calculator servers. The peptide sequences were entered in FASTA format (Kozlowski 2016).

2.10.7 Tertiary structure obtainment and prediction

In silico tertiary structure prediction was performed in the I-TASSER and RaptorX servers. The protein peptide sequence was entered in FASTA format and 3D models obtained (Morten Kallberg 2014; Yang et al. 2015). Other predicted protein structures were obtained from Alphafold, ID (P27835 (WZYE_ECOLI), without requiring the input of a sequence (Jumper et al. 2021) or from the protein database (PDB).

2.10.8 Protein structure modelling and manipulation

Three-dimensional protein structures were modelled and manipulated in Chimera software (1.15). Measurement of atomic distances between residues was performed using inbuilt tools within Chimera.

2.11 Statistical analysis

2.11.1 Student T-tests

Independent student T-tests were performed on triplicate experimental data values using the statistical analysis tool of Graph pad Prism 9. Graphs were plotted with the standard error of the mean (SEM) and statistical significance was displayed as the following: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$, and ns is a non-statistically significant.

2.11.2 One-way ANOVA

One-way ANOVA with multiple comparisons was performed on triplicate experimental data values using the statistical analysis tool of Graph pad Prism 9. Graphs were plotted with the standard error of the mean (SEM) and statistical significance was displayed as the following: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$, and ns is a non-statistically significant.

Chapter Three

ARTICLE I:

Topology of the Enterobacterial Common Antigen polymerase, WzyE.

Nicholas Tadeusz Maczuga, Elizabeth Tran and Renato Morona.

Chapter 3: Topology of the Enterobacterial Common Antigen polymerase, WzyE.

3.1 Statement of Authorship

Title of Paper	Topology of the Enterobacterial Common Antigen polymerase, WzyE
Status	Accepted (05/04/2022) – Society of Microbiology SGM
DOI	10.1099/mic.0.001183

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Nicholas Tadeusz Maczuga		
Contribution	Construction of strains and plasmids, conducted all experiments pertaining to all Figures and constructed these figures, data analysis, conception of model, corresponding author and writing of manuscript.		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/01/2022

Author	Elizabeth Ngoc Hoa Tran		
Contribution	Supervised development of work and provided editing and evaluation of the manuscript.		
Signature		Date	19/1/22

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, manuscript evaluation and editing, conception of model.		
Signature		Date	17/01/2022

3.2 Article Abstract

Enterobacteriales have evolved a specialized outer membrane polysaccharide (Enterobacterial Common Antigen (ECA)) which allows them to persist in various environmental niches. Biosynthesis of ECA initiates on the cytoplasmic leaflet of the inner membrane (IM) where glycosyltransferases assemble ECA repeat units (RUs). Complete RUs are then translocated across the IM and assembled into polymers by ECA specific homologs of the Wzy-dependent pathway. Consisting of the membrane proteins Wzx, Wzy and Wzz, the Wzy-dependent pathway is the most common polysaccharide biosynthetic pathway in Gram negatives where it is most notably involved in LPS O antigen (Oag) biosynthesis. As such, the majority of research directed towards these proteins have been orientated towards Oag biosynthetic homologs with little directed towards ECA homologs. Belonging to the Shape, Elongation, Division and Sporulation (SEDS) protein family, Wzy proteins are the polymerase, and are characterized as possessing little or no peptide homology among homologs as well as being polytopic membrane proteins with functionally relevant residues within periplasmic loops as defined by C-terminal reporter fusion topology mapping. Here, we present the first the first major study into the ECA polymerase WzyE. Multiple sequence alignments and topology mapping showed that WzyE is unlike WzyB proteins involved with Oag biosynthesis as they display high peptide conservation across Enterobacteriales. *In silico* structures and reporter mapping allowed us to identify possible functionally conserved residues with $WzyE_{SF}$'s periplasmic loops which we showed were crucial for its function. This work provides novel insight into Wzy proteins and suggests that WzyE is an optimal model to investigate Wzy proteins and the Wzy-dependent pathway.

3.3 Article Introduction

Enterobacteriacae have developed unique surface associated polysaccharides to allow them to persist in distinct niches. Enterobacterial common antigen (ECA) is a surface associated polysaccharide which is expressed on the surface of all *Enterobacteriacae* and is known to provide resistance to bile salts as well as maintain the integrity of the outer membrane (OM) (Mitchell, Srikumar & Silhavy 2018; Ramos-Morales et al. 2003). ECA comprises of trisaccharide repeating units (RUs) containing N-acetylglucosamine GlcNAc), N-acetyl-D-mannosaminuronic acid (ManNAcA) and 4-acetamido-4,6-dideoxy-D-galactose (Fuc4NAc) (Eade et al. 2021). ECA exists in two surface associated forms: ECA_{pg} and ECA_{lps} where ECA polysaccharides are attached to phosphatidylglycerol and the core-sugars of LPS, respectively as OM glycolipids (Rai & Mitchell 2020). ECA also exists as a cyclic, periplasmically restricted form ECA_{cyc}. Additionally to ECA, *Enterobacteriaceae* possess LPS O antigen (Oag) which is well characterized in the literature and is known to play key roles in viability, pathogenesis and immune evasion of *pathogens* (Günther et al. 2019). Both surface polysaccharides are biosynthesised by their distinct homologs of the Wzy-dependent pathway (Islam & Lam 2014).

ECA repeat units (RUs) are biosynthesised by the proteins of the *wec* operon and are assembled into polymers by ECA specific homologs of the Wzy-dependent pathway. Briefly, ECA biosynthesis begins on the cytoplasmic leaflet of the inner membrane (IM) where the glycosyltransferase WecA ligates UDP-GlcNAc onto undecaprenyl phosphate (Und-P) yielding ECA lipid-I (Eade et al. 2021). The glycosyltransferases WecG and WecF then sequentially ligate ManNAcA and Fuc4NAc yielding ECA lipid-II and ECA lipid-III respectively (Eade et al. 2021). The complete ECA RU is then translocated across the IM by WzxE where it is assembled into chains by WzyE and WzzE prior to ligation onto a final lipid anchor and export to the OM (Islam & Lam 2014). This process is illustrated in Figure 3.1a.

Until recently, research into ECA has been orientated towards understanding how and why ECA biosynthetic mutants cause pleiotropic mutant phenotypes affecting seemingly non-related processes (Jorgenson et al. 2016). In 2006 *Baba et al.* (Baba et al. 2006) released the KEIO collection, a library of single gene knockouts of all possible non-essential genes *in E. coli* K12 including all *wec* operon mutants, except for *wzyE,* which was consequently described as an essential gene. Since then research into WzyE has not been published.

WzyE, belonging to the Wzy protein family, is also defined as a member of a Shape, Elongation, Division and Sporulation (SEDS) protein family which are characterized as membrane proteins evolved in cell wall processes (Meeske et al. 2016). Research into Wzy proteins have

been orientated towards Wzy proteins involved in LPS Oag biosynthesis, and several of these WzyB proteins have been topologically defined and partially characterized where it has been found that WzyB protein sequences are highly variable (Daniels, Vindurampulle & Morona 1998; Islam et al. 2010; Kim et al. 2010). This characterization also included the identification of key residues within periplasmic loops as well as GX^nG motifs across WzyBs from *Shigella flexneri*, *Escherichia coli*, *Pseudomonas aeruginosa* and *Francisella tularensis* (Islam et al. 2011; Islam et al. 2013; Leo et al. 2021; Nath & Morona 2015b; Nath, Tran & Morona 2015).

Here we show that unlike WzyB proteins involved with LPS Oag biosynthesis, WzyE proteins are highly conserved throughout *Enterobacteriales* and possess a non-classical WzyB topology as determined by multiple sequence alignments and PhoA::LacZα C-terminal reporter fusions. Using both an *in silico* structure and an experimentally determined topology map, we identified conserved arginine residues and a possible positively charged region within WzyE suggestive of a catalytic site. Amino acid substitutions revealed that these arginine residues are all critical in WzyE's function and possibly have a role in maintaining the stability of the protein.

This study highlights the familiarity and differences between WzyE and WzyBs involved with Oag biosynthesis and in doing so present WzyE as a useful Wzy protein model for research into the Wzy-dependent pathway.

3.4 Article Methods

3.4.1 Ethics statement

The ECA antibodies were produced under the National Health and Medical Research Council Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and was approved by the University of Adelaide Animal Ethics Committee.

3.4.2 Bacterial strains, growth media and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 3.1. Bacteria were routinely grown at 37 ºC in Lysogeny-Bertani (LB) broth with aeration or on LB agar (LBA). Indicator plates for topology mapping consisted to LBA supplemented with isopropyl-β-D-thiogalactopyranoside (IPTG) (1 mM), 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (80 μg/ml) and 6-chloro-3-indolyl-β-D-galactoside (Red-Gal) (100 μg/ml). Antibiotics used were 50 μg kanamycin (Kan) ml^{-1}; 25 μg chloramphenicol (Cml) ml^{-1}; 100 μg ampicillin (Amp) ml^{-1}. Strains carrying either pBAD33 or pPLEO1 constructs requiring induction were grown in LB at 37 ºC with aeration for 16 hours, sub-cultured (1/20) into fresh broth and induced with either 0.2% (w/v) L-Arabinose or 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) respectively. Cultures were grown for a further for 4 hours.

3.4.3 DNA methods

The plasmids used in this study are shown in Table 3.1. Plasmid were purified from *E. coli* DH5α strains using a QIAprep Spin miniprep kit (Qiagen). Preparation of electrocompetent cells and electroporation methods were performed as described previously (Purins et al. 2008). DNA sequencing was performed by the Australian Genome Research Facility (AGRF).

3.4.4 Construct generation

Primers used for construct generation are listed in Table 3.2-(S1). Generation of pPLEO1-WzyE was performed by using primers (NM26/NM27) to PCR amplify a fragment of DNA containing *wzyE* with XbaI and SacI restriction enzyme sites from pBAD33WzyE3xFLAG. The resulting *wzyE* fragment was digested with XbaI and SacI and sub-cloned into likewise digested pPLEO1 to give pPLEO1-WzyEI. Site-directed mutagenesis with primers (NM28/NM29) were used to generate pPLEO1-$WzyE^{IIa1350t}$ with the QuikChange lightning kit (Agilent).

3.4.5 Generation of nested DNA deletions

Initially pPLEO1-$WzyE^{IIa1350t}$ was digested with restriction enzymes PstI and XbaI to generate 3’ and 5’ overhangs between *wzyE* and the *phoA::lacZ*α fusion sequences. ExoIII (NEB) was then used to digest *wzyE* 3’>5’ prior to quenching by heating at 75 ºC for 20 minutes followed

treatment with Mung bean nuclease (NEB) for 30 minutes at 30 ºC to remove 5' overhangs. DNA PolI (Klenow fragment) (NEB) was then used to remove remaining DNA overhangs prior to treatment with poly nuclease kinase and T4 DNA ligase. The resulting cocktail of truncated pPLEO1-WzyE^{IIa1350t} was then transformed into *E. coli* DH5α cells and plated onto LBA selection indicator plates. Pigmented colonies were then isolated and "truncated" pPLEO1-WzyE^{IIa1350t} plasmids extracted prior to DNA sequencing to identify the location of the truncation.

3.4.6 Generation of targeted DNA deletions

Generation of C-terminal truncations of WzyE were performed using inverse PCR and primers listed in Table 3.2-(S1). Briefly, primers were used to amplify DNA fragments using pPLEO1-WzyE^{IIa1350t} as a template at specific points to generate in-frame deletions. The linear fragment was 5' phosphorylated via polynucleotide kinase (NEB) before being circularized using T4 DNA ligase (NEB). Following transformation into *E. coli* DH5α cells, transformants were plated onto LBA selection indicator plates, and pigmented colonies were isolated and subjected to DNA sequencing to confirm the deletion.

3.4.7 Generation of arginine to glycine substitutions

Generation of WzyE R>G point mutations were performed via site-directed mutagenesis using a QuikChange lightning site-directed mutagenesis kit (Agilent). Primers pairs used to generate point mutants: (WzyE^{R204G}:NM153/NM154, WzyE^{R247G}:NM155/NM156, WzyE^{R266G}:NM157/NM158, WzyE^{R295G}:NM159/NM160, WzyE^{R309G}:NM161/NM162, WzyE^{R399G}:NM163/NM164, WzyE^{R408G}:NM165/NM166). DNA sequencing was used to confirm all constructs.

3.4.8 Whole cell protein sample preparation

Bacteria were grown and induced as described above before $5x10^8$ cells were collected by centrifugation (2000 x g) and resuspended in 100 µl of 2x sample buffer (Lugtenberg et al. 1975). Samples were heated at 56 ºC for 10 minutes prior to loading.

3.4.9 ECA and LPS sample preparation

Bacteria were grown and induced as described above before $1x10^9$ cells were collected by centrifugation (2000 x g), resuspended in 2x lysis buffer (Murray, Attridge & Morona 2003) and heated at 100 ºC for 10 minutes before incubation with 2.5 mg/ml proteinase K (Sigma-Aldrich) for 2 hours at 56 ºC. Samples were heated at 100 ºC for 1 minute prior to loading.

3.4.10 Western immunoblotting

Protein/ECA samples were loaded onto a 12% or 15% SDS-PAGE gel respectively and electrophoresed at 200 V for 1 hour. SDS-PAGE gels were then transferred onto a nitrocellulose membrane (Bio-Rad) at 400 mA for 1 hour prior to membranes being blocked with 5% (wt/vol) skim milk in Tris-Tween Buffer saline (TTBS). Membranes were then incubated overnight with either monoclonal mouse anti-FLAG antibodies (GenScript) (1:10,000) or polyclonal rabbit anti-ECA antibodies (1:500), diluted in 2.5% (wt/vol) skim milk in TTBS. Detection was performed with rabbit anti-mouse or goat anti-rabbit horseradish peroxidase-conjugated antibodies (KLP) and chemiluminesence reagent (Sigma). 5 µl of SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard.

3.4.11 Bioinformatic analysis

The peptide sequence of WzyE was obtained from NCBI, full list of GenBank accession numbers are located in Dataset 3.1-(S1). The WzyE peptide sequence was then submitted to TMHMM to *in silico* predict transmembrane helices (Krogh et al. 2001). The *in silico* predicted structure of WzyE was obtained from Alphafold (P27835 (WZYE_ECOLI) (Jumper et al. 2021). Multiple sequence alignments and PCA analysis were performed in Jalview (Waterhouse et al. 2009). In silico atomic measurements were performed using Chimera software (1.15). Isoelectric point analysis was performed in IPC isoelectric point calculator (Kozlowski 2016). Multiple sequence alignments were performed by Clustal Omega (Sievers et al. 2011).

3.4.12 Alkaline phosphatase assays

Alkaline phosphatase assays were performed as described in (Karimova & Ladant 2017). Briefly, sub-cultured cells were washed in cold wash buffer (10 mM Tris HCl pH 8.0; 10 mM $MgSO_4$) prior to suspension in cold PM1 buffer (1 M Tris HCl pH 8.0; 0.1 mM $ZnCl_2$; 1 mM Iodoacetamide). 100 µl of chloroform and 100 µl of 0.05% (w/v) SDS was added to 1 ml of cell suspension and incubated at 37 ºC for 5 minutes. 100 µl of the upper cell suspension was added to a 96 well tray followed by addition of 50 µls of pNPP solution (0.5% w/v pNPP (Thermofisher); 1 M Tris-HCl pH 8.0). The tray was then incubated at 37 ºC with OD_{405} and OD_{595} taken every 2 minutes.

3.4.13 β-galactosidase assays

β-galactosidase assays were performed as described in (Karimova & Ladant 2017). Briefly, sub-cultured cells were washed and resuspended in cold Z buffer (60 mM Na_2PO_4; 40 mM NaH_2PO_4; 10 mM KCl; 1 mM $MgSO_4$; 50 mM β-ME; pH 7.0). 100 µl of chloroform and 100 µl of 0.05% (w/v) SDS was added to 1 ml of cell suspension and incubated at 37 ºC for 5 minutes.

50 µl of the upper cell suspension was added to a 96 well tray followed by addition of 0.15% (w/v) ONPG in Z buffer without β-ME. The tray was then incubated at 37 ºC with OD_{405} and OD_{595} taken every 2 minutes.

3.4.14 Normalized Enzymatic Activity Ratio (NAR) calculation

The NAR value was determined as follows = (PhoA activity/Highest PhoA activity)/LacZ activity/Highest LacZ activity.

Table 3.1: Bacterial strains and plasmids

Strain or plasmid	Description	Source
Strains:		
RMA2162	*S. flexneri* PE860 Y serotype; strain lacks virulence plasmid and pHS-2 plasmid	Laboratory stock
NMRM120	*S. flexneri* PE860 Δ*wzyE*$_{10\text{-}440}$	*Chapter 5*
Plasmids*:*		
pPLEO1	Source of *phoA::lacZα* C-terminal reporter fusion, Amp^r	(Taylor, Véronique L. et al. 2016)
pPLEO1-WzyE^I	pPLEO1 encoding WzyE-PhoALacZα, Amp^r	This study
pPLEO1-WzyE^{II}	pPLEO1 encoding WzyE^{a1350t}-PhoALacZα, Amp^r	This study
pBAD33	Arabinose inducible, expression vector, Cml^r	(Guzman et al. 1995)
pWzyE	pBAD33 encoding WzyE3xFLAG, Cml^r	*Chapter 4*

3.5 Article Results

3.5.1 *In silico* analysis of WzyE

Wzy proteins have been extensively studied for their role in the biosynthesis of LPS O antigen (Oag) across multiple Gram negatives: *S. flexneri, P. aeruginosa,* and *E. coli* (Islam et al. 2011; Islam et al. 2013; Nath & Morona 2015b). As ECA remains a minor surface glycolipid specific to *Enterobacteriales*, little research has been reported regarding WzyE. Consequently, we set out to investigate WzyE and to characterize *wzyE* mutants, through the use of bioinformatic, topology mapping and mutagenesis.

Initially we performed a global multiple sequence alignment (MSA) of all available WzyE peptide sequences using the Clustal Omega servers. A total of 248 WzyE sequences from across Enterobacteriales were obtained and aligned revealing that unlike WzyB proteins (Figure 3.6-(S1), WzyE is highly conserved with 105 completely conserved residues identified and that the conservation was disproportionately distributed throughout WzyE (Figure 3.1b). Principle component analysis (PCA) was then used to investigate the variance within the MSA and showed that related species clustered into distinct groups based on their pairwise similarity scores defined using a BLOSUM62 matrix (Figure 3.1c). This showed that despite the high sequence conservation, WzyEs were distinct from one-another and did show some sequence variance which was phylogenetically defined. To investigate the differences in WzyE, separate MSAs were performed with WzyE peptide sequences utilizing genera which projected the furthest from each other in the PCA. This resulted in the generation of specific consensus sequences for the genus *Dickyea, Pantoea, Yersinea, Serratia, Shigella/Escherichia and Xenorhadus.* Subsequently the peptide consensus sequences were aligned to investigate where the variance between the WzyE sequences occured (Figure 3.2).

Differences were observed with *Xenorhadus* with the presence of a large insertion sequence from AA 100-130 as was with *Dickyea* which possessed multiple ‘GGGG’ motifs at 20 and 260. However only slight differences were observed for *Pantoea, Yersinea, Serratia* and *Shigella/Escherichia* observed as multiple conservative replacements.

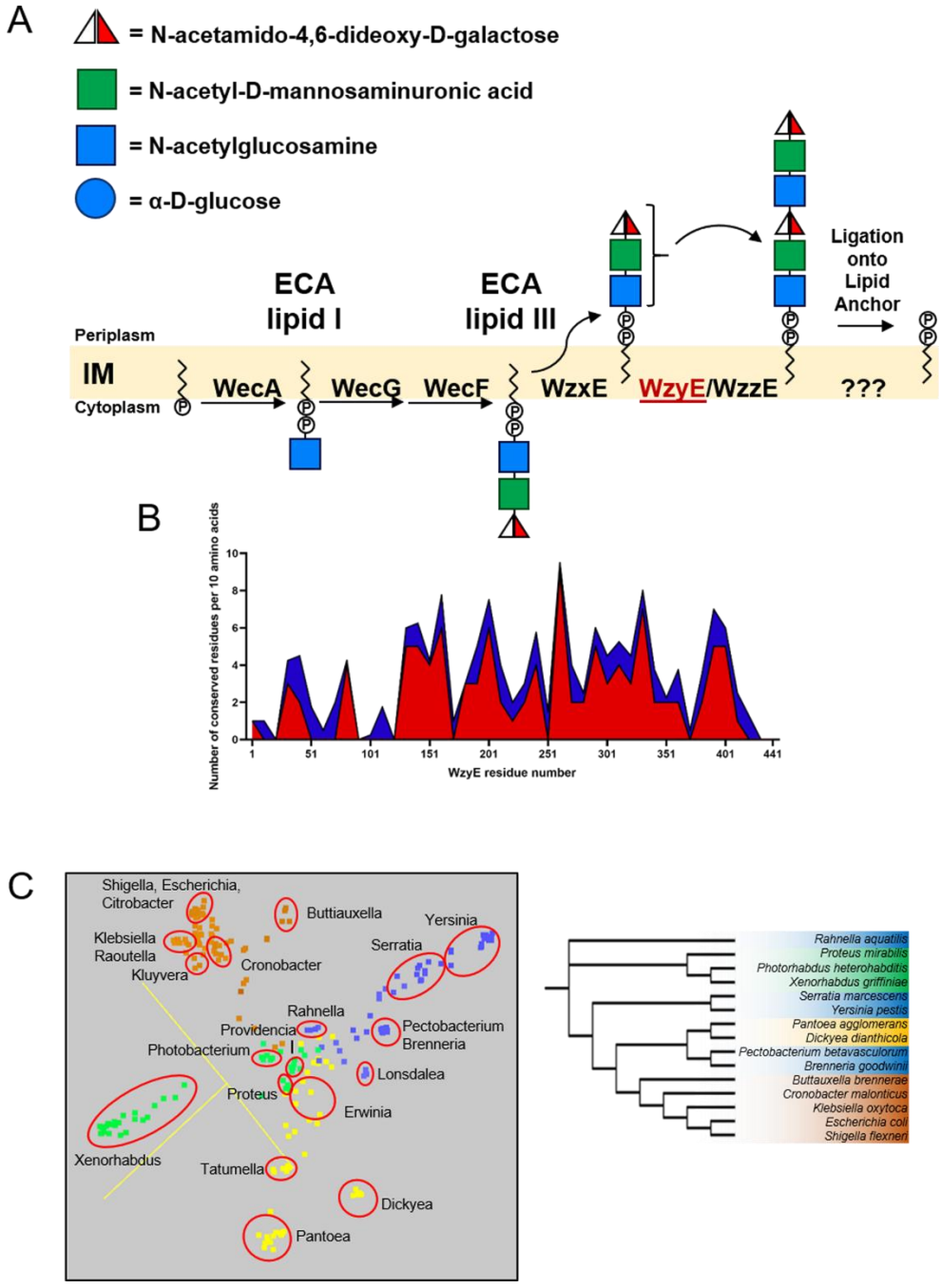


Figure 3.1: Bioinformatic analysis of WzyE peptide sequences.

(a) Simplified diagram of ECA biosynthesis in Enterobacteriales. Undecaprenyl phosphate (Und-P) is first acted upon by WecA to yield Und-PP-GlcNAcA ECA lipid-I (Al-Dabbagh, Mengin-Lecreulx & Bouhss 2008) following which the additions of ManNAcA and Fuc4NAc by the glycosyltransferases WecG and WecF which yield ECA lipid-II and ECA lipid-III. The complete ECA repeat unit (RU) which is then acted upon by WzxE which translocates the RU to the periplasmic leaflet of the IM (Eade et al. 2021). The polymerase WzyE and co-polymerase WzzE then polymerize a linear ECA polysaccharide of controlled length followed by ligation of the ECA polysaccharide onto its final lipid carrier prior to export to the outer membrane (Woodward et al. 2010). Figure adapted from (Maczuga et al. 2022). (b) Graph showing the number of conserved residues of WzyE per 10 amino acids. Peptide sequences were aligned using Clustal Omega and the number of 100% conserved residues were counted. Red shaded area = 100% conserved residues, blue shaded area = 90+ % conserved residues. (c) Principle Component Analysis (PCA) plot showing the distribution of the 248 WzyE peptide sequences. The MSA generated by Clustal Omega was inputted into Jalview where the PCA analysis was performed.

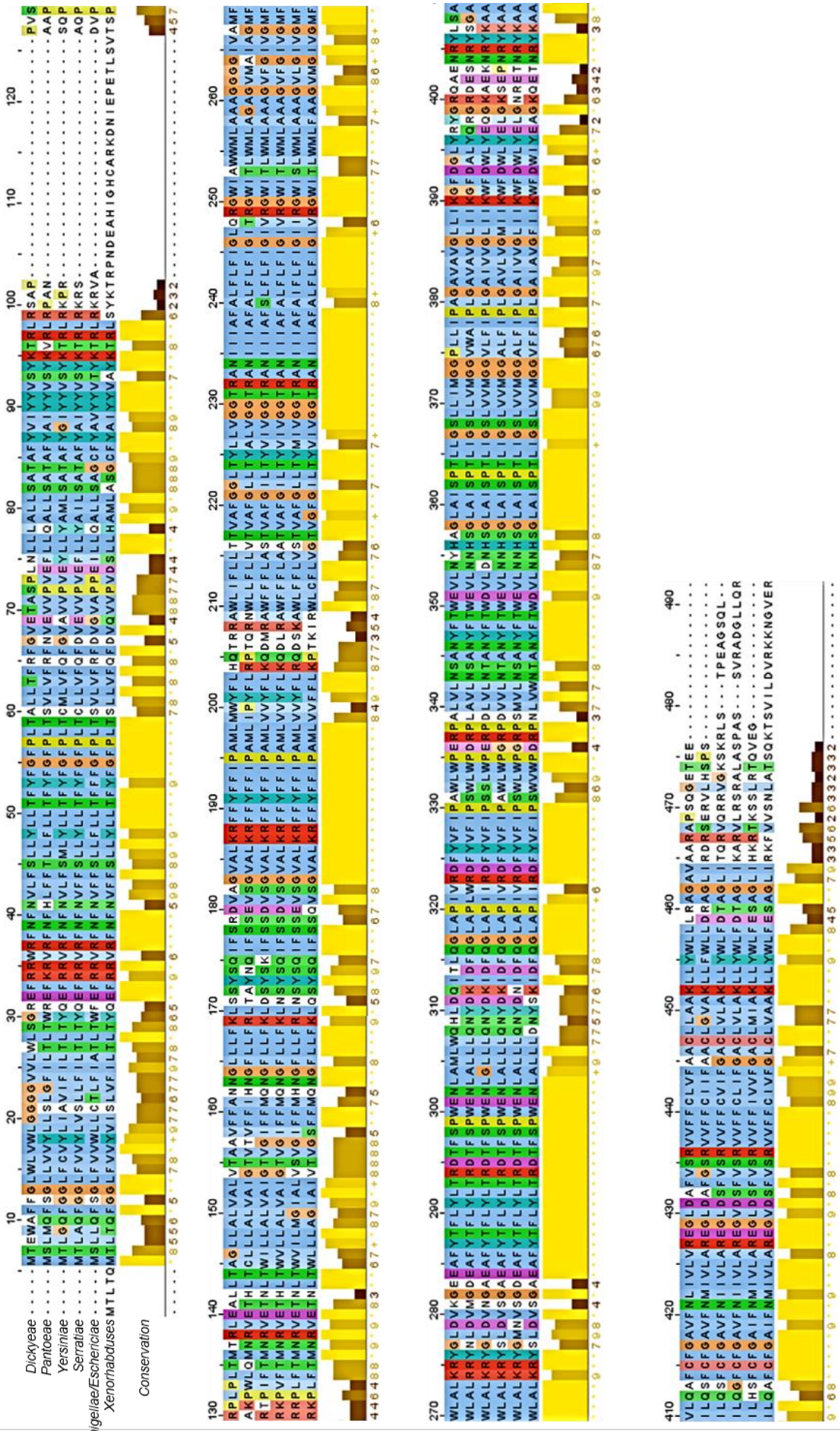

Dickeyeae
Pantoeae
Yersiniae
Serratiae
Shigellae/Escherichiae
Xenorhabduses
Conservation

Figure 3.2: Multiple sequence alignment of WzyE consensus sequences from *Dickyea, Pantoea, Yersinea, Serratia, Shigella/Escherichia and Xenorhadus.*

Separate multiple sequences alignments were generated using Clustal Omega to obtain genera specific consensus sequences for each genus. The resulting genera specific peptide sequences were aligned using Clustal Omega and residues coloured based on biochemical properties. Blue = Hydrophobic, Red = Positive charge, Magenta = Negative charge, Green = Polar, Cystines = Pink, Glycines = Orange, Prolines = Yellow, Cyan = Aromatic. Colouration intensity based on residue conservation.

To further investigate the *S. flexneri* WzyE ($WzyE_{SF}$), we generated an *in silico* predicted topology map and a structure of WzyE using the TMHMM servers and Alphafold, respectively, which was to be used as a reference point to validate the experimentally defined topology map. TMHMM predicted that WzyE consisted of 11 transmembrane segments with 5 periplasmic loops of which 2 and 4 were large and 4 cytoplasmic loops with an unusual periplasmically defined N-terminus and cytoplasmically defined C-terminus (Figure 3.3a). Alphafold predicted a tertiary structure of WzyE in which WzyE possessed 11 transmembrane segments, 5 periplasmic loops and 5 cytoplasmic loops with a topologically disordered region from AA 248 to 342 (Figure 3.3b). Additionally, like TMHMM, Alphafold predicted that $WzyE_{SF}$'s N-terminus and C-terminus were periplasmically and cytoplasmically located, respectively. As the alternating isoelectric values of the periplasmic loops of Wzy proteins have been shown to be functionally important (Islam et al. 2011), an *in silico* isoelectric point analysis of $WzyE_{SF}$'s predicted loops was performed and showed that the two largest predicted loops remain similarly charged (Figure 3.7-(S2).

3.5.2 Topology mapping of WzyE

To investigate the topology of $WzyE_{SF}$, C-terminal PhoA::LacZ reporter mapping was performed. *wzyE*$_{SF}$ was cloned into pPLEO1 such that *wzyE* reading frame became in-frame with the *phoA::lacZα* reporter fusion sequence present on the plasmid. The resulting construct pPLEO1-$WzyE^{I}$ was checked and confirmed via DNA sequencing where it became apparent that DNA methylation would inhibit the generation of a 3' overhang by *Xba*I which was required to generate nested deletions (data not shown). Site-directed mutagenesis was performed on pPLEO1-$WzyE^{I}$ to generate pPLEO1-$WzyE^{IIa1350t}$ which was subsequently confirmed by DNA sequencing. Targeted and nested C-terminal deletion were then performed on pPLEO1-$WzyE^{IIa1350t}$ to generate the $WzyE_{SF}$-PhoA::LacZ reporter library. Post identifying the point of truncations by DNA sequencing, bacterial cells harbouring these plasmids were assessed for alkaline phosphatase (PhoA) and β-galactosidase (LacZ) activity such that, the normalized activity ration (NAR) could be calculated to determine the sub-cellular localization of each reporter fusion (Karimova & Ladant 2017) (Table 3.3-(S2). An explanation to this process is described in (Figure 3.8-(S3). 110 reporter fusions were generated and enzymatically assessed to experimentally determine the topology of $WzyE_{SF}$, equalling 24.4% of the protein (Figure 3.4a). The reporter fusion mapping revealed that $WzyE_{SF}$ consists of 13 transmembrane segments, 4 periplasmic loops, 6 small cytoplasmic loops and a topologically ill-defined region between TM10 and TM11 referred to as RE1, which was reminiscent of a plausible periplasmic loop. Additionally, the reporter fusions showed that $WzyE_{SF}$ comprises of a periplasmically defined N-terminus and a cytoplasmically defined C-terminus (Figure 3.4a).

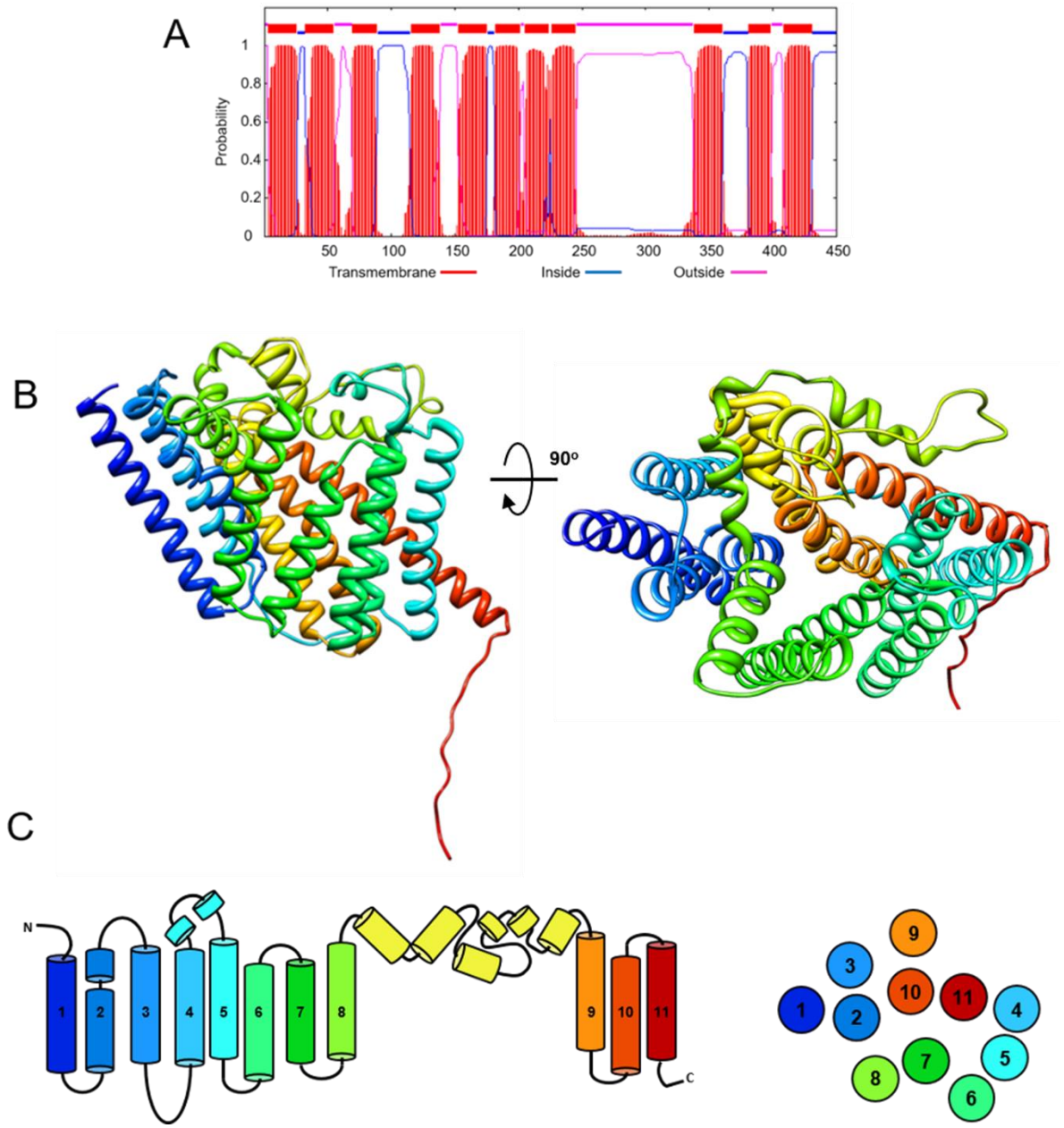


Figure 3.3: *In silico* structure analysis of WzyE.

(a) WzyE topology analysis from TMHMM. The WzyE peptide sequence was submitted to TMHMM which predicted that WzyE consisted of 11 transmembrane segments. (b) In silico predicted structure of WzyE. The WzyE structure was obtained from Alphafold and coloured blue to red from N-terminus to C-terminus respectively. (c) Cartoon representation of the topology and top down view of WzyE based on the structure obtained from Alphafold with transmembrane segments numbered 1-11. Coloured blue to red from N-terminus to C-terminus respectively.

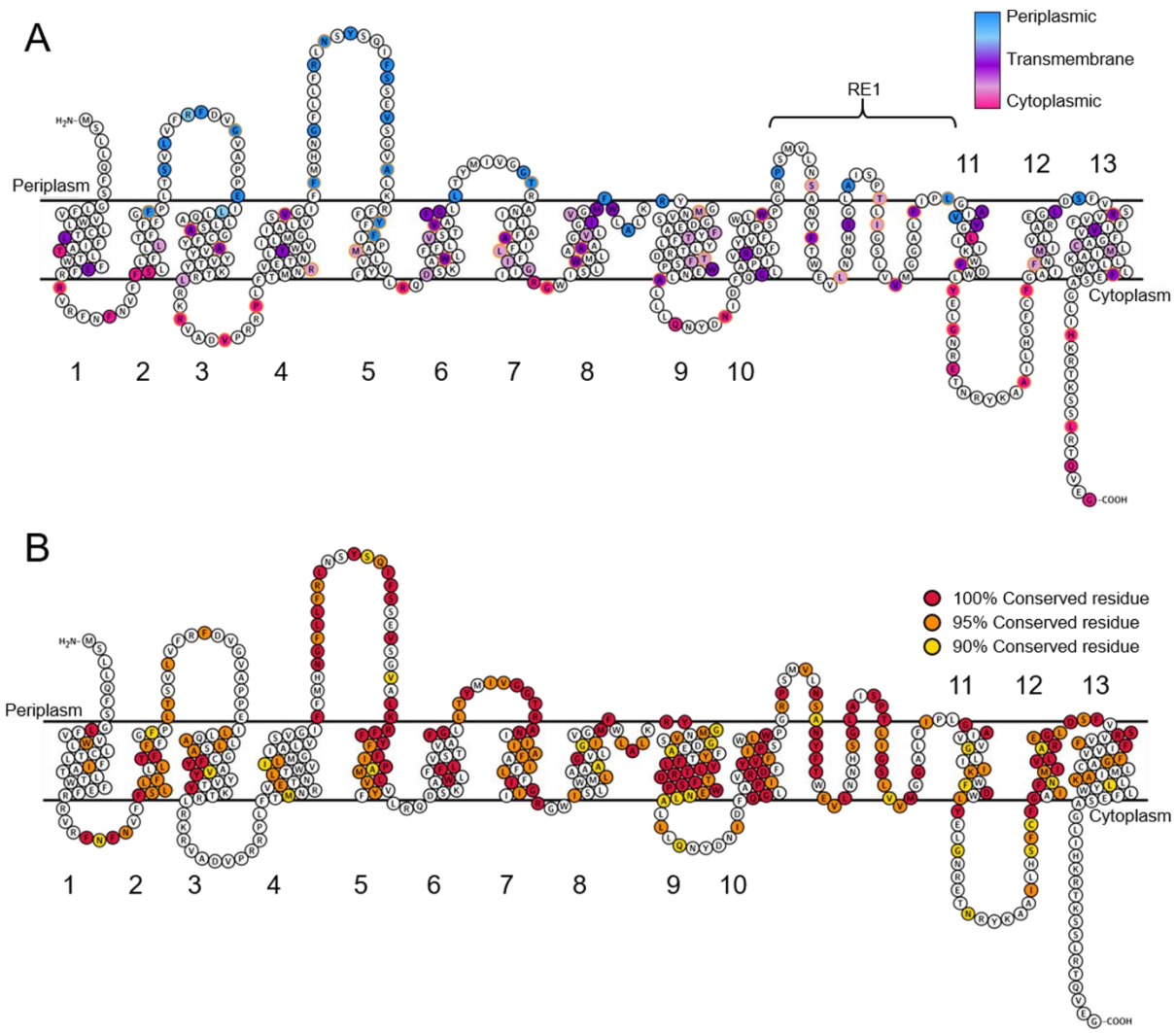


Figure 3.4: Enzymatically determined topology map of WzyE$_{SF}$ based on analysis of 110 *phoA::lacZα* reporter fusions.

(a) Topology map of WzyE$_{SF}$ based on analysis of 110 *phoA::lacZα* reporter fusions. Coloured residues denote the location of each truncation mapped. Subcellular localization of the reporter fusion at a given truncation is denoted by colour: periplasmic, blue; transmembrane, purple; cytoplasmic, red. Truncations which were generated by nested deletions have orange boarders. The representative NAR values are displayed in Table S2. The topologically ill-defined region RE1 is indicated between TM 10 and TM11. (b) WzyE amino acid conservation map using the experimentally determined topology map of WzyE$_{SF}$. Residues are coloured based on their conservation identified in the 248 MSA: red, 100 % conserved; orange, 95+ % conserved; yellow, 90+ % conserved.

Using the experimentally determined $WzyE_{SF}$ topology map, we then overlaid the conserved residues of WzyE determined earlier by MSA which showed that the conserved residues were unevenly distributed with respect to the topologically defined regions of $WzyE_{SF}$. Large clusters of conserved residues were observed in the transmembrane regions especially TM5, TM9, TM10 and TM12 as well as in the periplasmic loops PL2, PL3, PL5 with little homology within the cytoplasmic loops. Of the 105 completely conserved residues, 28 conserved residues were present in the periplasmic loops, 55 within the transmembrane regions, 5 in the cytoplasmic loops and 17 within RE1 which themselves clustered into groups (Figure 3.4b).

3.5.3 Characterization of WzyE conserved arginine residues

The regions of high conservation were indicative of functionally important regions of WzyE and as such, we decided to investigate the roles of selected residues in $WzyE_{SF}$'s function. Wzy proteins are characterized by containing catalytically conserved periplasmic arginine residues (Islam et al. 2011; Islam et al. 2013; Nath & Morona 2015b). We subsequently identified seven conserved arginine residues throughout the C-terminal half of WzyE where, according to the experimental map, three were present in periplasmic loops, R204, R247 and R309 and four were present in transmembrane segments, R266, R295, R399 and R408 (Figure 3.5a). We then referred to the *in silico* structure of WzyE from Alphafold and observed that R204, R247, R399 and R408 were predicted to form a central positively charged pocket within WzyE with the guanidinium groups of the arginines pointing inwards at close proximity (Figure 3.5b and c). This gave confidence that the conserved arginine residues may be important and as such, we preformed amino acid substitution mutagenesis of arginine to glycine to determine their importance in $WzyE_{SF}$'s function. The resulting mutants were transformed into a *S. flexneri* PE860 $\Delta wzyE_{10\text{-}440}$ mutant and assessed for their function.

Anti-ECA Western immunoblotting revealed that each substitutions prevented $WzyE_{SF}$'s ability to polymerize ECA as a lack of ECA banding was observed in all lanes (Figure 3.5d) despite WzyE3xFLAG being detectable via anti-FLAG Western immunoblotting (Figure 3.5e). Interestingly, the anti-FLAG immunoblot revealed that substitution of R309, R399 and R408 caused a secondary lower band of 22 kDa to appear indicating that that these residues may also be critical in maintaining the stability of $WzyE_{SF}$.

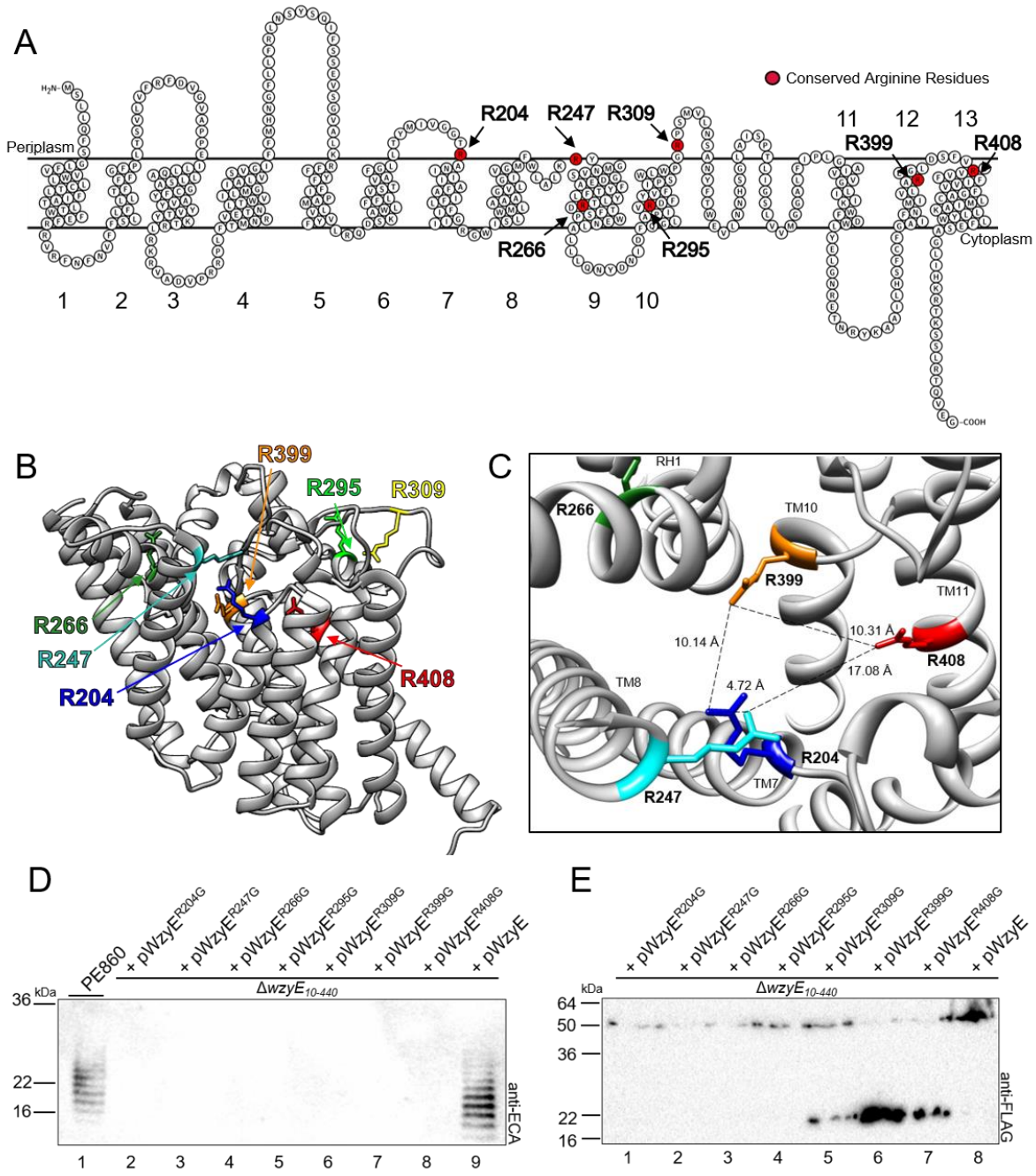


Figure 3.5: Analysis of conserved arginine residues of WzyE.

(a) Location of the seven conserved arginine residues displayed on a topological map. (b) Alphafold structure of WzyE showing the location of the seven conserved arginine residues. (c) Location of the positively charged central pocket in WzyE coordinated by R204, R247, R399 and R408. Atomic distance measured using Chimera software measuring from NH+ group of the guanidinium group of each arginine residue. (d) Anti-ECA Western immunoblot showing the ECA profiles of PE860 and PE860 Δ*wzyE*$_{10\text{-}440}$ denoted as (Δ*wzyE*$_{10\text{-}440}$) complemented with pWzyE^{R204G}, pWzyE^{R247G}, pWzyE^{R266G}, pWzyE^{R295G}, pWzyE^{R309G}, pWzyE^{R399G}, pWzyE^{R408G} and pWzyE. Mid-exponential phase cells were collected by centrifugation (1x10^{9} cells) and lysed with lysis buffer in the presence of Proteinase K. Samples were then electrophoresed on a SDS-15% (w/v) PAGE gel, transferred onto a nitrocellulose membrane and probed with polyclonal rabbit anti-ECA antibodies. (e) Anti-FLAG Western immunoblot of the above strains. Mid-exponential phase cells were collected by centrifugation (5x10^{8} cells) and lysed with 2x sample buffer. Samples were then electroporated on a SDS-15% (w/v) PAGE gel, transferred onto a nitrocellulose membrane and probed with monoclonal mouse anti-FLAG antibodies. SeeBlue Plus2 pre-stained protein ladder was used as a molecular mass standard.

3.6 Article Discussion

The majority of research performed investigating Wzy proteins have been orientated towards Wzy's role in the biosynthesis of LPS Oag (Islam et al. 2011; Islam et al. 2013; Islam et al. 2010; Nath & Morona 2015a, 2015b). Due to this, Wzys which have been topologically determined through an experimental approach; $WzyB_{SF}$, Wzy_{PAO1} and Wzy_{FT}, all belong to the LPS Oag biosynthesis homolog of the Wzy-dependent pathway (Daniels, Vindurampulle & Morona 1998; Islam et al. 2010; Kim et al. 2010). Here we report the first major study into WzyE, the Wzy-dependent homolog from the ECA biosynthetic pathway.

We showed that unlike Wzys involved in LPS Oag biosynthesis, WzyEs are highly conserved across all Enterobacterales (Figure 3.1b). Using C-terminal reporter fusions, we experimentally mapped 24.8% of $WzyE_{SF}$'s residues and showed that again, unlike other Wzys, WzyE consists of 13 transmembrane segments, 4 periplasmic loop, 6 cytoplasmic loops and a topologically ill-defined region (Daniels, Vindurampulle & Morona 1998; Islam et al. 2010; Kim et al. 2010) (Figure 3.4a). Combining the MSA and the experimental mapping data, we showed that the majority of the conserved residues observed were topologically defined, falling largely into the transmembrane segments and periplasmic loops of WzyE, with little conservation present in the cytoplasmic loops (Figure 3.4b). As WzyE polymerizes ECA from biosynthetic intermediates present on the periplasmic leaflet of the IM (Rai & Mitchell 2020), it is not surprising that the periplasmic loops and transmembrane segments of the protein hold the majority of the conservation.

Likewise, due to the variability of LPS Oag found across serotypes (Liu et al. 2008) and species (Wang, Wang & Reeves 2010), the variance observed between Wzy proteins involved in the biosynthesis of Oag is not surprising. Its reasonable to hypothesise that the peptide sequences and structures of Wzy proteins, have evolved to polymerize Oag repeat units (RUs) specific to different Gram-negative species and strains. Unlike Wzys involved in Oag biosynthesis, WzyE polymerizes ECA, which is conserved among Enterobacterales (Hella-Monika Kuhn 1988). Given the structure of ECA does not vary (Eade 2021), it suggests that the lack of evolutionary pressure to evolve the protein allows for WzyE to remain highly conserved and specialized. Consequently, WzyE may be the optimal Wzy protein to investigate Wzy-dependent biosynthesis and the distinct features of Wzys that are important for their function.

Wzy proteins belong to the Shape, Elongation, Division and Sporulation (SEDS) protein family (Meeske et al. 2016). This family consist of glycosyltransferases involved in cell wall biosynthetic processes such as FtsW, RodA and WaaL (Meeske et al. 2016; Sjodt et al. 2018;

Taguchi et al. 2019). Until recently, the majority of SEDS proteins were topologically characterized through the use of reporter fusion mapping where in general, they were found to possess multiple transmembrane helices with functionally important periplasmic loops which contained conserved residues (Daniels, Vindurampulle & Morona 1998; Gérard, Vernet & Zapun 2002; Islam et al. 2010).

Reporter mapping showed that the topology of $WzyE_{SF}$ is quite different from other Wzy proteins whose topology has been experimentally determined. While the number of transmembrane segments, periplasmic and cytoplasmic loops are known to vary among Wzys (Daniels, Vindurampulle & Morona 1998; Islam et al. 2010), the periplasmically location of the N-terminal domain is novel as is the presence of the topologically ill-defined region, RE1. The reporter-based topology map did not completely agree with TMHMM or Alphafold, that both predicted that $WzyE_{SF}$ possessed 11 transmembrane segments (Figure 3a and b and Figure 4a). The region RE1 however, did coincide with part of the small periplasmic secondary structures exhibited in Alphafold's predicted structure between TM8 and TM9. As this region of $WzyE_{SF}$ may consist of a highly dynamic region, it is plausible that C-terminal reporter fusions are unable to accurately report the topology at that specific point. This highlights potential issues with C-terminal reporter fusion mapping, as they may be inherently misleading due to the assumptions to structure of the remaining protein. It is plausible that large deletions of segments may not provide the native topology, as it is possible that some of the segments deleted may play important structural roles in maintaining the correct topology (Cymer, von Heijne & White 2015). This is demonstrated with membrane proteins as residues which are apart based on peptide sequence, interact with one another once the protein is folded (Li et al. 2021). Another shortcoming is the lack of consideration towards secondary structures and their impact on topology, as the protein itself does not exist as merely loops and transmembrane segments but as folded structures. A suitable alternative to C-terminal reporter fusions is the less disruptive probing of the full-length proteins using thiol probes such as PEG maleimide, which has been used multiple times to determine the topology of proteins (Howe & Brown 2017; Howe et al. 2015; Lamothe 2013; Tavares-Carreon et al. 2019).

Previous studies have indicated that SEDS proteins possess conserved residues and peptides in their periplasmic loops where, R215, R288 and H338 in WaaL and E240-A249 in FtsW were shown to be crucial in Oag ligation and Z ring formation respectively (Pastoret et al. 2004; Ruan et al. 2012). This is likewise observed in Wzy proteins as crucial periplasmic residues have been identified. *Islam et al. (2013*) and *Nath et al (2015)* showed that residues from the largest two periplasmic loops of Wzy_{PAO1} and Wzy_{SF}, respectively, were critical for the function of the protein

as their substitutions resulted in abnormal, or complete loss of Oag polymerization (Islam et al. 2011; Islam et al. 2013; Nath & Morona 2015b).

Using the reporter map with our initial MSA, we identified highly conserved arginine residues throughout the C-terminal half of $WzyE_{SF}$. These residues were then located in the Alphafold *in silico* structure which revealed that R204, R247, R399 and R408 potentially formed a central positively charged pocket within $WzyE_{SF,}$ with their guanidinium groups in close proximity to each other. Close to this pocket was R266 which was located on a re-entrant helix, as well as R295 and R309 which were present on nearby on loops (Figure 3.5b and c).

As ECA is negatively charged due to the incorporation of N-acetyl-D-mannosaminuronic acid (ManNAcA) in each RU (Rai & Mitchell 2020), we hypothesised that these arginine residues may be involved in the binding of both the RU as well as the growing polysaccharide chain. Individual substitutions of both the periplasmic and transmembrane arginine residues resulted in the complete loss of ECA polymerization which suggests, that like other SEDS proteins, the periplasmic arginine residues of $WzyE_{SF}$ are crucial for its function (Figure 3.5d). The single substitutions of R309G, R399G and R408G also resulted in a protein band being detected at 22 kDa which further suggested that arginine substitutions within the latter half of $WzyE_{SF}$ were crucial to main the stability of the protein. This may be plausible as R399 and R408 are predicted to be in close proximity to R204 and R247 (Figure 3.5e).

This study is the first to investigate a WzyE protein which is essential for ECA biosynthesis. Through experimental topology mapping, *in silico* analysis and MSAs we showed that WzyE is unique among Wzy proteins. High sequence conservation across a broad range of species allowed us to identify conserved arginine residues which we showed to be crucial for the function of $WzyE_{SF}$ and predict that they may form a binding pocket. We believe that due to the high peptide sequence conservation, WzyE may be the optimal Wzy protein to study mechanisms the Wzy proteins as well as the Wzy-dependent pathway, and shed further light onto these important proteins.

3.7 Article References

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008.

Cymer F, von Heijne G, White SH. 2015. Mechanisms of integral membrane protein insertion and folding. *J Mol Biol* 427:999-1022.

Daniels C, Vindurampulle C, Morona R. 1998. Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). *Mol Microbiol* 28:1211-22.

Eade CR, Wallen TW, Gates CE, Oliverio CL, Scarbrough BA, Reid AJ, Jorgenson MA, Young KD, Troutman JM. 2021. Making the enterobacterial common antigen glycan and measuring its substrate sequestration. *ACS Chem Biol 16* ,691-700.

Gérard P, Vernet T, Zapun A. 2002. Membrane topology of the *Streptococcus pneumoniae* FtsW division protein. *J Bacteriol* 184:1925-1931.

Günther SD, Fritsch M, Seeger JM, Schiffmann LM, Snipas SJ, Coutelle M, Kufer TA, Higgins PG, Hornung V, Bernardini ML, Höning S, Krönke M, Salvesen GS, Kashkar H. 2019. Cytosolic Gram-negative bacteria prevent apoptosis by inhibition of effector caspases through lipopolysaccharide. *Nat Microbiol* 5:354-367.

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

Hella-Monika Kuhn UM-D, Hubert Mayer. 1988. ECA, the enterobacterial common antigen. *FEMS Microbiol Rev*:195-222.

Howe V, Brown AJ. 2017. Determining the topology of membrane-bound proteins using PEGylation. *Methods in Mol Biol* 1583:201-210.

Howe V, Chua NK, Stevenson J, Brown AJ. 2015. The regulatory domain of squalene monooxygenase contains a re-entrant loop and senses cholesterol via a conformational change. *J Biol Chem* 290:27533-27544.

Islam ST, Gold AC, Taylor VL, Anderson EM, Ford RC, Lam JS. 2011. Dual conserved periplasmic loops possess essential charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in *Pseudomonas aeruginosa* PAO1. *J Biol Chem* 286:20600-5.

Islam ST, Huszczynski SM, Nugent T, Gold AC, Lam JS. 2013. Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1. *Sci Rep* 3:3441.

Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can J Microbiol* 60:697-716.

Islam ST, Taylor VL, Qi M, Lam JS. 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:e00189-10.

Jorgenson MA, Kannan S, Laubacher ME, Young KD. 2016. Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate. *Mol Microbiol* 100:1-14.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583-589.

Karimova G, Ladant D. 2017. Defining membrane protein topology using *pho-lac* reporter fusions, Bacterial Protein Secretion Systems: *Methods and Protocols* 129-142.

Kim TH, Sebastian S, Pinkham JT, Ross RA, Blalock LT, Kasper DL. 2010. Characterization of the O-antigen polymerase (Wzy) of *Francisella tularensis*. *J Biol Chem* 285:27839-49.

Kozlowski LP. 2016. IPC – Isoelectric Point Calculator. *Biol Direct* 11:55.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567-80.

Lamothe Sp. 2013. PEGylation as a novel tool to investigate the topology of *Escherichia coli* WecA, a membrane enzyme involved in lipopolysaccharide O antigen initiation.

Leo V, Teh MY, Tran ENH, Morona R. 2021. Identification of a region in *Shigella flexneri* WzyB disrupting the interaction with Wzz(pHS2). *J Bacteriol* 203:e0041321.

Li F, Egea PF, Vecchio AJ, Asial I, Gupta M, Paulino J, Bajaj R, Dickinson MS, Ferguson-Miller S, Monk BC, Stroud RM. 2021. Highlighting membrane protein structure and function: A celebration of the Protein Data Bank. *J Biol Chem* 296:100557.

Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SyN, Wang Q, Reeves PR, Wang L. 2008. Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* 32:627-653.

Lugtenberg B, Meijers J, Peters R, van der Hoek P, van Alphen L. 1975. Electrophoretic resolution of the 'major outer membrane protein' of *Escherichia coli* K12 into four bands. *FEBS Let* 58:254-258.

Maczuga, N, Tran, ENH, Qin, J & Morona, R 'Interdependence *of Shigella flexneri* O antigen and enterobacterial common antigen biosynthetic pathways', *J Bacteriol* e00546-00521.

Meeske AJ, Riley EP, Robins WP, Uehara T, Mekalanos JJ, Kahne D, Walker S, Kruse AC, Bernhardt TG, Rudner DZ. 2016. SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nat* 537:634-638.

Mitchell AM, Srikumar T, Silhavy TJ. 2018. Cyclic enterobacterial common antigen maintains the outer membrane permeability barrier of *Escherichia coli* in a manner controlled by YhdP. *mBio* 9:e01321-18.

Murray GL, Attridge SR, Morona R. 2003. Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol Microbiol* 47:1395-1406.

Nath P, Morona R. 2015. Detection of Wzy/Wzz interaction in *Shigella flexneri*. *Microbiology* 161:1797-805.

Nath P, Morona R. 2015. Mutational analysis of the major periplasmic loops of *Shigella flexneri* Wzy: identification of the residues affecting O antigen modal chain length control, and Wzz-dependent polymerization activity. *Microbiology* 161:774-85.

Nath P, Tran EN, Morona R. 2015. Mutational analysis of the *Shigella flexneri* O-antigen polymerase Wzy: identification of Wzz-dependent Wzy mutants. *J Bacteriol* 197:108-19.

Pastoret S, Fraipont C, den Blaauwen T, Wolf B, Aarsman MEG, Piette A, Thomas A, Brasseur R, Nguyen-Distèche M. 2004. Functional analysis of the cell division protein FtsW of *Escherichia coli*. *J Bacteriol* 186:8370-8379.

Purins L, Van Den Bosch L, Richardson V, Morona R. 2008. Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2. *Microbiology* 154:1104-1116.

Rai AK, Mitchell AM. 2020. Enterobacterial common antigen: Synthesis and function of an enigmatic molecule. *mBio* 11:e01914-20.

Ramos-Morales F, Prieto AI, Beuzon CR, Holden DW, Casadesus J. 2003. Role for *Salmonella enterica* enterobacterial common antigen in bile resistance and virulence. *J Bacteriol* 185:5328-32.

Ruan X, Loyola DE, Marolda CL, Perez-Donoso JM, Valvano MA. 2012. The WaaL O-antigen lipopolysaccharide ligase has features in common with metal ion-independent inverting glycosyltransferases. *Glycobiology* 22:288-99.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mole Syst Biol*.

Sjodt M, Brock K, Dobihal G, Rohs PDA, Green AG, Hopf TA, Meeske AJ, Srisuknimit V, Kahne D, Walker S, Marks DS, Bernhardt TG, Rudner DZ, Kruse AC. 2018. Structure of the peptidoglycan polymerase RodA resolved by evolutionary coupling analysis. *Nat* 556:118.

Taguchi A, Welsh MA, Marmont LS, Lee W, Sjodt M, Kruse AC, Kahne D, Bernhardt TG, Walker S. 2019. FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein. *Nat Microbiol* 4:587-594.

Tavares-Carreon F, Ruan X, Ford A, Valvano MA. 2019. Sulfhydryl labeling as a tool to investigate the topology of membrane proteins involved in lipopolysaccharide biosynthesis. *Methods Mol Biol* 1954:203-213.

Taylor VL, Hoage JFJ, Thrane SW, Huszczynski SM, Jelsbak L, Lam JS. 2016. A bacteriophage-acquired O-antigen polymerase (Wzyβ) from *P. aeruginosa* serotype O16 performs a varied mechanism compared to its cognate Wzyα. *Frontiers Microbiol* 7:393-393.

Wang L, Wang Q, Reeves PR. 2010. The variation of O antigens in gram-negative bacteria. *Subcell Biochem* 53:123-52.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189-91.

3.8 Article Supporting Information

3.8.1 Supporting Figures

WzyB peptide sequence alignment:

```
E. coli        MYILSMAIPLFFISLCWHRFSHQTQVAFRLLFIVFFSLSFFSYYTNGEDWSVYYLRFIED  60
S. typhi       ------------------------------------------------------------  0
S. flexneri    ------------------------------------------------------------  0

E. coli        EGLFTSFEFGFVIILKLLLIVS-----DDNFGLAILLYYFLCFVLLSLILKKYKVNEPLF 115
S. typhi       ------------------------------MLIISYIALCLLFIVYL---YTLSVRIE  25
S. flexneri    ---MNNINKIFITFLCIELIIGGGGRLLEPLGIFPLRYLLFVFSFILLIFNLVTFNFSIT  57
                                             :  : *  : : ::  :    ...  :

E. coli        LGCLLLLF--------------GYNLFLEQLRQ-------LLACIIVFYA----MLLYNQ 150
S. typhi       GKIINVMVPYLIITVPTLYVFEGIFVYLSEVQNYTVEYLFFYTCYITYIASFVISYLYTQ  85
S. flexneri    QKCVSLFIWL------LLFPFYGFFVGLLAGNK-INDILFDVQP---YLFMLSLIYLFTL 107
                  : ::.             *  : *   .:             :        *:.

E. coli        NKNL-------IESCILVITA-STFHVS-AVIILPTIFLISFRNVTTFIIITVSSITSIV 201
S. typhi       RKPIYNKSNTKNKPRYVFTSL--LFTFLAFIIYLPVLMEFREYILSPRRIYELTRTG--- 140
S. flexneri    RYTLKVFS------CEIFIKIVNAFALYGSLLYISYIILLNFGLLNFNLIYEHLSLTSEF 161
               .  :           :. .    * .   :: :  :: :    :.   *

E. coli        -----------VFMIALYSAISS--------LAAMS-FVFAKIDYYLHQNPVVLNFGWLN 241
S. typhi       ----YGIYFYPSLMFSLVASICAFFTYKKSKLFCISIVLFNCILIFLHGNKGPIFSIFIA 196
S. flexneri    FFRPDGAFFSKSFYFFGVGAIISFVD--KKYLKCLIIVL----AILLTESRGVLLFTTLS 215
                           : :   .:* :        * .:  .:       *  .   :    :

E. coli        ILDALFILFYMRYRNAIDRQININIL-TRIIFVGSVIHLFSGTITFLARVCFYFYFVGVY 300
S. typhi       -----FIL-YLSYI--ENKKIKFMFLVKSFAVIAVIVTAF----------FAYTFTDGNP 238
S. flexneri    -------L-LLASF-----KLHKLYLNTIIIILGSVL--F----------IIMLYMVGSR 250
                     *  :        :::   * . : .:. ::  *             :  *

E. coli        I-----------------------------YCYSLNDKTPRVFAI---KNFNTLVLAAFV 328
S. typhi       IENMANYSDYTRNAVLVASSNFDFMYGKLLMESEVYSRIPRAIWPDKPEDFGALYLAK-- 296
S. flexneri    S---------------------------------------------EDSDSVRFND-- 261
                                                           :: .:: :

E. coli        NVVLLLNFASYFRSEQAPVGFYNMNFQFFNTFNDNYLRTN-------------------- 368
S. typhi       ----VFFPDAFYRNQGAPAFGYGELYADFGLFT---------P-----------VWLVIS 332
S. flexneri    ----LY---FYYKNVDLATFLFGRGFGSFILDR---LRIEIVPLEILQKTGVIGVFISLV 311
                   :     :::.    .  :.  :  *

E. coli        -----AYDKYFKGMQGNVD---------------------------------------- 382
S. typhi       GVFKGVLAKYFSNKTQETKSAHYFIMFL-----------FCIGISVIPVSMGWLFPEHLM 381
S. flexneri    PMLLIFLKGYFLNSTKTSLMMSLILFFSITVSITNPFLFTPMGIFIIGVVVLWVFSIENI 371
                        ** .

E. coli        -------------------------- 382
S. typhi       IAFMVYIASSFVFSEHIRFVLLRNNK 407
S. flexneri    QI-----SNN----------LTSGAK 382
```

Figure 3.6-(S1): WzyB peptide sequence alignment.

WzyB peptide sequences were obtained from NCBI and aligned using Clustal Omega revealing the inability to globally align three WzyB peptide sequences from closely related species. * = identical amino acid; : = similar biochemical properties; . = loosely similar biochemical properties.

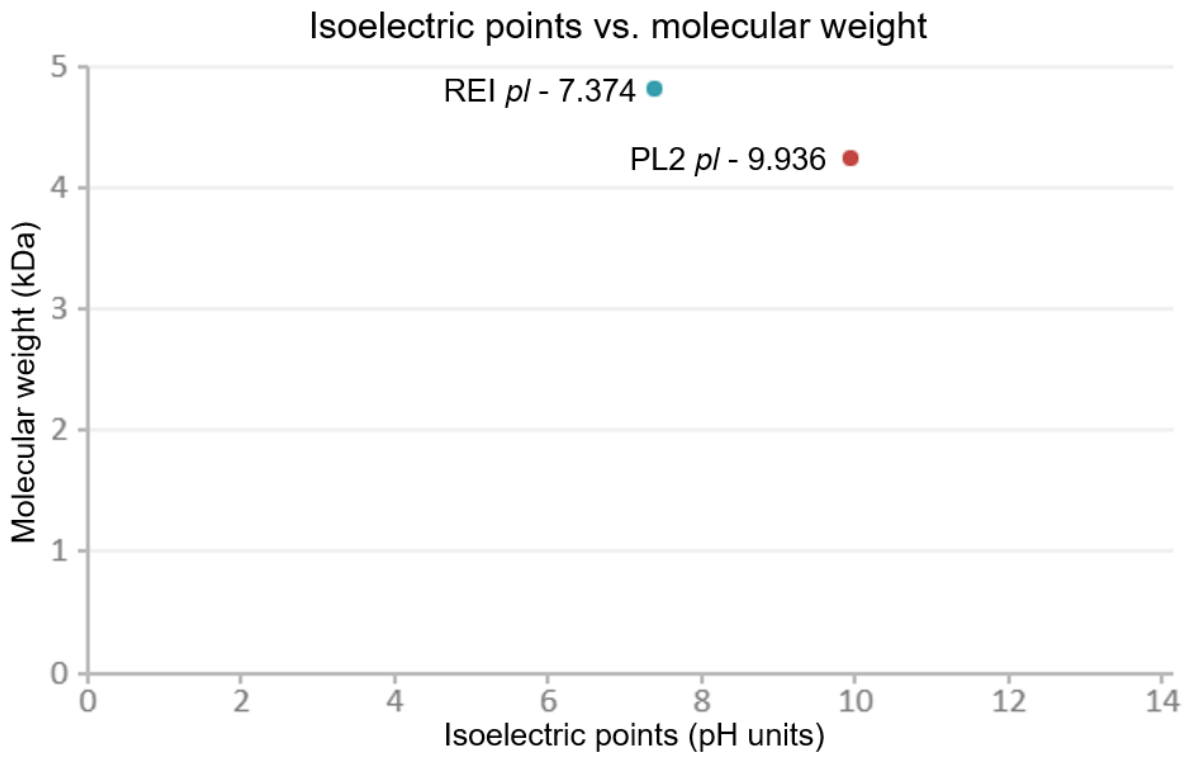
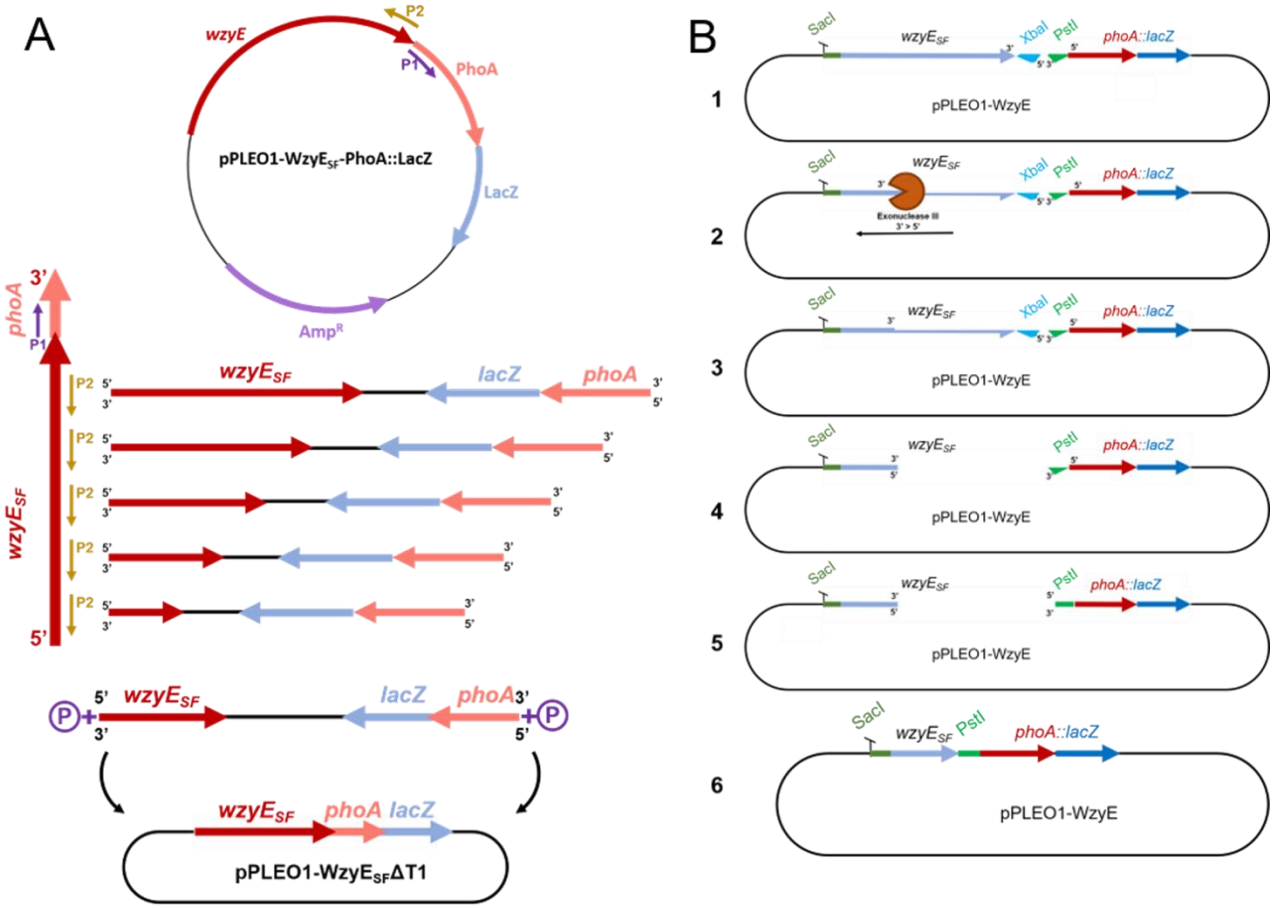


Figure 3.7-(S2): Isoelectric point analysis of WzyE experimentally determined periplasmic loops.

WzyE peptide sequences of the topologically defined regions PL2 and RE1 were submitted to IPC calculator and the output displayed in silico predicted isoelectric points of both regions at pH 7.4.


A
wzyE
P2
P1
PhoA
LacZ
Amp^R
pPLEO1-WzyE_SF-PhoA::LacZ
phoA
3'
P1
wzyE_SF
lacZ
phoA
P2
5'
pPLEO1-WzyE_SFΔT1
B
SacI
wzyE_SF
XbaI
PstI
phoA::lacZ
pPLEO1-WzyE
Exonuclease III
3' > 5'
1
2
3
4
5
6

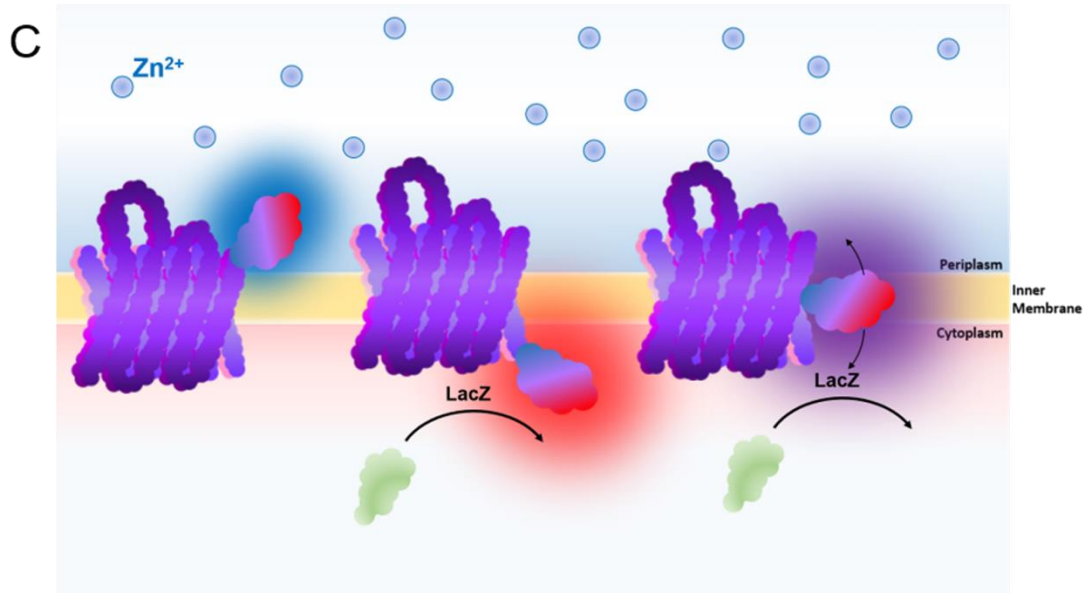

C
Zn2+
Periplasm
Inner Membrane
Cytoplasm
LacZ
LacZ

Figure 3.8-(S3): Generation of C-terminal deletions and mechanism of PhoA::LacZ C-terminal reporter fusions.

(a) Generation of targeted C-terminal reporter fusions. Upon cloning in the gene of interest, DNA primers are designed in an orientation as such that they PCR amplify the backbone vector. A universal P1 primer is designed to amplify the stop codon + 20~bp. Multiple P2 primers are then designed in reverse orientation along the gene of interest ensuring that any deletions made do not induce frameshift mutations. Post PCR, the linear DNA fragment is treated with poly nucleotide kinase and DNA ligase. (b) Generation of nested C-terminal reporter fusions. Upon cloning in the gene of interest, restriction enzymes are used to produce two DNA overhangs in both 3'>5' and 5'>3' directions between the gene of interest and the reporter fusion (1). Incubation with Exonuclease III results in the digestion of the DNA in a 3'>5' manner resulting in a random library of C terminal deletions (2). Post quenching (3), Mung bean nuclease (4) and DNA PolI (5) are used to remove all remaining DNA overhangs prior to treatment with DNA poly nucleotide kinase and DNA ligase (6). (c) Diagram explaining C-terminal reporter fusion activity. The dual reporter fusion displays dual enzyme activity dependent on the subcellular localisation of the fusion. If located in the periplasm, in the presence of Zn2+ PhoA is active where activity can be measured directly through alkaline phosphatase assays or through colorimetric means in which plated cells with PhoA activity cleave 5-bromo-4-chloro-3-indolyl phosphate (BCIP) on supplemented indicator plates resulting in a blue colony pigmentation. If located in the cytoplasm, the alpha fragment of LacZ recruits the remaining LacZ machinery whose activity can be measure directly through β-galactosidase assays or through colorimetric means in which plated cells with LacZ activity cleave 56-chloro-3-indolyl-β-D-galactoside (Red-Gal) on supplemented indicator plates resulting in a red colony pigmentation. If the reporter fusion in an transmembrane segment then the reporter shows both PhoA and LacZ activity which can be assessed directly, individually, as above and as such, colonies with both PhoA and LacZ activity produce purple pigmentation. The activity of the fusions can be assessed for both PhoA and LacZ activity where, the ratio of activity between the two is used to infer topological localization; this value is called the normalized activity ratio (NAR).

3.8.2 Supporting Tables

Table 3.2-(S1): DNA Oligonucleotides used in this study.

Primer	Oligonucleotide sequence (5'-3')	Target
Construct generation specific primers		
NM26	aatcttctagatccttcaacctgcgtccg	*wzyE* gene
NM27	tcaaagagctcatgagtctgctgcaattcagtg	*wzyE* gene
Site-directed mutagenesis specific primers		
NM28	gcgaggttgaagggtctaggactagtgg	*wzyE* gene a1350t
NM29	ccactagttctagacccttcaacctgcg	*wzyE* gene a1350t
NM153	atatgattgtcggcggcactggcgccaatatc	*wzyE* gene c610g
NM154	gataatggcgccagtgccgccgacaatcatat	*wzyE* gene c610g
NM155	gcatgttctggctggcactaaaaggctatggaatga	*wzyE* gene c739g
NM156	tcattccatagccttttagtgccagccagaacatgc	*wzyE* gene c739g
NM157	tctatacgtttctctatctcactggcgacaccttct	*wzyE* gene c796g
NM158	agaaggtgtcgccagtgagatagagaaacgtataga	*wzyE* gene c796g
NM159	gcctggctccaattgttcggcgatttctatgtcttt	*wzyE* gene c883g
NM160	aaagacatagaaatcgccgacaattggagccaggc	*wzyE* gene c883g
NM161	ggctgtggccgggtggcccgagt	*wzyE* gene c925g
NM162	actcgggccacccggccacagcc	*wzyE* gene c925g
NM163	atcgtgctggcgggtgaagggctgg	*wzyE* gene c1195g
NM164	ccagcccttcacccgccagcacgat	*wzyE* gene c1195g
NM165	gggctggattcgtttgtctcaggcgtggtcttt	*wzyE* gene c1222g
NM166	aaagaccacgcctgagacaaacgaatccagccc	*wzyE* gene c1222g
Targeted C-terminal fusion primers		
NM48	tctgatcacccgttaaacggcgag	*phoA* gene
NM30	gttctggaaaaccgggctgctc	*phoA::lacZα reporter*
NM31	tacatcagcaacgcgtttgcgtag	WzyE V100 truncation
NM32	ggtaagccggaacagcaaaaagcc	WzyE N143 truncation
NM33	gaccgtgctgacgaggaaaaacag	WzyE V190 truncation
NM34	aggtgtcgcgagtgagatagagaaac	WzyE T268 truncation
NM35	cagtacttcccaggtaaagtaattggctg	WzyE L325 truncation
NM36	gcccagctcatacagccagtcg	WzyE G371 truncation
NM37	acggcgaaactcaaaccaggtc	WzyE R30 truncation
NM38	accaacctcaaagcgaaataccagcac	WzyE G64 truncation
NM39	cgcgtagaagcagcccgc	WzyE A83 truncation
NM40	gaagaagatgccgacgcttaccagc	WzyE F132 truncation
NM41	gtcctggcgcagaaagtagaccac	WzyE D178 truncation
NM42	agtgccgccgacaatcatataagtcag	WzyE T203 truncation
NM43	gccgcgaataatgccaataaacaggaag	WzyE G222 truncation
NM44	cattccatagcgttttaacgccagcc	WzyE M250 truncation
NM45	gttgtcgtagttctgcaacagcaacg	WzyE N284 truncation
NM46	tataagcgtaggcgagatcgccagtc	WzyE I338 truncation
NM47	tccggcgctttcaaaaagccagta	WzyE G444 truncation
NM49	gagtgggatgaacaacgcgcc	WzyE L352 truncation
NM50	ggagatcgccccaaagcagaaactgt	WzyE F392 truncation
NM51	cgaatccagcccttcacgtgc	WzyE S404 truncation

NM52	agcgaatgcgatgatgatattggcg	WzyE A212 truncation
NM53	cgccgccagcatccaccac	WzyE A231 truncation
NM68	gaagccgaagaaaaaggtgagcaaaaacag	WzyE F51 truncation
NM69	cgcagaaagcaacacctgcaacaag	WzyE A78 truncation
NM70	taggcgggttttgtaggtgacatagtaaac	WzyE L93 truncation
NM71	cggacggcgcggtcaatcag	WzyE P104 truncation
NM72	gcggttcatggtaaacagcggac	WzyE R110 truncation
NM73	gacgcttaccagcgcgataccc	WzyE V128 truncation
NM74	cgccacgccggagacttca	WzyE A157 truncation
NM75	gtaaaagaagcgttttaacgccacgcc	WzyE Y163 truncation
NM76	catcgccgggatgaaaaagtaaaagaagc	WzyE M169 truncation
NM77	cgccagattctcccacggtgaa	WzyE A276 truncation
NM78	ccacagccaggaagggataaagacatag	WzyE W306 truncation
NM79	gaaccatttgatgatcagtccaaccacaatc	WzyE F364 truncation
NM80	aaagcagaaactgtgcaatatcgcagc	WzyE F388 truncation
NM81	gcgtgagacaaacgaatccagccc	WzyE R408 truncation
NM82	aaaaagccagtacaacagttttgcgatcatc	WzyE F430 truncation
NM83	gcgtttgcgtaggcgggttt	WzyE R96 truncation
NM84	gaaaaagtaaaagaagcgttttaacgccacgc	WzyE F165 truncation
NM85	aaacaggaagatagcgaatgcgatgatgata	WzyE F216 truncation
NM86	ccacaacgaaatccagccgcg	WzyE W227 truncation
NM87	tgagttcagcaccatactcgggc	WzyE S316 truncation
NM88	aaagtaattggctgagttcagcaccatact	WzyE F320 truncation
NM89	cgtaggcgagatcgccagtcc	WzyE T336 truncation
NM90	caccaccagtgagcctataagcgtag	WzyE V343 truncation
NM91	gaacaacgcgccgccca	WzyE F349 truncation
NM92	atacagccagtcgaaccatttgatgatcag	WzyE Y368 truncation
NM93	cgcagccttatagcgattagtctcgc	WzyE A381 truncation
NM94	atgaatgagtccggcgctttcaaaaag	WzyE H437 truncation
NM95	gcgcagaaagtagaccaccagca	WzyE R176 truncation

Table 3.3-(S2): Normalized activities of PhoA (AP) and LacZ (BG) WzyE C-terminal truncation fusions.

Residue [a]	Avg AP [b]	%AP [c]	Avg BG [d]	%BG [e]	AP:BG [f]	Location [g]	Colour [h]
Exocnuclease generated fusions							
E27	894.3	57.19	568.6	30.0	2:1	TM1	Purple
F39	57.8	3.64	766.5	40.46	1:10	CL1	Red
E69	896.5	57.34	157.9	8.29	7:1	PL1	Blue
T116	10.1	0.64	16.2	0.85	1:1	TM2	Purple
G136	382.4	24.44	68.2	3.6	7:1	PL2	Blue
S150	526.6	33.65	54.4	2.85	12:1	PL2	Blue
G193	85.4	5.44	96.2	5.07	1:1	TM6	Purple
L195	625.9	41.71	60.9	3.17	14:1	PL3	Blue
G202	586.5	37.49	38.8	2.01	19:1	PL3	Blue
I217	84.4	5.37	117.5	9.35	1:2	TM7	Purple
M240	465.2	29.75	216.2	11.4	3:1	TM8	Purple
F241	155.1	9.92	25.2	1.32	8:1	PL4	Blue
R247	448.5	28.66	123.8	6.49	5:1	PL4	Blue
T260	97.6	6.21	195.1	10.3	1:2	TM9	Purple
R295	220.1	14.08	285.9	15.06	1:1	TM10	Purple
P310	995.2	63.66	140.2	7.40	9:1	RE1	Blue
S329	82.7	5.25	95.6	5.02	1:1	RE1	Purple
A354	355.9	22.71	267.5	14.1	2:1	TM11	Purple
V357	129.6	8.25	109.7	5.76	2:1	TM11	Purple
L402	400.2	25.59	553.4	29.21	1:1	TM12	Purple
V414	212.1	13.56	341.4	18.01	1:1	TM13	Purple
Q447	14.6	0.90	572.2	30.22	1:34	CL	Red
PCR generated fusions							
R30	31.8	2.67	335.9	22.6	1:8	CL1	Red
F51	1084.4	93.29	268.1	17.1	5:1	PL1	Blue
G64	1132.6	97.41	317.5	20.23	6:1	PL1	Blue
A78	73.6	6.28	162.5	10.34	1:2	TM3	Purple
R96	33.2	2.84	770.2	49.14	1:24	CL2	Red
P104	60.2	5.15	1567.1	100	1:20	CL2	Red
V128	98.4	8.43	142.7	9.061	1:2	TM4	Purple
N143	1160.5	99.83	295.5	18.83	5:1	PL2	Blue
Y163	1148.6	98.795	125.6	7.98	12:1	PL2	Blue
R176	76.5	6.54	517.8	33.00	1:5	CL3	Red
V190	724.1	62.306	413.5	26.36	2:1	TM6	Purple

A212	415.6	35.71	226.9	16.98	2:1	TM7	Purple
G222	37.2	3.18	228.2	14.55	1:5	CL4	Red
W227	48.4	4.13	92.1	5.87	1:2	TM8	Purple
A276	85.9	7.31	110.2	7.02	1:1	TM9	Purple
N284	34.2	2.92	251.5	16.01	1:8	CL5	Red
I338	30.4	2.58	125.2	7.02	1:3	TM10	Purple
Y368	44.6	3.79	342.6	21.83	1:8	CL6	Red
A381	29.2	2.50	530.6	33.82	1:14	CL6	Red
F430	320.1	27.54	232.2	14.81	2:1	TM13	Purple
H437	43.5	3.70	943.1	60.17	1:16	CL	Red

The two sets of assays on the truncation fusions were independently normalized against the maximum activity within their respective subsets. [a] Position of the C-terminal fusion point to PhoA::LacZα. [b] and [d] Average of AP and BG activities across triplicates in Miller units. [c] and [e] Percentage of AP and BG activities relative to the maximum within each subset. [f] Normalized %AP/%BG activity ratio (NAR). [a] Colour of colony when grown on indicator plates. [a] Sub-cellular localization of C-terminal reporter fusion on topology map: PL, periplasmic loop; TM, transmembrane; CL, cytoplasmic loop; RE1, topologically ill-defined region.

3.8.3 Supporting Dataset.

Dataset (S1): Genebank and NCBI accession numbers of bacterial strains used for multiple sequence alignments.

Arsenophonus nasoniae - GenBank: FN545161.1

Atlantibacter hermannii - GenBank: GAB53594.1

Brenneria goodwinii - GenBank: CP014137.1

Brenneria nigrifluens - GenBank: QDKK01000006.1

Brenneria roseae subsp. americana - GenBank: PWC11794.1

Brenneria roseae subsp. roseae - GenBank: QDKI01000020.1

Budvicia aquatica - GenBank: PDDX01000001.1

Buttiauxella agrestis - GenBank: KFC84954.1

Buttiauxella brennerae - GenBank: OAT28234.1

Buttiauxella ferragutiae - GenBank: OAT31016.1

Buttiauxella gaviniae - GenBank: OAT20679.1

Buttiauxella noackiae - GenBank: OAT15398.1

Cedecea davisae - GenBank: KE161030.1

Cedecea lapagei - GenBank: PKA30084.1

Cedecea neteri strain - GenBank: CP023525.1

Citrobacter amalonaticus - GenBank: CP014070.2

Citrobacter braakii - GenBank: CP020448.2

Citrobacter europaeus - GenBank: SCA41670.1

Citrobacter farmeri - GenBank: CP022695.1

Citrobacter freundii - GenBank: CP016762.1

Citrobacter koseri - GenBank: CP026709.1

Citrobacter pasteurii - GenBank: CEJ63241.1

Citrobacter rodentium - GenBank: KIQ52202.1

Citrobacter sedlakii - NCBI Reference Sequence: NZ_BBNB01000018.1

Citrobacter werkmanii - GenBank: CP023504.1

Citrobacter youngae - GenBank: EFE07801.1

Cosenzaea myxofaciens - GenBank: OAT27236.1

Cronobacter condimenti - GenBank: CP012264.1

Cronobacter dublinensis - GenBank: CP012266.1

Cronobacter malonaticus - GenBank: PUX01266.1

Cronobacter muytjensii - GenBank: PUX14444.1

Cronobacter sakazakii - GenBank: CP012253.1

Cronobacter turicensis - GenBank: PUX35803.1

Cronobacter universalis - GenBank: CP012257.1

Dickeya chrysanthemi - GenBank: KGT98865.1

Dickeya dadantii - NCBI Reference Sequence: NZ_CM001978.1

Dickeya dadantii sp dieffenbachiae - NCBI Reference Sequence: WP_038924164.1

Dickeya dianthicola - GenBank: PWD75506.1

Dickeya fangzhongdai - GenBank: CP025003.1

Dickeya paradisiaca - GenBank: CP001654.1

Dickeya solani - GenBank: CP017454.1

Dickeya zeae - GenBank: CP025799.1

Edwardsiella hoshinae - GenBank: CP016043.1

Edwardsiella ictaluri - GenBank: CP020466.1

Edwardsiella piscicida - GenBank: PVD96881.1

Edwardsiella tarda - GenBank: CP023706.1

Enterobacter aerogenes - GenBank: FO203355.1

Enterobacter asburiae - GenBank: KVJ14514.1

Enterobacter cancerogenus - GenBank: OQD47239.1

Enterobacter cloacae - GenBank: CP022532.1

Enterobacter cloacae subsp. dissolvens - GenBank: KZQ40270.1

Enterobacter gergoviae - GenBank: OUF46421.1

Enterobacter hormaechei - GenBank: CP011662.1

Enterobacter kobei - GenBank: CP017181.1

Enterobacter ludwigii - GenBank: KIF84234.1

Enterobacter mori - GenBank: OXL41443.1

Erwinia amylovora - GenBank: CP024970.1

Erwinia billingiae - GenBank: PRB56403.1

Erwinia gerundensis - NCBI Reference Sequence: WP_067434625.1

Erwinia iniecta - GenBank: KOC88643.1

Erwinia mallotivora - GenBank: EXU74657.1

Erwinia oleae - NCBI Reference Sequence: WP_034948584.1

Erwinia persicina - NCBI Reference Sequence: WP_062749067.1

Erwinia piriflorinigrans - GenBank: CCG85588.1

Erwinia pyrifoliae - GenBank: CP023567.1

Erwinia tasmaniensis - GenBank: CU468135.1

Erwinia teleogrylli - NCBI Reference Sequence: WP_058911773.1

Erwinia toletana - NCBI Reference Sequence: WP_017800649.1

Erwinia tracheiphila - GenBank: JXNU01000003.1

Erwinia typographi - GenBank: KGT87794.1

Escherichia albertii - GenBank: CP025317.1

Escherichia coli K-12 - GenBank: AP009048.1

Escherichia fergusonii - GenBank: PQI99665.1

Escherichia marmotae - GenBank: CP025979.1

Escherichia vulneris - GenBank: GAL60088.1

Ewingella americana - GenBank: PKB91154.1

Franconibacter helveticus - NCBI Reference Sequence: WP_024551628.1

Franconibacter pulveris - GenBank: KMV32818.1

Gibbsiella quercinecans - GenBank: CP014136.1

Hafnia alvei - GenBank: CP015379.1

Hafnia paralvei - GenBank: KHS42153.1

Izhakiella australiensis - GenBank: OON38551.1

Izhakiella capsodis - GenBank: SFN51567.1

Klebsiella aerogenes - NCBI Reference Sequence: WP_043866160.1

Klebsiella grimontii - GenBank: SNU32815.1

Klebsiella michiganensis - GenBank: CP023185.1

Klebsiella oxytoca - GenBank: CP026285.1

Klebsiella pneumoniae - GenBank: FO834906.1

Klebsiella pneumoniae subsp. ozaenae - GenBank: CP027612.1

Klebsiella pneumoniae subsp. rhinoscleromatis - GenBank: EEW42825.1

Klebsiella quasipneumoniae - GenBank: CP014071.1

Klebsiella quasipneumoniae subsp. similipneumoniae - GenBank: OVW19422.1

Klebsiella variicola - GenBank: OZQ43466.1

Kluyvera ascorbata - GenBank: KFC99631.1

Kluyvera cryocrescens - NCBI Reference Sequence: WP_061283903.1

Kluyvera georgiana - GenBank: OAT51721.1

Kluyvera intermedia - GenBank: CP011602.1

Kosakonia arachidis - GenBank: SFU19740.1

Kosakonia cowanii - GenBank: CP022690.1

Kosakonia oryzae - GenBank: SFD33650.1

Kosakonia oryzendophytica - GenBank: SCC38348.1

Kosakonia oryziphila - GenBank: SCC36158.1

Kosakonia pseudosacchari - GenBank: PDO82947.1

Kosakonia radicincitans - GenBank: KDE34048.1

Kosakonia sacchari - GenBank: PDO82305.1

Leclercia adecarboxylata - GenBank: OOB85689.1

Lelliottia amnigena - GenBank: PEG64993.1

Lelliottia nimipressuralis - GenBank: OIR47920.1

Leminorella grimontii - GenBank: KFC93962.1

Lonsdalea britannica - GenBank: OSN06008.1

Lonsdalea iberica - GenBank: OSN04143.1

Lonsdalea populi - GenBank: OSM95825.1

Lonsdalea quercina - GenBank: SEB00508.1

Mangrovibacter phragmitis - GenBank: OAT77463.1

Mangrovibacter plantisponsor - GenBank: PWW02598.1

Moellerella wisconsensis - GenBank: KLN97999.1

Morganella morganii - GenBank: CP023505.1

Morganella psychrotolerans - GenBank: OBU09410.1

Obesumbacterium proteus - GenBank: OAT58704.1

Pantoea agglomerans - GenBank: PHP95306.1

Pantoea allii - GenBank: PBJ97671.1

Pantoea ananatis - GenBank: CP022427.1

Pantoea anthophila - GenBank: KKB06265.1

Pantoea coffeiphila - NCBI Reference Sequence: WP_105595140.1

Pantoea conspicua - NCBI Reference Sequence: WP_094120968.1

Pantoea cypripedii - GenBank: MLJI01000001.1

Pantoea deleyi - GenBank: ORM81447.1

Pantoea dispersa - GenBank: KTR90679.1

Pantoea eucrina - NCBI Reference Sequence: WP_065647810.1

Pantoea rodasii - NCBI Reference Sequence: WP_100701473.1

Pantoea rwandensis - NCBI Reference Sequence: WP_038643804.1

Pantoea septica - NCBI Reference Sequence: WP_033793101.1

Pantoea stewartii - GenBank: KHE02583.1

Pantoea theicola - NCBI Reference Sequence: WP_103060872.1

Pantoea vagans - GenBank: KGD73304.1

Pantoea wallisii - GenBank: ORM72246.1

Pectobacterium atrosepticum - GenBank: CP024956.1

Pectobacterium betavasculorum - GenBank: KFX01399.1

Pectobacterium carotovorum - GenBank: KHN55854.1

Pectobacterium parmentieri - NCBI Reference Sequence: WP_014701795.1

Pectobacterium polaris - NCBI Reference Sequence: WP_109412293.1

Pectobacterium wasabiae - GenBank: KGA27668.1

Phaseolibacter flectens - NCBI Reference Sequence: WP_051434129.1

Photorhabdus asymbiotica - GenBank: FM162591.1

Photorhabdus heterorhabditis - NCBI Reference Sequence: WP_054478930.1

Photorhabdus luminescens - GenBank: OWO80763.1

Photorhabdus luminescens subsp. namnaonensis - GenBank: OCA56468.1

Photorhabdus temperata - GenBank: OHV57103.1

Phytobacter ursingii - NCBI Reference Sequence: WP_047372306.1

Plesiomonas shigelloides - GenBank: EON89108.1

Pragia fontium - GenBank: SFD34613.1

Proteus cibarius - NCBI Reference Sequence: WP_109419157.1

Proteus hauseri - GenBank: EST57063.1

Proteus mirabilis - GenBank: CP026571.1

Proteus penneri - GenBank: GG661996.1

Proteus terrae - NCBI Reference Sequence: WP_109395458.1

Proteus vulgaris - GenBank: CP023965.1

Providencia alcalifaciens - GenBank: CP023536.1

Providencia burhodogranariea - GenBank: EKT61226.1

Providencia heimbachae - GenBank: OAT51901.1

Providencia rettgeri - GenBank: OZS73168.1

Providencia rustigianii - GenBank: ABXV02000023.1

Providencia sneebia - GenBank: EKT61628.1

Providencia stuartii - GenBank: CP014024.2

Pseudocitrobacter faecalis - GenBank: PUA64494.1

Rahnella - GenBank: PKE28247.1

Rahnella aquatilis - GenBank: KFD00377.1

Rahnella victoriana - NCBI Reference Sequence: WP_095924189.1

Raoultella ornithinolytica - GenBank: CP023888.1

Raoultella planticola - GenBank: OSU37269.1

Raoultella terrigena - GenBank: OMP89831.1

Rosenbergiella nectarea - NCBI Reference Sequence: WP_092676675.1

Rouxiella badensis - GenBank: ORJ25562.1

Rouxiella silvae - GenBank: ORJ22639.1

Salmonella bongori - NCBI Reference Sequence: WP_038393325.1

Salmonella enterica - GenBank: CP024169.1

Salmonella enterica sp typhii - GenBank: CQC04474.1

Salmonella enterica subsp. enterica serovar Enteritidis - GenBank: ELO83074.1

Salmonella enterica subsp. enterica serovar Montevideo - GenBank: OCI48034.1

Serratia ficaria - NCBI Reference Sequence: WP_061797957.1

Serratia fonticola - GenBank: CP013913.1

Serratia grimesii - GenBank: KFB87752.1

Serratia liquefaciens - GenBank: CP011303.1

Serratia marcescens - GenBank: CM008894.1

Serratia nematodiphila - GenBank: OQV32357.1

Serratia odorifera - GenBank: PNK91388.1

Serratia plymuthica - GenBank: LRQU01000001.1

Serratia proteamaculans - GenBank: SMB27943.1

Serratia rubidaea - NCBI Reference Sequence: WP_054306228.1

Serratia symbiotica - GenBank: CDS56727.1

Serratia ureilytica - GenBank: JSFB01000001.1

Shigella boydii - GenBank: CP011511.1

Shigella dysenteriae - NCBI Reference Sequence: WP_000055115.1

Shigella flexneri - GenBank: PAY77806.1

Shigella sonnei - GenBank: CP014099.2

Shimwellia blattae - GenBank: GAB82596.1

Siccibacter colletis - NCBI Reference Sequence: WP_031523759.1

Siccibacter turicensis - GenBank: PSN06044.1

Sodalis glossinidius - GenBank: LN854557.1

Sodalis praecaptivus - GenBank: CP006569.1

Tatumella citrea - NCBI Reference Sequence: WP_087489940.1

Tatumella morbirosei - NCBI Reference Sequence: WP_038015439.1

Tatumella ptyseos - NCBI Reference Sequence: WP_029991723.1

Tatumella saanichensis - NCBI Reference Sequence: WP_051150809.1

Thorsellia anophelis - NCBI Reference Sequence: WP_093317131.1

Trabulsiella guamensis GenBank: KFC00929.1

Trabulsiella odontotermitis - GenBank: KNC88106.1

Xenorhabdus beddingii - NCBI Reference Sequence: WP_086111756.1

Xenorhabdus bovienii - GenBank: CDG93513.1

Xenorhabdus budapestensis - NCBI Reference Sequence: WP_099134212.1

Xenorhabdus cabanillasii - NCBI Reference Sequence: WP_038268499.1

Xenorhabdus doucetiae - GenBank: FO704550.1

Xenorhabdus eapokensis - GenBank: OKP04648.1

Xenorhabdus ehlersii - NCBI Reference Sequence: WP_099133104.1

Xenorhabdus griffiniae - GenBank: KLU17421.1

Xenorhabdus hominickii - GenBank: CP016176.1

Xenorhabdus innexi - GenBank: PHM36349.1

Xenorhabdus ishibashii - GenBank: PHM62243.1

Xenorhabdus japonica - NCBI Reference Sequence: WP_092516713.1

Xenorhabdus khoisanae - GenBank: KMJ45648.1

Xenorhabdus koppenhoeferi - GenBank: SFU50053.1

Xenorhabdus kozodoii - GenBank: PHM74842.1

Xenorhabdus mauleonii - GenBank: PHM44571.1

Xenorhabdus miraniensis - GenBank: PHM50913.1

Xenorhabdus nematophila - NCBI Reference Sequence: WP_010845315.1

Xenorhabdus poinarii - GenBank: FO704551.1

Xenorhabdus stockiae - GenBank: PHM66775.1

Xenorhabdus szentirmaii - GenBank: PHM42927.1

Xenorhabdus thuongxuanensis - GenBank: OKP04708.1

Xenorhabdus vietnamensis - GenBank: OTA17820.1

Yersinia aldovae - NCBI Reference Sequence: NZ_CP009781.1

Yersinia aleksiciae - GenBank: CFQ50018.1

Yersinia bercovieri - GenBank: PHZ27340.1

Yersinia enterocolitica - GenBank: PNM12560.1

Yersinia entomophaga - GenBank: OWF85336.1

Yersinia frederiksenii - GenBank: OVZ86646.1

Yersinia intermedia - GenBank: OVZ85040.1

Yersinia kristensenii - GenBank: OVZ77395.1

Yersinia massiliensis - GenBank: PHZ22510.1

Yersinia mollaretii - GenBank: PJE87223.1

Yersinia nurmii - GenBank: CNF01149.1

Yersinia pekkanenii - GenBank: CRY67263.1

Yersinia pestis - GenBank: KZC52136.1

Yersinia pseudotuberculosis - NCBI Reference Sequence: WP_038823771.1

Yersinia rohdei - GenBank: CNE82538.1

Yersinia ruckeri - GenBank: OEU26259.1

Yersinia similis - GenBank: CNB40475.1

Yokenella regensburgei - GenBank: KFD24891.1

3.9 Article Acknowledgements

Funding for this work was provided by a Discovery Project Grant to R. Morona from the Australian Research Council (PROJECT ID: DP160103903). N. Maczuga is the recipient of a Research Training Program Stipend Research Scholarship from the University of Adelaide.

3.10 Article Conflicts of interest

The author(s) declare that there are no conflicts of interest.

3.11 Additional Results

Results presented in this chapter thus far were all included in the submission of the publication. The following additional results (Additional Results (section 3.11) were not in the submitted publication, but nevertheless continue on the topic of WzyE topology through the use of the thiol probe PEG maleimide.

3.11.1 Analysis of WzyE cystine residues

Although the reporter fusion topology mapping produced a map (Figure 3.4a) which correlated with the *in silico* predicted WzyE structure from Alphafold (Figure 3.3b), the possibility of an inaccurate protein topology being mapped due to the assumptions of C-terminal topology mapping remained. To independently verify the experimental topology map, I decided to use thiol reactive probes to probe the native topology of WzyE using PEG maleimide (mPEG) which induces a detectable band shift in protein weight ~ 5 kDa when Western immunoblotting. Due to the size of mPEG, it cannot cross cellular membranes and hence can only react with environmentally exposed cysteine residues which are not in disulphide bonds.

The location of WzyE's cysteine residues were assessed and I found that WzyE possesses four cysteine residues: C16, C80, C387 and C419 where, a pair of cysteine residues were localised to the N (C16/C80) and C-termini (C387/C419) of the WzyE peptide sequence (Figure 3.9a). Additionally, I observed that three of the four residues (C16, C80 and C419) were predicted to be transmembrane whereas C387 was predicted to be in the cytosol (Figure 3.9a). Upon consulting the Alphafold structure I saw that interestingly, the pairs of cysteine residues were predicted to be in close proximity of each other which could suggest possible disulphide bond formation between the pairs (Figure 3.9b).

Using the 248 WzyE peptide sequence MSA (Figure 3.1), I also assessed how conserved the cysteine residues were throughout WzyE peptide sequences which showed that the cysteines had varying levels of conservation. C16 had the least percentage of conservation with only 7.3% amongst the 248 peptide sequences along with C80, which showed 56.9% conservation (Figure 3.9c). The C-terminal cysteines however showed very high percentages of conservation with C387 and C419 having 92.3% and 99.6% each, respectively, which suggested that they play a more important role in WzyE proteins than the N-terminal pair (Figure 3.9c).

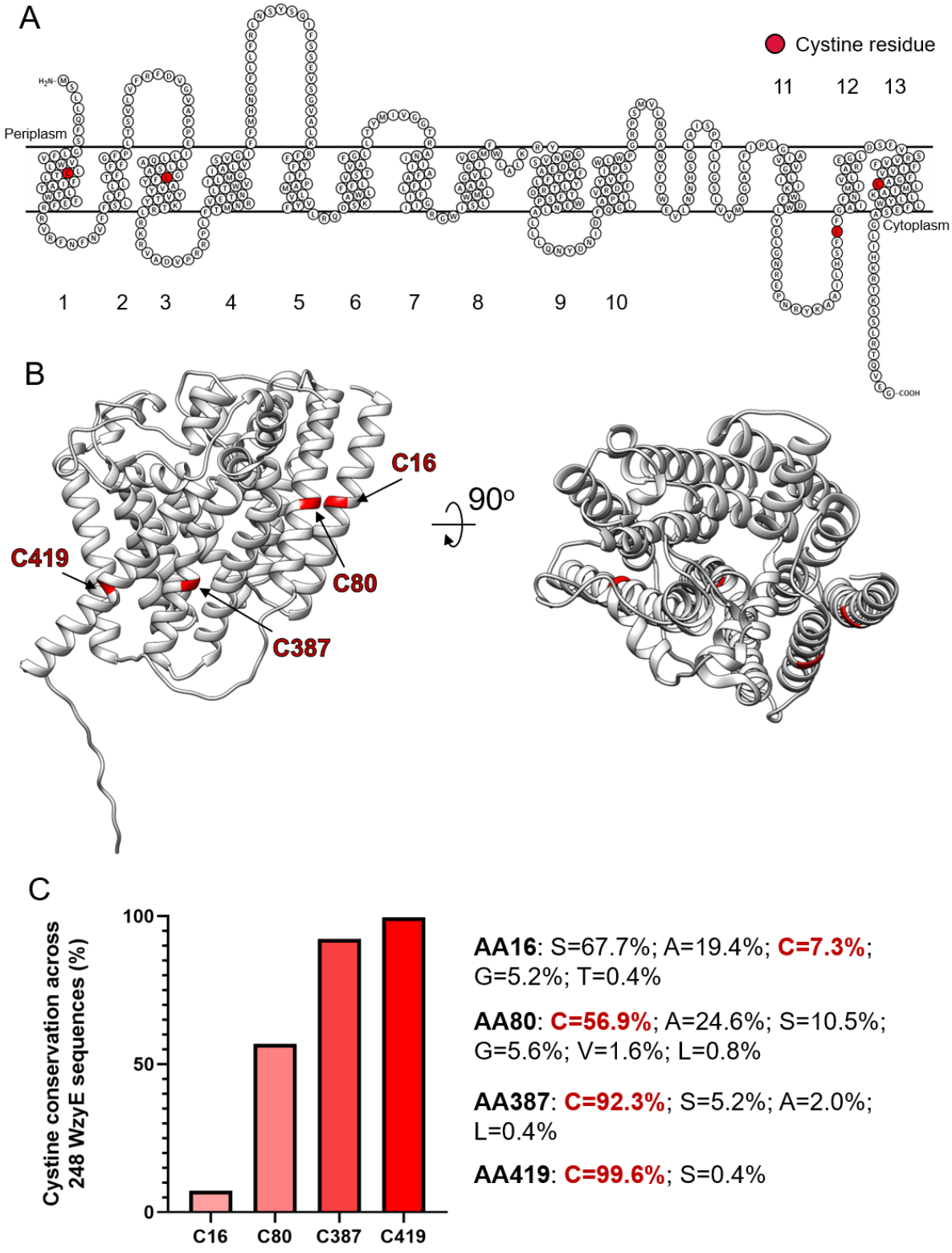


Figure 3.9: Analysis of $WzyE_{SF}$ cysteine residues.

a) Experimental topology map with cysteine residues indicated in red. b) Alphafold *in silico* predicted WzyE structure with cysteine residues indicated in red. c) Percentage of conservation of the cysteine residues C16, C80, C387 and C419 from the 248 WzyE peptide sequence MSA.

3.11.2 Probing of WzyE$_{SF}$ cysteine residues with mPEG

I then assessed the effect of PEGylation on WzyE through Western immunoblotting with and without mPEG. As WzyE was shown to possess a single cytosolic cysteine, I predicted that PEGylation of whole membrane samples would produce a single band shift. Subsequently, whole membrane samples from PE860 *ΔwzyE$_{10-440}$* harbouring pBAD33-WzyE3XFLAG were generated and PEGylated, as described in 2.5.5. Western immunoblotting revealed that the addition of mPEG had induced a second band shift to ~64 kDa in Figure 3.10 lane 2, which supported the experimental topology map as only a single residue would be PEGylated according to the map, likely being C387. Additional banding was observed at ~98 kDa size however, this was most likely non-specific banding due to the amount of sample loaded into the wells. The same bands are seen across both PEGylated and non-PEGylated lanes indicating that they are not due to protein PEGylation.

I then investigated whole cell PEGylation as PEGylation of whole membrane samples does not distinguish between periplasmic and cytoplasmic residues rather, any exposed cysteine residue would be PEGylated. Whole cells were PEGylated prior to disruption; a sample was PEGylated in the presence of EDTA to disrupt the OM theoretically exposing only periplasmic residues to mPEG, and another with 2% (w/v) SDS to completely disrupt the membrane which theoretically would lead to the PEGylation of all available cysteine residues in WzyE.

Whole cell PEGylation produced no additional mPEG bands as did the cells treated with EDTA (Figure 3.11a lanes 2, 5 & lanes 1,4), respectively, besides the non-specific protein bands seen in Figure 3.10. However, the sample treated with 2% (w/v) SDS produced only a single faint 50 kDa band (Figure 3.11a lanes 3, 6) which was inconsistent with all the other lanes which showed strong protein bands at 50 kDa. This may have been caused by experimental error as, once treated with 2% (w/v) SDS, the sample became highly viscous making proper pipetting extremely difficult and this may have led to incorrect volumes being load onto the SDS-PAGE gel. At minimum, the SDS-treated sample should have produced at least a single band shift as C387, presumably, would have theoretically been PEGylated.

Subsequently, the whole cell PEGylation was repeated in the presence of 2% (w/v) SDS with more sample loaded into the SDS-PAGE gel. Western immunoblotting revealed that the SDS-treated sample produced three bands which would indicate the PEGylation of two cysteine residues (Figure 3.11b lanes 4, 5, 6). Due a shift indicative of only two cysteine residues, it is likely that the other two cysteine residues of WzyE are in a disulphide bond and hence, are unable to be PEGylated.

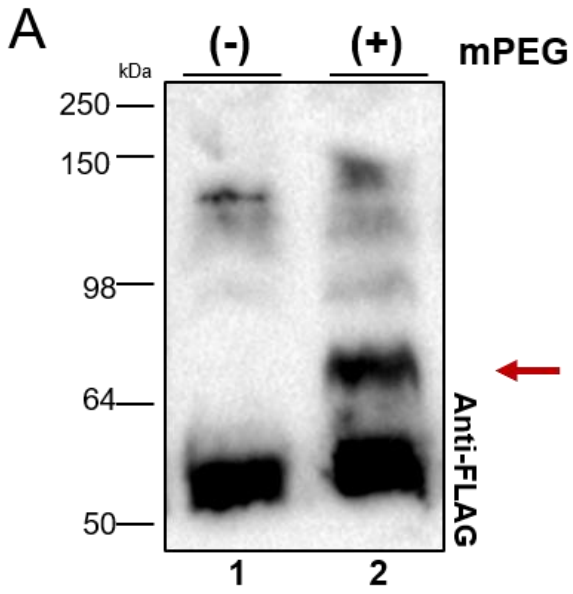


Figure 3.10: Whole membrane PEGylation of WzyE3XFLAG.

a) Anti-FLAG Western immunoblot of WzyE3XFLAG whole membrane PEGylation. PE860 *ΔwzyE*$_{10\text{-}440}$ harbouring pBAD33-WzyE3XFLAG mid-exponential phase cells were disrupted by sonication and WM were collected by ultracentrifugation. WM were then resuspended in PEGylation buffer and treated with 1 mM mPEG. Samples were then electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-FLAG antibodies. Arrows indicate protein PEGylation band shifts.

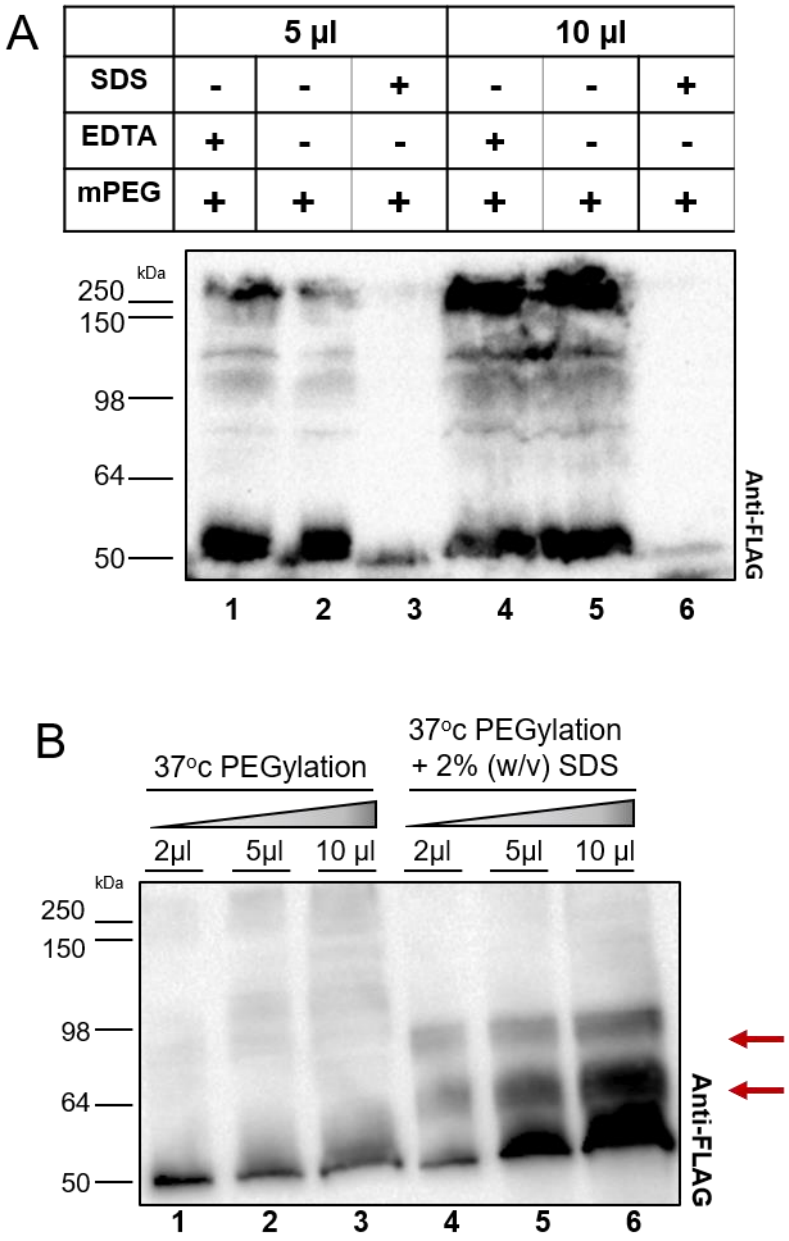


Figure 3.11: Whole cell PEGylation of WzyE3XFLAG.

a) Anti-FLAG Western immunoblot of WzyE3XFLAG whole cell PEGylation. PE860 *ΔwzyE*$_{10\text{-}440}$ harbouring pBAD33-WzyE3XFLAG mid-exponential phase cells were collected by centrifugation, resuspended in sonication buffer and 1 ml cell suspension aliquoted out. Samples were treated with EDTA (0.5 M), or SDS (2% w/v) with all receiving 1 mM mPEG and incubated for 1 hr at 37 °C. Cells were the disrupted by sonication and WM were collected by ultracentrifugation. Samples were then electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-FLAG antibodies. b) Anti-FLAG Western immunoblot of WzyE3XFLAG whole cell PEGylation SDS treated samples. Samples were prepared and Western immunoblotted as above. Arrows indicate protein PEGylation band shifts.

3.11.3 Summary of Additional Results

Due to the assumptions in the methodology used to generate the experimental topology map (Figure 3.4a), I attempted to independently verify the experimentally topology map through the use of the thiol reactive probe PEG maleimide. I assessed the localisation of the cysteine residues present in $WzyE_{SF}$ revealing that C16 and C80 and, C387 and C419 are predicted to be in close proximity with one another (Figure 3.9b) and that C16, C80 and C419 were transmembrane whereas C387 was cytosolically exposed, according to the experimental topology map (Figure 3.9a). Consulting the 248 WzyE peptide sequence MSA, I also investigated whether or not the cysteine residues were conserved in WzyE proteins which revealed that they show varying percentages of conservation where the two C-terminal cysteines, C387 and C419, showed 92.3% and 99.6% conservation. This suggested that they may be important due to their high level of conservation where interestingly, the cysteine residues were most commonly substituted with serine residues. Whole membrane PEGylation supported the experimental topology map by producing only a single band shift (Figure 3.10 lane 2) as the experimental map predicted only a single cysteine residue would be environmentally exposed. This was further supported by the whole cell PEGylation which showed that PEGylation in the presence of EDTA produced no band shifts whereas, once treated with SDS, Western immunoblotting revealed the presence of two band shifts, indicative of two separate cysteine residues being PEGylated (Figure 3.11 a & b). As treatment of SDS completely disrupts the membrane, PEGylation would react with every available cysteine residue where, due to only two band shifts, suggests that the other two cysteines of WzyE are in a disulphide bond and hence, cannot be PEGylated. The additional results here support the experimentally determined topology map and suggest the presence of a disulphide bond within WzyE. Substitutions of periplasmically exposed residues should be undertaken to probe the topology map further and the possibility of a disulphide bond in WzyE should be investigated starting with the substitution of cysteine residues to determine any impact on PEGylation band shifts.

Chapter Four

ARTICLE II:

Interdependence of *Shigella flexneri* surface polysaccharides ECA and LPS O-antigen.

Nicholas Tadeusz Maczuga, Elizabeth Tran, Jilong Qin and Renato Morona.

Chapter 4: Interdependence of *Shigella flexneri* surface polysaccharides ECA and LPS.

4.1 Statement of Authorship

Title of Paper	Interdependence of *Shigella flexneri* surface polysaccharides ECA and LPS O-antigen.
Status	Published
Citation	Maczuga, N, Tran, ENH, Qin, J & Morona, R 'Interdependence *of Shigella flexneri* O antigen and enterobacterial common antigen biosynthetic pathways', J Bacteriol e00546-00521.

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Nicholas Tadeusz Maczuga		
Contribution	Construction of strains and plasmids, conducted all experiments pertaining to all Figures and constructed these figures, data analysis, conception of model, corresponding author and writing of manuscript.		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/01/2022

Author	Elizabeth Ngoc Hoa Tran		
Contribution	Supervised development of work and provided editing and evaluation of the manuscript.		
Signature		Date	20/1/22

Author	Jilong Qin		
Contribution	Generation of mutant critical to the strain.		
Signature		Date	20/01/2022

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, manuscript evaluation and editing.		
Signature		Date	19/1/22

4.2 Article Abstract

Outer membrane (OM) polysaccharides allow bacteria to resist harsh environmental conditions and anti-microbial agents, traffic to and persist in pathogenic niches and evade immune responses. *Shigella flexneri* has two OM polysaccharide populations, being Enterobacterial Common Antigen (ECA) and lipopolysaccharide (LPS) O antigen (Oag); both are polymerised into chains by separate homologs of the Wzy-dependent pathway. The two polysaccharide pathways, along with peptidoglycan (PG) biosynthesis, compete for the universal biosynthetic membrane anchor, undecaprenyl phosphate (Und-P) as the finite pool of available Und-P is critical in all three cell wall biosynthetic pathways. Interactions between the two OM polysaccharide pathways have been proposed in the past where, through the use of mutants in both pathways, various perturbations have been observed. Here, we show for the first time that mutations in one of the two OM polysaccharide pathways can affect each other, dependent on where the mutation lies along the pathway, whilst the second pathway remains genetically intact. We then expand on this and show that the mutations also affect PG biosynthesis pathways as well, and provide data which supports that the classical mutant phenotypes of cell wall mutants are due to a lack of available Und-P. Our work here provides another layer in understanding the complex intricacies of the cell wall biosynthetic pathways and demonstrate their interdependence on Und-P, the universal biosynthetic membrane anchor.

4.3 Article Importance

Bacterial outer membrane polysaccharides play key roles in a range of bacterial activities from homeostasis to virulence. Two such OM polysaccharides populations are ECA and LPS Oag which are synthesised by separate homologs of the Wzy-dependent pathway. Both ECA and LPS Oag biosynthesis join with PG biosynthesis to form the cell wall biosynthetic pathways, which all are interdependent on the availability of Und-P for proper function. Our data show the direct effects of cell wall pathway mutations affecting all related pathways when they themselves remain genetically unchanged. This work furthers our understanding of the complexities and interdependence of the three cell wall pathways.

4.4 Article Introduction

Shigella flexneri is a Gram-negative bacteria which expresses two distinct populations of outer membrane (OM) bound polysaccharides; lipopolysaccharide (LPS) and enterobacterial common antigen (ECA). Comprising of three distinct domains that include lipid A, the proximal membrane anchor, the inner and outer core sugars, and a distal chain of O antigen (Oag) polysaccharides, LPS is known to play crucial roles in the viability, pathogenesis and immune evasion of Gram-negative pathogens (Günther et al. 2019; Whitfield, Wear & Sande 2020). ECA exists in two membrane associated forms, ECA_{pg} the most abundant and ECA_{lps} which occurs in some strains lacking O antigen, as well as a periplasmically restricted cyclic form ECA_{cyc} (Rai & Mitchell 2020). ECA is known to play roles in maintaining OM homeostasis and providing resistances to bile salt (Mitchell, Srikumar & Silhavy 2018; Ramos-Morales et al. 2003). Peptidoglycan (PG) biosynthesis, along with the two OM polysaccharide pathways, form the three cell wall biosynthetic pathways whose polymerases belong to the Shape, Elongation, Division and Sporulation (SEDS) protein family and all rely on undecaprenyl phosphate (Und-P) as their biosynthetic lipid linked anchor (Meeske et al. 2016).

The *S. flexneri* Y serotype Oag comprises of tetrasaccharide repeat units (RUs) which contain a single N-acetylglucosamine (GlcNAc) and three rhamnose (Rha) residues (Liu, B et al. 2008). ECA, universal in all Enterobacterales, comprises of trisaccharide RUs containing N-acetylglucosamine (GlcNAc), N-acetyl-D-mannosaminuronic acid (ManNAcA) and 4-acetamido-4,6-dideoxy-D-galactose (Fuc4NAc) (Eade et al. 2021). Once completed, RUs from both pathways are translocated across the inner membrane (IM) and polymerised into linear chains by their separate homologs of the Wzy-dependent pathway (Figure 4.1). The Wzy dependent pathway is the most common bacterial polysaccharide biosynthetic pathway which consists of three proteins: Wzx translocase, Wzy polymerase and Wzz polysaccharide co-polymerase which controls polymerization by Wzy (Islam & Lam 2014). In *S. flexneri*, the proteins associated with the Wzy-dependent pathway for ECA are denoted as WzxE, WzyE and WzzE, and for Oag these are termed WzxB, WzyB and WzzB.

Cell wall biosynthesis commences on the cytoplasmic side of the IM where the glycosyltransferases WecA and MraY act on Und-P to yield the foundations of OM polysaccharide biosynthesis and PG biosynthesis, respectively (Rai & Mitchell 2020; Typas et al. 2011; Whitfield & Trent 2014). Once acted upon, each Und-P molecule is committed to either one of the three biosynthetic pathways until completed RUs are cleaved off by FtsW/RodA, WzyE, WzyB, WaaL or the unknown final ECA glycosyltransferase in the process of polysaccharide biosynthesis (Meeske et al. 2016; Sham et al. 2014; Whitfield & Trent 2014). Lastly, undecaprenyl

pyrophosphate (Und-PP) is then recycled via dephosphorylation, primarily by UppP, to yield Und-P which then re-enters into the finite pool of available Und-P (Figure 4.1) (Workman & Strynadka 2020).

It is the requirement of Und-P in all of the three cell wall pathways which fundamentally interlinks them and causes their interdependence upon one another. The interdependence of each of the OM polysaccharide pathways with PG biosynthesis has been shown separately, where the majority of our current understandings originate from the research performed by *Jorgenson et al. (2016)* who showed that various mutations in the LPS Oag and ECA biosynthetic pathways of *Escherichia coli* can influence PG synthesis (Jorgenson et al. 2016; Jorgenson & Young 2016). In addition to this, *Marolda et al. (2006)* showed that components of the two Wzy-dependent pathways could supplement one-another, however only under very specific circumstances (Marolda et al. 2006). Recently, *Leo et al. (2020)* demonstrated that the Wzz proteins of both OM polysaccharide pathways could complement each other and partially restore polysaccharide modal length control (Leo 2020).

In this study, we demonstrate the interdependence of all three cell wall pathways with one another. Through the use of *wzy* mutants, we show for the first time that restricting the rate of Und-P recycling is sufficient to induce a classical mutant phenotype associated with cell wall mutants, resulting in an elongated cellular morphology (Jorgenson et al. 2016; Jorgenson & Young 2016). Additionally, we show the direct effects of sequestering intermediates from related OM polysaccharide biosynthetic pathways through the use of anti-ECA Western immunoblotting and LPS silver stained gels. This study reveals another layer of complexity in investigating cell wall mutants and reveals the extent of the interdependence between the three cell wall pathways.

4.5 Article Methods

4.5.1 Ethics statement

The ECA antibodies were produced under the National Health and Medical Research Council Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and was approved by the University of Adelaide Animal Ethics Committee.

4.5.2 Bacterial strains, growth media and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 4.1. Bacteria were routinely grown at 37 ºC in lysogeny broth (LB) with aeration or on LB agar (LBA) (Tran , Doyle & Morona 2013). Antibiotics used were as follows: 50 µg kanamycin (Kan) ml^{-1}; 100 µg ampicillin (Amp) ml^{-1}; 25 µg chloramphenicol (Cml) ml^{-1}; 10 ng tunicamycin ml^{-1} with 3 µg polymyxin B nonapeptide ml^{-1} (PBMN; Sigma). Strains carrying pBAD33, pBCKs+ or pWKS30 constructs requiring induction were grown in LB at 37 ºC with aeration for 16 hours, sub-cultured (1/20) into fresh broth and induced with either 1 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG) for pBCKs+ and pWKS30 constructs, or 0.2 % (w/v) L-arabinose for pBAD33 constructs. Cultures were grown for a further 4 hours.

4.5.3 DNA methods

Plasmid constructs were purified from *E. coli* DH5α strains using a QIAprep Spin miniprep kit (Qiagen). Preparation of electrocompetent cells and electroporation methods were performed as described previously (Purins et al. 2008).

4.5.4 Chromosomal mutagenesis

S. flexneri Y *ΔwzyE, ΔwzyB, ΔwecC, ΔwecA* strains were generated using λ Red mutagenesis as described previously (Datsenko & Wanner 2000). Briefly, primers were designed (*ΔwzyE:*NM1/NM2*, ΔwzyB*:NM136/NM137*, ΔwecC:*NM151/NM152*, ΔwecA:*NM134/NM135*)* to PCR amplify either a kanamycin or chloramphenicol resistance cassette flanked with 50 bp of the start or end of coding regions for the respective genes (Table 4.2). The purified PCR fragments were then electroporated into the parent PE860 strain carrying pKD46 to generate the mutant strains. The antibiotic resistance cassettes were eliminated by the introduction of pCP20 (Datsenko & Wanner 2000).

4.5.5 Generation of complementing constructs

Constructs for ectopic protein expression were generated via PCR and restriction cloning. Primers used for construct generation used in this study are listed in Table 4.2. Generation of pWzyE was performed by using primers (VL70/NM25) to PCR amplify a fragment of DNA

containing *wzyE* with *Xma*I and *Sph*I restriction enzyme sites from pWALDO-WzyE-GFP-His_8. The resulting *wzyE* fragment was digested with *Xma*I and *Sph*I and sub-cloned into likewise digested pBAD33 to give pBAD33-WzyE-3xFLAG. Generation of pUppS and pWecA was performed using primers (NM122/NM123) and (NM128/129) to PCR amplify a fragment of DNA from *S. flexneri* 2457T containing *uppS* or *wecA* respectively, with *EcoR*I and *Xba*I restriction enzyme sites. The resulting DNA fragments were digested with *EcoR*I and *Xba*I and cloned into likewise digested pWKS30 to give pWKS30-UppS-HA and pWKS30-WecA-HA. DNA sequencing was used to confirm that no mutations had been introduced during the PCR amplification and to ensure the presence of the C-terminal epitope tags.

4.5.6 LPS/ECA sample preparation

Bacteria were grown and induced as described above before $1x10^9$ cells were collected by centrifugation (2,000 x g), resuspended in 2x lysis buffer (Murray, Attridge & Morona 2003) and heated at 100 ºC for 10 minutes before incubation with 2.5 mg/ml proteinase K (Sigma-Aldrich) for 2 hours at 56 ºC (Murray, Attridge & Morona 2003).

4.5.7 LPS SDS-PAGE and silver staining

LPS samples were heated at 100 ºC for 10 minutes before being loaded and electrophoresed on SDS-15% PAGE gels for 13 hours at 12 mA. Silver staining of LPS was performed as described previously (Murray, Attridge & Morona 2003).

4.5.8 ECA PAGE and Western immunoblotting

ECA samples were heated at 100 ºC for 10 minutes before being loaded and electrophoresed on SDS-15% PAGE gels at 200V for 1 hour. SDS PAGE gels were transferred onto nitrocellulose membranes (Bio-Rad) at 400 mA for 1 hour. Membranes were then blocked with 5% (w/v) skim milk in Tris-Tween Buffer Saline (TTBS) (Purins et al. 2008) followed by overnight incubation with polyclonal rabbit anti-ECA antibodies at 1:500 dilution in 2.5% (w/v) skim milk in TTBS. Detection was performed with goat anti-rabbit horseradish peroxidase-conjugated antibodies (KLP) and chemiluminesence reagent (Sigma). 5 µl of SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard.

4.5.9 Measuring OM polysaccharide abundance

Densitometry was performed on three biological replicates from silver-stained SDS-PAGE gels for LPS Oag quantification and anti-ECA Western immunoblots. The degree of polymerization was calculated by normalizing the densitometry of S-Oag and ECA molecules to

the parent where the relative abundance of each sample to the parent was presented as a scatter plot.

4.5.10 Growth curves

Bacteria were grown for 16 hours as described above and $1x10^7$ cells were collected via centrifugation and resuspended in 1 ml of LB. Cells were then sub-cultured (1/10) into 135 µl of fresh LB media in a 96 well tray. The tray was then incubated at 37 ºC for 10 hours with aeration. OD_{600} absorbance readings were taken every 20 minutes (BioTek PowerWave XS2). Experiments were performed in independent and technical triplicate where three cultures were grown for 16 hours and subsequently sub-cultured.

4.5.11 Colony Forming Units (CFU) counting

Bacteria were grown for 16 hours as described above. Cells were normalized and subcultured into 10 ml of LB. 20 µl of cell culture was taken at time points 0, 2, 4, 6 and 8 hours and serially diluted 1:10 with PBS prior to spotting 10 µl of cell suspension from the range of 10^{-5} to 10^{-8} in triplicate onto LBA plates. Plates were incubation O/N at 37 ºC and colonies counted the next day.

4.5.12 Microscopy

Bacteria were grown as described above and 10 µl of cultures were spotted onto glass slides. The cultures were then allowed to dry followed by mounting with Moviol (Calbiochem) and sealing with nail polish. Cells were observed via phase contrast microscopy (Olympus IX70) under a 100x oil objective lens (Tran, ENH, Doyle & Morona 2013). Images of cells were acquired and cellular lengths were then manually measured using Metamorph 7.5.6.

4.5.13 Statistical analysis

Independent student t-tests were performed on triplicate experimental data values using the statistical analysis tool of Graph pad Prism 9. Graphs were plotted with the standard error of the mean (SEM) and statistical significance was displayed as the following: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$, and ns is a non-statistically significant.

Table 4.1: Bacterial strains and plasmids

Strain or plasmid	Description	Source
Strains:		
RMA2162	*S. flexneri* PE860 Y serotype; strain lacks virulence plasmid and pHS-2 plasmid	Laboratory stock
RMA2171	*S. flexneri* PE860 *ΔrmlD*	Laboratory Stock
NMRM55	*S. flexneri* PE860 *ΔwzyE*	This study
NMRM339	*S. flexneri* PE860 *ΔwzyB*	This Study
NMRM345	*S. flexneri* PE860 *ΔwecA*	This Study
NMRM348	*S. flexneri* PE860 *ΔwecC*	This Study
Plasmids*:*		
pBAD33	Arabinose inducible, expression vector, Cml^r	(Guzman et al. 1995)
pWzyE	pBAD33 encoding WzyE3xFLAG, Cml^r	This Study
pBCKs+	pBluescript KS +, IPTG inducible, expression vector, Cml^r	Stratagene
pWzyB	pBCKs+ encoding WzyB3xFLAG, Cml^r	Laboratory Stock
pWKS30	IPTG inducible, expression vector, Amp^r	(Wang, RF & Kushner 1991)
pUppS	pWKS30 encoding UppS-HA, Amp^r	This Study
pWecA	pWKS30 encoding WecA-HA, Amp^r	This Study
pWALDO-WzyE-GFP-His_8	Cloning vector expressing $WzyE_{EC}$-GFP-His_8	(Rapp et al. 2004)

Table 4.2: DNA oligonucleotides used in this study

Primer	Oligonucleotide sequence (5'-3')	Target
Construct generation specific primers		
VL70	gtaacacccgggttgtttaactttaagaaggagactcg	pWALDO-WzyE-GFP-His_8
NM25	tcttcgcatgctcacttgtcatcgtcatccttgtagtcgatgtcatgatctttataatcaccgtcatggtctttgtagtctccttcaacctgcgtccgg	*wzyE* gene
NM122	cgatagaattcgtagggcttcagtgatatagtctgcgcc	*uppS* gene
NM123	attagtctagattcaagcgtaatctggaacatcgtatgggtaggctgtttcatcaccgggctc	*uppS* gene
NM128	ttagagaattcgggttcggaacggactttcccttc	*wecA* gene
NM129	attagtctagattcaagcgtaatctggaacatcgtatgggtatttggttaaattggggctgccacca	*wecA* gene
λ Red mutagenesis specific primers		
NM1	caatcaactgtaagccacgcagcgtataggttggtgccgtggtgttgttattcattgatgggaattagccatggtcc	*wzyE* gene
NM2	tatctacaaggctggcagcgggcgttggcgattgccgccagggaggtcgcgtgtaggctggagctgcttc	*wzyE* gene
NM134	ggttatacttctgctaataattttctctgagagcatgcattgtgaatttagtgtaggctggagctgcttc	*wecA* gene
NM135	tttcccaggcattggttgtgtcatcacatcctcatttatttggttaaattatgggaattagccatggtcc	*wecA* gene
NM136	tgttataaaaattttatttatatttttcatattcgtaaggtgatgttttgtgtaggctggagctgcttc	*wzyB* gene
NM137	agtaataacctcacttctggagcaaaataaaggatcttaaaaatagggaaatgggaattagccatggtcc	*wzyB* gene
NM151	aaaataatcggatatcactatgagttttgcgaccatttctgttatcggactgggttacatgtgtaggctggagctgcttc	*wecC* gene
NM152	ttttctcatcagcgccagactcctttggcatcgacgacatactgctgatatgggaattagccatggtcc	*wecC* gene

4.6 Article Results

4.6.1 Investigating the biosynthetic effects of the *ΔwzyE* mutation

Research has shown that the sequestering of undecaprenyl phosphate (Und-P) in the form of dead-end biosynthetic precursors of interrupted polysaccharide pathways leads to altered cellular morphology (Jorgenson et al. 2016; Jorgenson & Young 2016). As the two major OM polysaccharides of *Shigella flexneri,* Oag and ECA, share the same initial biosynthetic precursor, Und-PP-GlcNAc (Figure 4.1), the OM polysaccharide profile of *S. flexneri* PE860 *ΔwzyE* was analysed by anti-ECA Western immunoblotting and silver-stained SDS PAGE (Figure 4.2a and b). The *wzyE* mutant was unable to polymerize any detectable ECA as seen by a lack of a ladder banding pattern (Figure 4.2a, lane 2), and when complemented with pWzyE, the mutant was able to partially polymerize ECA, observed as an ECA ladder banding pattern (Figure 4.2a, lane 4).

Surprisingly, a decrease in the smooth Oag (S-Oag) polymerization was observed when comparing the LPS of the PE860 parent to the *wzyE* mutant (Figure 4.2b, lanes 1 and 2), and the mutant phenotype was rescued to wildtype when complemented with pWzyE (Figure 4.2b, lane 4). This reduction in S-Oag polymerization was further quantified via densitometry and showed a 61% statistically significant decrease in abundance between the parent and the *wzyE* mutant, as well as between the *wzyE* mutant and the complemented mutant with pWzyE ($p < 0.001$). No statistically significance difference in Oag between PE860 and the *wzyE* mutant complemented with pWzyE was observed (Figure 4.2c).

As blocking ECA production leads to an altered cellular morphology due to PG biosynthesis disruption (Jorgenson et al. 2016), we investigated whether *wzyE* mutants also exhibited altered morphologies. Cellular lengths were measured by phase-contrast microscopy and a statistically significantly difference was observed in cell size when comparing PE860 and the *wzyE* mutant, as well as between the *wzyE* mutant complemented with pWzyE, which reverted back to wildtype length ($p < 0.001$) (Figure 4.2d and e).

This suggested that the theoretical lack of available Und-P in the *wzyE* mutants affected PG biosynthesis in a similar manner as previously described (Jorgenson et al. 2016). Growth curves were then performed to investigate if the *wzyE* mutant exhibited altered growth presumably due to the lack of available Und-P compared to the parent (Figure 4.2f). The *wzyE* mutants showed reduced growth, reaching a lower OD_{600} when compared to the parent, however this phenotype was rescued when complemented with pWzyE. As absorbance readings are dependent on particle size, a CFU count was performed and showed that in addition to having a smaller cell length, the

wzyE mutant produced less CFUs than PE860 at each time point investigated (Figure 4.6-(S1).

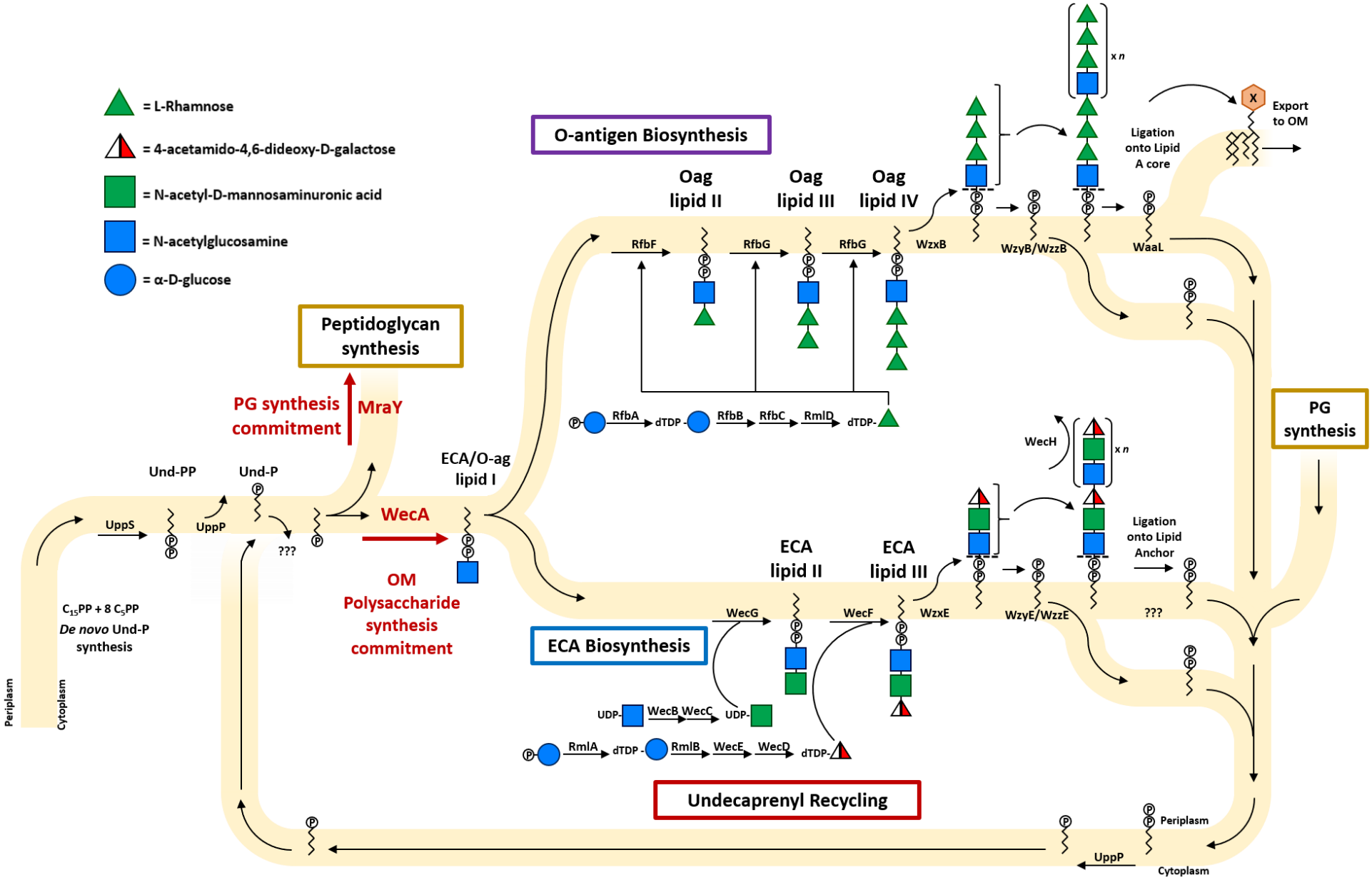

= L-Rhamnose
= 4-acetamido-4,6-dideoxy-D-galactose
= N-acetyl-D-mannosaminuronic acid
= N-acetylglucosamine
= α-D-glucose
O-antigen Biosynthesis
Oag lipid II
Oag lipid III
Oag lipid IV
RfbF
RfbG
RfbG
WzxB
WzyB/WzzB
WaaL
x n
Ligation onto Lipid A core
X
Export to OM
RfbA
dTDP -
RfbB
RfbC
RmlD
dTDP-
Peptidoglycan synthesis
PG synthesis commitment
MraY
PG synthesis
Und-PP
Und-P
ECA/O-ag lipid I
UppS
UppP
???
WecA
OM Polysaccharide synthesis commitment
C15PP + 8 C5PP
De novo Und-P synthesis
Periplasm
Cytoplasm
WecH
x n
Ligation onto Lipid Anchor
ECA lipid II
ECA lipid III
WecG
WecF
WzxE
WzyE/WzzE
???
ECA Biosynthesis
UDP-
WecB WecC
UDP-
RmlA
dTDP -
RmlB
WecE
WecD
dTDP-
Undecaprenyl Recycling
Periplasm
UppP
Cytoplasm

Figure 4.1: OM polysaccharide biosynthesis in *Shigella flexneri*.

Mutations in the genes post the generation of ECA and Oag lipid II, lead to the accumulation of lipid-linked intermediates and Und-P sequestration. Once translocated across the IM, Und-P is released by either Wzy or the final glycosyltransferase from each respective pathway. Undecaprenyl is firstly generated through *de novo* synthesis and then acted upon by UppS to yield Und-PP (Jukič et al. 2019). UppP then removes a phosphate group to yield Und-P, releasing Und-P on the periplasmic side of the IM (Tatar et al. 2007). Und-P is translocated back across the IM to the cytoplasmic side where then it can be acted on by MraY or WecA which commits Und-P to either PG or OM polysaccharide biosynthesis respectively until it is released from the pathways (Al-Dabbagh et al. 2016). WecA yields Und-P-GlcNAc which is a common lipid-linked intermediated for both ECA and Oag biosynthesis. It is acted upon by WecG or RfbF to yield ECA lipid-II or Oag lipid-II, committing the Und-P moiety to ECA or Oag biosynthesis respectively (Eade et al. 2021; Morona et al. 1994). Addition of sugar residues by WecF and RfbG yield ECA lipid-III, Oag lipid-III and lipid-IV respectively completing the ECA and Oag RU (Eade et al. 2021; Morona et al. 1994). RUs are then acted upon by separate homologs of the Wzy-dependent pathway (Wzx flippase, Wzy polymerase and Wzz polymerase co-polymerase denoted WzxB, WzyB, WzzB and WzxE, WzyE, WzzE for Oag and ECA biosynthesis respectively) at which point Und-PP is predominately released from the OM polysaccharide pathways (Islam & Lam 2014). Ligation of the complete polysaccharide chain onto final lipid carrier molecule also releases Und-PP from these pathways. A similar series of biosynthetic steps also occur in PG biosynthesis where MurJ and FtsW release PG RUs from Und-PP (Egan, Errington & Vollmer 2020). Und-PP is again acted upon by UppP to yield Und-P which is available for PG or OM polysaccharide biosynthesis once again (Tatar et al. 2007). Abbreviations: Oag, O antigen; UDP, uridine diphosphate; dTDP, thymidine diphosphate, X, LPS core-sugars; Und-PP, undecaprenyl pyrophosphate. Symbol nomenclature for glycans adapted from (Neelamegham et al. 2019; Varki et al. 2015).

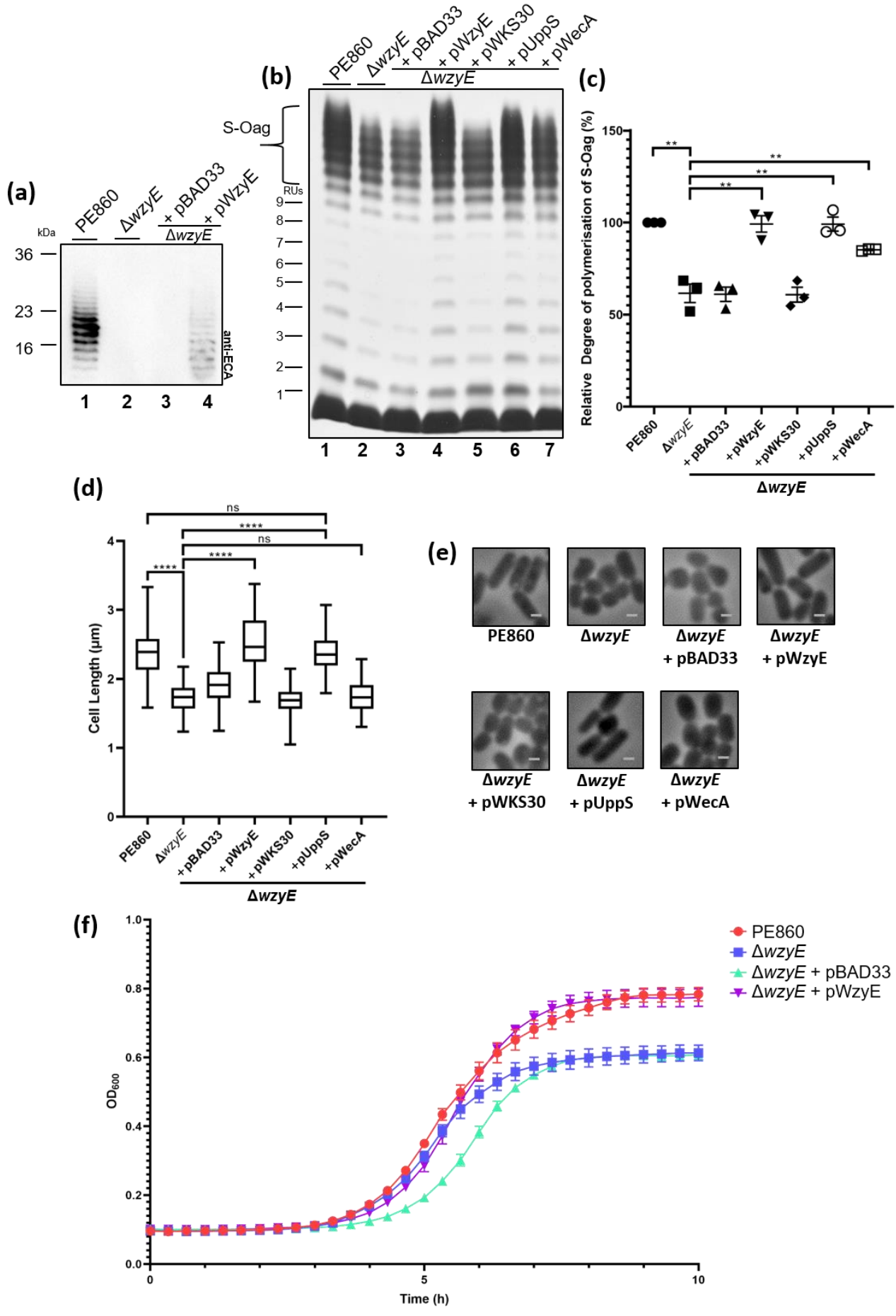

(a)
PE860
ΔwzyE
+ pBAD33
+ pWzyE
ΔwzyE
kDa
36
23
16
anti-ECA
1 2 3 4
(b)
PE860
ΔwzyE
+ pBAD33
+ pWzyE
+ pWKS30
+ pUppS
+ pWecA
ΔwzyE
S-Oag
RUs
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7
(c)
Relative Degree of polymerisation of S-Oag (%)
150
100
50
0
**
**
**
**
PE860
ΔwzyE
+pBAD33
+pWzyE
+pWKS30
+pUppS
+pWecA
ΔwzyE
(d)
Cell Length (µm)
4
3
2
1
0
ns
****
ns
****
****
PE860
ΔwzyE
+pBAD33
+pWzyE
+pWKS30
+pUppS
+pWecA
ΔwzyE
(e)
PE860
ΔwzyE
ΔwzyE
+ pBAD33
ΔwzyE
+ pWzyE
ΔwzyE
+ pWKS30
ΔwzyE
+ pUppS
ΔwzyE
+ pWecA
(f)
OD600
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
Time (h)
PE860
ΔwzyE
ΔwzyE + pBAD33
ΔwzyE + pWzyE

Figure 4.2: Analysis of the PE860 *ΔwzyE* phenotype.

(a) Anti-ECA Western immunoblot showing ECA banding profiles of PE860, an isogenic *ΔwzyE* mutant denoted as (*ΔwzyE*), and that mutant harbouring pBAD33 or pWzyE as indicated. Mid-exponential phase cells were collected ($1x10^9$ cells) and lysed in lysis buffer in the presence of proteinase K and following SDS-PAGE and Western transfer, the membrane was probed with polyclonal rabbit anti-ECA antibodies. SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard. (b) Analysis of LPS profiles of PE860, an isogenic *ΔwzyE* mutant denoted as (*ΔwzyE*), and that mutant harbouring pBAD33, pWzyE, pWKS30, pUppS or pWecA as indicated. Samples were made as above and electrophoresed on a SDS-15% (w/v) PAGE gel and silver stained. Number of Oag RUs are indicated on the left-hand side. (c) Analysis of polymerization of smooth Oag (S-Oag) via densitometry (Image lab). Degree of polymerization of S-Oag is represented as the desitometry of Oag RUs from 10-17 as a percentage relative to the parent. Data represents 3 independent experiments with SEM shown, and significance was calculated with independent student t-tests comparing the *ΔwzyE* mutant with relevant strains. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001, and ns is a non-statistically significant result. (d) and (e) Cellular measurements via Phase-contrast microscopy. 10 µl of mid-exponential phase culture was dried onto a microscope slide, mounted with Moviol and sealed. Cells were imaged and cell lengths measured. Scale bars = 1 µm. Data represents 150 individual cells with SEM shown, and significance was calculated with independent student t-tests comparing the PE860 *ΔwzyE* mutant with relevant strains. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001, and ns is a non-statistically significant result. The box covers the upper and lower quartiles of the data, with the bisecting bar indicating the median cell length, whereas the whiskers indicate the maximum and minimum cell lengths. (f) Analysis of growth of PE860, an isogenic *ΔwzyE* mutant denoted as (*ΔwzyE*), and that mutant harbouring harbouring pBAD33 or pWzyE as indicated. Overnight culture was sub-cultured 1/20 into a 96 well tray with OD600 absorbance readings taken every 20 minutes. Data represents averages of time points from 3 independent biological replicates.

4.6.2 Investigating the biosynthetic effects of the *ΔwzyB* mutation

Considering the pleiotropic phenotype observed for the *wzyE* mutant, we decided to investigate if the Oag biosynthetic pathway, through the use of a *wzyB* mutant, was similarly affected. Firstly, the OM polysaccharide profile of the *wzyB* mutant was assessed by silver-stained SDS PAGE where as expected, the *wzyB* mutant had LPS with a single Oag RU, seen as a single band (Figure 4.3a, lane 2). Introduction of pWzyB was able to partially complement the *wzyB* mutant (Figure 4.3a, lane 4).

ECA immunoblotting showed a reduction of ECA banding intensity between PE860 and the *wzyB* mutant (Figure 4.3b, lanes 1 and 2). When complemented with pWzyB, the *wzyB* mutant displayed an ECA banding pattern similar to the parent (Figure 4.3b, lanes 2 and 4). We then sought to investigate if similar to the *wzyE* mutant, the *wzyB* mutant also had altered cellular morphologies that could be rescued to wildtype morphology by complementation. Cell length measurements made by phase-contrast microscopy showed that the *wzyB* mutant cells were significantly shorter than the parent ($p < 0.001$). When complemented with pWzyB, their lengths reverted back to wildtype ($p < 0.001$) (Figure 4.3d and e).

Growth was then assessed by growth curves which showed that as similar to the *wzyE* mutant (Figure 4.2f), the *wzyB* mutant grew to a lower OD_{600} compared to the parent (Figure 4.3f). Complementing the *wzyB* mutant with pWzyB restored the growth of the *wzyB* mutant to near parental levels (Figure 4.3f). We likewise performed a CFU count on the *wzyB* mutant which showed that as with the *wzyE* mutant, the *wzyB* mutant produced less CFUs than PE860 at each time point investigated (Figure 4.6-(S1).

This supported the hypothesis that the two major OM polysaccharide populations of *S. flexneri*, ECA and LPS Oag, are fundamentally linked at the biosynthetic level where, disruption of one of the biosynthetic pathways interferes with the other, most likely due to the sequestration of Und-P. Additionally, the results also suggested that PG biosynthesis was also affected as Und-P is required in the biosynthesis of PG biosynthetic intermediates (Figure 4.1).

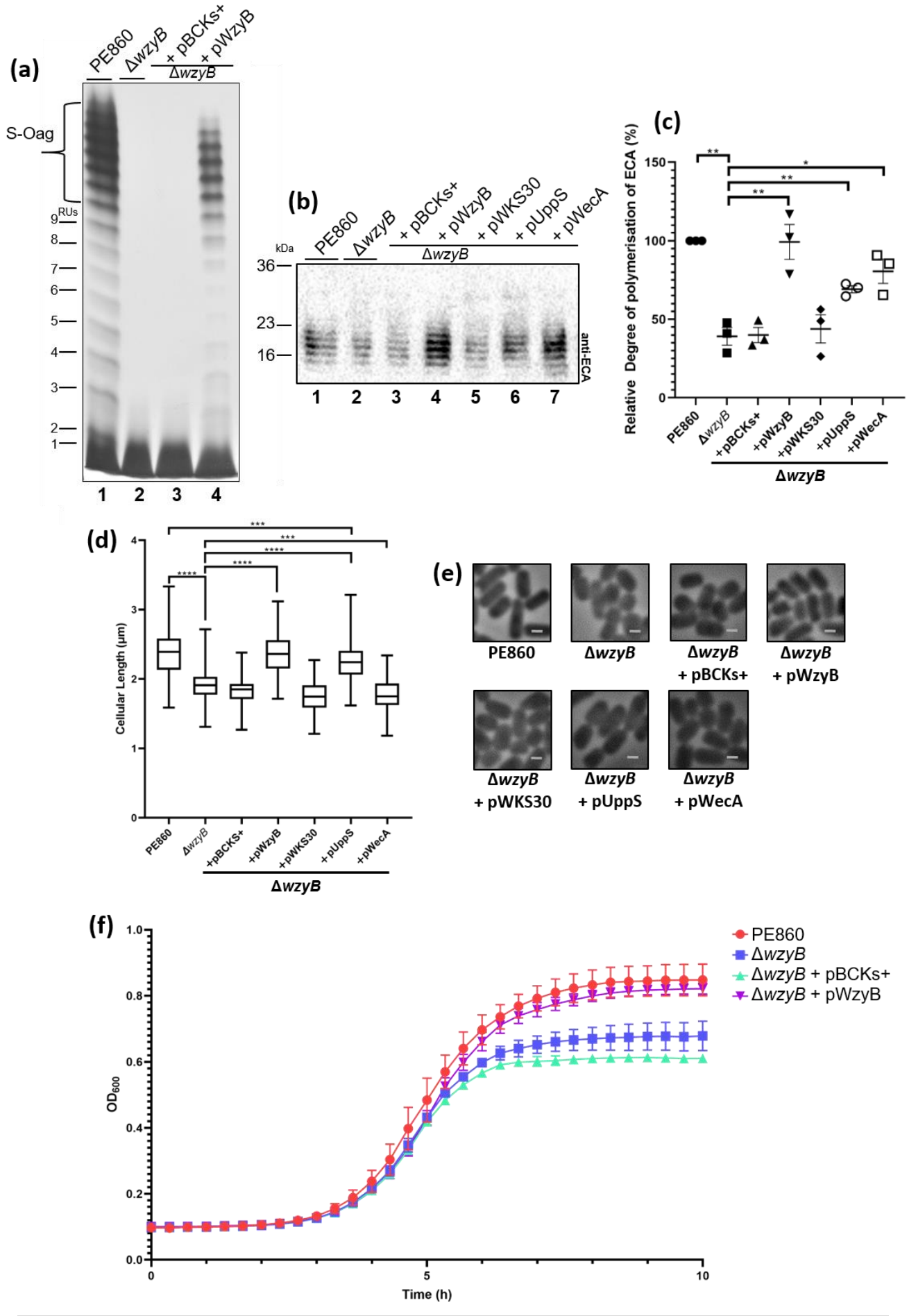

(a)
PE860
ΔwzyB
+ pBCKs+
+ pWzyB
ΔwzyB
S-Oag
RUs
9
8
7
6
5
4
3
2
1
1 2 3 4
(b)
PE860
ΔwzyB
+ pBCKs+
+ pWzyB
+ pWKS30
+ pUppS
+ pWecA
ΔwzyB
kDa
36
23
16
anti-ECA
1 2 3 4 5 6 7
(c)
Relative Degree of polymerisation of ECA (%)
150
100
50
0
**
**
**
*
PE860
ΔwzyB
+pBCKs+
+pWzyB
+pWKS30
+pUppS
+pWecA
ΔwzyB
(d)
Cellular Length (µm)
4
3
2
1
0
***
***
****
****
****
PE860
ΔwzyB
+pBCKS+
+pWzyB
+pWKS30
+pUppS
+pWecA
ΔwzyB
(e)
PE860
ΔwzyB
ΔwzyB
+ pBCKs+
ΔwzyB
+ pWzyB
ΔwzyB
+ pWKS30
ΔwzyB
+ pUppS
ΔwzyB
+ pWecA
(f)
OD600
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
Time (h)
PE860
ΔwzyB
ΔwzyB + pBCKs+
ΔwzyB + pWzyB

Figure 4.3: Analysis of the PE860 *ΔwzyB* phenotype.

(a) Analysis of LPS profiles of PE860, an isogenic *ΔwzyB* mutant denoted as (*ΔwzyB*), and that mutant harbouring pBCKs+ or pWzyB as indicated. Mid-exponential phase cells were collected ($1x10^9$ cells) and lysed in lysis buffer in the presence of proteinase K. Samples were then electrophoresed on a SDS-15% (w/v) gel and silver stained. Number of Oag RUs are indicated on the left-hand side. (b) Anti-ECA Western immunoblot of PE860, an isogenic *ΔwzyB* mutant denoted as (*ΔwzyB*), and that mutant harbouring pBCKs+, pWzyB, pWKS30, pUppS or pWecA as indicated. Samples were made as above and following SDS-PAGE and Western transfer, the membrane was probed with polyclonal rabbit anti-ECA antibodies. SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard. (c) Analysis of polymerization of ECA via densitometry. Degree of polymerization of ECA is represented as the desitometry of ECA RUs as a percentage relative to the parent. Data represents 9 independent experiments with SEM shown, and significance was calculated with independent student t-tests comparing the *ΔwzyB* mutant with relavent strains. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001, and ns is a non-statistically significant result. (d) and (e) Cellular measurements via Phase-contrast microscopy. 10 µl of mid-exponential phase culture was dried onto a microscope slide, mounted with Moviol and sealed. Cells were imaged and cell lengths measured. Scale bars = 1 µm. Data represents 150 individual cells with SEM shown, and significance was calculated with independent student t-tests comparing the ΔwzyE mutant with relavent strains. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001, and ns is a non-statistically significant result. The box covers the upper and lower quartiles of the data, with the bisecting bar indicating the median cell length, whereas the whiskers indicate the maximum and minimum cell lengths. (f) Analysis of growth of PE860, an isogenic *ΔwzyB* mutant denoted as (*ΔwzyB*), and that mutant harbouring pBCKs+ or pWzyB as indicated. Overnight culture was sub-cultured 1/20 into a 96 well tray with OD600 absorbance readings taken every 20 minutes. Data represents averages of time points from 3 independent biological replicates.

4.6.3 Improving cellular undecaprenyl pool rescues mutant *wzy* phenotypes

To investigate if the pleiotropic *wzy* mutant phenotypes were due to a lack in available Und-P, *wecA,* which generates ECA/Oag lipid I and commits Und-P to OM polysaccharide biosynthesis in *S. flexneri,* and *uppS,* which is responsible for *de novo* synthesis of Und-PP, were ectopically expressed. The constructs pWKS30-UppS-HA and pWKS30-WecA-HA, denoted as pUppS and pWecA were generated, transformed into the *wzyE* and *wzyB* mutants and assessed to determine if either theoretically increasing or decreasing the cellular pool of Und-P through the expression of *uppS* either alleviated or exacerbated their pleiotropic phenotypes.

In the *wzyE* mutant, LPS analysis showed a statistically significant restoration in Oag banding intensity with either pUppS or pWecA ($p < 0.0001$ and $p < 0.01$ respectively) (Figure 4.2b, lanes 2, 6 & 7, Figure 4.2c). However, cell measurements showed that only complementation with pUppS which would theoretically increase the cellular pool of Und-P, and not pWecA, was able to restore the *wzyE* mutant length back to a wildtype ($p < 0.0001$ and non-significant, respectively) (Figure 4.2d and e). Furthermore, growth curves showed that upon introduction of pUppS, the *wzyE* mutant was able reach a higher OD_{600} of 0.7 whereas, the introduction of pWecA resulted in the *wzyE* mutant reaching a lower OD_{600} of 0.5 after 10 hours of growth, with the *wzyE* mutant reaching an OD_{600} of 0.6 (Figure 4.4a).

Similar to the *wzyE* mutant, the *wzyB* mutant also showed partial rescue from its mutant *wzy* phenotypes upon introduction of pUppS and pWecA. The *wzyB* mutant displayed a statistically significant restoration in ECA banding intensity when transformed with either pUppS or pWecA (Figure 4.3b, lanes 2, 6 & 7 and Figure 4.3c). Cell measurements revealed that, similar to the *wzyE* mutant, the *wzyB* mutant length was only restored back to a wildtype length by pUppS and not pWecA ($p < 0.0001$ and $p < 0.001$ respectively) (Figure 4.3d and e). Growth analysis further showed that the addition of pUppS did not improve the growth of the *wzyB* mutant, and introduction of pWecA led to a decrease in growth relative to the *wzyB* mutant, reaching an OD_{600} of 0.68 and 0.58 after 10 hours of growth respectively compared to the *wzyB* mutant reaching an OD_{600} of 0.67 (Figure 4.4b).

These results further supported the hypothesis that the three pathways are fundamentally linked through the common use of Und-P in their biosynthetic pathways. However, a plausible alternative explanation to the cause of the phenotypes was that the loss of OM ECA or LPS Oag was indirectly impacting the other non-affected polysaccharide pathway. To investigate this, we subsequently analysed whether the theoretical reduction in available Und-P was the cause of the observed phenotypes, or if the loss of the OM polysaccharides themselves were responsible.

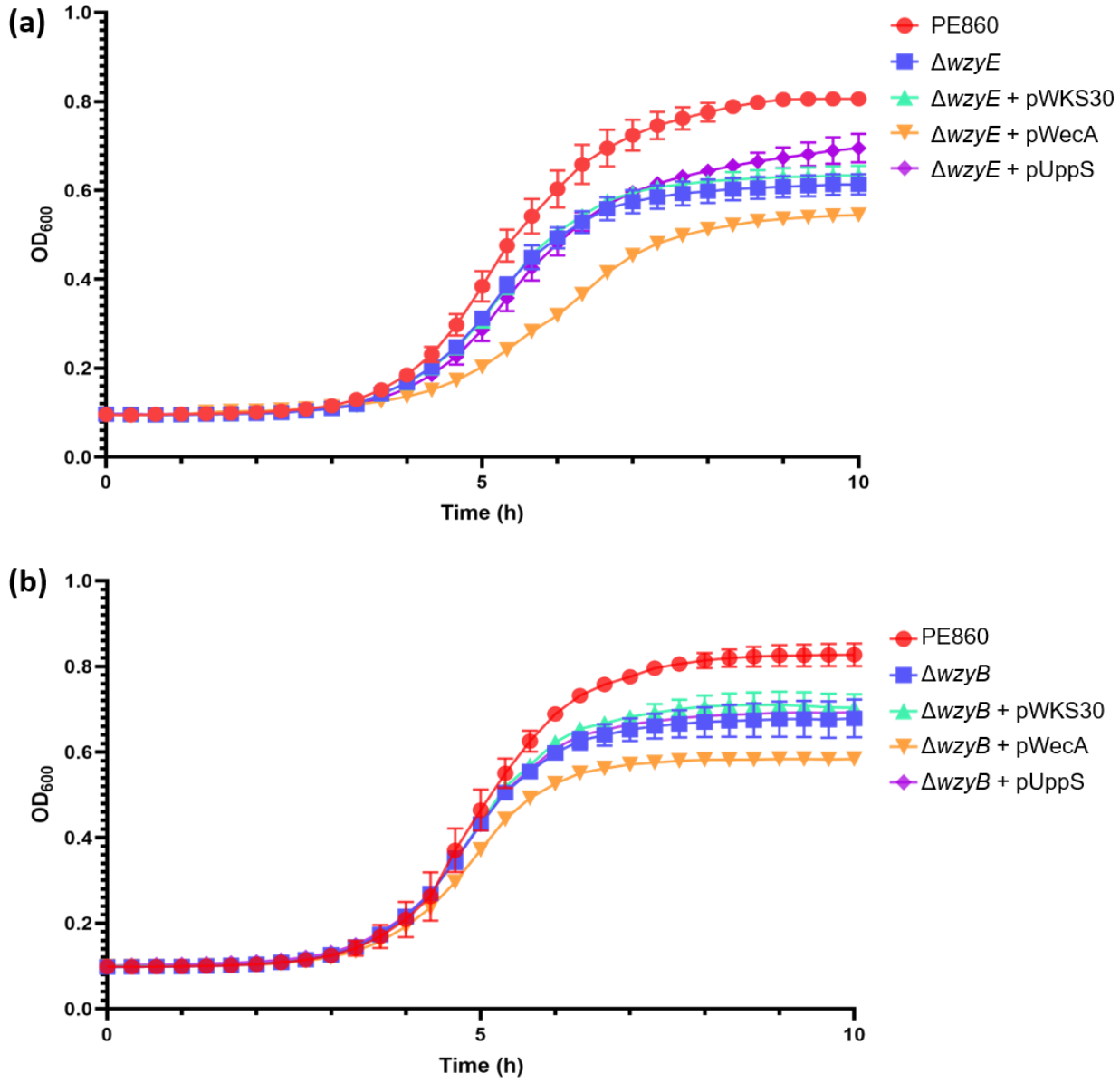


Figure 4.4: Effects of increasing or decreasing the cellular availability of Und-P on *Δwzy* mutant growth.

Analysis of growth of strains (a) PE860, an isogenic *ΔwzyE* mutant denoted as (*ΔwzyE*), and that mutant harbouring pWKS30, pWecA or pUppS and (b) PE860, an isogenic *ΔwzyB* mutant denoted as (*ΔwzyB*), and that mutant harbouring pWKS30, pWecA or pUppS as indicated. Overnight culture of the above strains were sub-cultured 1/20 into a 96 well tray with OD600 absorbance readings taken every 20 minutes. Data represents averages of time points from 3 independent biological replicates.

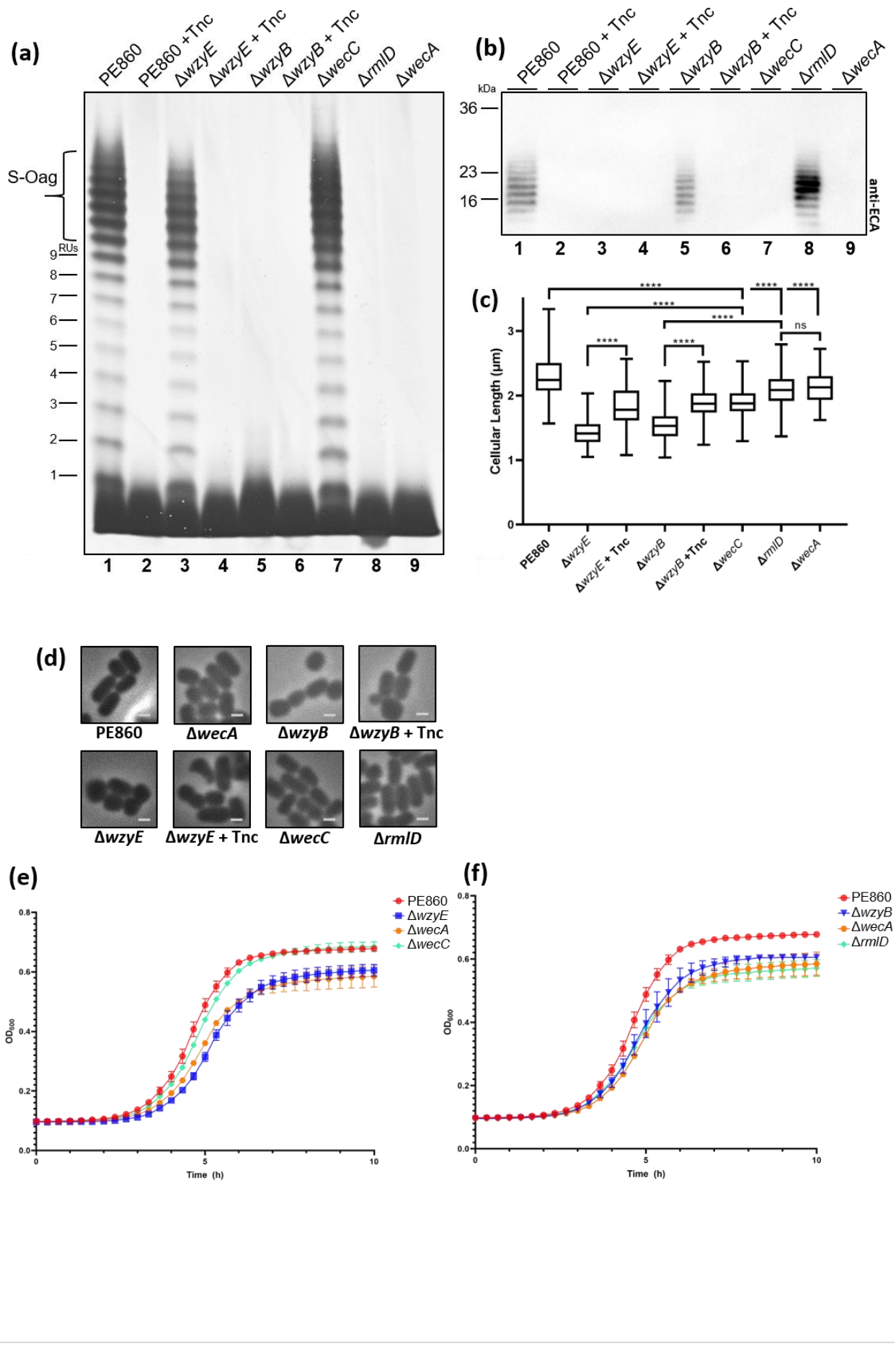

(a)
PE860
PE860 +Tnc
ΔwzyE
ΔwzyE + Tnc
ΔwzyB
ΔwzyB + Tnc
ΔwecC
ΔrmlD
ΔwecA
S-Oag
RUs
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7 8 9
(b)
kDa
36
23
16
anti-ECA
1 2 3 4 5 6 7 8 9
(c)
Cellular Length (μm)
ns
PE860
ΔwzyE
ΔwzyE + Tnc
ΔwzyB
ΔwzyB +Tnc
ΔwecC
ΔrmlD
ΔwecA
(d)
PE860
ΔwecA
ΔwzyB
ΔwzyB + Tnc
ΔwzyE
ΔwzyE + Tnc
ΔwecC
ΔrmlD
(e)
PE860
ΔwzyE
ΔwecA
ΔwecC
OD600
Time (h)
(f)
PE860
ΔwzyB
ΔwecA
ΔrmlD
OD600
Time (h)

Figure 4.5: The impact of *wecA*, *wecC* and *rmlD* mutations and tunicamycin treatment on ECA, Oag and growth.

(a) Analysis of LPS profiles of strains PE860, treated with and without tunicamycin, PE860 *ΔwzyE* denoted as (*ΔwzyE*) treated with and without tunicamycin, PE860 *ΔwzyB* denoted as (*ΔwzyB*) treated with and without tunicamycin and PE860 mutants *ΔwecC*, *ΔrmlD* and *ΔwecA* as indicated. Cells were grown to mid-exponential phase and collected ($1x10^9$ cells) and lysed in lysis buffer in the presence of proteinase K. Samples were then electrophoresed on a SDS-15% (w/v) gel and silver stained. Treatment with tunicamycin denoted as (+ Tnc). Number of Oag RUs are indicated on the left-hand side. (b) Anti-ECA Western immunoblot of the same strains as above as indicated. Samples were made as above and following SDS-PAGE and Western transfer, the membrane was probed with polyclonal rabbit anti-ECA antibodies. SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard. (c) and (d) Cellular length measurements via Phase-contrast microscopy. 10 µl of mid-exponential phase culture was dried onto a microscope slide, mounted with Moviol and sealed. Cells were imaged and cell lengths measured. Scale bars = 1 µm. Data represents 150 individual cells with SEM shown, and significance was calculated with independent student t-tests comparing the relavent strains. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$, and ns is a non-statistically significant result. The box covers the upper and lower quartiles of the data, with the bisecting bar indicating the median cell length, whereas the whiskers indicate the maximum and minimum cell lengths. Treatment with tunicamycin denoted as (+ Tnc). Analysis of growth of strains (e) PE860 and PE860 mutants *ΔwzyE*, *ΔwecA* and *ΔwecC* and (f), PE860 and PE860 mutants *ΔwzyB, ΔwecA* and *ΔrmlD* as indicated. Overnight cultures of the above strains were sub-cultured 1/20 into a 96 well tray with OD600 absorbance readings taken every 20 minutes. Data represents averages of time points from 3 independent biological replicates.

4.6.4 The sequestration of Und-P and not the loss of OM polysaccharides is correlated to the *wzy* mutant phenotypes

We investigated the impact of polysaccharide specific mutations in additional genes involved in the biosynthesis of ECA and LPS Oag specific biosynthetic intermediates, *wecC* and *rmlD,* respectively (Figure 4.1). Both genes were selected as they are involved in the biosynthesis of precursors for ECA and Oag lipid II respectively and hence, once mutated, would prevent the formation of the substrates required by the second glycosyltransferases of each pathways, WecG and RfbF, which consequently would prevent the sequestration of Und-P. Additionally, we investigated the impact of a *wecA* mutant which is involved in the biosynthesis of lipid I, the common precursor for both OM polysaccharide biosynthetic pathways (Figure 4.1).

The *wecC* mutant, as expected, was unable to produce ECA (Figure 4.5b, lane 7) however, unlike the *wzyE* mutant (Figure 4.5a, lane 3), it had an increase in Oag banding intensity compared to the parent PE860 (Figure 4.5, lane 7). The *rmlD* mutant showed a total loss of LPS Oag banding (Figure 4.5a, lane 8) but, unlike the *wzyB* mutant (Figure 4.5b, lane 5), showed an increased ECA banding intensity compared to PE860 (Figure 4.5b, lane 8). As expected, the *wecA* mutant showed a total loss of both OM polysaccharides, seen as a lack of banding in both the ECA immunoblot and the SDS PAGE silver stain gel, respectively (Figure 4.5a and Figure 4.5b, lane 9). The cellular morphologies of the *wecC, rmlD* and *wecA* mutants were then assessed. This showed the *wecA, wecC* and *rmlD* mutants were significantly longer than the *wzyE* and *wzyB* mutants (Figure 4.5c and d). Additionally, when comparing the parent with the *wecC, rmlD* and *wecA* mutants, a statistically significant reduction in cell length ($p < 0.0001$) was observed with no statistically significant difference in cell length observed when comparing the *rmlD* mutant with the *wecA* mutant (Figure 4.5c and d). Growth analysis showed that the *wecC* mutant grew at a rate similar to the parental strain reaching an OD_{600} of 0.65 after 10 hours of growth, whereas the *rmlD* and *wecA* mutants were only able to reach an OD_{600} of 0.51 and 0.55, respectively, after 10 hours of growth (Figure 4.5e and f).

Overall, the data supported the hypothesis that if one OM polysaccharide was absent, the reciprocal OM polysaccharide system could function without interference and, that the cause of the *wzy* pleiotropic mutant phenotypes was likely not due to a lack in reciprocal OM polysaccharide (Figure 4.1). Additionally, growth curve data suggested that there was an effect on the growth of strains lacking Oag as the *rmlD* mutant had similar growth profile as the *wecA* mutant, presumably due to the activation of the Rcs pathway (Meng et al. 2020)

Therefore, to further investigate if the lack of Und-P was the cause of the *wzy* mutant phenotypes, tunicamycin, an inhibitor of WecA (Heifetz, Keenan & Elbein 1979), was used to treat both *wzy* mutants to prevent lipid I production (Figure 4.1). Following treatment with tunicamycin, the parent and the mutants, *wzyE* or *wzyB*, showed no detectable OM polysaccharide production (Figure 4.5a and b) as expected. However, microscopy analysis revealed that post treatment with tunicamycin, the two *wzy* mutants showed a statistically significant restoration in cellular length ($p < 0.0001$) (Figure 4.5c and d) which supported our hypothesis that the lack of available Und-P was most likely the cause of the pleiotropic phenotypes of the *wzy* mutants.

4.7 Article Discussion

Jorgenson et al. (2016) demonstrated that *E. coli* K-12 mutants of the OM polysaccharide biosynthetic pathways LPS Oag and ECA, displayed morphological abnormalities due to the sequestration of dead-end biosynthetic intermediates within the cell (Jorgenson et al. 2016; Jorgenson & Young 2016). Here, we demonstrate that the two OM polysaccharide pathways of *S. flexneri* are intimately linked, through their use of the universal biosynthetic membrane anchor, Und-P.

Our study demonstrates novel indirect cross-talk between the two OM polysaccharide biosynthetic pathways, LPS Oag and ECA. Direct interactions between the two Wzy-dependent pathways has been previously speculated, due to the structural similarities between the Wzz protomers from each pathway as well as molecular similarities in the substrates they translocate and polymerize (Kalynych et al. 2015; Marolda et al. 2006). Most recently it was shown that interchanging the transmembrane (TM) regions of WzzB and WzzE, most importantly TM2, was sufficient to restore polysaccharide modal length control in *S. flexneri wzzB* and *wzzE* mutants complemented with $WzzB_{ETM2}$ or $WzzE_{BTM2}$ respectively (Leo 2020). *Leo et al. (2020)* further demonstrated that WzyB could natively pull down WzzE, as well as the chimeric Wzz proteins, which had only been demonstrated with WzyB and WzzB at the time of publication (Leo 2020).

In our study, we have shown that the two OM polysaccharide pathways interact indirectly through perceived competition for Und-P, where for the first time a drop in abundance of the reciprocal OM polysaccharide, ECA and LPS Oag was observed via anti-ECA Western immunoblotting and LPS silver staining (Figure 4.2b and Figure 4.3b) when the *wzyB/wzyE* homologs were deleted, respectively. The abundance of OM polysaccharide was restored upon complementation of the *wzy* mutants or via ectopic expression of pUppS or pWecA which is in accordance with experimental observations from other OM biosynthetic mutants (Jorgenson et al. 2016; Jorgenson & Young 2016; Marolda et al. 2006).

It is well documented that mutations affecting the ECA and Oag biosynthetic pathways across Enterobacteriales produce abnormal cellular morphologies (Castelli & Vescovi 2011; Jorgenson et al. 2016; Jorgenson & Young 2016; Liu, D, Cole & Reeves 1996; Rick et al. 1988; Yuasa, Levinthal & Nikaido 1969). Specifically in these studies, mutations *in E. coli* K-12 which led to biosynthetic dead-ends were used. Similarly, to the mutant phenotypes observed in *E. coli* K-12, the *S. flexneri wzy* mutants displayed altered cell morphology albeit a shorter cell length (Figure 4.2d and e and Figure 4.3d and e), unlike those previously observed mutants which produced elongated cells (Jorgenson et al. 2016; Jorgenson & Young 2016).

A plausible explanation as to why the *wzy* mutants did not display the same phenotypes observed by *Jorgenson et al.* is that instead of trapping Und-P in dead-end intermediates, *wzy* mutants only greatly restrict the recycling of Und-P from the lipid-linked intermediates as they produce OM polysaccharides which consist of a single RU (Figure 4.1). This is seen in Figure 4.5a as when treated with tunicamycin, *wzyB* mutants lose their single RU and express a LPS profile similar to that of a *wecA* and or *rmlD* mutant (Figure 4.5a lanes 8 and 9). This is because Wzys are not the only enzymes which release Und-P from the intermediates, but do so alongside with WaaL and the final unknown ECA ligase equivalent. As both WaaL and the putative ECA ligase are present in the cell, the complete sequestering of Und-P does not occur. Additionally, as stated above, the strains commonly used when these phenotypes were investigated are derived from *E. coli* K-12 (Castelli & Vescovi 2011; Jorgenson et al. 2016; Jorgenson & Young 2016; Liu, D, Cole & Reeves 1996; Rick et al. 1988; Yuasa, Levinthal & Nikaido 1969).

Isolated in 1922, *E. coli* K-12 is a cycled, laboratory strain which historically was exposed to mutagens and unintentionally became rough, and has existed as a rough strain for an extended period of time (Bachmann 1972; Hobman, Penn & Pallen 2007; Liu, D & Reeves 1994; Stevenson et al. 1994). As secondary mutations occur frequently in OM polysaccharide mutants (Jorgenson et al. 2016; Jorgenson & Young 2016), it is plausible that studies in *E. coli* K-12 are unable to fully represent the intricacy of the OM biosynthetic pathways which thus cautions direct comparisons of these OM pathways between rough and smooth LPS backgrounds such as that of *S. flexneri*.

The restriction of Und-P from related pathways is a common cause of the pleiotropic phenotypes seen in cell wall mutants and are frequently corrected by the supplementation of Und-PP via expression of Undecaprenyl pyrophosphotase synthase (UppS) (Jorgenson et al. 2016; Jorgenson & Young 2016). UppS synthesizes Und-PP in cells which is then acted upon by UppP to yield Und-P where, its over expression was shown to correct the abnormal morphologies of *E. coli* K-12 *waaC*, *waaL* and *wecE* mutants (Jorgenson et al. 2016; Jorgenson & Young 2016; Jukič et al. 2019; Tatar et al. 2007). This is consistent with our findings where the expression of pUppS was able restore wildtype ECA and LPS banding (Figure 4.2b and c & Figure 4.3b and c), reverting the *wzy* mutants back to wildtype length (Figure 4.2d and e & Figure 4.3d and e) as well as improving the growth of the *wzyE* mutant (Figure 4.4a).

Over expression of WecA is also a common tool to drive OM polysaccharide intermediates into dead-end pathways to exacerbate and investigate mutant phenotypes however, distinctions cannot be made that the affects seen are due to the accumulation of intermediates or sequestering

Und-P; rather that both affects are responsible (Jorgenson et al. 2016; Jorgenson & Young 2016; Marolda et al. 2006; Rick et al. 2003). The addition of pWecA did not decrease the cellular lengths of the *wzy* mutants (Figure 4.2 and e & Figure 4.3d and e). However, it did alter the growth characteristics of the cells whereby they were unable to reach a similar OD_{600} as the *wzy* mutants (Figure 4.4a and b) which is presumably due to Und-P being sequestered away from PG synthesis. ECA and LPS Oag production was restored to near wildtype which was expected as *wecA* drives the OM polysaccharide biosynthetic pathways forward (Figure 4.2a and b & Figure 4.3a and b) (Al-Dabbagh et al. 2016).

Additionally, we used tunicamycin to inhibit WecA and block OM polysaccharide biosynthesis. Tunicamycin is a known cell wall glycosyltrasferase inhibitor, preventing the transfer of GlcNAc and or MurNAc-pentapeptides onto polyprenylphosphate lipid carriers (Elbein, Gafford & Kang 1979; Heifetz, Keenan & Elbein 1979). Despite targeting MraY as well as WecA, the inhibitory concentration for WecA was shown to be 1000x less than that of MraY, allowing us to selectively inhibit WecA by using MraY sub-inhibitory concentrations (Al-Dabbagh, Mengin-Lecreulx & Bouhss 2008). As expected, treatment with tunicamycin inhibited the production of ECA or Oag in the *wzy* mutants (Figure 4.5a and b). Interestingly, the *wzy* mutants displayed longer cell lengths when treated with tunicamycin than when untreated which supports our hypothesis that the *wzy* mutant phenotypes may be partially caused by a lack of available of Und-P for PG biosynthesis (Figure 4.5c and d). Overall, as the results from the theoretical expression of UppS was sufficient to rescue the pleiotropic phenotypes of the *wzy* mutants and, as a similar phenotypic rescue was observed when the *wzy* mutants were treated with tunicamycin, we believe the results support our hypothesis that the pleiotropic phenotypes of the *wzy* mutants was due to a lack of available Und-P as it was sequestered in biosynthetic intermediates.

Lastly, we showed that not all OM polysaccharide mutations lead to the mutant phenotypes associated with cell wall mutants. The deletion of *wecC* and *rmlD*, which cause the loss of ECA and LPS Oag respectively, appeared to increase the OM polysaccharide production of the other pathway when comparing to the parent (Figure 4.5a and b). Length measurements also supported our hypothesis that both the *wecC* and *rmlD* mutants possibly did not sequester Und-P from PG synthesis, as each mutant possessed a statistically significantly longer cell length compared to the *wzy* mutants ($P<0.0001$) (Figure 4.5c and d). These results suggested that if one of the OM polysaccharide pathways is blocked in such a way that the cell does not theoretically sequester Und-P and/or biosynthetic intermediates, then Und-P can be funnelled into the other non-affected OM polysaccharide pathway, leading to an increase in that OM polysaccharide population.

Additionally, we found a reduction in final OD_{600} by the *rmlD* and *wecA* mutants when compared to the *wecC* mutant (Figure 4.5e and f) as well by the *wzyB* mutant expressing pUppS (Figure 4.4b), we speculate that this can be explained by the induction of the Rcs and Cpx pathways.

The Rcs and Cpx membrane stress response pathways are known to contribute to the phenotypes of cell wall mutants (Klein & Raina 2019; Meng, Young & Chen 2021; Raivio, Popkin & Silhavy 1999; Wall, Majdalani & Gottesman 2018) and have been implicated in their pleiotropic phenotypes across various backgrounds: *Serratia marcescens* (Castelli et al. 2008; Castelli & Vescovi 2011), *E. coli* K-12 (Danese et al. 1998; Evans et al. 2013; Jorgenson et al. 2016; Jorgenson & Young 2016; Laubacher & Ades 2008) and *Proteus mirabilis* (Morgenstein, R. M., Clemmer & Rather 2010; Morgenstein, Randy M. & Rather 2012). It has been shown that the lack of Oag due to a *waaL* mutation in *P. mirabilis* was sufficient to induce the Rcs pathway (Morgenstein, R. M., Clemmer & Rather 2010; Morgenstein, Randy M. & Rather 2012). Therefore it is plausible that in *S. flexneri*, the lack of Oag itself can induce the Rcs pathway leading to the observed reduced growth in the *rmlD*, *wecA* and *wzyB* mutants, including the *wzyB* mutant when complemented with pUppS (Figure 4.4b and Figure 4.5f). Whereas it has been indicated specifically that the mutations which accumulate lipid linked intermediates, and not the lack of ECA itself, stimulate the Rcs pathway (Jorgenson et al. 2016). Therefore, as *wecC* mutants do not accumulate lipid linked intermediates (Figure 4.1), the Rcs pathway is not stimulated in the *wecC* mutant allowing for growth comparable to the parent (Figure 4.5e).

In conclusion, the results here reveal another layer of complexity in our understanding of the interactions between the cell wall biosynthetic pathways. The critical link between LPS, ECA and PG biosynthesis lies in their combined reliance on Und-P as the cells universal membrane anchor (Figure 4.1). As such, cell wall mutants derived from these pathways are likely to provide unclear mutant phenotypes, influenced by their related and non-related pathways such as Rcs and Cpx pathways. This also calls into question previously published data in which cell wall mutants were investigated, as alternative plausible explanations may now better explain the phenomenona observed. These findings provide a greater understanding of the complexities and interlacings of cell wall biosynthetic pathways, and provide an awareness of the possible entanglements which can occur when investigating cell wall mutants.

4.8 Article Reference

Al-Dabbagh B, Mengin-Lecreulx D, Bouhss A. 2008. Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA. *J Bacteriol* 190:7141-7146.

Al-Dabbagh B, Olatunji S, Crouvoisier M, El Ghachi M, Blanot D, Mengin-Lecreulx D, Bouhss A. 2016. Catalytic mechanism of MraY and WecA, two paralogues of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily. *Biochimie* 127:249-257.

Bachmann BJ. 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol Rev* 36:525-557.

Castelli ME, Fedrigo GV, Clementin AL, Ielmini MV, Feldman MF, Garcia Vescovi E. 2008. Enterobacterial common antigen integrity is a checkpoint for flagellar biogenesis in *Serratia marcescens*. *J Bacteriol* 190:213-20.

Castelli ME, Vescovi EG. 2011. The Rcs signal transduction pathway is triggered by enterobacterial common antigen structure alterations in *Serratia marcescens*. *J Bacteriol* 193:63-74.

Danese PN, Oliver GR, Barr K, Bowman GD, Rick PD, Silhavy TJ. 1998. Accumulation of the enterobacterial common antigen Lipid II biosynthetic intermediate Stimulates *degP* transcription in *Escherichia coli*. *J Bacteriol* 180:5875.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-5.

Eade CR, Wallen TW, Gates CE, Oliverio CL, Scarbrough BA, Reid AJ, Jorgenson MA, Young KD, Troutman JM. 2021. Making the enterobacterial common antigen glycan and measuring its substrate sequestration. *ACS Chem Biol* 16, 691-700.

Egan AJF, Errington J, Vollmer W. 2020. Regulation of peptidoglycan synthesis and remodelling. *Nat Rev Microbiol* 18:446-460.

Elbein AD, Gafford J, Kang MS. 1979. Inhibition of lipid-linked saccharide synthesis: comparison of tunicamycin, streptovirudin, and antibiotic 24010. *Arch Biochem Biophys* 196:311-8.

Evans KL, Kannan S, Li G, de Pedro MA, Young KD. 2013. Eliminating a set of four penicillin binding proteins triggers the Rcs phosphorelay and Cpx stress responses in *Escherichia coli*. *J Bacteriol* 195:4415-24.

Günther SD, Fritsch M, Seeger JM, Schiffmann LM, Snipas SJ, Coutelle M, Kufer TA, Higgins PG, Hornung V, Bernardini ML, Höning S, Krönke M, Salvesen GS, Kashkar H. 2019. Cytosolic Gram-negative bacteria prevent apoptosis by inhibition of effector caspases through lipopolysaccharide. *Nat Microbiol* 5:354-367.

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

Heifetz A, Keenan RW, Elbein AD. 1979. Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate GlcNAc-1-phosphate transferase. *Biochemistry* 18:2186-2192.

Hobman JL, Penn CW, Pallen MJ. 2007. Laboratory strains of *Escherichia coli*: model citizens or deceitful delinquents growing old disgracefully? *Mol Microbiol* 64:881-5.

Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can J Microbiol* 60:697-716.

Jorgenson MA, Kannan S, Laubacher ME, Young KD. 2016. Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate. *Mol Microbiol* 100:1-14.

Jorgenson MA, Young KD. 2016. Interrupting biosynthesis of O antigen or the lipopolysaccharide core produces morphological defects in *Escherichia coli* by sequestering undecaprenyl phosphate. *J Bacteriol* 198:3070-3079.

Jukič M, Rožman K, Sova M, Barreteau H, Gobec S. 2019. Anthranilic acid inhibitors of undecaprenyl pyrophosphate synthase (UppS), an essential enzyme for bacterial cell wall biosynthesis. *Frontiers in Microbiol* 9:3322.

Kalynych S, Cherney M, Bostina M, Rouiller I, Cygler M. 2015. Quaternary structure of WzzB and WzzE polysaccharide copolymerases. *Protein Sci* 24:58-69.

Klein G, Raina S. 2019. Regulated assembly of LPS, its structural alterations and cellular rResponse to LPS defects. *Int Mol Sci* 20:356.

Laubacher ME, Ades SE. 2008. The Rcs phosphorelay is a cell envelope stress response activated by peptidoglycan stress and contributes to intrinsic antibiotic resistance. *J Bacteriol* 190:2065-74.

Leo V, Tran E, Morona R. 2020. Polysaccharide co-polymerase WzzB/WzzE chimeras reveal transmembrane 2 region of WzzB is important for interaction with WzyB. *J Bacteriol* 203:598-20.

Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SyN, Wang Q, Reeves PR, Wang L. 2008. Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* 32:627-653.

Liu D, Cole RA, Reeves PR. 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. *J Bacteriol* 178:2102-2107.

Liu D, Reeves PR. 1994. *Escherichia coli* K12 regains its O antigen. *Microbiology* 140):49-57.

Marolda CL, Tatar LD, Alaimo C, Aebi M, Valvano MA. 2006. Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide o antigen. *J Bacteriol* 188:5124-35.

Meeske AJ, Riley EP, Robins WP, Uehara T, Mekalanos JJ, Kahne D, Walker S, Kruse AC, Bernhardt TG, Rudner DZ. 2016. SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature* 537:634-638.

Meng J, Young G, Chen J. 2021. The Rcs system in *Enterobacteriaceae*: Envelope stress responses and virulence regulation. *Frontiers Microbiol* 12:627104.

Mitchell AM, Srikumar T, Silhavy TJ. 2018. Cyclic Enterobacterial common antigen maintains the outer membrane permeability barrier of *Escherichia coli* in a manner controlled by YhdP. *MBio* 9:e01321-18.

Morgenstein RM, Clemmer KM, Rather PN. 2010. Loss of the *waaL* O-antigen ligase prevents surface activation of the flagellar gene cascade in *Proteus mirabilis*. *J Bacteriol* 192:3213-21.

Morgenstein RM, Rather PN. 2012. Role of the Umo proteins and the Rcs phosphorelay in the swarming motility of the wild type and an O-antigen (*waaL*) mutant of *Proteus mirabilis*. *J Bacteriol* 194:669-676.

Morona R, Mavris M, Fallarino A, Manning PA. 1994. Characterization of the *rfc* region of *Shigella flexneri*. *J Bacteriol* 176:733.

Murray GL, Attridge SR, Morona R. 2003. Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol Microbiol* 47:1395-1406.

Neelamegham S, Aoki-Kinoshita K, Bolton E, Frank M, Lisacek F, Lütteke T, O'Boyle N, Packer NH, Stanley P, Toukach P, Varki A, Woods RJ. 2019. Updates to the symbol nomenclature for glycans guidelines. *Glycobiology* 29:620-624.

Purins L, Van Den Bosch L, Richardson V, Morona R. 2008. Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2. *Microbiology* 154:1104-1116.

Rai AK, Mitchell AM. 2020. Enterobacterial common antigen: Synthesis and function of an enigmatic molecule. *MBio* 11:e01914-20.

Raivio TL, Popkin DL, Silhavy TJ. 1999. The Cpx envelope stress response is controlled by amplification and feedback inhibition. *J Bacteriol* 181:5263-5272.

Ramos-Morales F, Prieto AI, Beuzon CR, Holden DW, Casadesus J. 2003. Role for *Salmonella enterica* enterobacterial common antigen in bile resistance and virulence. *J Bacteriol* 185:5328-32.

Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melén K, De Gier JW, Von Heijne G. 2004. Experimentally based topology models for E. coli inner membrane proteins. *Protein Sci* 13:937-45.

Rick PD, Barr K, Sankaran K, Kajimura J, Rush JS, Waechter CJ. 2003. Evidence that the *wzxE* gene of *Escherichia coli* K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen. *J Biol Chem* 278:16534-42.

Rick PD, Wolski S, Barr K, Ward S, Ramsay-Sharer L. 1988. Accumulation of a lipid-linked intermediate involved in enterobacterial common antigen synthesis in *Salmonella typhimurium* mutants lacking dTDP-glucose pyrophosphorylase. *J Bacteriol* 170:4008-14.

Sham LT, Butler EK, Lebar MD, Kahne D, Bernhardt TG, Ruiz N. 2014. Bacterial cell wall MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science* 345:220-2.

Stevenson G, Neal B, Liu D, Hobbs M, Packer NH, Batley M, Redmond JW, Lindquist L, Reeves P. 1994. Structure of the O antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. *J Bacteriol* 176:4144-56.

Tatar LD, Marolda CL, Polischuk AN, van Leeuwen D, Valvano MA. 2007. An *Escherichia coli* undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling. *Microbiology* 153:2518-29.

Tran ENH, Doyle MT, Morona R. 2013. LPS Unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PLOS ONE* 8:e70508.

Typas A, Banzhaf M, Gross CA, Vollmer W. 2011. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* 10:123-36.

Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, Prestegard JJ, Schnaar RL, Freeze HH, Marth JD, Bertozzi CR, Etzler ME, Frank M, Vliegenthart JF, Lütteke T, Perez S, Bolton E, Rudd P, Paulson J, Kanehisa M, Toukach P, Aoki-Kinoshita KF, Dell A, Narimatsu H, York W, Taniguchi N, Kornfeld S. 2015. Symbol nomenclature for graphical representations of glycans. *Glycobiology* 25:1323-4.

Wall E, Majdalani N, Gottesman S. 2018. The complex Rcs regulatory cascade. *Annual Rev Microbiol* 72:111-139

Wang RF, Kushner SR. 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 100:195-9.

Whitfield C, Trent MS. 2014. Biosynthesis and export of bacterial lipopolysaccharides. *Annual Rev Biochem* 83:99-128.

Whitfield C, Wear SS, Sande C. 2020. Assembly of bacterial capsular polysaccharides and exopolysaccharides. *Ann Rev Microbiol* 74:521-543.

Workman SD, Strynadka NCJ. 2020. A slippery scaffold: Synthesis and recycling of the bacterial cell wall carrier lipid. *J Mol Biol* 432:4964-4982.

Yuasa R, Levinthal M, Nikaido H. 1969. Biosynthesis of cell wall lipopolysaccharide in mutants of *Salmonella*. V. A mutant of *Salmonella typhimurium* defective in the synthesis of cytidine diphosphoabequose. *J Bacteriol* 100:433-444.

4.9 Article Supporting Information

4.9.1 Supporting Figures

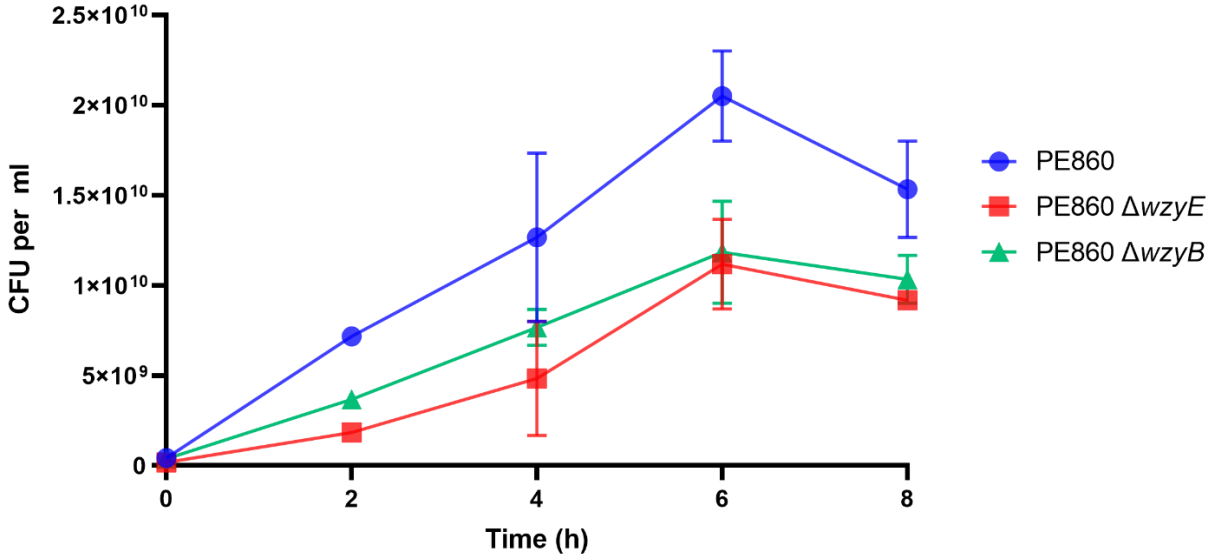


Figure 4.6-(S1): CFU counting of PE860 and *wzy* mutants show difference in cell amount at different time points.

Analysis of CFU of strains PE860, PE860 *ΔwzyE* and PE860 *ΔwzyB* from normalized subcultures at time points 0, 2, 4, 6 and 8 hours. Overnight culture of the above strains were normalized, sub-cultured and 20 µl of culture was collected and serially diluted 1:10 at the time points above. 10 µl of cell suspension from the range of 10^{-5} to 10^{-8} was spotted to LBA plates and colonies were counted the following day. Data represents averages of CFU from 2 independent biological replicates.

4.10 Article Acknowledgements

Funding for this work was provided by a Discovery Project Grant to R. Morona from the Australian Research Council (PROJECT ID: DP160103903). N. Maczuga is the recipient of a Research Training Program Stipend Research Scholarship from the University of Adelaide.

4.11 Article Conflicts of interest

The author(s) declare that there are no conflicts of interest.

Chapter Five

ARTICLE III:

Subcellular localization of WecG in *Shigella flexneri*.

Nicholas Tadeusz Maczuga, Elizabeth Tran and Renato Morona.

Chapter 5: Subcellular localization of WecG in *Shigella flexneri.*

5.1 Statement of Authorship

Title of Paper	Subcellular localization of WecG in *Shigella flexneri*.
Status	Under review

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Nicholas Tadeusz Maczuga		
Contribution	Construction of strains and plasmids, conducted all experiments pertaining to all Figures and constructed these figures, data analysis, conception of model, corresponding author and writing of manuscript.		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/01/2022

Author	Elizabeth Ngoc Hoa Tran		
Contribution	Supervised development of work and provided editing and evaluation of the manuscript.		
Signature		Date	20/1/22

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, manuscript evaluation and editing.		
Signature		Date	19/1/22

5.2 Article Abstract

Enterobacteriales have developed a specialized outer membrane polysaccharide (Enterobacterial Common Antigen (ECA)). ECA biosynthesis begins on the cytoplasmic side of the inner membrane (IM) where glycosyltransferases sequentially add sugar moieties to form a complete repeat unit which is then translocated across the IM by WzxE before being polymerized into short linear chains by WzyE/WzzE. Research into WecG, the enzyme responsible for generating ECA lipid-II, has not progressed beyond *Barr et al. (1988)* who described WecG as a membrane protein. Here we show that WecG is not a membrane protein but one which is peripherally associated with the IM. Through the use of Western immunoblotting we show that WecG is maintained to the IM via its three C-terminal helices and further identify key residues in helix II which are critical for this interaction and has allowed us to identify WecG as a GT-E glycosyltransferase. We investigate the possibility of protein complexes and ultimately show that ECA lipid-I maintains WecG to the membrane which is crucial for its function. This research is the first since *Barr et al. (1988)* to investigate the biochemistry of WecG and reveals possible novel drug targets to inhibit WecG and thus ECA function and cell viability.

5.3 Article Importance

ECA has been shown to play major roles in the homeostasis of cells as mutants along its biosynthetic pathway, particularly its glycosyltransferases, have been shown to induce pleiotropic phenotypes as well as induce cell wall stress pathways (Castelli & Vescovi 2011; Jorgenson et al. 2016). Due to this, research into the roles that the ECA glycosyltransferases play in these pleiotropic phenotypes has dominated ECA research with little research being performed in understanding the enzymes themselves. WecG research has not progressed since 1988 when it was first identified by *Barr et.al* who described WecG as a membrane protein. Here, we show that WecG is not a membrane protein, but is maintained to the IM via interactions facilitated by its C-terminal tail. We describe a new model of how WecG is maintained to the IM and present WecG as the second protein in the novel glycosyltransferase fold family, GT-E.

5.4 Article Introduction

Enterobacteriales, a family of bacteria belonging to Gram-negatives, have developed unique adaptations which allow them to persist and cause disease in their specific environmental niches; one such adaptations is the production of the outer membrane (OM) polysaccharide Enterobacterial Common Antigen (ECA). ECA is presented on the OM in two membrane associated forms, ECA_{pg} and ECA_{lps}, where ECA polysaccharide chains are directly linked to either phosphatidylglycerol (pg) or the core-sugars of LPS, respectively (Gozdziewicz, Lukasiewicz & Lugowski 2015; Maciejewska et al. 2020). The membrane associated forms of ECA has been shown to play crucial roles in maintaining OM homeostasis as well as providing resistance to bile salts (Castelli & Vescovi 2011; Ramos-Morales et al. 2003). ECA additionally presents in a periplasmically restricted form, ECA_{cyc}, which contains no lipid anchor and but also acts to maintain the permeability barrier of the OM (Kajimura, Rahman & Rick 2005; Mitchell, Srikumar & Silhavy 2018).

Universal in all Enterobacteriales, ECA comprises of a trisaccharide repeating units (RUs) containing N-acetylglucosamine (GlcNAc), N-acetyl-D-mannosaminuronic acid (ManNAcA) and 4-acetamido-4,6-dideoxy-D-galactose (Fuc4NAc) with the biosynthesis of ECA commencing on the cytoplasmic side of the inner membrane (IM) (Eade et al. 2021). Briefly, UDP-GlcNAc is transferred onto the polyprenyl phosphate lipid carrier undecaprenyl phosphate (Und-P) by the glycosyltransferase WecA, generating ECA lipid-I (Al-Dabbagh, Mengin-Lecreulx & Bouhss 2008). Sequential additions of ManNAcA and Fuc4Nac facilitated by WecG and WecF yield the complete ECA RU which is then translocated across the IM by WzxE (Eade et al. 2021). Lastly, ECA polysaccharide chains are assembled by WzyE and WzzE before the complete polysaccharide chain is ligated onto a final lipid anchor prior to export to the OM (Figure 5.1) (Islam & Lam 2014).

Research into the glycosyltransferases (GT) of the ECA biosynthetic pathway has been orientated towards understanding the effects of mutations in their coding genes in causing pleiotropic mutant phenotypes through the accumulation of biosynthetic intermediates and sequestering Und-P from related pathways, with little work orientated in furthering our understanding of the enzymes themselves (Jorgenson et al. 2016; Ramos-Morales et al. 2003). WecG is one such glycosyltransferase which was extensively investigated for its role in cell wall mutants however, research specifically into the enzyme has not progressed since *Barr et al. (1988)* at the time when the genes responsible for ECA biosynthesis were first being discovered and described (Barr et al. 1988; Danese et al. 1998; Jiang et al. 2020; Jorgenson et al. 2016).

WecG, a UDP-N-acetyl-D-mannosaminuronic acid transferase, facilitates the transfer of UDP-ManNAcA onto ECA lipid-I, generating ECA lipid-II and in the process of doing so, commits the biosynthetic intermediate to ECA biosynthesis (Eade et al. 2021). Initial research into understanding WecG placed it as a membrane protein, as whole membranes from a $wecG^{+}$ strain were able to *in vitro* polymerise ECA using ECA lipid-I supplied from a $wecG^{-}$ mutant, observed as a restoration of ECA banding by anti-ECA Western immunoblotting (Barr et al. 1988). In this study we show that WecG is not a membrane protein, but one which is strongly associated with the membrane through interactions facilitated with its *in silico* predicted C-terminal helices.

High peptide sequence homology with a novel GT-E fold protein TagA, a UDP-N-acetyl-D-mannosaminuronic acid transferase from Gram positives, suggested to us that WecG is a glycosyltransferase belonging to the novel GT fold family, GT-E (Kattke et al. 2019). Western immunoblotting of carboxy terminal (C-terminal) deletions of WecG showed that WecG was not a membrane protein but one which is strongly associated with the IM via interactions facilitated by its C-terminal helices. Multiple sequence alignments identified conserved residues throughout the C-terminal helices which we showed are crucial for maintaining WecG to the membrane and ECA production establish that WecG's functionality is dependent on its membrane association. Lastly we demonstrate that WecG's membrane association is dependent on the presence of ECA lipid-I, and present a hypothesis that WecG is primarily maintained to the membrane through interactions facilitated between crucial residues in its second C-terminal helix and ECA lipid-I.

This study provides a new understanding of the ECA biosynthetic pathway and the key enzyme which specifically initiates it, WecG. We clarify the long held model of ECA biosynthesis and show the extent of WecG's dependence on its membrane association and elucidate possible targets to block ECA production in Enterobacteriales.

5.5 Article Methods

5.5.1 Ethics statement

The ECA and WzzE antibodies were produced under the National Health and Medical Research Council Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and was approved by the University of Adelaide Animal Ethics Committee.

5.5.2 Bacterial strains, growth media and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 5.1. Bacteria were routinely grown at 37 ºC in Lysogeny-Bertani (LB) broth with aeration or on LB agar (LBA). Antibiotics used were 100 µg ampicillin (Amp) ml^{-1} and 10 ng tunicamycin ml^{-1} with 3 µg polymyxin B nonapeptide ml^{-1} (PBMN; Sigma). Strains carrying pWKS30 requiring induction were grown in LB at 37 ºC with aeration for 16 hours, sub-cultured (1/20) into fresh broth and induced with 1 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG). Cultures were grown for a further for 4 hours.

5.5.3 DNA methods

The plasmids used in this study are shown in Table 5.1. Plasmid were purified from *E. coli* DH5α strains using a QIAprep Spin miniprep kit (Qiagen). Preparation of electrocompetent cells and electroporation methods were performed as described previously (Purins et al. 2008).

5.5.4 Chromosomal mutagenesis

The *S. flexneri* Y PE860 Δ$wzyE_{10\text{-}440}$ strain was generated using λ Red mutagenesis as described in *Datsenko et al. (2000)* (Datsenko & Wanner 2000). Briefly, DNA primers were designed to PCR amplify a chloramphenicol resistance cassette flanked with 50 bp of homologous sequence to *wzyE* such that once deleted, 30 nucleotides of the 5' and 3' coding regions of *wzyE* would remain. The purified PCR fragment was then electroporated into the parent PE860 strain carrying pKD46 to generate the mutant strains. The antibiotic resistance cassette was removed by the introduction of pCP20.

5.5.5 Construct generation/DNA sequencing

Primers used for construct generation are listed in Table 5.2. Generation of pWKS30-His_{10}WecG, denoted as pHis_{10}WecG was performed by inverse PCR using primers (NM126/127) to amplify a fragment of DNA containing *wecG* with the coding sequence for a amino-terminal His_{10} epitope tag using pWKS30-WecG as a template. The linear fragment was 5' phosphorylated via poly nucleotide kinase (Genebank) before being circularized using T4 DNA ligase (Genebank). Generation of C-terminal deletions of WecG were performed as above using inverse PCR and

primers (WecG$^{\Delta III}$:NM130.1/NM132, WecG$^{\Delta II+III}$:NM130.1/NM133, WecG$^{\Delta A}$:NM130.1/NM131.1) to amplify DNA fragments containing *his$_{10}$wecG* deletions using pWKS30-His$_{10}$WecG as a template. The linear fragments were then treated as above. Generation of WecG helix II point mutations were performed via site-directed mutagenesis by inverse PCR and primers (WecG^{L215E}:NM140/NM141, WecG^{L218E}:NM142/NM143, WecG^{L222E}:NM146/NM147) were used to PCR amplify DNA fragments using pWKS30-His$_{10}$WecG as a template. The linear fragments were treated as above. DNA sequencing was used to confirm all constructs and that post ligation, the *wecG* coding sequence remained in-frame.

5.5.6 Whole cell protein sample preparation

Bacteria were grown and induced as described above before 5x10^8 cells were collected by centrifugation (2000 x g) and resuspended in 100 µl of 2x sample buffer (Lugtenberg et al. 1975). Samples were heated at 56 ºC for 10 minutes prior to loading.

5.5.7 ECA sample preparation

Bacteria were grown and induced as described above before 1x10^9 cells were collected by centrifugation (2000 x g), resuspended in 2x lysis buffer (Murray, Attridge & Morona 2003) and heated at 100 ºC for 10 minutes before incubation with 2.5 mg/ml proteinase K (Sigma-Aldrich) for 2 hours at 56 ºC. Samples were heated at 100 ºC for 1 minute prior to loading.

5.5.8 Membrane fractionation

Overnight cultures were subcultured into 200 ml LB and induced as above. Cultures were then pelleted via centrifugation (Beckman Coulter Avanti J-26 XPI centrifuge; 9,600 x *g*, 10 minutes, 4 ºC) before resuspension in 10 ml of sonication buffer (100 mM $NaCO_3$ pH 7.0) and disrupted by sonication (Branson B15). Cellular debris was then collected and removed via centrifugation (Thermo Scientific Labofuge 400 R centrifuge; 3,500 x *g*, 10 minutes, 4 ºC) prior to whole membrane (WM) collection by ultracentrifugation (Beckman Coulter Optima L-100 XP ultracentrifuge; 250,000 x *g*, 45 minutes, 4° C). WM were then resuspended in 5 ml of MQ at 4 ºC.

5.5.9 Dissociation assays

500 µl of WMs were aliquoted into 1.5 ml reaction tubes to which 500 µl of various solutions were added (3 M NaI 200 mM $NaCO_3$ pH 7.0; 4 M NaCl 200 mM $NaCO_3$ pH 7.0; 200 mM $NaCO_3$ pH 12.0; 200 mM $NaCO_3$ pH 7.0 or PBS). The reaction tubes were then incubated at room temperature with for 1 hour with agitation before the insoluble faction membrane was collected via ultracentrifugation as above.

5.5.10 *In vivo* DSP Crosslinking

Overnight bacterial cultures were sub-cultured (1/20) into 200 ml fresh LB broth, induced as above and grown for a further 4 hours. Cells were then collected via centrifugation as above, washed and resuspended in 5 ml DSP-Crosslinking buffer (150 mM NaCl, 20 mM Na_2PO_4/NaH_2PO_4, pH 7.2) prior to incubation with 1 mM DSP (Thermo fisher Scientific) for 30 minutes at 37 ºC (+ DSP samples). A duplicate sample was also incubated for 30 minutes at 37 ºC without treatment (- DSP samples). Excess DSP was then quenched using 20 mM Tris-HCl, pH 7.5 for 10 minutes at room temperature prior to collection and resuspension in sonication buffer as above.

5.5.11 Western immunoblotting

Protein/ECA samples were loaded onto a 12% or 15% SDS-PAGE gel respectively and electrophoresed at 200 V for 1 hour. SDS-PAGE gels were then transferred onto a nitrocellulose membrane (Bio-Rad) at 400 mA for 1 hour prior to membranes being blocked with 5% (wt/vol) skim milk in Tris-Tween Buffer saline (TTBS). Membranes were then incubated overnight with either monoclonal mouse anti-His antibodies (GenScript) (1:50,000), polyclonal rabbit anti-WzzE (1:500) or polyclonal rabbit anti-ECA antibodies (1:500), diluted in 2.5% (wt/vol) skim milk in TTBS. Detection was performed with rabbit anti-mouse or goat anti-rabbit horseradish peroxidase-conjugated antibodies (KLP) and chemiluminesence reagent (Sigma). 5 µl of SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard.

5.5.12 Bioinformatic analysis

The peptide sequence of WecG was obtained from NCBI (GenBank: QEO94344.1). The WecG peptide sequence was then submitted to: TMHMM to *in silico* predict transmembrane helices, Jpred to *in silico* predict secondary structures and I-TASSER, Alphafold and RaptorX servers to *in silico* predict the tertiary structure of WecG (Drozdetskiy et al. 2015; Jumper et al. 2021; Källberg et al. 2012; Krogh et al. 2001; Yang et al. 2015). Multiple sequence alignments were performed in Jalview (Amar et al. 2018).

Table 5.1: Bacterial strains and plasmids used in this study

Strain or plasmid	Description	Source
Strains:		
RMA2162	*S. flexneri* PE860 Y serotype; strain lacks virulence plasmid and pHS-2 plasmid	Laboratory stock
NMRM120	*S. flexneri* PE860 Δ*wzyE*$_{10\text{-}440}$	This study
NMRM348	*S. flexneri* PE860 Δ*wecC*	*Chapter 4*
RMA4622	*E. coli* K-12 BW25113	Laboratory stock
JW3770-1	*E. coli* K-12 BW25113 Δ*wecG::Km*	(Baba et al. 2006)
Plasmids:		
pKD3	Source of Chloramphenicol resistance cassette for λ Red mutagenesis, Cml^r	(Datsenko & Wanner 2000)
pKD46	Source of λ phage recombinase for homology recombination, 30 ^oC growth for maintenance, 42 ^oC for expression and cure, Amp^r	(Datsenko & Wanner 2000)
pCP20	Source of yeast FRT specific recombinase for λ red mutagenesis, 30 ^oC growth for maintenance, 42 ^oC for expression and cure, Amp^r	(Datsenko & Wanner 2000)
pWKS30	IPTG inducible, expression vector, Amp^r	(Wang, RF & Kushner 1991)
pWKS30-WecG	pWKS30 encoding WecG, Amp^r	This study
pHis$_{10}$WecG	pWKS30 encoding His$_{10}$WecG, Amp^r	This study
pHis$_{10}$WecG$^{\Delta A}$	pWKS30 encoding His$_{10}$WecG$^{\Delta A}$, Amp^r	This study
pHis$_{10}$WecG$^{\Delta II+III}$	pWKS30 encoding His$_{10}$WecG $^{\Delta II+III}$, Amp^r	This study
pHis$_{10}$WecG$^{\Delta III}$	pWKS30 encoding His$_{10}$WecG $^{\Delta III}$, Amp^r	This study
pHis$_{10}$WecG^{L215E}	pWKS30 encoding His$_{10}$WecG^{L215E}, Amp^r	This study
pHis$_{10}$WecG^{L218E}	pWKS30 encoding His$_{10}$WecG^{L218E}, Amp^r	This study
pHis$_{10}$WecG^{L222E}	pWKS30 encoding His$_{10}$WecG^{L222E}, Amp^r	This study

Table 5.2: DNA oligonucleotides used in this study

Primer	Oligonucleotide sequence (5'-3')	Target
Construct generation specific primers		
NM126	gtgatgatggtggtgatgatggtgatgatgcattgttatccttcaacctgcgtccgg	pWKS30-WecG
NM127	aataacaacaacacggcaccgacc	pWKS30-WecG
NM130.1	tgactgcagcccggggg	pWKS30-His_{10}WecG
NM131.1	accggtgaaaacatcgtaagtcccg	pWKS30-His_{10}WecG
NM132	aatgcggctcggctgcg	pWKS30-His_{10}WecG
NM133	ttgccagattttcggtgcgcg	pWKS30-His_{10}WecG
NM140	tcccagcgtttgccagattttcg	pWKS30-His_{10}WecG
NM141	gaggagtggctctaccgcctgct	pWKS30-His_{10}WecG
NM142	ccactccagtcccagcgtttg	pWKS30-His_{10}WecG
NM143	gagtaccgcctgctttcgcagc	pWKS30-His_{10}WecG
NM146	caggcggtagagccactcca	pWKS30-His_{10}WecG
NM147	gagtcgcagccgagccgcattaa	pWKS30-His_{10}WecG
λ Red mutagenesis specific primers		
NM58	gattgccgccggggaggtcgcatgagtctgctgcaattcagtggcctgtttgtgtaggctggagctgcttc	*wzyE* gene
NM59	cgtggtggtggtattcattgttatccttcaacctgcgtccggagcgatgaatgggaattagccatggtcc	*wzyE* gene

5.6 Article Results

5.6.1 Bioinformatic analysis of WecG

Initial research on the localization of WecG performed by *Barr et al. (1988)* (Barr et al. 1988) stated that WecG was a membrane protein. Since that initial study, there has been little research on the structure and subcellular localization of WecG, hence we first applied a bioinformatic approach to investigate the biophysical characteristics of WecG and its structure. The TMHMM (Krogh et al. 2001), Jpred (Drozdetskiy et al. 2015), I-TASSER (Yang et al. 2015), Alphafold (Jumper et al. 2021) and RaptorX (Källberg et al. 2012) servers were used to investigate the probability of transmembrane helices, the general secondary structure of WecG as well as to generate *in silico* predicted structures of WecG. Firstly TMHMM predicted that WecG consisted of no transmembrane helices and Jpred predicted that WecG consists of 11 alpha helices and 8 beta sheets but none long enough to traverse a membrane as well as a possible single Rossmann fold with its characteristic alternating α-helices and β-sheets (Figure 5.8-(S1)) (Bottoms, Smith & Tanner 2002; Saidijam, Azizpour & Patching 2018).

I-TASSER, Alphafold and RaptorX gave a predicted protein structure of a globular protein consisting of a single Rossmann fold but no parallel alpha helices which is characteristic of transmembrane segments (Figure 5.2a, b & c). Additionally, once structurally compared it was observed that I-TASSER and RaptorX predicted a disordered C terminal domain (CTD) whereas Alphafold predicted three C terminal helices (Figure 5.2d) which was consistent with the prediction provided by Jpred (Figure 5.8-(S1)). Interestingly, I-TASSER and RaptorX both independently produced a model using the template protein TagA, with a TM score of 0.62+/-0.14 and p-value of 8.03e-30 respectively, which indicated a high structural similarity to TagA. TagA, a UDP-N-acetyl-D-mannosaminuronic acid transferase present in Gram positive *Thermoanaerobacter italicus* and *Staphylococcus aureus* is a member of the Pfam glycosyltransferase WecG/TagA/CpsF family and is involved in the biosynthesis of wall teichoic acids (WTAs) (Swoboda et al. 2010).

5.6.2 WecG associates with the membrane

To investigate the localization of WecG, a poly-His epitope tagged WecG expression construct was made. Initial attempts to generate a poly-his tagged WecG proved difficult as a $WecG_{His12}$ C-terminal His_{12} tag construct could complement the *E. coli* BW25113 *ΔwecG::km* strain but could not be detected (Figure 5.9-(S2)). An alternative N-terminal domain His_{10} tagged WecG construct, $pHis_{10}WecG$, was made and transformed into *E. coli*

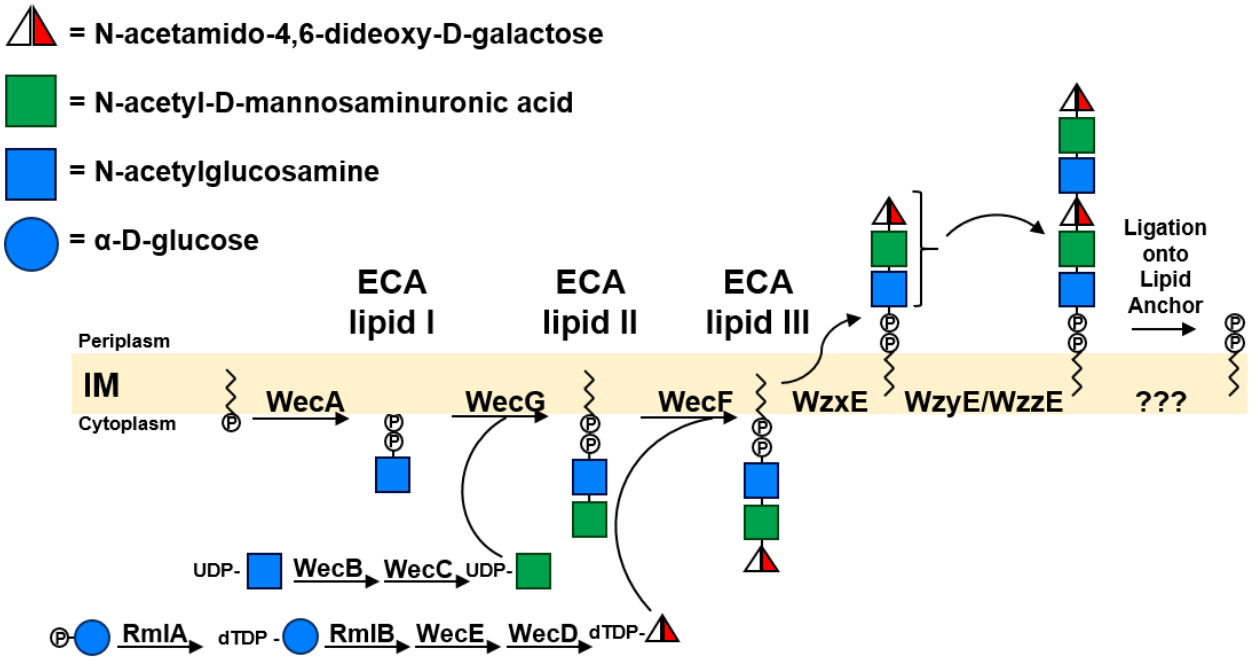


Figure 5.1: Enterobacterial common antigen biosynthesis in Enterobacteriales.

ECA biosynthesis begins on the cytoplasmic side of the inner membrane (IM) where undecaprenyl phosphate (Und-P) is acted upon by WecA to yield Und-PP-GlcNAcA ECA lipid-I (Al-Dabbagh, Mengin-Lecreulx & Bouhss 2008). The sequential addition of ManNAcA and Fuc4Nac by the glycosyltransferases WecG and WecF yield ECA lipid-II and ECA lipid-III respectively at which point the complete ECA repeat unit (RU) which is subsequently flipped onto the periplasmic leaflet of the IM by the flippase WzxE (Eade et al. 2021). The polymerase WzyE and co-polymerase WzzE then polymerize a linear ECA polysaccharide of controlled length followed by ligation of the ECA polysaccharide onto its final lipid carrier prior to export to the outer membrane (Woodward et al. 2010).

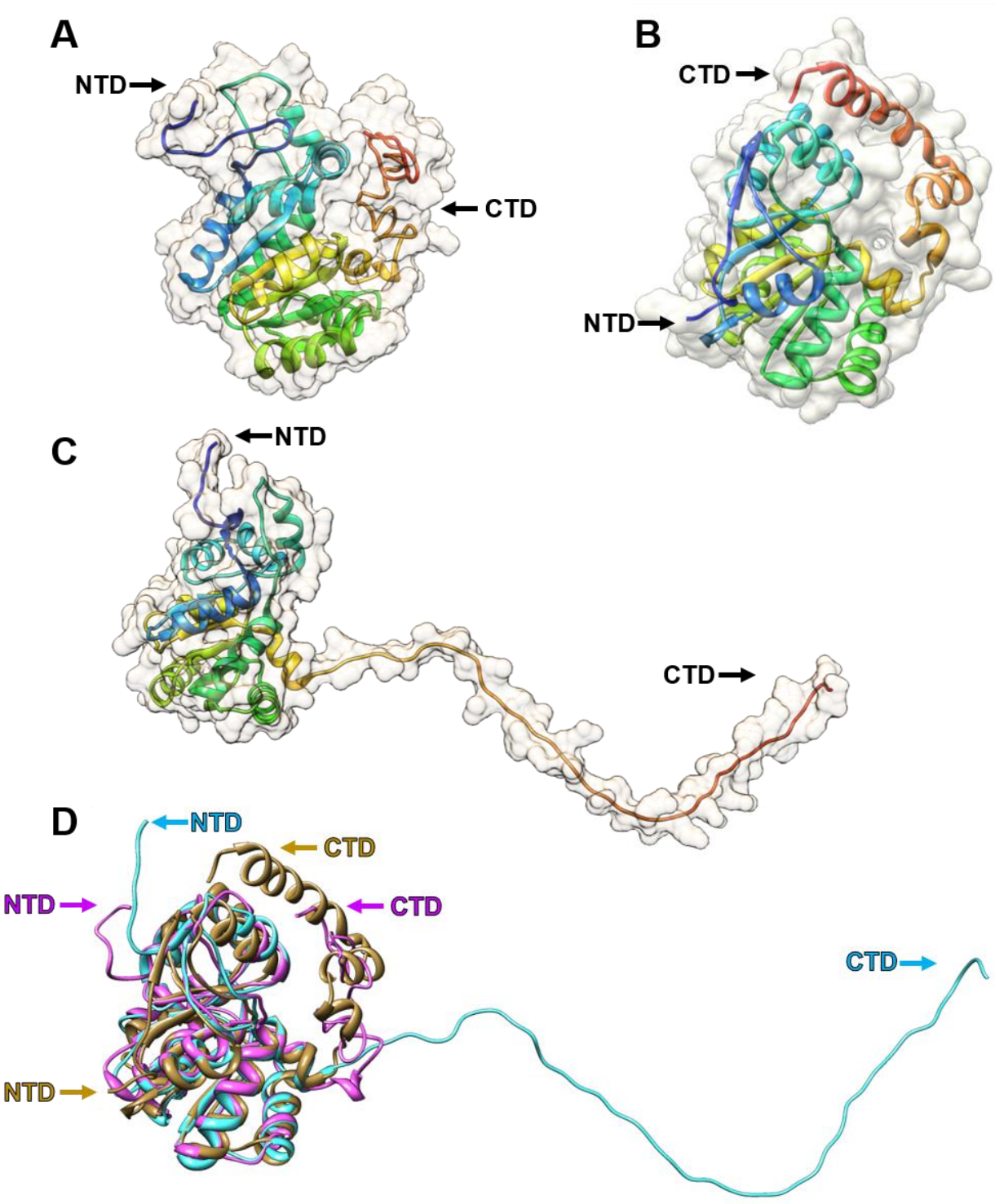


Figure 5.2: *In silico* predicted tertiary structures of WecG.

a) I-TASSER predicted structure of WecG suggests that WecG is a globular protein with an unorganised CTD. B) Alphafold predicted structure of WecG shows that WecG is a globular protein with three CTD helicies. C) RaptorX predicted structure of WecG indicates that the structure of WecG consists of a globular NTD and a protruding CTD. D) Structural overlay comparison of WecG predicted models from I-TASSER in pink, Alphafold in gold and RaptorX in blue. NTD = amino terminal domain, CTD = carboxyl terminal domain.

BW25113 *ΔwecG::km* where it showed complementation of ECA and the protein was detected in whole cells (WC) via anti-His Western immunoblotting (Figure 5.9-(S2)).

To determine whether or not WecG was a membrane protein, membrane fractionation was performed on induced PE860 carrying pHis$_{10}$WecG (Figure 5.3a). A band of ~30 kDa which correlated to the expected mass of His$_{10}$WecG was detected in both the WC **(**Figure 5.3a, lane 1) and the whole membrane (WM) samples (Figure 5.3a, lane 3), which indicated that His$_{10}$WecG associated with the membrane.

Peripherally associated membrane proteins can be dissociated from the membrane by treatment with chaotropic agents, high salt, and high pH solutions (Smith 2017). Therefore, to further investigate WecG's localization, WM samples were treated with a range of chemicals and conditions including 1.5 M NaI, 1 M NaCl and pH 12 in an attempt to dissociate WecG from the WM. As seen in Figure 5.3b, His$_{10}$WecG consistently associated with the WM as the treatments used did not dissociate His$_{10}$WecG from the WM samples.

5.6.3 Peripheral association of WecG with the membrane is facilitated via CTD helices.

A plausible explanation as to how WecG is associated with the membrane, whilst displaying no typical membrane protein predicted secondary folding and structure, could be that WecG is maintained at the membrane through interactions facilitated between its three predicted C-terminal alpha helices in a similar way to TagA (Kattke et al. 2019). To investigate this, three C-terminal domain (CTD) deletions were generated in pHis$_{10}$WecG; pHis$_{10}$WecG$^{\Delta A}$ which consisted of WecG with no CTD alpha helices, pHis$_{10}$WecG$^{\Delta II+III}$ which consisted of WecG and its first CTD alpha helix, and pHis$_{10}$WecG$^{\Delta III}$ which consisted of WecG and its first and second CTD alpha helix (Figure 5.4a). The constructs were then tested to determine what effect the CTD deletions had on WecG's membrane association by membrane fractionation. As seen in Figure 5.4c, deletion of WecG CTD helix I and II caused WecG to dissociate from the membrane as His$_{10}$WecG bands were only present in the supernatant (SN) samples (Figure 5.4c, lanes 4 & 6), however deletion of WecG's third CTD helix did not affect WecG's association with the membrane fraction as a His$_{10}$WecG band was present in the WM sample (Figure 5.4c, lane 7). This supported the hypothesis that WecG's CTD play an important role in maintaining WecG membrane association and revealed that helix II plays an essential role as its deletion appeared to cause WecG to dissociate from the membrane.

We further investigated what impact the CTD deletions had on WecG's function in ECA biosynthesis by complementing the *E. coli* BW25113 *ΔwecG::km* mutant with pHis$_{10}$WecG$^{\Delta A}$, pHis$_{10}$WecG$^{\Delta II+III}$, and pHis$_{10}$WecG$^{\Delta III}$ and determining ECA production via Western

immunoblotting. The three WecG CTD deletions mutants were unable to complement the BW25113 *ΔwecG::km* mutant (Figure 5.4d, lanes 5-7) leading us to hypothesise that key catalytic domains are at least partially formed by the three CTD helices of WecG. The deletion not only disrupts this catalytic domain but also causes WecG to dissociate from the membrane.

5.6.4 Key hydrophobic residues in WecG CTD helix II facilitate membrane association.

During the bioinformatic analysis of WecG, it was observed that a series of conserved hydrophobic residues lie along the same face of WecG's CTD helix II, taking 3.6 amino acid residues per turn (Figure 5.5a). This led us to hypothesise that the association of WecG to the membrane could be facilitated by these residues. Amino acid substitutions were generated in $pHis_{10}WecG$ to determine the importance of the residues L215, L218 and L222, as described in the methods.

The resulting constructs ($pHis_{10}WecG^{L215E}$, $pHis_{10}WecG^{L218E}$ and $pHis_{10}WecG^{L222E}$) were transformed into *S. flexneri* PE860 and *E. coli* BW25113 *ΔwecG::km* to assess the effect on WecG's subcellular localization as well as any effect on ECA biosynthesis (Figure 5.1). As seen in Figure 5.5b & c, substitutions of L215E did affect the subcellular localization of $His_{10}WecG^{L215E}$ as a protein band was present in both the WM and SN samples (Figure 5.5b, lane 2 & 3, &, c, lanes 8 & 9). However the substitution of L218E caused the complete dissociation of $His_{10}WecG^{L218E}$ from the membrane as no protein bands were present in the membrane fraction (MF) sample (Figure 5.5b, lane 5, &, c, lane 11) but were present in the S/N sample (Figure 5.5b, lane 6, &, c, lane 12). The substitution of L222E resulted in $His_{10}WecG^{L222E}$ partially dissociating from the membrane as protein bands were detected in both the WM and S/N samples (Figure 5.5b, lanes 8 & 9, &, c, lanes 14 and 15).

To assess what effect these mutational alterations had on ECA biosynthesis, an anti-ECA Western immunoblot was performed (Figure 5.5d). Mutant $His_{10}WecG^{L215E}$ had minimal effect on ECA banding (Figure 5.5d, lane 5) as the ECA banding profile showed slightly less intensity when compared to the $His_{10}WecG$ complemented strain (Figure 5.5d, lane 4). The mutants $His_{10}WecG^{L218E}$ and $His_{10}WecG^{L222E}$ however had a drastic effect on ECA biosynthesis as the $His_{10}WecG^{L218E}$ mutant was unable to complement the *wecG* mutant, seen as the lack of ECA banding (Figure 5.5d, lane 6) and the $His_{10}WecG^{L222E}$ mutant was only partially able to complement the *wecG* mutant (Figure 5.5d, lane 7), detected as low intensity reduced size ECA banding. This profile data supported the hypothesis that residues L218 and L222 were crucial for WecG membrane association which in turn appeared to be crucial for ECA biosynthesis.

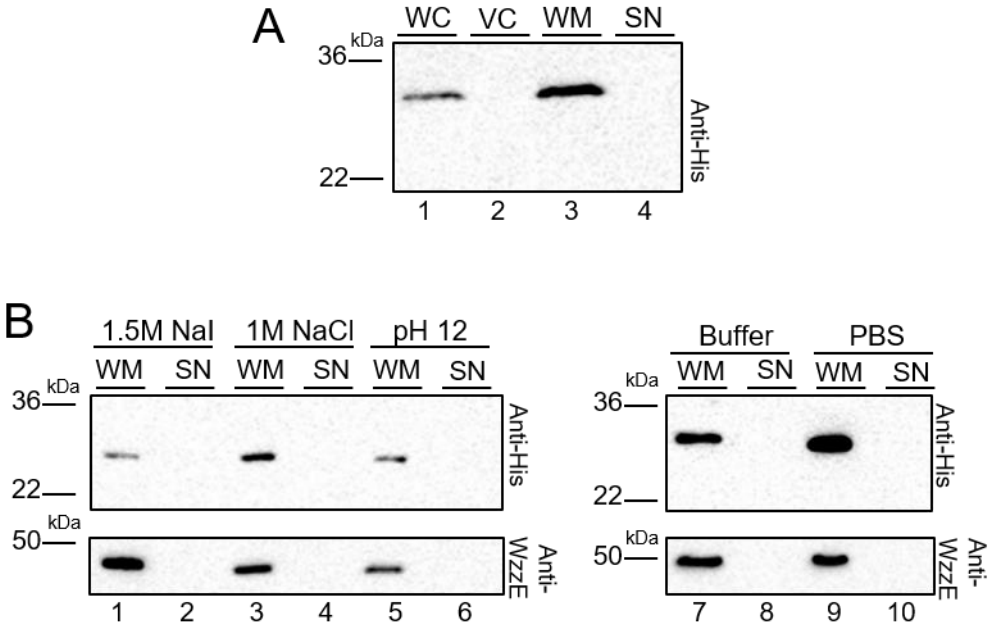


Figure 5.3: Subcellular localization of His_{10}WecG in *Shigella flexneri* PE860.

a) Anti-His Western immunoblot of His_{10}WecG subcellular localization expressed from PE860. Mid-exponential phase cells were disrupted by sonication and WM were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. b) Anti-His Western immunoblot of chemical treated WMs from PE860 cells expressing $pHis_{10}$WecG. WM were collected as above and were incubated 1:1 with 3 M NaI, 2 M NaCl or pH 12 buffer before membranes were collected by ultracentrifugation. Samples were electrophoresed as above prior to immunoblotting with anti-His or anti-WzzE antibodies. WC=whole cell, VC=vector control, WM=whole membrane, SN=supernatant. SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard.

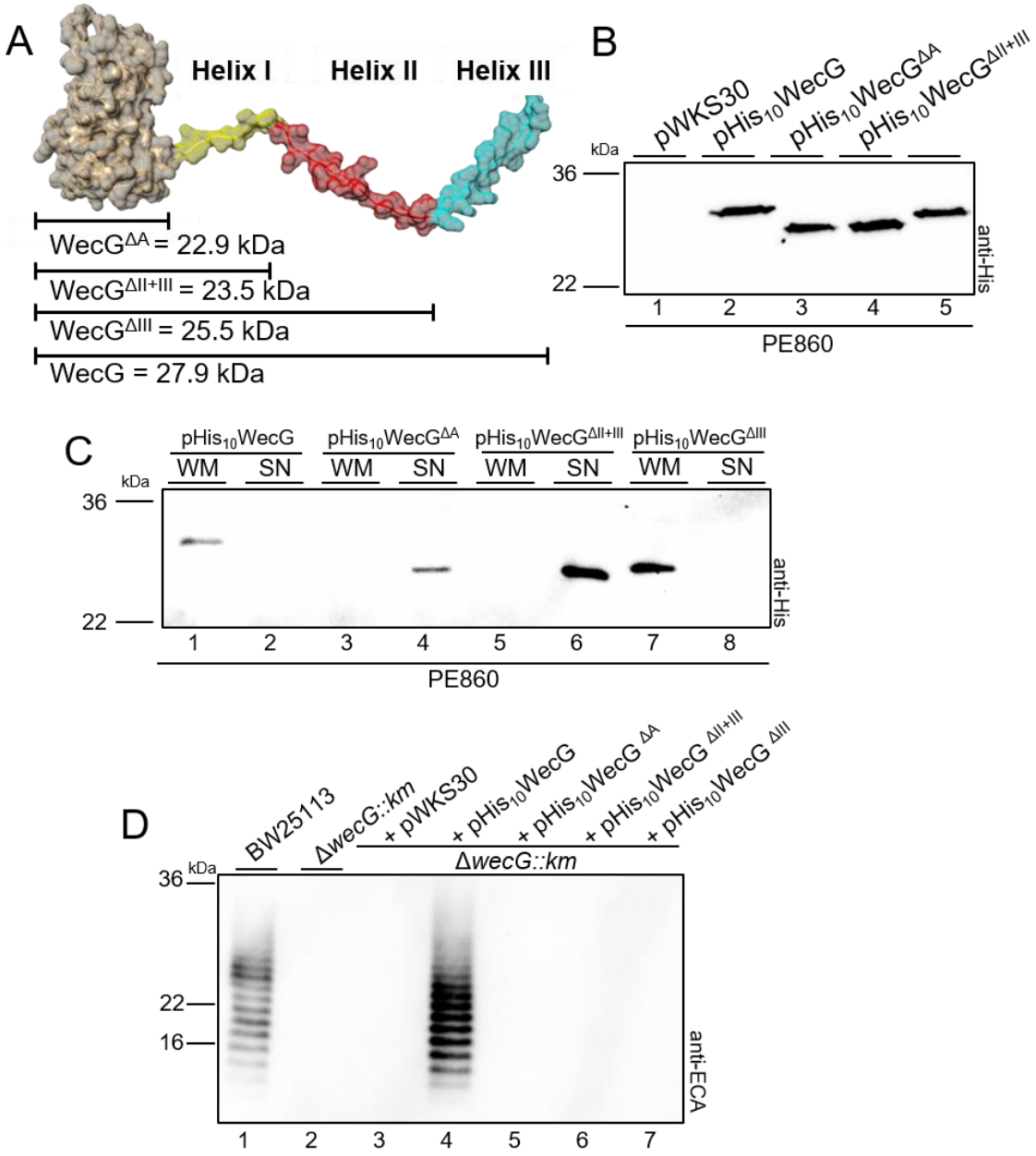


Figure 5.4: WecG CTD helices are crucial in maintaining WecG's peripheral association with the membrane and ECA production.

a) Model of WecG based on TagA derived by RaptorX showing the location and structure of WecG CTD deletion mutations. B) Anti-His Western immunoblot of whole cell samples from PE860 strains harbouring pHis$_{10}$WecG, pHis$_{10}$WecG$^{\Delta A}$, pHis$_{10}$WecG$^{\Delta II+III}$ and pHis$_{10}$WecG$^{\Delta III}$. Mid-exponential phase cells were collected (5x10^8 cells) and lysed in lysis buffer. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel, transferred onto a nitrocellulose membrane and probed with anti-His antibodies. c) Anti-His Western immunoblot of the subcellular location of the WecG CTD mutants. PE860 cells harbouring pWKS30, pHis$_{10}$WecG, pHis$_{10}$WecG$^{\Delta A}$, pHis$_{10}$WecG$^{\Delta II+III}$ and pHis$_{10}$WecG$^{\Delta III}$ were disrupted and membranes were collected by ultracentrifugation. WM samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. WM=Whole membrane, SN=Supernatant. D) Anti-ECA Western immunoblot showing ECA production from *E. coli K-12* BW25113 *ΔwecG::km* strains harbouring pWKS30, pHis$_{10}$WecG, pHis$_{10}$WecG$^{\Delta A}$, pHis$_{10}$WecG$^{\Delta II+III}$ and pHis$_{10}$WecG$^{\Delta III}$. Mid-exponential phase cells were collected (1x10^9 cells) by centrifugation and lysed in lysis buffer in the presence of proteinase K. Samples were electrophoresed on a SDS-15% (w/v) PAGE gel followed by immunoblotting with anti-ECA antibodies. SeeBlue Plus2 Pre-stained molecular weight ladder (Invitrogen) was used as a molecular weight marker for (b) to (d).

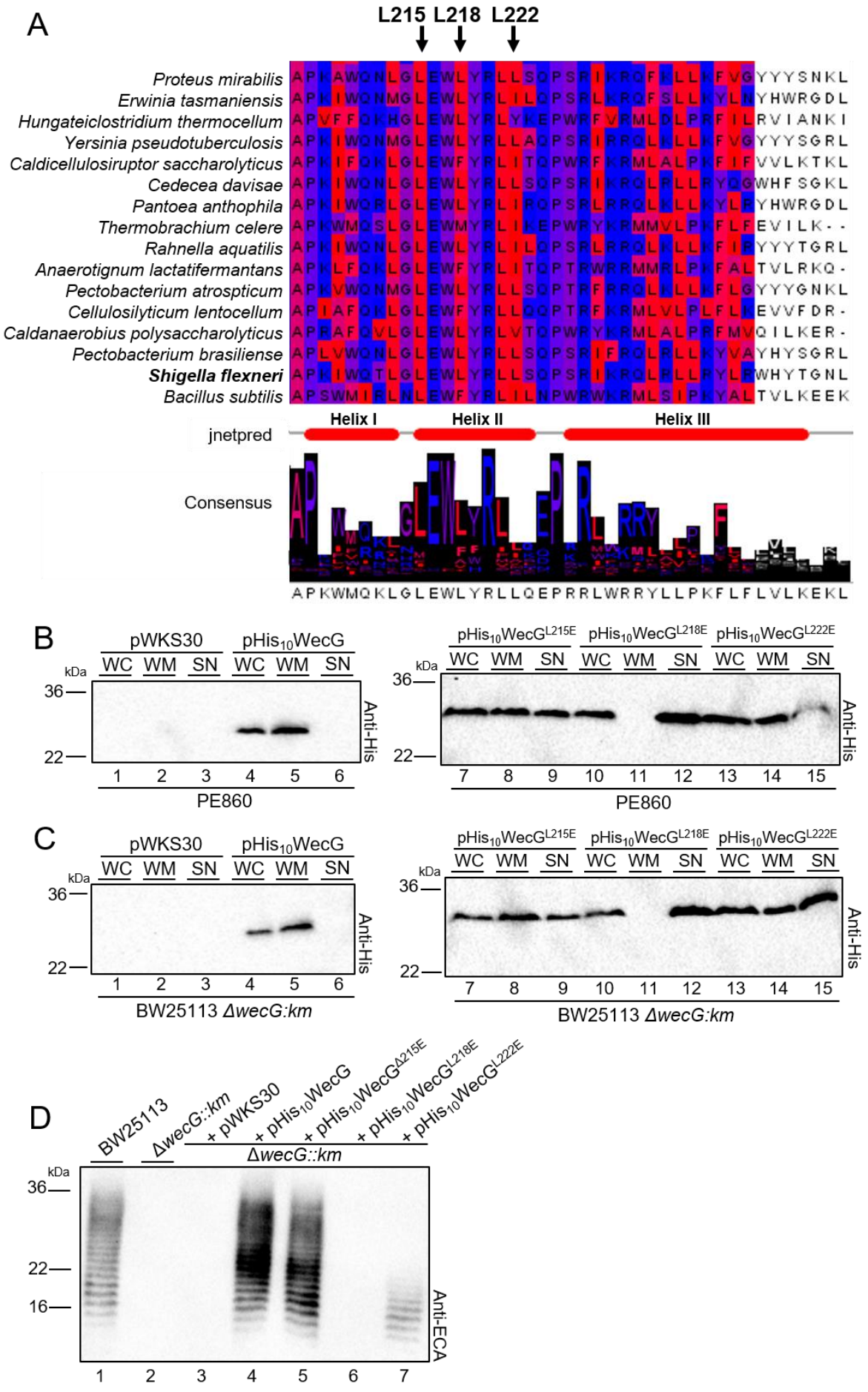

A
L215 L218 L222
Proteus mirabilis
Erwinia tasmaniensis
Hungateiclostridium thermocellum
Yersinia pseudotuberculosis
Caldicellulosiruptor saccharolyticus
Cedecea davisae
Pantoea anthophila
Thermobrachium celere
Rahnella aquatilis
Anaerotignum lactatifermantans
Pectobacterium atrospticum
Cellulosilyticum lentocellum
Caldanaerobius polysaccharolyticus
Pectobacterium brasiliense
Shigella flexneri
Bacillus subtilis
jnetpred
Helix I
Helix II
Helix III
Consensus
APKWMQKLGLEWLYRLLQEPRRLWRRYLLPKFLFLVLKEKL
B
pWKS30
pHis10WecG
pHis10WecGL215E
pHis10WecGL218E
pHis10WecGL222E
WC WM SN
kDa
36
22
Anti-His
PE860
C
BW25113 ΔwecG:km
D
BW25113
ΔwecG::km
+ pWKS30
+ pHis10WecG
+ pHis10WecGΔ215E
+ pHis10WecGL218E
+ pHis10WecGL222E
16
Anti-ECA

Figure 5.5: Key residues within WecG's CTD helix II play crucial roles in WecG's peripheral association with the membrane and ECA production.

a) Local multiple sequence alignment from Jalview (Waterhouse et al. 2009) of WecG with other UDP-N-acetyl-D-mannosaminuronic acid transferases. The three CTD helices show a high level of sequence identity with multiple residues remaining entirely conserved. Colouration based on residue hydrophobicity where red indicates hydrophobic and blue indicates hydrophilic b) Anti-His Western immunoblot of the subcellular localization of His_{10}WecG substitution mutants from PE860 cells harbouring pWKS30, $pHis_{10}WecG^{L215E}$, $pHis_{10}WecG^{L218E}$ and $pHis_{10}WecG^{L222E}$. Mid-exponential phase cells were disrupted and membranes were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibody. C) Anti-His Western immunoblot of the subcellular localization of His_{10}WecG substitution mutants from BW25113 *ΔwecG::km* cells harbouring pWKS30, $pHis_{10}WecG^{L215E}$, $pHis_{10}WecG^{L218E}$ and $pHis_{10}WecG^{L222E}$. Mid-exponential phase cells were disrupted and membranes were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibody. D) Anti-ECA Western immunoblot of His_{10}WecG substitution mutant's effect on ECA biosynthesis in BW25113 *ΔwecG::km* strains harbouring pWKS30, $pHis_{10}WecG^{L215E}$, $pHis_{10}WecG^{L218E}$ and $pHis_{10}WecG^{L222E}$. Mid-exponential phase cells were collected ($1x10^9$ cells) by centrifugation and lysed in lysis buffer in the presence of proteinase K. Samples were electrophoresed on a SDS-15% (w/v) PAGE gel followed by immunoblotting with anti-ECA antibodies. SeeBlue Plus2 Pre-stained molecular weight ladder (Invitrogen) was used as a molecular weight marker for (b) to (d). WC=Whole cell, WM=Whole membrane, SN=Supernatant.

5.6.5 WzyE polymerase impacts WecG's membrane association.

WecG is translationally coupled to WzyE as *wecG*'s ribosomal binding site (RBS) lies within the 3' end of *wzyE*'s coding region (Figure 5.6a). It has been proposed in the past that the proteins involved in the Wzy-dependent pathway form a protein complex (Islam & Lam 2014; Marolda et al. 2006), hence the impact of WzyE on WecG's membrane association was investigated. The pHis$_{10}$WecG plasmid was transformed into PE860 and PE860 *ΔwzyE$_{10-440}$* following which cell fractionations were performed (Figure 5.6b). His$_{10}$WecG was detected in the WM samples of PE860 *ΔwzyE$_{10-440}$* (Figure 5.6b, lane 3) indicating that in the absence of WzyE, WecG still associates with the membrane. To further investigate the impact of WzyE, WMs from PE860 *wzyE$_{10-440}$* mutant expressing pHis$_{10}$WecG were chemically treated post fractionation as above to attempt to dissociate His$_{10}$WecG from the membrane. Unexpectedly, unlike in PE860 (Figure 5.3b, lane 2), the presence of 1.5 M NaI induced the dissociation of His$_{10}$WecG from the WM of the PE860 *wzyE$_{10-440}$* mutant strain (Figure 5.6c, lane 3).

To investigate the possibility that WzyE physically interacts with WecG to maintain WecG's peripheral membrane association, DSP crosslinking was performed on whole cells. As seen in (Figure 5.6d), crosslinking with DSP showed no band shifts for His$_{10}$WecG in both the PE860 and *wzyE$_{10-440}$* mutant strains indicating that WecG does not detectably interact with WzyE or any other protein.

5.6.6 Lipid interactions are crucial for WecG's association with the membrane.

As no physical protein-protein interaction was detected via DSP crosslinking as described above, it was hence hypothesised that WecG may be primarily maintained to the membrane through lipid-mediated interactions between itself and ECA lipid-I and/or lipid-II (Figure 5.1). To investigate if ECA biosynthetic intermediates instead affected WecG's membrane association, treatment with tunicamycin was used to block the production of ECA lipid-I by inhibiting WecA (Heifetz, Keenan & Elbein 1979). Wildtype PE860 cells expressing pHis$_{10}$WecG were treated with 10 ng/ml^{-1} tunicamycin, fractionated and WMs chemically treated as above. Unlike the attempts at dissociation of His$_{10}$WecG by chemical treatments (Figure 5.3b), treatment with tunicamycin caused the dissociation of His$_{10}$WecG from the MF in all treatment samples (Figure 5.7a, lanes 1-9). This strongly supported the hypothesis that WecG is indeed peripherally maintained to the membrane through strong, lipid mediated interactions. To determine whether this peripheral interaction is facilitated by ECA lipid-I or ECA lipid-II, pHis$_{10}$WecG was transformed into PE860 *ΔwecC* as *wecC* mutants are known to interrupt the ECA biosynthetic pathway and prevent the biosynthesis of ECA lipid-II (Figure 5.1) (Jiang et al. 2018). As seen in Figure 5.7b, none of the

treatments were able to dissociate His_{10}WecG from the whole membrane of the *wecC* mutant, leading us to hypothesise that WecG is maintained to the membrane through lipid interactions involving ECA lipid-I.

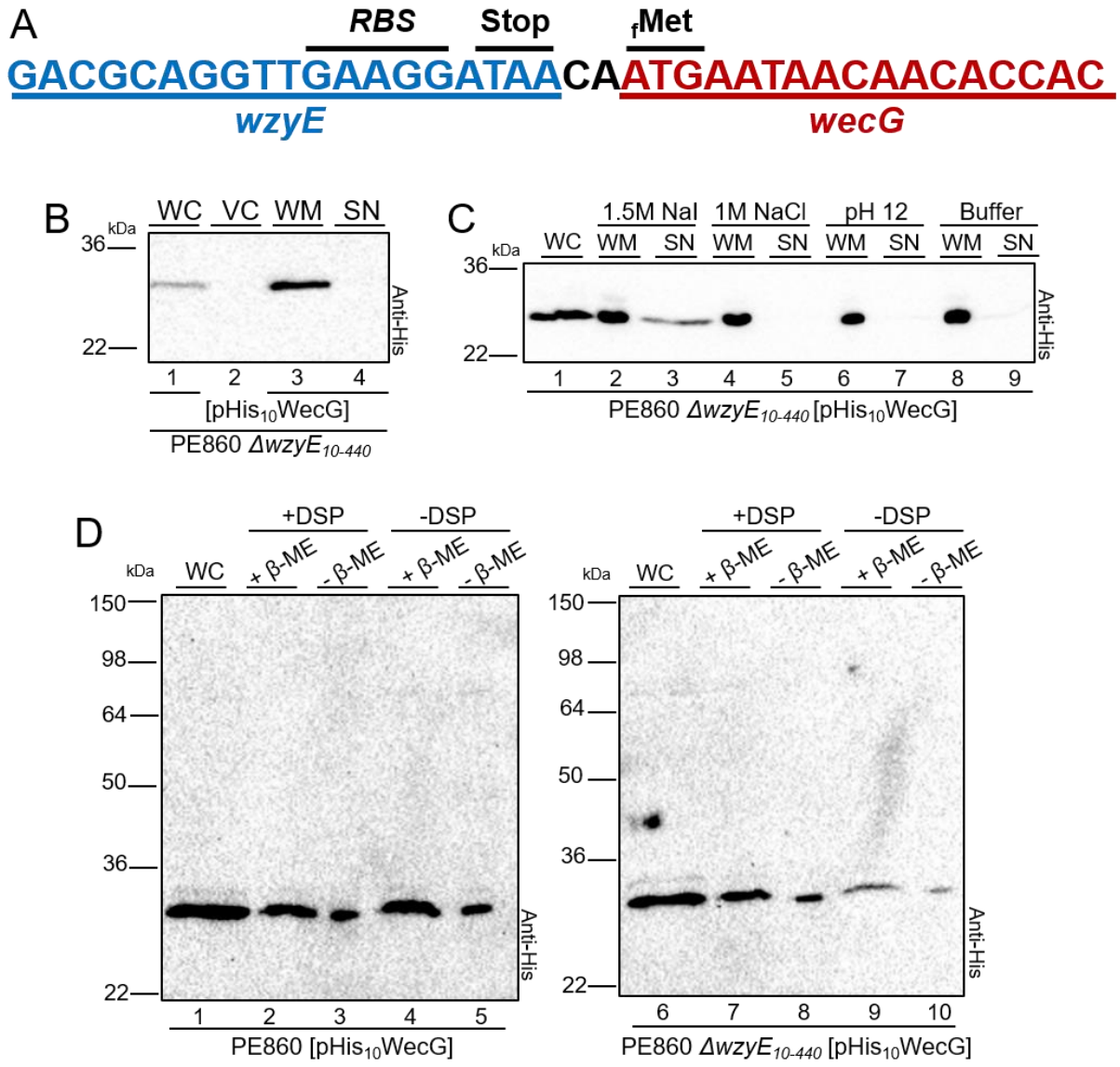


Figure 5.6: Effect of WzyE on WecG's peripheral membrane association.

a) Genetic arrangement of *wzyE* and *wecG* in *S. flexneri*. B) Anti-His Western immunoblot of the subcellular localization of His_{10}WecG in PE860 $\Delta wzyE_{10\text{-}440}$. Mid-exponential phase cells were disrupted by sonication and WM were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. WC=Whole cell, VC=Vector control, WM=Whole membrane, SN=Supernatant. C) Western immunoblot of chemical treated WMs from PE860 $\Delta wzyE_{10\text{-}440}$ cells expressing $pHis_{10}$WecG. WM were collected as above and were incubated 1:1 with 3 M NaI, 2 M NaCl or pH 12 buffer before membranes were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. d) Anti-His Western immunoblot of DSP treated PE860 and PE860 $\Delta wzyE_{10\text{-}440}$ cells expressing $pHis_{10}$WecG. Mid-exponential phase cells were treated with 0.1 M DSP and quenched with 1M Tris HCl prior to the addition of sample buffer with and without β-ME. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. SeeBlue Plus2 Pre-stained molecular weight ladder (Invitrogen) was used as a molecular weight marker for (b) to (d).

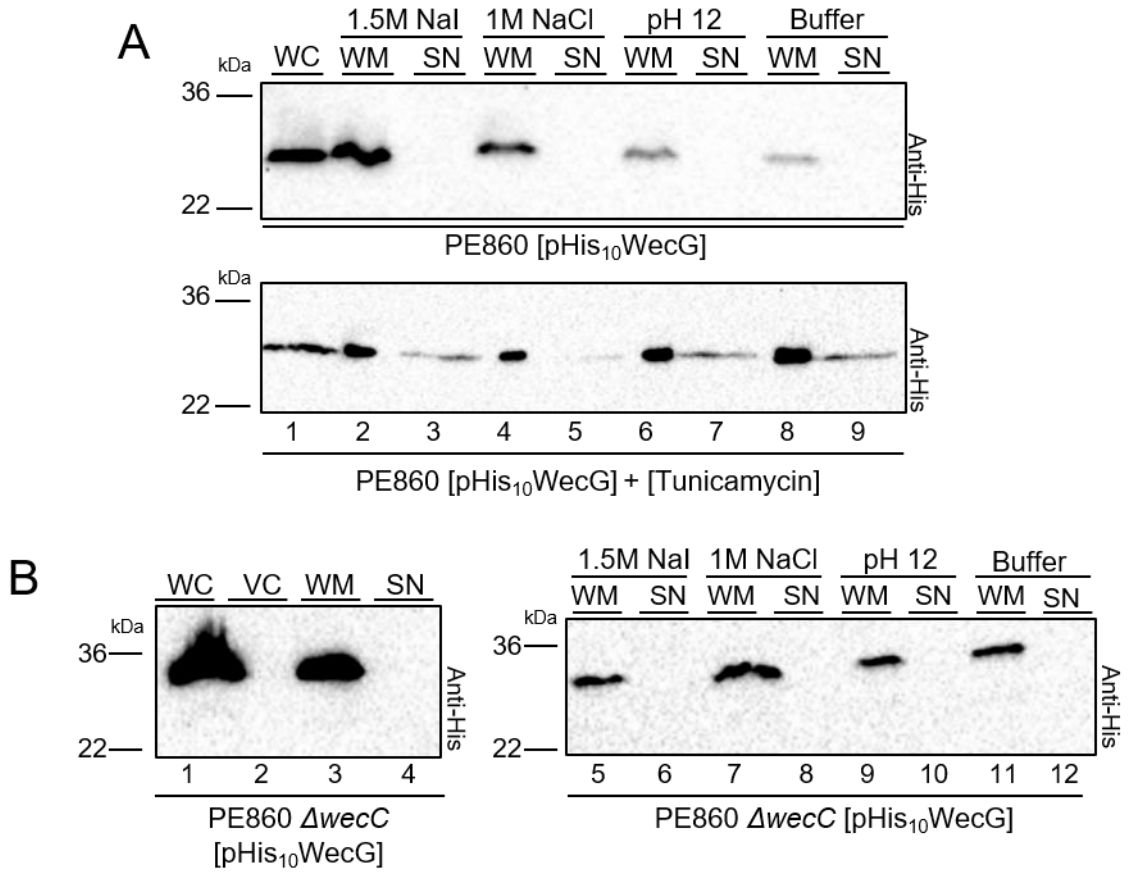


Figure 5.7: Investigating the role of lipids in WecG's peripheral membrane association.

a) Anti-His Western immunoblot of PE860 cells expressing $pHis_{10}$WecG treated with and without tunicamycin. Mid-exponential phase cells were grown in the presence of 10 ng/ml^{-1} tunicamycin and 3 $\mu g/ml^{-1}$ PMBN. Cells were then disrupted and membranes collected by ultracentrifugation, following which WMs were incubated 1:1 with 3 M NaI, 2 M NaCl or pH 12 buffer before membranes were collected by ultracentrifugation. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel followed by immunoblotting with anti-His antibodies. b) Anti-His Western immunoblot of PE860 *ΔwecC* cells expressing $pHis_{10}$WecG. Mid-exponential phase cells were disrupted and membranes collected by ultracentrifugation. Chemical treatment of WM samples were performed as above. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel prior to immunoblotting with anti-His antibodies. WC=Whole cell, VC=Vector control, WM=Whole membrane, SN=Supernatant. SeeBlue Plus2 Pre-stained molecular weight ladder (Invitrogen) was used as a molecular weight marker for (a) and (b).

5.7 Article Discussion

Glycosyltransferases play crucial roles in the biogenesis of the cellular envelope. WecG, which is involved in ECA biosynthesis, facilitates the addition of UDP-N-acetyl-D-mannosaminuronic acid to ECA lipid-I, generating ECA lipid-II (Meier-Dieter et al. 1990). Despite historically being thought to be a membrane protein (Barr et al. 1988), here we demonstrate that WecG is in fact a protein which is strongly associated with the membrane and present WecG as the second protein belonging to the novel glycosyltransferase fold family, GT-E (Mestrom et al. 2019).

Our study shows for the first time the presence and importance of WecG's C-terminal helices in maintaining WecG to the membrane. The importance of the C-terminal helices in GT-E fold proteins was previously demonstrated by *Kattke et al. 2019* who similarly demonstrated their role in maintaining TagA's membrane association as well as in forming TagA's catalytic domain (Kattke et al. 2019). Their importance was shown through the use of chemical treatments, CTD deletions and nucleotide substitutions to dissociate TagA from the membrane (Kattke et al. 2019).

Employing a similar approach in this study, we demonstrated that unlike TagA which was dissociated from membranes using 1.5 M NaI or 0.1 N KOH (Kattke et al. 2019), WecG cannot be dissociated from the membrane through the use of chemical treatments (Figure 5.3b). CTD deletions revealed that WecG's helix II is crucial for maintaining WecG's membrane association however, all three helices are required for ECA production (Figure 5.4c & d). Further, using a local multiple sequence alignment from a diverse range of phylogenetically distinct species, we focused on WecG's helix II and identified conserved leucine residues L215, L218 and L222. We then demonstrated their crucial roles in WecG's membrane association of as well as ECA production (Figure 5.5b, c & d) where, $His_{10}WecG^{L218E}$ completely dissociated from the membrane and the substitutions of $His_{10}WecG^{L215E}$ and $His_{10}WecG^{L222E}$ induced a partial dissociation from the membrane (Figure 5.5b & c).

Complementation assays using the *E. coli* K-12 BW25113 *ΔwecG::km* strain revealed the inability of the L218E mutant and the partial ability of the L222E mutant to complement the ECA production in the *wecG* mutant (Figure 5.5d). These results suggested that WecG's CTD helices are involved in two distinct roles; in WecG's membrane association and also in WecG's catalytic activity. The formation of the catalytic domain facilitated by the CTD helices is a common feature in GT-E fold glycosyltransferases as TagA's catalytic domain also is facilitated by its CTD helices (Kattke et al. 2019).

We also investigated other possible factors which could contribute to maintaining WecG's membrane association. In the absence of WzyE, WecG could be dissociated from the WM by treatment with 1.5 M NaI (Figure 5.6c). This dissociation could have been induced due to the absence of a physical interaction between WzyE and WecG. Alternatively, it is known that mutants of the *wec* operon induce morphological abnormalities in cells due to the build-up of dead-end biosynthetic precursors (Jorgenson et al. 2016). As such, it is plausible that a build-up of biosynthetic precursors due to the reduced ECA lipid-III usage in *wzyE* mutant could have interrupted lipid mediated interactions between WecG and ECA biosynthetic precursors, namely ECA lipid-I or ECA lipid-II (Figure 5.1) (Maczuga et al. 2022).

To investigate these possibilities, DSP crosslinking and chemical treatment of cells with 10 ng/ml^{-1} tunicamycin and 3 µg PMBN to inhibit ECA biosynthesis was performed (Al-Dabbagh, Mengin-Lecreulx & Bouhss 2008). DSP crosslinking showed WecG did not physically interact with any other protein and, unlike TagA, did not dimerise (Figure 5.6d). This supported the hypothesis that perhaps WecG is instead maintained to the membrane through lipid mediated interactions rather than a physical interaction with WzyE. Membrane fractionation of cells treated with tunicamycin to block ECA lipid-I production (Figure 5.7a) revealed that WecG could not be maintained exclusively at the membrane, and suggested that the biosynthetic intermediates themselves played a critical role in maintaining WecG's membrane association. It is known that some glycosyltransferase's membrane association; such as MurM, are facilitated via direct interactions between the biosynthetic intermediate and the protein, where the biosynthetic intermediate becomes the tether facilitating membrane association (York et al. 2021).

To determine if WecG was maintained at the membrane with the aid of biosynthetic intermediates, chemical treatment of whole cells was performed in the PE860 *wecC* mutant. *wecC* mutants prevent the generation of ECA lipid-II (Figure 5.1) and it was showed that WecG remained associated with the membrane (Figure 5.7b) whereas, when cells were treated with tunicamycin it caused the dissociation of WecG from MFs (Figure 5.7b). These results suggest that WecG is indeed maintained to the membrane with aid of its biosynthetic, membrane embedded substrate, ECA lipid-I. Overall, the results support the hypothesis that WecG is maintained to the membrane through interactions with ECA lipid-I which is facilitated through its unusual, protruding CTD helices which are reminiscent GT-E fold glycosyltransferases.

The GT-E fold is a novel glycosyltransferase family where unlike GT-A and GT-B folds, they contain a single Rossmann-fold domain followed by a series of three α-helices (O'Toole, Imperiali & Allen 2021). The fold was first exhibited and described by *Kattke et al.* (*2019)*, where

they presented TagA as the first protein to be denoted with the novel fold. Despite the lack of an experimentally determined structure, WecG contains elements which are inconsistent with the biophysical definitions of GT-A, GT-B, GT-C or GT-D fold glycosyltransferases and hence, due to WecG's similarities with TagA we characterize it as a GT-E fold glycosyltransferase.

WecG lacks the canonical Asp-X-Asp motif which is used to coordinate Mg^{2+} or Mn^{2+}ions where upon binding, induce local conformational changes which facilitate substrate binding and catalysis in GT-A class glycosyltransferases (Liang et al. 2015). Due to the relative length of WecG's peptide sequence, it is unlikely to form two Rossmann folds which are characteristic of GT-B fold glycosyltransferases where the binding pockets reside deep within the two Rossmann folds (Liang et al. 2015). Additionally, despite being an *in silico* generated structure, the structures predicted by I-TASSER, RaptorX and Alphafold all independently predict that WecG consists of a single Rossmann fold in which, the remaining residues would be insufficient to fold into a second Rossmann fold (Figure 5.2a, b & c). GT-C glycosyltransferases commonly contain a N-terminal TM domain and a C-terminal globular domain. Common motifs of GT-C family glycosyltransferases are the WWDYG motif which is believed be to be the catalytic site of the glycosyltransferase with two additional, smaller motifs the DK (DXXK) and or MI (MXXI) motifs which are spatially adjacent to the WWDYG motif (Igura et al. 2008). Similar to the coordination motifs of GT-A family glycosyltransferases, WecG does not contain either of the three motifs and does not, indicated through the bioinformatic analysis, possess any TM segments hence, it is improbable that WecG is a GT-C family glycosyltransferase. Lastly, GT-D fold glycosyltransferases were first described by *Zhang et al, (2014)* (Zhang et al. 2014) with DUF1792 (PDB ID:4PFX) and are characterized by the presence of a DXE motif that facilitates Mn^{2+} ion binding which is required for catalysis. Like TagA, WecG shares limited predicted structural homology with DUF1792 however, this limited structural identity is most likely due to TagA being used as a structural template by I-TASSER and RaptorX during the *in silico* generation of the WecG structural model, and hence this limited similarity shared between DUF1792 and WecG is uncertain and speculative at best.

Due to the broad similarities between TagA and WecG, the GT-E family of glycosyltransferases best describe WecG. WecG, like TagA, possess a single *in silico* predicted, N-terminal Rossmann fold followed by three CTD α-helices which we have shown to be crucial to maintaining WecG's membrane association (Figure 5.4). Similarly to TagA, substitutions of conserved residues along CTD helix II of WecG prevent membrane association (Figure 5.5b & c) and, un-investigated in TagA, these CTD helix II substitutions were shown to greatly impact polysaccharide biosynthesis (Figure 5.5d). From these observations it is clear that WecG is

peripherally associated with the membrane via its CTD helices as is TagA. These similarities between TagA and the disparities between WecG and the structural requirements of the other GT fold families (A,B,C,D) place WecG as the second glycosyltransferase to be identified as a GT-E fold glycosyltransferase.

In conclusion, this study reveals the true nature of WecG as not a membrane protein, but a protein which is peripherally associated with the membrane and in doing so, classifies WecG as a GT-E fold glycosyltransferase. Additionally, the identification of WecG's reliance on its membrane association for ECA production provides new insight into our understanding of the biosynthesis of ECA, as well as provides novel drug targets to inhibit ECA biosynthesis.

5.8 Article References

Al-Dabbagh B, Mengin-Lecreulx D, Bouhss A. 2008. Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA. *J Bacteriol* 190:7141-7146.

Amar A, Pezzoni M, Pizarro RA, Costa CS. 2018. New envelope stress factors involved in σE activation and conditional lethality of *rpoE* mutations in *Salmonella enterica*. *Microbiology* 164:1293-1307.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008.

Barr K, Ward S, Meier-Dieter U, Mayer H, Rick PD. 1988. Characterization of an *Escherichia coli rff* mutant defective in transfer of N-acetylmannosaminuronic acid (ManNAcA) from UDP-ManNAcA to a lipid-linked intermediate involved in enterobacterial common antigen synthesis. *J Bacteriol* 170:228-33.

Bottoms CA, Smith PE, Tanner JJ. 2002. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Prot Sci* 11:2125-2137.

Castelli ME, Vescovi EG. 2011. The Rcs signal transduction pathway is triggered by enterobacterial common antigen structure alterations in *Serratia marcescens*. *J Bacteriol* 193:63-74.

Danese PN, Oliver GR, Barr K, Bowman GD, Rick PD, Silhavy TJ. 1998. Accumulation of the enterobacterial common antigen lipid II biosynthetic intermediate stimulates *degP* transcription in *Escherichia coli*. *J Bacteriol* 180:5875.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-5.

Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. Jpred4: a protein secondary structure prediction server. *Nucleic Acids Research* 43:389-394.

Eade CR, Wallen TW, Gates CE, Oliverio CL, Scarbrough BA, Reid AJ, Jorgenson MA, Young KD, Troutman JM. 2021. Making the enterobacterial common antigen glycan and measuring its substrate sequestration. *ACS Chem Biol* 16, 691-700.

Gozdziewicz TK, Lukasiewicz J, Lugowski C. 2015. The structure and significance of enterobacterial common antigen (ECA). *Postepy Hig Med Dosw* 69:1003-12.

Heifetz A, Keenan RW, Elbein AD. 1979. Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate GlcNAc-1-phosphate transferase. *Biochemistry* 18:2186-2192.

Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K, Kohda D. 2008. Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* 27:234-243.

Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can J Microbiol* 60:697-716.

Jiang X, Tan WB, Shrivastava R, Seow DCS, Chen SL, Guan XL, Chng SS. 2020. Mutations in enterobacterial common antigen biosynthesis restore outer membrane barrier function in *Escherichia coli tol-pal* mutants. *Mol Microbiol* 2020;114:991–1005.

Jiang XE, Shrivastava R, Tan WB, San Seow DC, Chen SL, Guan XL, Chng S-S. 2018. Biosynthetic intermediates of the enterobacterial common antigen overcome outer membrane lipid dyshomeostasis in *Escherichia coli*. *BioRxiv* 480533:480533.

Jorgenson MA, Kannan S, Laubacher ME, Young KD. 2016. Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate. *Mol Microbiol* 100:1-14.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583-589.

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

Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. 2012. Template-based protein structure 168odelling using the RaptorX web server. *Nat Prot* 7:1511.

Kattke MD, Gosschalk JE, Martinez OE, Kumar G, Gale RT, Cascio D, Sawaya MR, Philips M, Brown ED, Clubb RT. 2019. Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *PloS Pathog* 15:e1007723.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567-80.

Liang DM, Liu JH, Wu H, Wang BB, Zhu HJ, Qiao JJ. 2015. Glycosyltransferases: mechanisms and applications in natural product development. *Chem Soc Rev* 44:8350-74.

Lugtenberg B, Meijers J, Peters R, van der Hoek P, van Alphen L. 1975. Electrophoretic resolution of the 'major outer membrane protein' of *Escherichia coli* K12 into four bands. *FEBS Let* 58:254-258.

Maciejewska A, Kaszowska M, Jachymek W, Lugowski C, Lukasiewicz J. 2020. Lipopolysaccharide-linked enterobacterial common antigen (ECA(LPS)) occurs in rough strains of *Escherichia coli* R1, R2, and R4. *Int J Mol Sci* 21.

Marolda CL, Tatar LD, Alaimo C, Aebi M, Valvano MA. 2006. Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide o antigen. *J Bacteriol* 188:5124-35.

Meier-Dieter U, Starman R, Barr K, Mayer H, Rick PD. 1990. Biosynthesis of enterobacterial common antigen in Escherichia coli. Biochemical characterization of Tn10 insertion mutants defective in enterobacterial common antigen synthesis. *J Biol Chem* 265:13490-7.

Mestrom L, Przypis M, Kowalczykiewicz, Pollender A, Kumpf, Marsden, Bento I, Jarzębski, Szymańska, Chruściel A, Tischler D, Schoevaart, Hanefeld, Hagedoorn P. 2019. Leloir glycosyltransferases in applied biocatalysis: A multidisciplinary approach. *Int J Mol Sci* 20:5263.

Mitchell AM, Srikumar T, Silhavy TJ. 2018. Cyclic enterobacterial common antigen maintains the outer membrane permeability barrier of *Escherichia coli* in a manner controlled by YhdP. *Mbio* 9:e01321-18.

Murray GL, Attridge SR, Morona R. 2003. Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol Microbiol* 47:1395-1406.

O'Toole KH, Imperiali B, Allen KN. 2021. Glycoconjugate pathway connections revealed by sequence similarity network analysis of the monotopic phosphoglycosyl transferases. *Pro Nat Aca Sci* 118:e2018289118.

Purins L, Van Den Bosch L, Richardson V, Morona R. 2008. Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2. *Microbiology* 154:1104-1116.

Ramos-Morales F, Prieto AI, Beuzon CR, Holden DW, Casadesus J. 2003. Role for *Salmonella enterica* enterobacterial common antigen in bile resistance and virulence. *J Bacteriol* 185:5328-32.

Saidijam M, Azizpour S, Patching SG. 2018. Comprehensive analysis of the numbers, lengths and amino acid compositions of transmembrane helices in prokaryotic, eukaryotic and viral integral membrane proteins of high-resolution structure. *J Biomol Struct Dyn* 36:443-464.

Smith SM. Strategies for the purification of membrane proteins. *Methods Mol Biol*. 2017;1485:389-400.

Swoboda JG, Campbell J, Meredith TC, Walker S. 2010. Wall teichoic acid function, biosynthesis, and inhibition. *Chem Biochem*: a European journal of chemical biology 11:35-45.

Wang RF, Kushner SR. 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 100:195-9.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2: A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189-91.

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

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12:7-8.

York A, Lloyd AJ, del Genio CI, Shearer J, Hinxman KJ, Fritz K, Fulop V, Dowson CG, Khalid S, Roper DI. 2021. Structure-based 170odelling and dynamics of MurM, a *Streptococcus pneumoniae* penicillin resistance determinant present at the cytoplasmic membrane. *Structure* 29, 1-12.

Zhang H, Zhu F, Yang T, Ding L, Zhou M, Li J, Haslam SM, Dell A, Erlandsen H, Wu H. 2014. The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat Commun* 5:4339.

5.9 Article Supporting Information

5.9.1 Supporting Figures

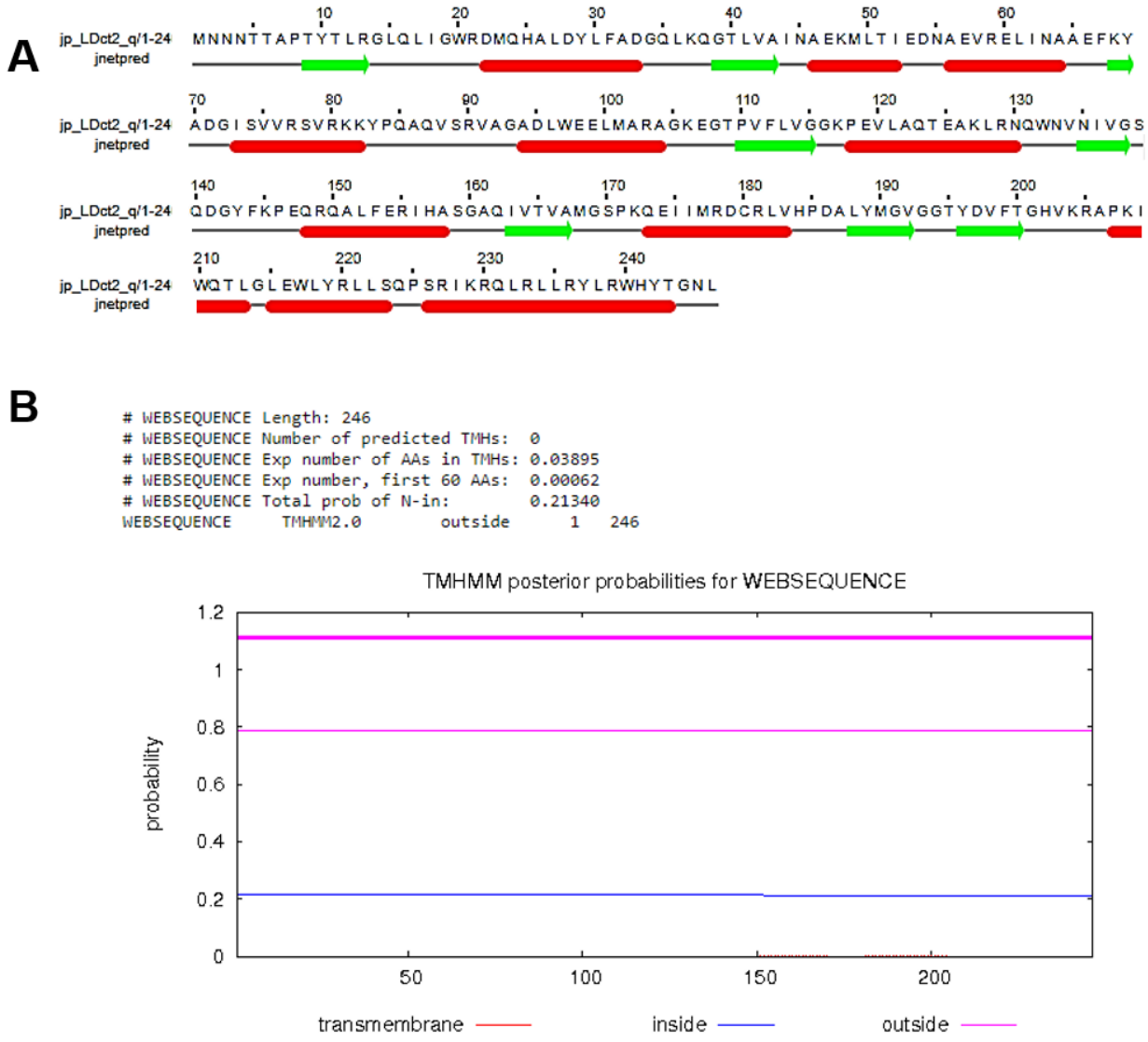


Figure 5.8-(S1): Bioinformatic analysis of WecG.

a) Jpred analysis of WecG's secondary structure shows that WecG consists of 11 alpha helices and 8 small beta sheets. Of relevance here are the three carboxyl terminal domain alpha helices predicted by Jpred. B) TMHMM analysis of the probability of transmembrane segments in WecG based on peptide sequence.

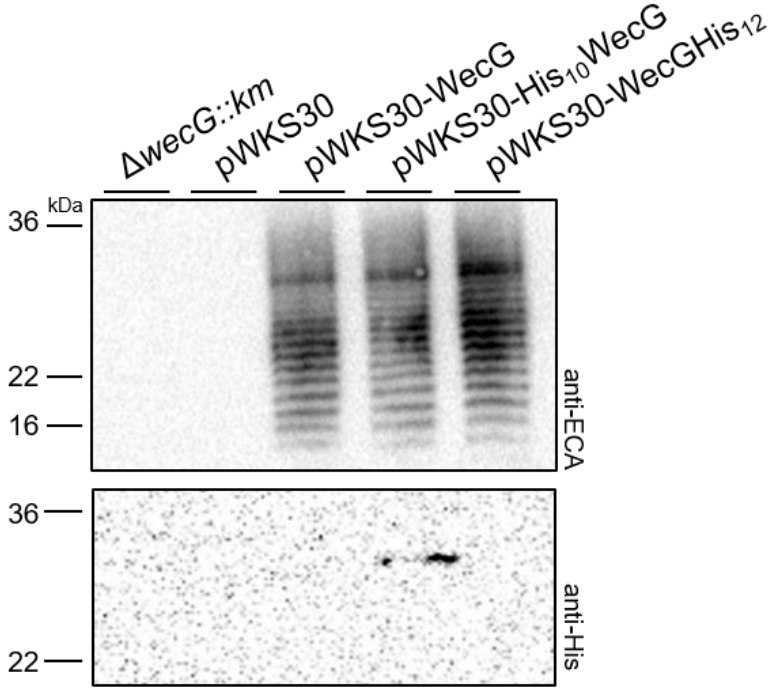


Figure 5.9-(S2): Detection and complementation of *E. coli* K-12 BW25113 *ΔwecG::km* using pWKS30 plasmids.

Anti-ECA Western immunoblot showing ECA profiles of *E. coli* K-12 BW25113 *ΔwecG::km* cells expressing pWKS30 complementation plasmids. Mid-exponential phase cells were collected ($1x10^9$ cells) by centrifugation and lysed in lysis buffer in the presence of proteinase K. Samples were electrophoresed on a SDS-15% (w/v) PAGE gel followed by immunoblotting with anti-ECA antibodies. B) Anti-His Western immunoblot of *E. coli* K-12 BW25113 *ΔwecG::km* expressing pWKS30-His_{10}WecG. Mid-exponential phase cells were collected ($5x10^8$ cells) and lysed in lysis buffer. Samples were electrophoresed on a SDS-12% (w/v) PAGE gel, transferred onto a nitrocellulose membrane and probed with anti-His antibodies. SeeBlue Plus2 pre-stained protein ladder (Invitrogen) was used as a molecular mass standard

5.10 Article graphical abstract

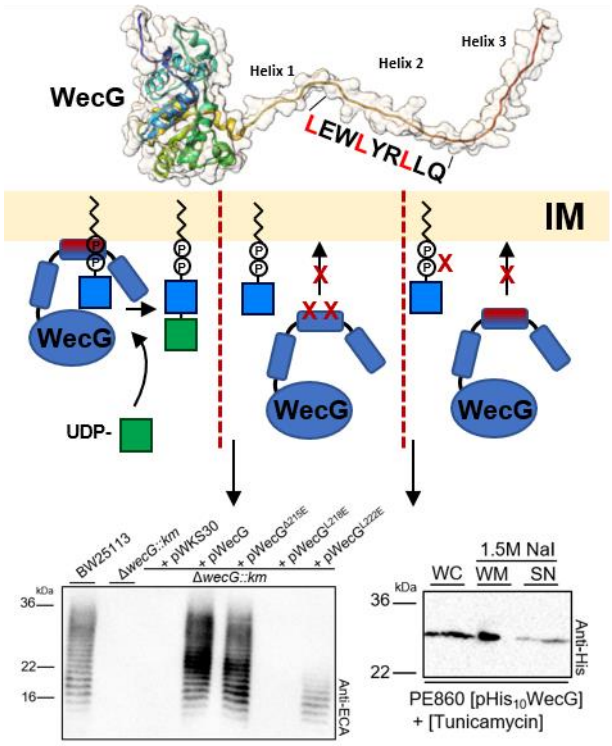


Figure 5.10: Graphical abstract for submission to Molecular Microbiology.

Conserved leucine residues on novel WecG structure allow for WecG to associate with the inner membrane. Mutations of the residues or, the lack of ECA lipid-I prevents WecG's association with the membrane

5.11 Article Acknowledgments

Funding for this work was provided by a Discovery Project Grant to R. Morona from the Australian Research Council (PROJECT ID: DP160103903). N. Maczuga is the recipient of a Research Training Program Stipend Research Scholarship from the University of Adelaide.

5.12 Article Conflicts of interest

The author(s) declare that there are no conflicts of interest.

Chapter Six

ADDITIONAL PHENOTYPES OF *wzyE* MUTANTS

Chapter 6: Additional phenotypes of *wzyE* mutant.

6.1 Introduction.

While many *wec* operon mutants have been generated and investigated, single mutations in *wzyE* remain unreported and uncharacterised. This is due to the concept which was introduced by Baba et al. (2006) wherein a BW25113 *wzyE::kan*^{R} mutant was unable to be generated in the KEIO collection, and subsequently *wzyE* was deemed an essential gene (Baba et al. 2006). Researchers since have not reported on *wzyE* mutants until Chapter 3 and 5, and thus they remain completely un-investigated and uncharacterised. This is also seen in research into WzyE proteins as a whole as to date, besides Chapter 5, there have been no studies directed to investigating the WzyE, one of the two major Wzy proteins used in the Wzy-dependent pathway (Whitfield, Wear & Sande 2020).

The aim of this chapter is to phenotypically characterise *wzyE* mutants beyond their outer membrane polysaccharide profiles in *Shigella flexneri* and to investigate any potential additional phenotypes.

6.2 Generation of *ΔwzyE* and *ΔwzyE*$_{10\text{-}440}$ mutants by λ Red mutagenesis.

The investigation into *wzyE* required the generation of *wzyE* mutants in *S. flexneri* PE860 which were generated through λ Red mutagenesis (Datsenko & Wanner 2000). The *wec* operon contains multiple overlapping genetic elements across its twelve genes. This includes *wzyE* where the 3' end of *wecF*, the upstream gene to *wzyE*, lies with in the 5' coding region of *wzyE*. Additionally, this is likewise observed with the downstream gene *wecG*, whose RBS is predicted to lie within the 3' end of *wzyE*'s coding region.

With this in-mind, two mutants were generated: PE860 Δ*wzyE* which deleted *wzyE* entire coding region, and PE860 Δ*wzyE*$_{10\text{-}440}$ which left 30 nucleotides (nts) of *wzyE*'s 5' and 3' coding region intact. The methodology to generate the *wzyE* mutants is described in 2.3.16**.** Briefly, two DNA oligomers were designed with homology (50 bps) to the 5' and 3' regions of *wzyE* and were used to amplify a *frt::cml*^{R}*::frt* resistance cassette from pKD3. The resulting fragment of DNA was electroporated into PE860 cells harbouring pKD46 where the chromosomal homology based recombination occurred. The resulting transformants were plated onto selective LBA and confirmed by PCR. The removal of the resistance cassette was performed through the transformation and induction of pCP20 which removed the *cml*^{R} cassette leaving a single *frt* scar sequence on the chromosome.

6.3 Characterization of the outer membrane polysaccharides of *wzyE* mutants.

The full PE860 *wzyE* mutant with a complete deletion (Δ*wzyE*) has been shown to lack the ability to assemble linear ECA until complemented, as well as have reduced Oag biosynthesis Chapter 4. With this in mind, we set out to determine whether the different form of *wzyE* mutant, PE860 Δ$wzyE_{10\text{-}440}$, would display the same OM polysaccharide phenotypes as the PE860 complete deletion *wzyE* mutant. This was investigated through the use of anti-ECA Western immunoblotting to determine if the PE860 Δ$wzyE_{10\text{-}440}$ mutant could be complemented like the PE860 Δ*wzyE* mutant, as well as by LPS silver staining to determine if the PE860 Δ$wzyE_{10\text{-}440}$ mutant likewise had a reduction in Oag biosynthesis as observed for the PE860 Δ*wzyE* mutant Chapter 4.

Cells from mid-exponential phase cultures were collected to generate polysaccharide samples prepared (2.6.1) for anti-ECA Western immunoblotting (2.6.3) as well as for LPS silver staining (2.6.2). As expected, neither mutant was capable of polymerizing ECA, seen as a lack of a banding pattern, until complementation with pBAD33-WzyE3XFLAG (pWzyE) Chapter 4 which restored their function (Figure 6.1 A and B lanes 1 and 4). While both mutants could be complemented with pWzyE, only the $wzyE_{10\text{-}440}$ mutant was able to be completely complemented, observed as a darker ECA banding pattern (Figure 6.1 B lane 4). This is most likely due to the disruption in *wecG*'s expression as the full deletion in PE860 *ΔwzyE* results in the deletion of *wecG*'s *in silico* predicted RBS; the *wecG* product is the gene responsible for the biosynthesis of ECA lipid-II (Barr et al. 1988). Disruption of *wecG*'s RBS does not occur in the $wzyE_{10\text{-}440}$ mutant.

We then investigated the ability of both *wzyE* mutants to polymerise LPS Oag by LPS silver staining (Figure 6.2 A). Again, as expected, both *wzyE* mutants produced less Oag than parent PE860 (**, $p < 0.01$) which is consistent with previous findings Chapter 4. Additionally, there was no observable statistical difference (ns) between the amount of Oag polymerized between the two mutants (Figure 6.2 B).

6.4 Cell sensitivities of *wzyE* mutants to deoxycholate (DOC) and Colicin E2.

As it was clear that the two *wzyE* mutants possessed similar OM polysaccharide phenotypes, we assessed if they likewise showed similar sensitivities to extracellular toxins. Cells lacking ECA are known to become sensitive to the bile salt DOC (Jorgenson et al. 2016; Ramos-Morales et al. 2003) where they are commonly observed to be unable to grow in the presence of 1% DOC (Jorgenson et al. 2016).

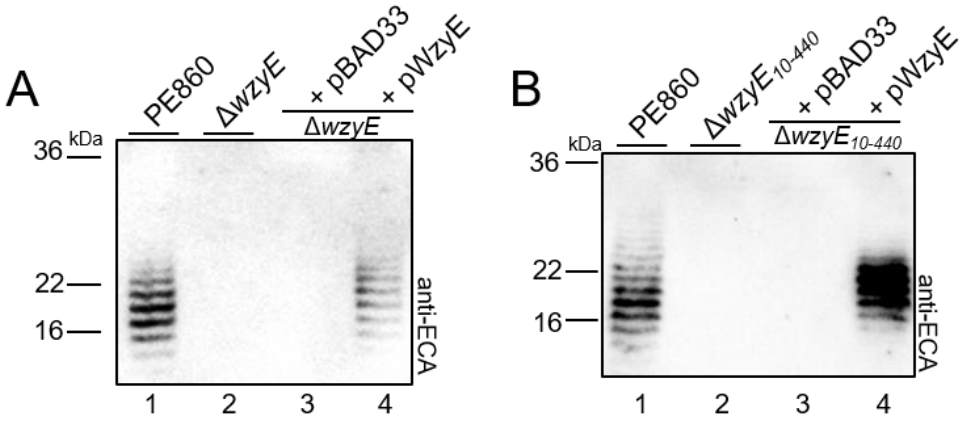


Figure 6.1: Complementation of *S. flexneri wzyE* mutants restores ECA assembly.

Anti-ECA Western blot showing complementation of (a) PE860 Δ*wzyE* and (b) PE860 Δ$wzyE_{10\text{-}440}$ mutants. Mid-exponential phase cells were collected by centrifugation ($1x10^9$ cells) and lysed with lysis buffer in the presence of Proteinase K. Samples were then electroporated on a SDS-15% (w/v) PAGE gel, transferred onto a nitrocellulose membrane and probed with polyclonal rabbit anti-ECA antibodies.

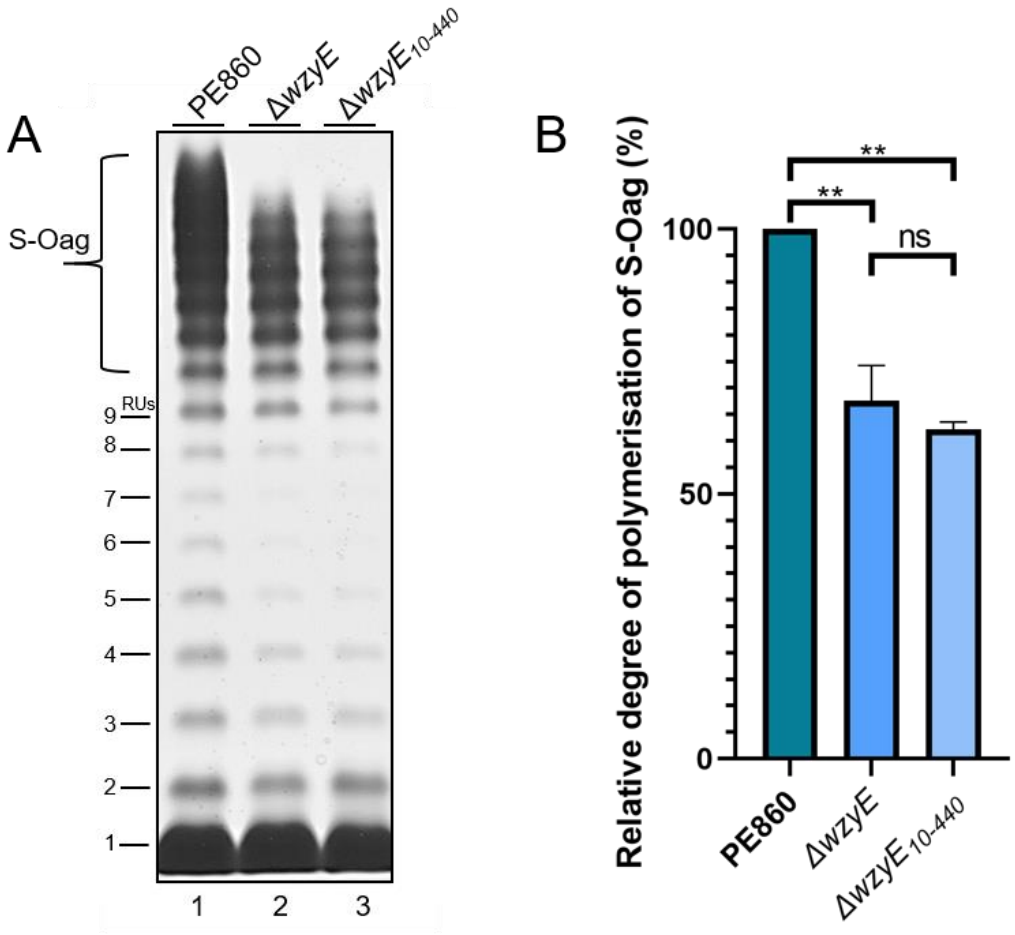


Figure 6.2: *S. flexneri wzyE* mutants show reduced amounts of LPS O antigen.

(a) Analysis of LPS profiles of PE860, PE860 Δ*wzyE* denoted as (Δ*wzyE*) and PE860 Δ$wzyE_{10\text{-}440}$ denoted as (Δ$wzyE_{10\text{-}440}$). Samples were made as above and electrophoresed on a SDS-15% (w/v) PAGE gel and silver stained. Number of RUs are indicated on the left hand side. (b) Analysis of polymerization of smooth Oag (S-Oag) via densitometry (ImageLab). Degree of polymerization of S-Oag is represented as the densitometry of Oag RUs from 10-17 as a percentage relative to the parent. Data represents 3 independent experiments with One-way ANOVA statistical analysis performed with SEM shown;**, $P<0.01$.

Similarly, cells lacking Oag are known to become sensitive to the bacteriocin DNase, Colicin E2 (Cramer, Sharma & Zakharov 2018). Upon interacting with BtuB, Colicin E2 is processed and enters the cytoplasm and performs its DNase activity; the steric hindrance from Oag is known to confer resistance to Colicins (Kim et al. 2014); (Tran, Papadopoulos & Morona 2014). As such, we assessed the sensitivities of the two *wzyE* mutants to DOC and Colicin E2 through the use of DOC resistance assays and Colicin E2 resistance assays.

To investigate any potential DOC sensitivity, cells from mid-exponential phase cultures were collected and serially diluted in LB and spotted onto LB agar plates supplemented with 1% DOC (2.7.1). This assay revealed that the PE860 *ΔwzyE* mutant was indeed sensitive to DOC with sensitivity being observed at dilution 10^{-3}. However, unlike the PE860 *ΔwzyE* mutant, the PE860 $\Delta wzyE_{10\text{-}440}$ showed no sensitivity to DOC, as its growth was comparable to the PE860 WT at all dilutions (Figure 6.3, A). Likewise, we tested the two mutants' sensitivity to Colicin E2 where cells from mid-exponential phase cultures were collected and spread onto LB agar plates, on top of which serially diluted Colicin E2 toxin (7.5 mg/ml) was spotted **(2.7.2)**. The PE860 *ΔwzyE* was more sensitive to colicin E2 compared to the WT as clearing were seen at 2 μg/ml of Colicin E2. However, the PE860 $\Delta wzyE_{10\text{-}440}$ mutant showed the opposite phenotype as a clearing was observed at 20 μg/ml revealing that the strain was hyper-resistant to Colicin E2 despite displaying less Oag compared to the WT as observed in the LPS silver staining (Figure 6.3, B). This data suggested that despite sharing similar OM polysaccharide phenotypes, the two *wzyE* mutants possessed very different sensitivities to extracellular toxins where it appeared that the PE860 $\Delta wzyE_{10\text{-}440}$ mutant was hyper-resistant to both DOC and Colicin E2.

Why the two *wzyE* mutants displayed such different phenotypes is unknown but likely complex. Plausible explanations can be made by accounting for the state of the remaining *wec* operon in both mutants with the key difference being the disruption of *wecG* expression Figure 6.4. In the PE860 *ΔwzyE* mutant, *wecG*'s expression is disrupted due to the loss of *wecG*'s RBS during the deletion of *wzyE*, resulting in reduced WecG activity, and hence less ECA biosynthetic intermediate accumulation in the IM. Due to the loss of ECA and reduced Oag, the PE860 *ΔwzyE* mutant is more sensitive to DOC and Colicin E2.

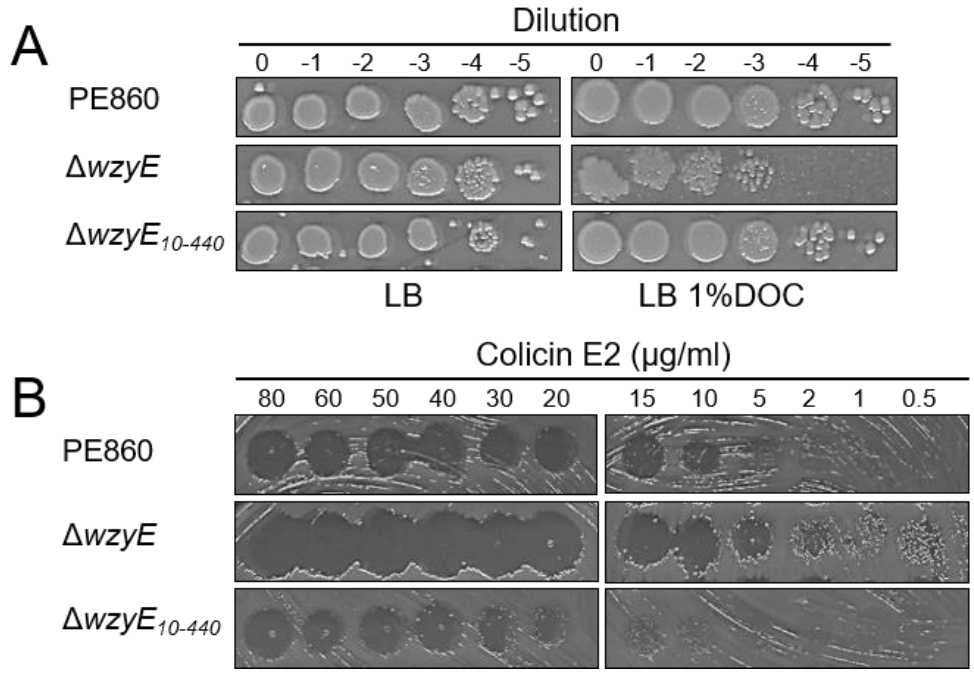


Figure 6.3: Sensitivities of *S. flexneri wzyE* mutants to DOC and Colicin E2.

(a) Analysis of PE860, PE860 Δ*wzyE* denoted as (Δ*wzyE*) and PE860 Δ*wzyE*$_{10\text{-}440}$ denoted as (Δ*wzyE*$_{10\text{-}440}$) sensitivity to 1% deoxycholate. Mid-exponential phase cells were collected (1x10^8 cells) and serially diluted 1:10 following which 3 µl of cellular suspensions were spotted onto pre-dried LBA plates supplemented with 1% DOC. (b) Analysis of PE860, PE860 Δ*wzyE* denoted as (Δ*wzyE*) and PE860 Δ*wzyE*$_{10\text{-}440}$ denoted as (Δ*wzyE*$_{10\text{-}440}$) sensitivity to Colicin E2. 100 µl of mid-exponential phase cells were spread onto pre-dried LBA plates. Colicin E2 protein was then serially diluted and 3 µls was spotted upon the plated bacteria. n=3.

	WecG activity	ECA lipid-III accumulation	Deoxycholate^{R}	Colicin E2^{R}
ΔwzyE	-	+	- - -	- - - -
ΔwzyE$_{10\text{-}440}$	+	+ + +	+ + +	+ + + +

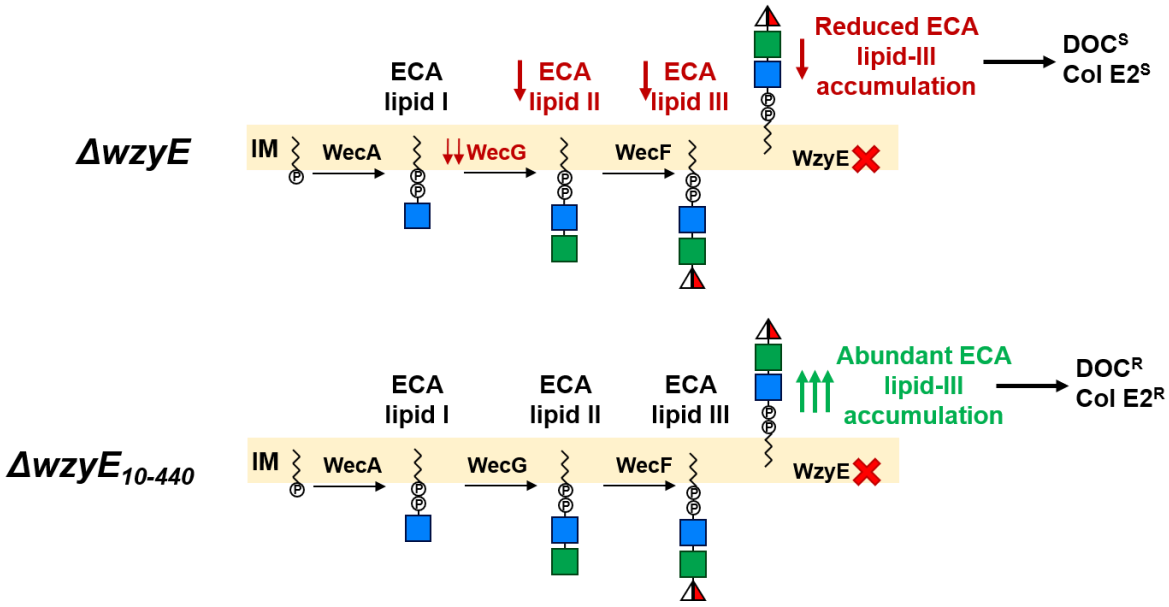


Figure 6.4: Summary of *wzyE* mutant phenotypes.

The phenotypes of the *wzyE* mutants are plausibly explained through the difference in the accumulation of ECA lipid-III in the IM due to the disruption of WecG expression in the *wzyE* mutant but not in the *ΔwzyE10-440* mutant.

However, in the PE860 *ΔwzyE*$_{10\text{-}440}$ mutant, *wecG*'s RBS remains unaltered leading to the undisrupted expression of WecG and subsequently, the accumulation of lipid linked ECA biosynthetic intermediates in the IM.

The accumulation of ECA biosynthetic intermediates themselves has been shown to provide resistances to a range of extracellular toxins. In *E. coli* K12, the accumulation of lipid linked intermediates due to disruptions in *wecG* and *wecF* restored vancomycin resistance in *tolA* mutants; the authors concluded that it was unclear why the accumulation of the biosynthetic intermediates provided resistance to vancomycin (Jiang et al. 2020). Similar examples are seen again in *E. coli* K12 with mutations in *wecC* and *wecF* restoring resistances to SDS and vancomycin (Mitchell, Srikumar & Silhavy 2018) as well as mutations in *wecE* and *wzxE* providing resistance to gentamycin and nalidixic and amikacin acid, respectively (Girgis, Hottes & Tavazoie 2009; Tamae et al. 2008). In all of these examples, the mutations described lead to the accumulation of lipid linked intermediates (Figure 1.9).

A plausible explanation of why the PE860 *wzyE*$_{10\text{-}440}$ mutant showed resistances to DOC and Colicin E2 may be due to the accumulation of ECA lipid-III where it may augment the permeability barrier of the IM causing the strain to become resistant to extracellular cytoplasmic toxins as in the studies above. Deoxycholate has been shown to cause unfolding and aggregation of cytosolic proteins (Cremers et al. 2014) which would be prevented in such a case. As for Colicin E2, perhaps a similar phenomenon may occurs in which the biosynthetic intermediates themselves alter the activity of proteins such as TolA that Colicin E2 requires to translocate across the IM (Cramer, Sharma & Zakharov 2018). As the Rcs stress response pathway is likely activated in both mutants (Castelli & Vescovi 2011; Wall, Majdalani & Gottesman 2018), it is unlikely to be responsible for the resistant phenotype observed in the *wzyE*$_{10\text{-}440}$ mutant.

Chapter Seven

CONCLUSIONS

Chapter 7: Conclusion

7.1 Novel Wzy topology characteristics are found in WzyE.

Complex OM polysaccharides such as ECA and LPS Oag are polymerized by the Wzy-dependent pathway (Whitfield, Wear & Sande 2020). Due to the importance of Oag in the pathogenesis and virulence of multiple Gram negative pathogens, the majority of research directed towards the Wzy-dependent pathway proteins has been directed towards the LPS Oag homologs. Subsequently, research towards the ECA homologs is limited with no studies directed towards WzyE, the ECA polymerase. In Chapter 3 I addressed this by investigating WzyE genetically, structurally and functionally in the first study directed towards understanding WzyE. Results in Chapter 3 revealed that WzyE is uncharacteristically highly conserved among *Enterobacterials*, displaying 23% global sequence conservation across 248 sequences. This was unexpected as Wzy proteins typically share little global sequence homology; it is possible that WzyE can display high levels of sequence homology as ECA is a conserved polysaccharide and thus WzyE has remained conserved and specialised. Topology mapping using C-terminal reporter fusions with PhoA::LacZα showed that WzyE, like other Wzy proteins, consisted of multiple transmembrane segments, periplasmic and cytoplasmic loops, with 13 transmembrane segments, 4 periplasmic loops and 6 cytosolic loops being identified along with an uncharacteristic topologically, ill-defined region. Cross referencing with an *in silico* predicted structure from Alphafold showed that the topologically ill-defined region coincided with a portion of WzyE predicted to consist of multiple small alpha helices and short loops where this region is exposed on the periplasmic leaflet of the IM. Chapter 3 also revealed that like other Wzy proteins, WzyE contains functionally important arginine residues where seven conserved arginine residues were identified throughout the C-terminal half of WzyE, and when mutated caused the loss of ECA polymerisation despite WzyE being detectable via anti-FLAG western Immunoblotting. Subsequently the localization of the arginine residues within the Alphafold predicted structure revealed a possible binding cavity within WzyE, coordinated by R204, R247, R399 and R408. I believe that this pocket facilitates the binding of the ECA polysaccharide chain through electrostatic interactions between the positively charged arginine residues and the negatively charged mannosaminuronic acid moieties which are components of the ECA polysaccharide RUs. The conserved arginine residues R295 and R309 were also shown to be present within the topologically ill-defined region and I suggest that the region may also be involved in the binding of the polysaccharide chains.

I also believe that due to the high peptide sequence conservation, WzyE may be the optimal Wzy protein to study mechanisms of Wzy protein function as well as the Wzy-dependent pathway.

Other key interacting regions of WzyE are yet to be identified such as binding regions with WzzE and as such, the highly conserved regions identified in here present a starting point for future studies into characterizing WzyE and Wzy proteins in general.

7.2 ECA and LPS Oag pathways are indirectly interdependent on one another in *Shigella flexneri.*

ECA and LPS Oag are both synthesised by separate homologs of the Wzy-dependent pathway (Whitfield, Wear & Sande 2020). The two Wzy-dependent pathways have traditionally been thought as being separate and independent from one another, with cross-complementation occurring between the two Wzy-dependent pathways only under specific genetic circumstances (Leo et al. 2020; Marolda et al. 2006). In Chapter 4 my findings show that in *S. flexneri,* ECA and LPS Oag biosynthetic pathways are indirectly dependent on one another where, if one pathway is biosynthetically blocked, this results in the hinderance of the other pathway and the reduction in OM polysaccharide biosynthesis. Furthermore, the findings show that this phenomenon is caused by the sequestration of Und-P in biosynthetic intermediates due to late stage biosynthetic blockages which starves Und-P from other key cell wall biosynthetic processes, i.e., peptidoglycan, LPS Oag and ECA biosynthesis. In these late stage biosynthetic mutants, I believe that the observed reduction in OM polysaccharide biosynthesis is due to the lack of free Und-P in the cell. While it has been demonstrated that mutations within the ECA and LPS Oag biosynthetic pathways separately impact peptidoglycan biosynthesis, the findings in Chapter 4 provide the first evidence that this same phenomenon occurs between ECA and LPS Oag as well in *S. flexneri* (Jorgenson et al. 2016; Jorgenson & Young 2016). Additionally, these findings show that not all ECA and LPS Oag biosynthetic mutations lead to the sequestration of Und-P from other related pathways; the mutations of *wecC* and *rmlD,* genes which commit Und-P to ECA and LPS Oag biosynthesis, respectively, do not lead to the sequestration of Und-P. I hypothesise that this is because, in these mutants, Und-P is not yet committed to either ECA or LPS Oag biosynthesis, respectively, hence Und-P is not sequestered in-dead end biosynthetic intermediates and subsequently does not impact other cell wall biosynthetic systems. The results of Chapter 4 clarify some of the pleotropic phenotypes of cell wall mutants which have often been attributed to the induction of cell stress response pathways such as the Rcs pathway (Castelli & Vescovi 2011).

I believe that this phenomenon may also occur in other bacteria, in which the two OM polysaccharides share common biosynthetic intermediates such as in other *Enterobacterials* or other Gram negative species. The findings of Chapter 4 also heed caution to researchers using cell wall mutants as to the possible pleotropic effects they may present and questions previously

published data in which cell wall mutants were investigated, as alternative plausible explanations may now better explain the phenomena observed.

7.3 WecG is a GT-E fold protein and is maintained to the membrane by ECA lipid-I.

As stated throughout the chapters of this thesis, the majority of research conducted towards *wec* operon proteins and genes have been directing in investigating their roles in causing pleotropic phenotypes (Castelli et al. 2008; Jorgenson et al. 2016). An example of this is WecG, a protein which was first and only described in 1988 and has subsequently only been investigated for inducing pleotropic phenotypes when mutated (Barr et al. 1988; Jiang et al. 2020; Jorgenson et al. 2016). In chapter 5, I addressed this and investigated WecG's subcellular localisation and key domains when I observed that WecG was predicted to fold in an uncharacteristic membrane protein structure. Prior to my findings, WecG had been thought of as a membrane protein due to initial studies performed by (Barr et al. 1988). However, through the use of C-terminal deletions and chaotropic chemical treatments, I showed that WecG is in-fact not a membrane protein but a protein which is peripherally associated with the membrane through interactions facilitated through it's C-terminal helices. Furthermore, I revealed the presence of conserved leucine residues along helix II which were shown to be critical for WecG's function and membrane association. I hypothesise that these leucine residues, once folded as a helix, provide a hydrophobic face on helix II which allows WecG to embed helix II into the IM and subsequently maintain its peripheral association with the membrane. Due to this unusual structure and the function of the C-terminal alpha helices as well as WecG's dissimilarities with other glycosyltransferase protein family folds, WecG was classified as the second protein in the novel glycosyltransferase fold family GT-E (Kattke et al. 2019). Proteins from this family are known to possess C-terminal alpha helices which are catalytically important and maintain the protein to the membrane where TagA, the other protein in this family, is also a N-acetyl-D-mannosaminuronic acid transferase involved in wall teichoic acid biosynthesis (Kattke et al. 2019). Additionally in Chapter 5, I showed the reliance of WecG's membrane association with the presence of biosynthetic intermediates revealing that they are crucial in maintaining WecG to the membrane. Through the use of *wecC* and *wzyE* mutants, and treatment with tunicamycin, I narrowed down the pool of available biosynthetic intermediates which showed that ECA lipid-I is required to maintain WecG's membrane association. Subsequently I hypothesise that ECA lipid-I is likely the glycolipid which WecG primarily uses to associate with the membrane as the loss ECA lipid-I causes the loss of WT WecG's membrane association.

The findings in Chapter 5 clarify our understandings of the nature of WecG as well as gain insight into it's mechanistic workings. I believe that the understanding of how critical WecG's

membrane association is to its function may allow the development of novel drugs to disrupt this interaction and hence, attenuate *Enterobacterial* pathogens. Furthermore, the results in Chapter 5 show the importance of understanding the *wec* operon genes and proteins biochemically, to clarify our understanding of how the system functions as opposed to investigating them to understand why they cause pleotropic phenotypes when mutated.

7.4 The accumulation of biosynthetic intermediates provides strong resistances in *wzyE* mutants.

While many *wec* operon mutants have been generated and investigated, *wzyE* remains un-investigated where a consensus has been formed that *wzyE* is an essential gene in *E. coli* and cannot be mutated. This is because of the inability for a *wzyE* mutant to be generated in the KEIO collection (Baba et al. 2006) which has subsequently lead to this assumption. Throughout Chapters 3, 4 and 5, *wzyE* mutants were generated and used to investigate various aims and showed that in *S. flexneri*, *wzyE* is not an essential gene. Due to the possibility of disrupting *wecG* expression by deleting the entire *wzyE* encoding region, as *wecG*'s RBS is predicted to lie within the 3' region of *wzyE*, two *wzyE* mutants were generated: PE860 Δ*wzyE* complete mutant, and PE860 Δ$wzyE_{10\text{-}440}$ mutant in which the first and last 30 bps of *wzyE* remained on the chromosome. Whilst both *wzyE* mutants shared similar Oag phenotypes, the $wzyE_{10\text{-}440}$ mutant seemed to be better complemented than the full *wzyE* mutant. The similarities in the Oag phenotype were expected and followed results previously observed in Chapter 4. However, the two mutants showed a large difference in their sensitivities to the extracellular toxins deoxycholate and Colicin E2, where the full *ΔwzyE* mutant was sensitive to both whereas the $wzyE_{10\text{-}440}$ mutant was hyper-resistant. I believe that the reasoning why the $wzyE_{10\text{-}440}$ mutant is resistant is due to the accumulation of biosynthetic intermediates which augment the permeability barrier of the cells. While the accumulation of biosynthetic intermediates does occur in the *ΔwzyE* mutant, due to the disruption in *wecG* expression, I believe that there is a reduction in the accumulation of the intermediates which is insufficient to protect the cells against deoxycholate and Colicin E2 unlike in the $wzyE_{10\text{-}440}$ mutant, which has undisrupted *wecG* expression. Similar resistant phenotypes have been observed in other *wec* operon mutants where the gained resistances have been attributed to the accumulation of lipid linked intermediates (Girgis, Hottes & Tavazoie 2009; Jiang et al. 2020; Mitchell, Srikumar & Silhavy 2018; Tamae et al. 2008). However, further research is required to substantiate these claims that the resistances seen in the $wzyE_{10\text{-}440}$ mutant are due to the accumulation of lipid linked intermediates. A suitable experiment would be the quantification of the intermediates through LC-MS.

The work performed in Chapter 6 showcases the potential ability for the ECA biosynthetic intermediates to directly provide strong resistances to stresses and toxins and, to my knowledge, it is the first to reveal this potential using mutants of the same gene. It is remarkable that such a strong difference in the phenotypes observed in the same gene mutant can occur, simply due to the potential disruption of gene expression of neighbouring genes. The genes of the *wec* operon are known to have overlapping genetic elements which includes overlapping reading frames and hence, the findings here caution researchers who investigate *wec* operon mutants as to the potential repercussions of disrupting neighbouring gene expression. Lastly, the findings here show that *wzyE* is not an essential gene in *S. flexneri* and may present *S. flexneri* as a model organism to investigate *wzyE* mutants, if the gene is indeed essential in *E. coli*.

7.5 Summary

The theme of this thesis was to explore Enterobacterial Common Antigen biosynthesis in *Shigella flexneri* with emphasis placed on exploring WzyE, the Wzy-dependent polymerase from the ECA biosynthetic system. The aspects of WzyE subsequently investigated generated novel understandings of WzyE which had previously not been investigated. The major findings revealed were that WzyE is highly conserved among Enterobacteriales, WzyE possesses an atypical Wzy protein topology with the possibility of a binding cavity within and this cavity is coordinated by conserved arginine residues which are functionally important. In the process of generating *wzyE* mutants, strong resistant phenotypes were observed when characterising them which was attributed to the accumulation of lipid linked biosynthetic intermediates in the $wzyE_{10\text{-}440}$ mutant and highlighted the importance of WecG in the system. WecG was subsequently investigated and the findings revolutionise our understanding of WecG and how the protein functions revealing it as a protein which is peripherally associated with the membrane. My findings showed how WecG is maintained through interactions between it's C-terminal helix II and the membrane, how conserved leucine residues affect this association and WecG's function, and hypothesise that WecG is primarily maintained to the membrane through interactions with ECA lipid-I. Ultimately the findings allowed me to characterise WecG as the second protein in a novel glycosyltransferase fold family, GT-E. Throughout the work investigating WzyE, additional phenomena were observed and subsequently investigated. Such work revealed that the two OM polysaccharide biosynthetic systems of *S. flexneri*, ECA and LPS Oag, are interdependent on one another where the deletion of *wzyE* and/or *wzyB* caused the decreased in OM polysaccharide abundance. This was subsequently shown to be due to the sequestration of Und-P from cell wall pathways and revealed an indirect link between ECA and Oag biosynthesis. The work presented in this thesis ultimately provides insight into multiple factors of the ECA biosynthetic pathway and not solely

on WzyE. The work here adds to establish concepts of cell wall pathways and revolutionises our understanding of neglected ECA biosynthetic proteins. Ultimately, the findings here greatly advances our understanding of the ECA biosynthetic pathway, reveals the interdependence of the OM polysaccharide pathways of *S. flexneri* and adds to our understanding of the Wzy-dependent pathway and Wzy proteins as a whole.

Future directions of studies from the work described here can be directed at answering multiple questions. First and foremost would be the quantification of Und-P linked intermediates in *wzyE* and $wzyE_{10\text{-}440}$ mutants as to discern if the accumulation of lipid linked intermediates truly do provide resistances to toxins and chemicals non-specifically. Simple experiments such as LC-MS would provide such answers. Other questions of interest include discerning the catalytic activity of WzyE proteins. A plausible catalytic chamber was described in *Chapter 3* and it would be interesting to classify the impact of amino acid substitutions of other nearby residues on WzyE's catalytic functions. And subsequently, as a follow up to *Chapter 3*, investigating the structure of WzyE would be very informative. As I believe WzyE may be the best Wzy model to understand the functions of Wzy proteins, an experimentally derived structure would provide great insight into the functions of Wzy proteins and the Wzy-dependent pathway. As described throughout the thesis, the protein responsible for the transfer of ECA to phosphatidylglycerol still remains unknown. An effort to discern the protein responsible should be undertaken so that we can gain a more complete understanding of ECA biosynthesis. Lastly, an effort should be made to characterise and re-examine underappreciated proteins of the *wec* operon. Besides WecG, multiple other *wec* proteins remain not fully characterised where, this includes the wide majority of *wec* proteins. It may be found that the original classifications do not suit the proteins any longer as more powerful and insightful techniques have been created.

Appendices

STRAINS LIST

CONSTRUCT LIST

OLIGONUCLEOTIDE LIST

BIBLIOGRAPHY

Appendix A - Bacterial Laboratory Strains

Strain ID	Description/Phenotype
***E. coli* K-12**	
DH5α	F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$ (NEB)
W3110	F$^-$ λ$^-$ rph-1 INV(*rrnD*, *rrnE*) (Laboratory stock)
BW25113	*lacI*$^+$*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR*514 Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$ *rph-1* Δ*(araB–D)567* Δ*(rhaD–B)568* Δ*lacZ4787*(::*rrnB-3*) *hsdR514 rph-1*
BW25113 *ΔwecG*	BW25113 *ΔwecG::kan*^{R} (Baba et al. 2006)
GM2929	*dam*-13::*Tn9* (Cm^R) *dcm*-6 *hsdR*2 *mcrA mcrB rpsL*136 (Str^R) (Laboratory stock)
C43	F$^-$ *ompT gal dcm hsdS*$_B$(r_B^- m_B^-)(DE3) (NEB)
BL21	F$^-$ *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) [*malB*$^+$]$_{K-12}$(λ^S) (NEB)
BL21 DE3	F$^-$ *ompT gal dcm lon hsdS*$_B$($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S) (NEB)
LEMO DE3	*fhuA2 [lon] ompT gal* (λ DE3) *[dcm]* Δ*hsdS/ pLemo(CamR)* (NEB)
Shigella flexneri	
RMA2162	PE860, *Shigella flexneri* Y (laboratory strain)
RMA4199	PE638, *Shigella flexneri* Y (Leo et al. 2021)
RMA4258	2457T, *Shigella flexneri* 2a (Laboratory strain)
MYRM1020	2457T Vp- *ΔgtrII::strep*^{R}
MYRM1034	2457T Vp- *Δgtr II::FRT ΔwzyB::strep*^{R}
MYRM1287	2457T Vp- *ΔwzyB::strep*^{R}

Appendix B - Bacterial strains generated

Strain ID	Parent (plasmid (s))	Description/Phenotype
NMRM51	RMA2162 [pKD46]	PE860 Amp^R, 30 ^oC sensitive
NMRM52	RMA2162 *ΔwzzE::cml*^{R}	PE860 *ΔwzzE*, Cml^R
NMRM53	RMA2162 *ΔwzyE::cml*^{R}	PE860 *ΔwzyE*, Cml^R
NMRM54	RMA2162 *ΔwzzE::frt*	PE860 *ΔwzzE*, (RMA5112)
NMRM55	RMA2162 *ΔwzyE::frt*	PE860 *ΔwzyE*, (RMA5113)
NMRM56	RMA2162 *ΔwzyE::frt ΔwzzE::cml*^{R}	PE860 *ΔwzyE ΔwzzE*, Cml^R
NMRM57	RMA2162 *ΔwzyE::frt ΔwzzE::frt*	PE860 *ΔwzyE ΔwzzE*, (RMA5115)
NMRM58	BL21 DE3 LEMO [pMA632]	BL21 DE3 LEMO, Amp^R
NMRM59	W3110 [pAC/BAD T-7] [pWALDO-WzyE-GFP-His_8]	W3110 *ΔwzyE::frt*, Kan^R, Cml^R
NMRM60	DH5α [pWALDO-WzyE-GFP-His_8]	DH5α, Kan^R
NMRM61	DH5α [pPLEO1-$WzyE_{SF}$]	DH5α, Amp^R
NMRM62	DH5α [pBAD33-WzyE-GFP-His_8]	DH5α, Cml^R
NMRM63	DH5α [pBAD33-WzyE3xFLAG]	DH5α, Cml^R
NMRM64	RMA5113 [pPLEO1^I]	PE860 *ΔwzyE::frt*, Amp^R
NMRM65	RMA5113 [pPLEO1-WzyE^I]	PE860 *ΔwzyE::frt*, Amp^R
NMRM66	RMA5113 [pBAD33]	PE860 *ΔwzyE::frt*, Cml^R
NMRM67	RMA5113 [pBAD33-WzyE-GFP-His_8]	PE860 *ΔwzyE::frt*, Cml^R
NMRM68	RMA5113 [pBAD33-WzyE3xFLAG]	PE860 *ΔwzyE::frt*, Cml^R
NMRM69	Not used	Not used
NMRM70	Not used	Not used
NMRM71	GM2929 TN9 [pPLEO1-WzyE^I]	GM2929 Dam^- Dcm^-, Cml^R, Amp^R
NMRM72	DH5α [pPLEO1-WzyE^{II}]	DH5α [pPLEO1-WzyE^{a1350t}], Amp^R
NMRM73	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V100, Amp^R
NMRM74	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at N143, Amp^R
NMRM75	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V190, Amp^R
NMRM76	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at T268, Amp^R
NMRM77	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L325, Amp^R
NMRM78	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G371, Amp^R
NMRM79	BL21 [pBAD33-WzyE3xFLAG]	BL21, Cml^R
NMRM80	BL21 DE3 [pBAD33-WzyE3xFLAG]	BL21 DE3, Cml^R
NMRM81	C43 [pBAD33-WzyE3xFLAG]	C43, Cml^R

NMRM82	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R30, Amp^{R}
NMRM83	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G64, Amp^{R}
NMRM84	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A83, Amp^{R}
NMRM85	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F132, Amp^{R}
NMRM86	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at D178, Amp^{R}
NMRM87	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at T203, Amp^{R}
NMRM88	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G222, Amp^{R}
NMRM89	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at M250, Amp^{R}
NMRM90	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at N284, Amp^{R}
NMRM91	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at I338, Amp^{R}
NMRM92	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G444, Amp^{R}
NMRM93	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L352, Amp^{R}
NMRM94	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F392, Amp^{R}
NMRM95	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at S404, Amp^{R}
NMRM96	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A212, Amp^{R}
NMRM97	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A231, Amp^{R}
NMRM98	BL21 [pBAD33]	BL21, Cml^{R}
NMRM99	BL21 DE3 [pBAD33]	BL21 DE3, Cml^{R}
NMRM100	C43 [pBAD33]	C43, Cml^{R}
NMRM101	RMA5113 [pBAD33-WzyE-His$_{10}$]	PE860 *ΔwzyE::frt*, Cml^{R}
NMRM102	RMA5113 [pBAD33-WzyE]	PE860 *ΔwzyE::frt*, Cml^{R}
NMRM103	RMA5113 [pBAD33-GFP-His8] [pWSK29]	PE860 *ΔwzyE::frt*, Cml^{R}, Amp^{R}
NMRM104	RMA5113 [pBAD33-WzyE-GFP-His$_{8}$] [pWSK29-WecG]	PE860 *ΔwzyE::frt*, Cml^{R}, Amp^{R}
NMRM105	RMA5113 [pBAD33-WzyE3XFLAG] [pWSK29]	PE860 *ΔwzyE::frt*, Cml^{R}, Amp^{R}
NMRM106	RMA5113 [pBAD33-WzyE3XFLAG] [pWSK29-WecG]	PE860 *ΔwzyE::frt*, Cml^{R}, Amp^{R}
NMRM107	DH5α [pBAD33-WzyE-His$_{12}$]	DH5α, Cml^{R}
NMRM108	RMA5113 [pBAD33-WzyE-His$_{12}$]	PE860 *ΔwzyE::fr*, Cml^{R}
NMRM109	RMA5113 [pBAD33-WzyE-WecG]	PE860 *ΔwzyE::fr*, Cml^{R}
NMRM110	DH5α [pBAD33-WzyE-WecG]	PE860 *ΔwzyE::fr*, Cml^{R}

NMRM111	RMA5113 [pBAD33-WzyE-GFP-His_8] [pWKS30-WecG]	PE860 *ΔwzyE::frt*, Cml^R, Amp^R
NMRM112	RMA5113 [pBAD33-WzyE-GFP-His_8] [pWKS30]	PE860 *ΔwzyE::frt*, Cml^R, Amp^R
NMRM113	RMA5113 [pBAD33-WzyE3XFLAG] [pWKS30-WecG]	PE860 *ΔwzyE::frt*, Cml^R, Amp^R
NMRM114	RMA5113 [pBAD33-WzyE3XFLAG] [pWKS30]	PE860 *ΔwzyE::frt*, Cml^R, Amp^R
NMRM115	DH5α [pWKS30-WecG]	DH5α, Amp^R
NMRM116	RMA5113 *Δ$wzyE_{10\text{-}440}$::cml^R*	PE860 *Δ$wzyE_{10\text{-}440}$*, Cml^R
NMRM117	BL21 [pBAD33-WzyE-His_{12}]	BL21, Cml^R
NMRM118	BL21 DE3 [pBAD33-WzyE-His_{12}]	BL21 DE3, Cml^R
NMRM119	C43 [pBAD33-WzyE-His_{12}]	C43, Cml^R
NMRM120	RMA2162 *Δ$wzyE_{10\text{-}440}$::frt*	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, (RMA5114)
NMRM121	RMA5114 [pBAD33]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM122	RMA5114 [pBAD33-WzyE3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM123	RMA5114 [pBAD33-WzyE-WecG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM124	RMA5114 [pBAD33-WzyE^{R204G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM125	RMA5114 [pBAD33-WzyE^{R247G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM126	RMA5114 [pBAD33-WzyE^{R266G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM127	RMA5114 [pBAD33-WzyE^{R295G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM128	RMA5114 [pBAD33-WzyE^{R309G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM129	RMA5114 [pBAD33-WzyE^{R399G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM130	RMA5114 [pBAD33-WzyE^{R408G}3XFLAG]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Cml^R
NMRM131	DH5α [pBAD33]	DH5α, Cml^R
NMRM132	DH5α [pBAD33-ScarS]	DH5α, Cml^R scar sequence from NMRM120
NMRM133	DH5α [pBAD33-ScarB]	DH5α, Cml^R scar sequence from NMRM120
NMRM134	DH5α [pWQ552-WecG-HA]	DH5α, Amp^R, Uses Ah-Tet inducible promoter
NMRM135	RMA2162 [pBAD33]	PE860, Cml^R
NMRM136	RMA2162 [pBAD33-ScarS]	PE860, Cml^R
NMRM137	RMA2162 [pBAD33-ScarB]	PE860, Cml^R
NMRM138	RMA5114 [pBAD33-WzyE3XFLAG] [pWQ552]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Amp^R, Cml^R
NMRM139	RMA5114 [pBAD33-WzyE3XFLAG] [pWQ552-WecG-HA]	PE860 *Δ$wzyE_{10\text{-}440}$::frt*, Amp^R, Cml^R
NMRM140	RMA2266 [pWQ552]	BW25113 *Δ$wecG::kan^R$*, Kan^R, Amp^R, Ah-Tet inducible
NMRM141	RMA2266 [pWQ552-WecG-HA]	BW25113 *Δ$wecG::kan^R$*, Kan^R, Amp^R, Ah-Tet inducible
NMRM142	DH5α [pBAD18-WzyE3XFLAG]	DH5α, Cml^R

NMRM143	RMA5114 [pBAD18]	PE860 *ΔwzyE$_{10-440}$::frt*, Cml^R
NMRM144	RMA5114 [pBAD18-WzyE3XFLAG]	PE860 *ΔwzyE$_{10-440}$::frt*, Cml^R
NMRM145	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at E40 (N2.17), Amp^R
NMRM146	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L57 (N14), Amp^R
NMRM147	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at N71 (N4), Amp^R
NMRM148	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R141 (N13), Amp^R
NMRM149	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V153 (N2.6), Amp^R
NMRM150	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F192 (2.9), Amp^R
NMRM151	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R221 (N2.14), Amp^R
NMRM152	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V238 (N2.18), Amp^R
NMRM153	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L18 (N7), Amp^R
NMRM154	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at E27 (N5), Amp^R
NMRM155	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G202 (N12), Amp^R
NMRM156	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at T116 (N10), Amp^R
NMRM157	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G136 (N1), Amp^R
NMRM158	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at P310 (N2.2), Amp^R
NMRM159	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W242 (N2.4), Amp^R
NMRM160	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Y145 (N2.7), Amp^R
NMRM161	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F241 (N2.8), Amp^R
NMRM162	RMA2266 [pWKS30]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R
NMRM163	DH5α [pWKS30-WecGHis$_{12}$]	DH5α, Amp^R
NMRM164	DH5α [pWQ552-WecGHis$_{12}$]	DH5α, Amp^R, Ah-Tet inducible
NMRM165	RMA2266 [pWKS30-WecGHis$_{12}$]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R
NMRM166	RMA2266 [pWQ552-WecGHis$_{12}$]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R
NMRM167	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30]	PE860 *ΔwzyE$_{10-440}$::frt*, Cml^R, Amp^R
NMRM168	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30-WecGHis$_{12}$]	PE860 *ΔwzyE$_{10-440}$::frt*, Cml^R, Amp^R
NMRM169	Not used	Not used
NMRM170	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F51 (3T68), Amp^R
NMRM171	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A78 (3T69), Amp^R

NMRM172	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L93 (3T70), Amp^{R}
NMRM173	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at P104 (3T71), Amp^{R}
NMRM174	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R110 (3T72), Amp^{R}
NMRM175	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V128 (3T73), Amp^{R}
NMRM176	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at M169 (3T76), Amp^{R}
NMRM177	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A276 (3T77), Amp^{R}
NMRM178	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F364 (3T79), Amp^{R}
NMRM179	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F388 (3T80), Amp^{R}
NMRM180	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F430 (3T82), Amp^{R}
NMRM181	RMA5115 [pBAD33-WzyE3XFLAG]	PE860 *ΔwzyE ΔwzzE*, Cml^{R}
NMRM182	RMA5115 [pQE30-WzzE]	PE860 *ΔwzyE ΔwzzE*, Amp^{R}
NMRM183	RMA5115 [pBAD33-WzyE3XFLAG] [pQE30-WzzE]	PE860 *ΔwzyE ΔwzzE*, Cml^{R}, Amp^{R}
NMRM184	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R96 (4T1), Amp^{R}
NMRM185	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F165 (4T2), Amp^{R}
NMRM186	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F216 (4T3), Amp^{R}
NMRM187	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W227 (4T4), Amp^{R}
NMRM188	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at S316 (4T5), Amp^{R}
NMRM189	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F320 (4T6), Amp^{R}
NMRM190	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at T336 (4T7), Amp^{R}
NMRM191	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V343 (4T8), Amp^{R}
NMRM192	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F349 (4T9), Amp^{R}
NMRM193	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Y368 (4T10), Amp^{R}
NMRM194	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A381 (4T11), Amp^{R}
NMRM195	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at H437 (4T12), Amp^{R}
NMRM196	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R176 (4T13), Amp^{R}
NMRM197	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V234 (2.2.21), Amp^{R}
NMRM198	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at I237 (2.2.10), Amp^{R}

NMRM199	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at T260 (2.2.75), Amp^R
NMRM200	RMA2266 [pWKS30-WecG]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R
NMRM201	RMA5114 [pWKS30-WecG]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R
NMRM202	RMA5114 [pBAD33-WzyE^{R160K}3XFLAG]	PE860 *ΔwzyE$_{10-440}$::frt*, Cml^R
NMRM203	RMA5114 [pBAD18-WzyE3XFLAG] [pWKS30-WecG]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R, Cml^R
NMRM204	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30-WecG]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R, Cml^R
NMRM205	RMA5113 [pBAD18-WzyE3XFLAG]	PE860 *ΔwzyE::frt*, Cml^R
NMRM206	RMA5113 [pBAD18]	PE860 *ΔwzyE::frt*, Cml^R
NMRM207	RMA5114 [pBAD33-WzyE3XFLAG] [pQE30]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R, Cml^R
NMRM208	RMA5114 [pBAD33-WzyE3XFLAG] [pQE30-WzzE]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R, Cml^R
NMRM209	RMA2266 [pBAD33] [pWKS30-WecG]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R, Cml^R
NMRM210	RMA2266 [pBAD33-WzyE3XFLAG] [pWKS30-WecG]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R, Cml^R
NMRM211	RMA2162 [pWKS30]	PE860, Amp^R
NMRM212	RMA2162 [pWKS30-WecG]	PE860, Amp^R
NMRM213	RMA2162 [pBAD33-WzyE3XFLAG]	PE860, Cml^R
NMRM214	RMA5114 [pWKS30]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R
NMRM215	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30]	PE860 *ΔwzyE$_{10-440}$::frt*, Amp^R, Cml^R
NMRM216	RMA2266 [pSUP203-WbbL]	BW25113, Amp^R
NMRM217	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecA::kan^R*, Kan^R, Amp^R
NMRM218	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecB::kan^R*, Kan^R, Amp^R
NMRM219	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecC::kan^R*, Kan^R, Amp^R
NMRM220	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecD::kan^R*, Kan^R, Amp^R
NMRM221	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecE::kan^R*, Kan^R, Amp^R
NMRM222	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecF::kan^R*, Kan^R, Amp^R
NMRM223	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwecG::kan^R*, Kan^R, Amp^R
NMRM224	RMA2266 [pSUP203-WbbL]	BW25113 *ΔrmlA::kan^R*, Kan^R, Amp^R
NMRM225	RMA2266 [pSUP203-WbbL]	BW25113 *ΔrmlB::kan^R*, Kan^R, Amp^R
NMRM226	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwzxE::kan^R*, Kan^R, Amp^R
NMRM227	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwzyE::kan^R*, Kan^R, Amp^R
NMRM228	RMA2266 [pSUP203-WbbL]	BW25113 *ΔwzzE::kan^R*, Kan^R, Amp^R
NMRM229	RMA5113 [pWSK29]	PE860 *ΔwzyE::frt*, Amp^R
NMRM230	RMA5113 [pWKS30]	PE860 *ΔwzyE::frt*, Amp^R
NMRM231	RMA5113 [pWKS30-WecG]	PE860 *ΔwzyE::frt*, Amp^R

NMRM232	DH5α [pWKS30-WecA-HA]	DH5α, Amp^R
NMRM233	DH5α [pWKS30-UppS-HA]	DH5α, Amp^R
NMRM234	RMA2162 [pWKS30-UppS-HA]	PE860, Amp^R
NMRM235	RMA2162 [pWKS30-WecA-HA]	PE860, Amp^R
NMRM236	RMA5113 [pWKS30-UppS-HA]	PE860 *ΔwzyE::frt*, Amp^R
NMRM237	RMA5113 [pWKS30-WecA-HA]	PE860 *ΔwzyE::frt*, Amp^R
NMRM238	RMA5114 [pWKS30-UppS-HA]	PE860 $\Delta wzyE_{10\text{-}440}$*::frt*, Amp^R
NMRM239	RMA5114 [pWKS30-WecA-HA]	PE860 $\Delta wzyE_{10\text{-}440}$*::frt*, Amp^R
NMRM240	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30-UppS-HA]	PE860 $\Delta wzyE_{10\text{-}440}$*::frt*, Cml^R, Amp^R
NMRM241	RMA5114 [pBAD33-WzyE3XFLAG] [pWKS30-WecA-HA]	PE860 $\Delta wzyE_{10\text{-}440}$*::frt*, Cml^R, Amp^R
NMRM242	RMA4199 [pWKS30-UppS-HA]	PE638, Amp^R
NMRM243	RMA4199 [pWKS30-WecA-HA]	PE638, Amp^R
NMRM244	RMA2608 [pWKS30-UppS-HA]	PE638 *ΔwzyB::frt*, Amp^R
NMRM245	RMA2608 [pWKS30-WecA-HA]	PE638 *ΔwzyB::frt*, Amp^R
NMRM246	RMA2608 [pBCKs+::*wzyB*-LacZα-FLAG] [pWKS30-UppS-HA]	PE638 *ΔwzyB::frt*, Cml^R, Amp^R
NMRM247	RMA2608 [pBCKs+::*wzyB*-LacZα-FLAG] [pWKS30-WecA-HA]	PE638 *ΔwzyB::frt*, Cml^R, Amp^R
NMRM248	DH5α [pBAD18-WecA-HA]	DH5α, Cml^R
NMRM249	DH5α [pBAD18-$WecA^{G173C}$-HA]	DH5α, Cml^R
NMRM250	DH5α [pBAD18-WecG]	DH5α, Cml^R
NMRM251	DH5α [pBAD18-UppS-HA]	DH5α, Cml^R
NMRM252	DH5α [pWKS30-His_{10}WecG]	DH5α, Amp^R
NMRM253	RMA2162 [pBAD18-UppS-HA]	PE860, Cml^R
NMRM254	RMA2162 [pWKS30-His_{10}WecG]	PE860, Amp^R
NMRM255	RMA2608 [pWKS30]	PE638 *ΔwzyB::frt*, Amp^R
NMRM256	RMA2608 [pWKS30-WecA-HA]	PE638 *ΔwzyB::frt*, Amp^R
NMRM257	RMA2608 [pWKS30-UppS-HA]	PE638 *ΔwzyB::frt*, Amp^R
NMRM258	RMA2266 [pWKS30-His_{10}WecG]	BW25113 *ΔwecG::kan*^{R}, Kan^R, Amp^R
NMRM259	RMA2266 [pBAD18-WecG]	BW25113 *ΔwecG::kan*^{R}, Kan^R, Cml^R
NMRM260	RMA2266 [pWKS30]	BW25113 *ΔwecA::kan*^{R}, Kan^R, Amp^R
NMRM261	RMA2266 [pWKS30-WecA-HA]	BW25113 *ΔwecA::kan*^{R}, Kan^R, Amp^R
NMRM262	RMA2266 [pBAD18-WecA-HA]	BW25113 *ΔwecA::kan*^{R}, Kan^R, Cml^R
NMRM263	RMA2266 [pBAD18]	BW25113 *ΔwecA::kan*^{R}, Kan^R, Cml^R
NMRM264	RMA2266 [pBAD18]	BW25113 *ΔwecG::kan*^{R}, Kan^R, Cml^R
NMRM265	RMA5108 [pBCKs+]	2457T Vp- *ΔwzyB::strep*^{R}, $Strep^R$, Cml^R (MYRM1287)

NMRM266	RMA5108 [pBCKs+::*wzyB*-LacZα-FLAG]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Cml^{R} (MYRM1287)
NMRM267	RMA5108 [pWKS30]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1287)
NMRM268	RMA5108 [pWKS30-WecA-HA]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1287)
NMRM269	RMA5108 [pWKS30-WecG]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1287)
NMRM270	RMA5108 [pWKS30-UppS-HA]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1287)
NMRM271	RMA5108 [pKD46]	2457T Vp- *ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1287) 30 ºC sensitive
NMRM272	RMA4258 *ΔwzyE::cml*^{R}	2457T Vp- *ΔwzyE::cml*^{R}, Cml^{R}
NMRM273	RMA4258 *Δ*$wzyE_{10\text{-}440}$*::cml*^{R}	2457T Vp- *Δ*$wzyE_{10\text{-}440}$*::cml*^{R}, Cml^{R}
NMRM274	RMA4258 *ΔwzyE::frt*	2457T Vp- *ΔwzyE::frt*, (RMA5109)
NMRM275	RMA4258 *Δ*$wzyE_{10\text{-}440}$*::frt*	2457T Vp- *Δ*$wzyE_{10\text{-}440}$*::frt*, (RMA5110)
NMRM276	RMA5107 [pWKS30]	2457T Vp- *ΔgtrII ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1034)
NMRM277	RMA5107 [pWKS30-WecA-HA]	2457T Vp- *ΔgtrII ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1034)
NMRM278	RMA5107 [pWKS30-WecG]	2457T Vp- *ΔgtrII ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1034)
NMRM279	RMA5107 [pWKS30-UppS-HA]	2457T Vp- *ΔgtrII ΔwzyB::strep*^{R}, Strep^{R}, Amp^{R} (MYRM1034)
NMRM280	DH5α [pWKS30-$His_{10}WecG^{\Delta A}$]	DH5α, Amp^{R}, WecG truncation missing CTD helices
NMRM281	DH5α [pWKS30-$His_{10}WecG^{\Delta II+III}$]	DH5α, Amp^{R}, WecG truncation missing CTD helices
NMRM282	DH5α [pWKS30-$His_{10}WecG^{\Delta III}$]	DH5α, Amp^{R}, WecG truncation missing CTD helices
NMRM283	RMA2266 [pWKS30-$His_{10}WecG^{\Delta A}$]	BW25113 *ΔwecG::kan*^{R}, Kan^{R}, Amp^{R}
NMRM284	RMA2266 [pWKS30-$His_{10}WecG^{\Delta II+III}$]	BW25113 *ΔwecG::kan*^{R}, Kan^{R}, Amp^{R}
NMRM285	RMA2266 [pWKS30-$His_{10}WecG^{\Delta III}$]	BW25113 *ΔwecG::kan*^{R}, Kan^{R}, Amp^{R}
NMRM286	RMA5113 [pWKS30-His_{10}WecG]	PE860 *ΔwzyE::frt*, Amp^{R}
NMRM287	RMA2162 [pWKS30-$His_{10}WecG^{\Delta A}$]	PE860, Amp^{R}
NMRM288	RMA2162 [pWKS30-$His_{10}WecG^{\Delta II+III}$]	PE860, Amp^{R}
NMRM289	RMA2162 [pWKS30-$His_{10}WecG^{\Delta III}$]	PE860, Amp^{R}
NMRM290	RMA5114 [pWKS30-His_{10}WecG]	PE860 *Δ*$wzyE_{10\text{-}440}$*::frt*, Amp^{R}
NMRM291	RMA5114 [pWKS30-$His_{10}WecG^{\Delta A}$]	PE860 *Δ*$wzyE_{10\text{-}440}$*::frt*, Amp^{R}
NMRM292	RMA5114 [pWKS30-$His_{10}WecG^{\Delta II+III}$]	PE860 *Δ*$wzyE_{10\text{-}440}$*::frt*, Amp^{R}
NMRM293	RMA5114 [pWKS30-$His_{10}WecG^{\Delta III}$]	PE860 *Δ*$wzyE_{10\text{-}440}$*::frt*, Amp^{R}
NMRM294	RMA5113 [pWKS30-$His_{10}WecG^{\Delta A}$]	PE860 *ΔwzyE::frt*, Amp^{R}
NMRM295	RMA5113 [pWKS30-$His_{10}WecG^{\Delta II+III}$]	PE860 *ΔwzyE::frt*, Amp^{R}
NMRM296	RMA5113 [pWKS30-$His_{10}WecG^{\Delta III}$]	PE860 *ΔwzyE::frt*, Amp^{R}

NMRM297	RMA5106 *ΔwzyE::cml*^{R}	2457T Vp- *ΔgtrII::strep*^{R} *ΔwzyE::cml*^{R}, Strep^{R}, Cml^{R} (MYRM1020)
NMRM298	RMA5106 *ΔwzyE*$_{10\text{-}440}$*::cml*^{R}	2457T Vp- *ΔgtrII::strep*^{R} *ΔwzyE*$_{10\text{-}440}$*::cml*^{R}, Strep^{R}, Cml^{R} (MYRM1020)
NMRM299	Not used	Not used
NMRM300	RMA5106 *ΔwzyE*$_{10\text{-}440}$*::frt*	2457T Vp- *ΔgtrII::strep*^{R} *ΔwzyE*$_{10\text{-}440}$*::frt*^{R}, Strep^{R} (MYRM1020)
NMRM301	RMA2266 [pWKS30]	BW25113 *ΔwecB::kan*^{R}, Kan^{R}, Amp^{R}
NMRM302	RMA2266 [pWKS30-His$_{10}$WecG]	BW25113 *ΔwecB::kan*^{R}, Kan^{R}, Amp^{R}
NMRM303	RMA2266 [pWKS30]	BW25113 *ΔwecC::kan*^{R}, Kan^{R}, Amp^{R}
NMRM304	RMA2266 [pWKS30-His$_{10}$WecG]	BW25113 *ΔwecC::kan*^{R}, Kan^{R}, Amp^{R}
NMRM305	RMA2266 [pWKS30]	BW25113, Amp^{R}
NMRM306	RMA2266 [pWKS30- His$_{10}$WecG]	BW25113, Amp^{R}
NMRM307	DH5α [pWKS30- His$_{10}$WecG^{L215E}]	DH5α, Amp^{R}
NMRM308	DH5α [pWKS30- His$_{10}$WecG^{L218E}]	DH5α, Amp^{R}
NMRM309	DH5α [pWKS30- His$_{10}$WecG^{L222E}]	DH5α, Amp^{R}
NMRM310	RMA2162 [pWKS30- His$_{10}$WecG^{L215E}]	PE860, Amp^{R}
NMRM311	RMA2162 [pWKS30- His$_{10}$WecG^{L218E}]	PE860, Amp^{R}
NMRM312	RMA2162 [pWKS30- His$_{10}$WecG^{L222E}]	PE860, Amp^{R}
NMRM313	RMA2266 [pWKS30- His$_{10}$WecG^{L215E}]	BW25113 *ΔwecG::kan*^{R}, Amp^{R}
NMRM314	RMA2266 [pWKS30- His$_{10}$WecG^{L218E}]	BW25113 *ΔwecG::kan*^{R}, Amp^{R}
NMRM315	RMA2266 [pWKS30- His$_{10}$WecG^{L222E}]	BW25113 *ΔwecG::kan*^{R}, Amp^{R}
NMRM316	RMA2162 *ΔwzyB::kan*^{R}	PE860 *ΔwzyB::kan*^{R}, Kan^{R} Clone#8
NMRM317	RMA2162 *ΔwzyB::kan*^{R}	PE860 *ΔwzyB::kan*^{R}, Kan^{R} Clone#10
NMRM318	DH5α [pNTR-WecA]	DH5α, Amp^{R}
NMRM319	DH5α [pNTR-WecC]	DH5α, Amp^{R}
NMRM320	DH5α [pNTR-WecD]	DH5α, Amp^{R}
NMRM321	DH5α [pNTR-WecE]	DH5α, Amp^{R}
NMRM322	DH5α [pNTR-WecF]	DH5α, Amp^{R}
NMRM323	DH5α [pNTR-WzxE]	DH5α, Amp^{R}
NMRM324	DH5α [pNTR-RmlA]	DH5α, Amp^{R}
NMRM325	DH5α [pNTR-RmlB]	DH5α, Amp^{R}
NMRM326	RMA2162 [pNTR-WecA]	PE860, Amp^{R}
NMRM327	RMA2162 [pNTR-WecC]	PE860, Amp^{R}
NMRM328	RMA2162 [pNTR-WecD]	PE860, Amp^{R}
NMRM329	RMA2162 [pNTR-WecE]	PE860, Amp^{R}
NMRM330	RMA2162 [pNTR-WecF]	PE860, Amp^{R}

NMRM331	RMA2162 [pNTR-WzxE]	PE860, Amp^R
NMRM332	RMA2162 [pNTR-RmlA]	PE860, Amp^R
NMRM333	RMA2162 [pNTR-RmlB]	PE860, Amp^R
NMRM334	RMA2162 [pNTR]	PE860, Amp^R
NMRM335	RMA2162 [pBAD33-WzyE3XFLAG] [pWKS30]	PE860, Cml^R, Amp^R
NMRM336	RMA2162 [pBAD33-WzyE3XFLAG] [pWKS30-WecG]	PE860, Cml^R, Amp^R
NMRM337	RMA2266 [pWKS30]	BW25113 *ΔwecA::kan*^{R}, Amp^R, Kan^R
NMRM338	RMA2266 [pWKS30-His_{10}WecG]	BW25113 *ΔwecA::kan*^{R}, Amp^R, Kan^R
NMRM339	RMA2162 *ΔwzyB::frt*	PE860 *ΔwzyB::frt*, (RMA5155)
NMRM340	RMA5155 [pBCKs+]	PE860 *ΔwzyB::frt*, Cml^R
NMRM341	RMA5155 [pBCKs+::*wzyB*-LacZα-FLAG]	PE860 *ΔwzyB::frt*, Cml^R
NMRM342	RMA5155 [pWKS30]	PE860 *ΔwzyB::frt*, Amp^R
NMRM343	RMA5155 [pWKS30-UppS-HA]	PE860 *ΔwzyB::frt*, Amp^R
NMRM344	RMA2162 *ΔwecA::cml*^{R}	PE860 *ΔwecA::cml*^{R}, Cml^R
NMRM345	RMA2162 *ΔwecA::frt*	PE860 *ΔwecA::frt*
NMRM346	NMRM345 [pWKS30-WecA-HA]	PE860 *ΔwecA::frt*, Amp^R
NMRM347	RMA2162 *ΔwecC::cml*^{R}	PE860 *ΔwecC::cml*^{R}, Cml^R
NMRM348	RMA2162 *ΔwecC::frt*	PE860 *ΔwecC::frt*
NMRM349	NMRM345 [pWKS30]	PE860 *ΔwecA::frt*, Amp^R
NMRM350	NMRM345 [pWKS30-WecA-HA]	PE860 *ΔwecA::frt*, Amp^R
NMRM351	NMRM345 [pWKS30-His_{10}WecG]	PE860 *ΔwecA::frt*, Amp^R
NMRM352	NMRM348 [pWKS30]	PE860 *ΔwecC::frt*, Amp^R
NMRM353	NMRM348 [pWKS30-His_{10}WecG]	PE860 *ΔwecC::frt*, Amp^R
NMRM354	NMRM348 [pNTR]	PE860 *ΔwecC::frt*, Amp^R
NMRM355	NMRM348 [pNTR-WecC]	PE860 *ΔwecC::frt*, Amp^R
NMRM356	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at T22 (2.2.59), Amp^R
NMRM357	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at F35 (2.2.12), Amp^R
NMRM358	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at F39 (2.2.60), Amp^R
NMRM359	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at L45 (2.2.37), Amp^R
NMRM360	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at S55 (2.2.4), Amp^R
NMRM361	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at I60 (2.2.1), Amp^R
NMRM362	DH5α [pPLEO1-$WzyE^{II}$]	DH5α Reporter fusion fused at F61 (2.2.5), Amp^R

NMRM363	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L69 (2.2.35), Amp^{R}
NMRM364	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F149 (2.2.41), Amp^{R}
NMRM365	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at S150 (2.2.9), Amp^{R}
NMRM366	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A157 (2T10), Amp^{R}
NMRM367	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Y163 (2T12), Amp^{R}
NMRM368	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W182, Amp^{R}
NMRM369	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L183 (2.2.86), Amp^{R}
NMRM370	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V187 (2.2.52), Amp^{R}
NMRM371	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G193 (2.2.2), Amp^{R}
NMRM372	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L195 (2.2.14), Amp^{R}
NMRM373	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L215 (4T3-2), Amp^{R}
NMRM374	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at I217 (2.2.17), Amp^{R}
NMRM375	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G218 (2.2.46), Amp^{R}
NMRM376	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at M240 (2.2.6), Amp^{R}
NMRM377	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A244 (2.2.8), Amp^{R}
NMRM378	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R247 (2.2.85), Amp^{R}
NMRM379	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at F269 (2.2.15), Amp^{R}
NMRM380	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W272 (2.2.80), Amp^{R}
NMRM381	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Q280 (2.2.68), Amp^{R}
NMRM382	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G289 (2.2.42), Amp^{R}
NMRM383	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R295 (2.2.82), Amp^{R}
NMRM384	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W306, Amp^{R}
NMRM385	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Q280 (2.2.68), Amp^{R}
NMRM386	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G289 (2.2.42), Amp^{R}
NMRM387	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R295 (2.2.82), Amp^{R}
NMRM388	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W306, Amp^{R}
NMRM389	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at Q280 (2.2.68), Amp^{R}

NMRM390	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G289 (2.2.42), Amp^{R}
NMRM391	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R295 (2.2.82), Amp^{R}
NMRM392	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at W306, Amp^{R}
NMRM393	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at S329 (2.2.34), Amp^{R}
NMRM394	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A332 (N2.13), Amp^{R}
NMRM395	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at A354 (N2.11), Amp^{R}
NMRM396	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V356 (2.2.24), Amp^{R}
NMRM397	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V357 (2.2.11), Amp^{R}
NMRM398	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L359 (2.2.36), Amp^{R}
NMRM399	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at E374 (2.2.71), Amp^{R}
NMRM400	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at M394 (2.2.26), Amp^{R}
NMRM401	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at L402 (2.2.40), Amp^{R}
NMRM402	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at R408, Amp^{R}
NMRM403	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at V414 (2.2.25), Amp^{R}
NMRM404	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at C419 (2.2.13), Amp^{R}
NMRM405	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at M421 (N2.5), Amp^{R}
NMRM406	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at G447 (2.2.93), Amp^{R}
NMRM407	DH5α [pPLEO1-WzyE^{II}]	DH5α Reporter fusion fused at E450 (N2.12), Amp^{R}

Appendix C – Plasmids

Construct ID	Description
Laboratory plasmids used in this thesis:	
pKD3	Chloramphenicol cassette for mutagenesis, (Datsenko & Wanner 2000)
pKD4	Kanamycin cassette for mutagenesis, (Datsenko & Wanner 2000)
pKD46	Lambda red recombinase construct, (Datsenko & Wanner 2000)
pCP20	Yeast recombinase recognising *frt* sequences, (Datsenko & Wanner 2000)
pWALDO-WzyE$_{EC}$-GFP-His$_8$	pWALDO expression construct encoding *wzyE$_{EC}$-gfp-his$_8$*, Kan^R, (laboratory stock)
pBAD33	Expression construct on P$_{BAD}$ promotor, Arabinose inducible, Cml^R, (Guzman et al. 1995)
pWKS30	Expression construct on T7 and T3 promotors, IPTG inducible, Amp^R, (Wang, RF & Kushner 1991)
pPLEO1	Expression construct expressing *phoA::lacZα*, IPTG inducible, Amp^R, (Taylor, V. L. et al. 2016)
pBCKs+	Expression construct derived from pBluescript II KS(+), Cml^R, (Stratagene)
pBCKs+ WzyB3XFLAG	pBCKs+ expression construct encoding *wzyB3xFLAG*, Cml^R, (laboratory stock)
Laboratory plasmids generated in this thesis:	
pNM1	pBAD33-WzyE-GFP-His$_8$, Cml^R
pNM2	pBAD33-WzyE-3XFLAG, used in *Chapter 3, 4, 6,* Cml^R
pNM3	pBAD33-WzyE^{R204G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM4	pBAD33-WzyE^{R247G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM5	pBAD33-WzyE^{R266G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM6	pBAD33-WzyE^{R295G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM7	pBAD33-WzyE^{R309G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM8	pBAD33-WzyE^{R399G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM9	pBAD33-WzyE^{R408G}-3XFLAG, arginine substitution library, used in *Chapter 3*, Cml^R
pNM10	pBAD33-WzyE^{R160K}-3XFLAG, arginine substitution library, Cml^R
pNM11	pBAD33-WzyE-His$_{10}$, Cml^R
pNM12	pBAD33-WzyE, Cml^R
pNM13	pBAD33-WzyE-His$_{12}$, Cml^R
pNM14	pBAD33-WzyE-WecG, Cml^R
pNM15	pBAD33-ScarS, pBAD33 expressing *frt* scar sequence from NMRM50, Cml^R
pNM16	pBAD33-ScarB, pBAD33 expressing *frt* scar sequence from NMRM120, Cml^R
pNM17	pBAD18-WzyE-3XFLAG, Cml^R
pNM18	pBAD18-WecA-HA, Cml^R
pNM19	pBAD18-WecA^{G173C}-HA, Cml^R
pNM20	pBAD18-WecG, Cml^R
pNM21	pBAD18-UppS-HA, Cml^R

pNM22	pWSK29-WecG, Amp^R
pNM23	pWKS30-WecG, Amp^R
pNM24	pWKS30-WecG-His_{12}, Amp^R
pNM25	pWKS30-His_{10}-WecG, used in *Chapter 5*, Amp^R
pNM26	pWKS30-His_{10}-$WecG^{\Delta A}$, WecG missing all terminal helices, used in *Chapter 5*, Amp^R
pNM27	pWKS30-His_{10}-$WecG^{\Delta II+III}$, WecG missing terminal helices II & III, used in *Chapter 5*, Amp^R
pNM28	pWKS30-His_{10}-$WecG^{\Delta III}$, WecG missing terminal helix III, used in *Chapter 5*, Amp^R
pNM29	pWKS30-His_{10}-$WecG^{L215E}$, used in *Chapter 5*, Amp^R
pNM30	pWKS30-His_{10}-$WecG^{L218E}$, used in *Chapter 5*, Amp^R
pNM31	pWKS30-His_{10}-$WecG^{L222E}$, used in *Chapter 5*, Amp^R
pNM32	pWKS30-WecA-HA, used in *Chapter 4*, Amp^R
pNM33	pWKS30-UppS-HA, used in *Chapter 4*, Amp^R
pNM34	pWQ522-WecG-HA, anhydrous tetracycline inducible, Amp^R
pNM35	pWQ522-WecG-His_{12}, anhydrous tetracycline inducible, Amp^R
pNM36	pPLEO1-$WzyE^{I}$, used in *Chapter 3*, Amp^R
pNM37	pPLEO1-$WzyE^{IIa1350t}$, used in *Chapter 3*, Amp^R

Appendix D – Oligonucleotides

ID	Sequence (5' – 3')	Description
Oligonucleotides from previous studies:		
NM1	caatcaactgtaagccacgcagcgtataggttggtgccgtggtgttgttattcattgatgggaattagccatggtcc	Mutagenesis of *wzyE*, complete ORF deletion
NM2	tatctacaaggctggcagcgggcgttggcgattgccgccagggaggtcgcgtgtaggctggagctgcttc	Mutagenesis of *wzyE*, complete ORF deletion
NM153	atatgattgtcggcggcactggcgccaatatc	Mutagenesis primer WzyE^{R204G}, Quik change kit
NM154	gataatggcgccagtgccgccgacaatcatat	Mutagenesis primer WzyE^{R204G}, Quik change kit
NM155	gcatgttctggctggcactaaaaggctatggaatga	Mutagenesis primer WzyE^{R247G}, Quik change kit
NM156	tcattccatagccttttagtgccagccagaacatgc	Mutagenesis primer WzyE^{R247G}, Quik change kit
NM157	tctatacgtttctctatctcactggcgacaccttct	Mutagenesis primer WzyE^{R266G}, Quik change kit
NM158	agaaggtgtcgccagtgagatagagaaacgtataga	Mutagenesis primer WzyE^{R266G}, Quik change kit
NM159	gcctggctccaattgttcggcgatttctatgtcttt	Mutagenesis primer WzyE^{R295G}, Quik change kit
NM160	aaagacatagaaatcgccgacaattggagccaggc	Mutagenesis primer WzyE^{R295G}, Quik change kit
NM161	ggctgtggccgggtggcccgagt	Mutagenesis primer WzyE^{R309G}, Quik change kit
NM162	actcgggccacccggccacagcc	Mutagenesis primer WzyE^{R309G}, Quik change kit
NM163	atcgtgctggcgggtgaagggctgg	Mutagenesis primer WzyE^{R399G}, Quik change kit
NM164	ccagcccttcacccgccagcacgat	Mutagenesis primer WzyE^{R399G}, Quik change kit
NM165	gggctggattcgtttgtctcaggcgtggtcttt	Mutagenesis primer WzyE^{R408G}, Quik change kit
NM166	aaagaccacgcctgagacaaacgaatccagccc	Mutagenesis primer WzyE^{R408G}, Quik change kit
VL70	gtaacacccgggttgtttaactttaagaaggagactcg	Binds before RBS of pWALDO

Oligonucleotides generated in this thesis:		
NM17.3	aatctggatcctccttcaacctgcgtccg	Adds *Xho*I-*wzyE*
NM18.3	tcaaactcgagatgagtctgctgcaattcagtg	Binds *wzyE* + *BamH*I
NM19	tgaaagcatgctcagtggtggtggtggt	Binds his_8 + *Sph*I
NM20	catttgcagtctcagtggtggtggtggtgatgatgatgtccttcaacctgcgtccg	Binds *wzyE* + 15bp-*Pst*I-*Stu*I
NM21	agtttggatcctccttcaacctgcgtccg	Adds *Pst*I-*wzyE*
NM22	ccagtaggcctctgcagaatctccgatggtatctcgagtccttcaacctgcgtccg	Binds *wzyE* + *BamH*I-15bp-*Pst*I-*Stu*I
NM23	tcaaacccgggatgagtctgctgcaattcagtg	Adds *Sma*I-*wzyE*
NM24	cagcccggttttccagaac	Binds 3'-5' on *phoA*
NM25	tcttcgcatgtccacttgtcatcgtcatccttgtagtcgatgtcatgatctttataatcaccgtcatggtctttgtagtctccttcaacctgcgtccgg	Binds *wzyE* + 3xFLAG-*Sph*I
NM26	aatcttctagatccttcaacctgcgtccg	Binds *wzyE* + *Xba*I
NM27	tcaaagagctcatgagtctgctgcaattcagtg	Adds *Sac*I-*wzyE*
NM28	gcgaggttgaagggtctaggactagtgg	Mutagenesis primer $wzyE^{a1350t}$, Quik change kit
NM29	ccactagttctagacccttcaacctgcg	Mutagenesis primer $wzyE^{a1350t}$, Quik change kit
NM30	gttctggaaaaccgggctgctc	Binds to start of *phoA::lacZα*
NM31	tacatcagcaacgcgtttgcgtag	Truncation primer WzyE V100, Inverse PCR
NM32	ggtaagccggaacagcaaaaagcc	Truncation primer WzyE N143, Inverse PCR
NM33	gaccgtgctgacgaggaaaaacag	Truncation primer WzyE V190, Inverse PCR
NM34	aggtgtcgcgagtgagatagagaaac	Truncation primer WzyE T268, Inverse PCR
NM35	cagtacttcccaggtaaagtaattggctg	Truncation primer WzyE L325, Inverse PCR
NM36	gcccagctcatacagccagtcg	Truncation primer WzyE G371, Inverse PCR
NM37	acggcgaaactcaaaccaggtc	Truncation primer WzyE R30, Inverse PCR
NM38	accaacctcaaagcgaaataccagcac	Truncation primer WzyE G64, Inverse PCR
NM39	cgcgtagaagcagcccgc	Truncation primer WzyE A83, Inverse PCR
NM40	gaagaagatgccgacgcttaccagc	Truncation primer WzyE F132, Inverse PCR
NM41	gtcctggcgcagaaagtagaccac	Truncation primer WzyE D178, Inverse PCR

NM42	agtgccgccgacaatcatataagtcag	Truncation primer WzyE T203, Inverse PCR
NM43	gccgcgaataatgccaataaacaggaag	Truncation primer WzyE G222, Inverse PCR
NM44	cattccatagcgttttaacgccagcc	Truncation primer WzyE M250, Inverse PCR
NM45	gttgtcgtagttctgcaacagcaacg	Truncation primer WzyE N284, Inverse PCR
NM46	tataagcgtaggcgagatcgccagtc	Truncation primer WzyE I338, Inverse PCR
NM47	tccggcgctttcaaaaagccagta	Truncation primer WzyE G444, Inverse PCR
NM48	tctgatcacccgttaaacggcgag	Binds 100 bp into *phoA* 3'-5' for sequencing
NM49	gagtgggatgaacaacgcgcc	Truncation primer WzyE L352, Inverse PCR
NM50	ggagatcgccccaaagcagaaactgt	Truncation primer WzyE F392, Inverse PCR
NM51	cgaatccagcccttcacgtgc	Truncation primer WzyE S404, Inverse PCR
NM52	agcgaatgcgatgatgatattggcg	Truncation primer WzyE A212, Inverse PCR
NM53	cgccgccagcatccaccac	Truncation primer WzyE A231, Inverse PCR
NM49.1	tgattgcatgcttatccttcaacctgcgtccg	Binds *wzyE* + *Sph*I
NM50.1	gcattgcatgctaaatgatgatgatgatgatgatgatgatgtccttcaacctgcgtccgg	Binds *wzyE* + his_{10}-*Sph*I
NM51.1	tcattctgcagtcagaggtggccggtgtagtg	Binds *wecG* + *Pst*I
NM52.1	gtaaactcgagtcgacttccagggcctggc	Adds *XhoI* 500 bp upstream of *wecG*
NM54	gacttgcatgcttaatgatgatgatgatgatgatgatgatgatgatgtccttcaacctgcgtccgga	Binds *wzyE* + his_{12}-*Sph*I
NM55	attgacccggggaccgcagcaatgagcatgtc	Adds *Sma*I 500 bp upstream of *wzyE*
NM56	gtagaaagcttcaccgccatccacataaaccag	Adds *Hind*III 500 bp downstream of *wecG*
NM57	tattgtgcgtgaagcgcagcg	Binds 120 bp upstream of *wzyE*
NM58	gattgccgccggggaggtcgcatgagtctgctgcaattcagtggcctgtttgtgtaggctggagctgcttc	Mutagenesis of *wzyE*, deletion leaving 30 bp at the 3' & 5' end
NM59	cgtggtggtggtattcattgttatccttcaacctgcgtccggagcgatgaatgggaattagccatggtcc	Mutagenesis of *wzyE*, deletion leaving 30 bp at the 3' & 5' end
NM60	attgacccgggacgtacaaaatcatcgctccgg	Binds before *wecG* RBS + *Sma*I
NM61	caaagctgcagtcaagcgtaatctggaacatcgtatgggtagaggttgccggtgtagtgc	Binds *wecG* before stop + HA-stop-*Pst*I
NM62	gttcctattctctagaaagtataggaacttcgaagcagctccagcctacacaaacaggccactgaattgcagcagactcatcgagctcgaattc gctagc	Binds *wecG* before stop + his_{12}-stop-*Pst*I

NM63	ttcggaataggaactaaggggatattcatatggaccatggctaattcccattcatcgctccggacgcaggttgaaggataacctctagagtcg acctgcagg	Binds 120 bp into *wecG* 3'-5'
NM64	tatctaaccggctggcagcgggcgttggcgattgccgccggggaggtcgctgtgtaggctggagctgcttc	Scarless mutant, amplifies cassette from pSIM6
NM65	acgcagcgtataggtcggtgccgtggtgttgttattcattgttatccttccctccttagttcctattccgaag	Scarless mutant, amplifies cassette from pSIM6
NM66	tatctacaaggctggcagcgggcgttggcgattgccgccggggaggtcgcatgagtctgctgcaattcagtgg	Scarless mutant, binds *wzyE* 5'
NM67	acgcagcgtataggtcggtgccgtggtgttgttattcattgttatccttctcacttgtcatcgtcatccttgtag	Scarless mutant, binds *wzyE* 3'
NM68	gaagccgaagaaaaaggtgagcaaaaacag	Truncation primer WzyE F51, Inverse PCR
NM69	cgcagaaagcaacacctgcaacaag	Truncation primer WzyE A78, Inverse PCR
NM70	taggcgggttttgtaggtgacatagtaaac	Truncation primer WzyE L93, Inverse PCR
NM71	cggacggcgcggtcaatcag	Truncation primer WzyE P104, Inverse PCR
NM72	gcggttcatggtaaacagcggac	Truncation primer WzyE R110, Inverse PCR
NM73	gacgcttaccagcgcgataccc	Truncation primer WzyE V128, Inverse PCR
NM74	cgccacgccggagacttca	Truncation primer WzyE A157, Inverse PCR
NM75	gtaaaagaagcgttttaacgccacgcc	Truncation primer WzyE Y163, Inverse PCR
NM76	catcgccgggatgaaaaagtaaaagaagc	Truncation primer WzyE M169, Inverse PCR
NM77	cgccagattctcccacggtgaa	Truncation primer WzyE A276, Inverse PCR
NM78	ccacagccaggaagggataaagacatag	Truncation primer WzyE W306, Inverse PCR
NM79	gaaccatttgatgatcagtccaaccacaatc	Truncation primer WzyE F364, Inverse PCR
NM80	aaagcagaaactgtgcaatatcgcagc	Truncation primer WzyE F388, Inverse PCR
NM81	gcgtgagacaaacgaatccagccc	Truncation primer WzyE R408, Inverse PCR
NM82	aaaaagccagtacaacagttttgcgatcatc	Truncation primer WzyE F430, Inverse PCR
NM83	gcgtttgcgtaggcgggttt	Truncation primer WzyE R96, Inverse PCR
NM84	gaaaaagtaaaagaagcgttttaacgccacgc	Truncation primer WzyE F165, Inverse PCR
NM85	aaacaggaagatagcgaatgcgatgatgata	Truncation primer WzyE F216, Inverse PCR
NM86	ccacaacgaaatccagccgcg	Truncation primer WzyE W227, Inverse PCR
NM87	tgagttcagcaccatactcgggc	Truncation primer WzyE S316, Inverse PCR

NM88	aaagtaattggctgagttcagcaccatact	Truncation primer WzyE F320, Inverse PCR
NM89	cgtaggcgagatcgccagtcc	Truncation primer WzyE T336, Inverse PCR
NM90	caccaccagtgagcctataagcgtag	Truncation primer WzyE V343, Inverse PCR
NM91	gaacaacgcgccgccca	Truncation primer WzyE F349, Inverse PCR
NM92	atacagccagtcgaaccatttgatgatcag	Truncation primer WzyE Y368, Inverse PCR
NM93	cgcagccttatagcgattagtctcgc	Truncation primer WzyE A381, Inverse PCR
NM94	atgaatgagtccggcgctttcaaaaag	Truncation primer WzyE H437, Inverse PCR
NM95	gcgcagaaagtagaccaccagca	Truncation primer WzyE R176, Inverse PCR
NM96	ggcgtggcgttaaaaaaattcttttacttttttc	Mutagenesis primer WzyE^{R160K}, Quik change kit
NM97	gaaaaagtaaaagaatttttttaacgccacgcc	Mutagenesis primer WzyE^{R160K}, Quik change kit
NM98	ggcgtggcgttaaaagcgttcttttacttttttc	Mutagenesis primer WzyE^{R160A}, Quik change kit
NM99	gaaaaagtaaaagaacgcttttaacgccacgcc	Mutagenesis primer WzyE^{R160A}, Quik change kit
NM100	cctggctccaattgtcaaagatttctatgtc	Mutagenesis primer WzyE^{R295K}, Quik change kit
NM101	gacatagaaatctttgacaattggagccagg	Mutagenesis primer WzyE^{R295K}, Quik change kit
NM102	gatattgcacagtttcgcgtttggggcgatc	Mutagenesis primer WzyE^{C387A}, Quik change kit
NM103	gatcgccccaaacgcgaaactgtgcaatatc	Mutagenesis primer WzyE^{C387A}, Quik change kit
NM104	gcccgagtatgtgcctgaactcagc	Mutagenesis primer WzyE^{V313C}, Quik change kit
NM105	gctgagttcaggcacatactcgggc	Mutagenesis primer WzyE^{V313C}, Quik change kit
NM106	ggactggcgatctgccctacgcttatag	Mutagenesis primer WzyE^{S334C}, Quik change kit
NM107	ctataagcgtagggcagatcgccagtcc	Mutagenesis primer WzyE^{S334C}, Quik change kit
NM108	gttgttcatctgcctcggggcgattg	Mutagenesis primer WzyE^{P351C}, Quik change kit
NM109	caatcgccccgaggcagatgaacaac	Mutagenesis primer WzyE^{P351C}, Quik change kit
NM110	gcgttcttttacttttttcatcccggcgatgc	Mutagenesis primer WzyE^{R160A}, Inverse PCR
NM111	ttttaacgccacgccggagacttc	Mutagenesis primer WzyE^{R160A}, Inverse PCR
NM112	aaagatttctatgtctttatcccttcctggctgtg	Mutagenesis primer WzyE^{R295K}, Inverse PCR
NM113	gacaattggagccaggccctg	Mutagenesis primer WzyE^{R295K}, Inverse PCR

NM114	tgcctgaactcagccaattactttacctgggaag	Mutagenesis primer WzyE^{V313C}, Inverse PCR
NM115	catactcgggcgccccg	Mutagenesis primer WzyE^{V313C}, Inverse PCR
NM116	tgccctacgcttataggctcactggtggt	Mutagenesis primer WzyE^{S334C}, Inverse PCR
NM117	gatcgccagtccggagtggtt	Mutagenesis primer WzyE^{S334C}, Inverse PCR
NM118	tgcctcggggcgattgtggttgga	Mutagenesis primer WzyE^{P351C}, Inverse PCR
NM119	gatgaacaacgcgccgccc	Mutagenesis primer WzyE^{P351C}, Inverse PCR
NM120	acctttggggcgatcttcaatatgatcgtgc	Mutagenesis primer WzyE^{C387T}, Inverse PCR
NM121	gaaactgtgcaatatcgcagccttatagcg	Mutagenesis primer WzyE^{C387T}, Inverse PCR
NM122	cgatagaattcgtagggcttcagtgatatagtctgcgcc	Binds 150 bp upstream *uppS* + *Eco*RI
NM123	attagtctagattcaagcgtaatctggaacatcgtatgggtaggctgtttcatcaccgggctc	Binds to *uppS* + HA-Stop-*Xba*I
NM124	gtactgaattcgtctgatgatcgcaaaactgttgtactggc	Binds 100 bp upstream *wecG* + *Eco*RI
NM125	taggttctagattcagaggttgccggtgtagtgc	Binds to *wecG* + HA-Stop-*Xba*I
NM126	gtgatgatggtggtgatgatggtgatgatgcattgttatccttcaacctgcgtccgg	Binds to 'ATG' of *wecG* + his_{10}
NM127	aataacaacaacacggcaccgacc	Binds to 3' end of *wecG*
NM128	ttagagaattcgggttcggaacggactttcccttc	Binds 82 bp upstream of *wecA* + *Eco*RI
NM129	attagtctagattcaagcgtaatctggaacatcgtatgggtatttggttaaattggggctgccacca	Binds *wecA* + HA-Stop-*Xba*I
NM130	tgactgcagcccggggg	Cloning *wecG* into pQE30
NM131	accggtgaaaacatcgtaagtcccg	Cloning *wecG* into pQE30
NM130.1	tgactgcagcccggggg	Binds *wecG* 3' including stop, Inverse PCR
NM131.1	accggtgaaaacatcgtaagtcccg	Binds *wecG* deleting all CTD helices, Inverse PCR
NM132	aatgcggctcggctgcg	Binds *wecG* deleting helix III, Inverse PCR
NM133	ttgccagattttcggtgcgcg	Binds *wecG* deleting helix II+III, Inverse PCR
NM134	ggttatacttctgctaataattttctctgagagcatgcattgtgaatttagtgtaggctggagctgcttc	Mutagenesis of *wecA*, deletion leaving 82 bp at the 5' end
NM135	tttcccaggcattggttgtgtcatcacatcctcatttatttggttaaattatgggaattagccatggtcc	Mutagenesis of *wecA*, deletion leaving 82 bp at the 5' end
NM136	tgttataaaaattttatttatatttttcatattcgtaaggtgatgttttgtgtaggctggagctgcttc	Mutagenesis of *wzyB*, deletion leaving 30 bp at the 3' & 5' end
NM137	agtaataacctcacttctggagcaaaataaaggatcttaaaaatagggaaatgggaattagccatggtcc	Mutagenesis of *wzyB*, deletion leaving 30 bp at the 3' & 5' end

NM138	attttaataaaaatctcacatgaaaatacttttaaagtatatcttagtgtgtgtaggctggagctgcttc	Mutagenesis of *wzyB*, deletion leaving 365 bp at the 3' end
NM139	gtattaccgtttctataactaatccatttctttttacacccatgggaattatgggaattagccatggtcc	Mutagenesis of *wzyB*, deletion leaving 134 bp at the 3' end
NM140	tcccagcgtttgccagattttcg	Mutagenesis primer WecG^{L215E}, Inverse PCR
NM141	gaggagtggctctaccgcctgct	Mutagenesis primer WecG^{L215E}, Inverse PCR
NM142	ccactccagtcccagcgtttg	Mutagenesis primer WecG^{L218E}, Inverse PCR
NM143	gagtaccgcctgctttcgcagc	Mutagenesis primer WecG^{L218E}, Inverse PCR
NM144	gcggtagagccactccagtcc	Mutagenesis primer WecG^{L221E}, Inverse PCR
NM145	gagctttcgcagccgagccgc	Mutagenesis primer WecG^{L221E}, Inverse PCR
NM146	caggcggtagagccactcca	Mutagenesis primer WecG^{L222E}, Inverse PCR
NM147	gagtcgcagccgagccgcattaa	Mutagenesis primer WecG^{L222E}, Inverse PCR
NM148	gcccataacccgtatggtgatggtc	Binds 100 bp upstream *wecC*
NM149	tgatataacgcaccagcgccgag	Binds 100 bp downstream *wecC*
NM150	gataccactgagcgtccggaagc	Binds 300 bp upstream *wecC*
NM151	aaaataatcggatatcactatgagttttgcgaccatttctgttatcggactgggttacatgtgtaggctggagctgcttc	Mutagenesis of *wecC*, deletion leaving 41 bp at the 3'& 5' end
NM152	ttttctcatcagcgccagactcctttggcatcgacgacatactgctgatatgggaattagccatggtcc	Mutagenesis of *wecC*, deletion leaving 41 bp at the 3'& 5' end

Appendix E – Thesis bibliography

Al-Dabbagh, B, Mengin-Lecreulx, D & Bouhss, A 2008, 'Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA', *J Bacteriol* 190:7141-7146.

Al-Dabbagh, B, Olatunji, S, Crouvoisier, M, El Ghachi, M, Blanot, D, Mengin-Lecreulx, D & Bouhss, A 2016, 'Catalytic mechanism of MraY and WecA, two paralogues of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily', *Biochimie* 127:249-257.

Albertson, TE, Panacek, EA, MacArthur, RD, Johnson, SB, Benjamin, E, Matuschak, GM, Zaloga, G, Maki, D, Silverstein, J, Tobias, JK, Haenftling, K, Black, G & Cowens, JW 2003, 'Multicenter evaluation of a human monoclonal antibody to Enterobacteriaceae common antigen in patients with Gram negative sepsis', *Crit Care Med*, 2:419-427.

Amar, A, Pezzoni, M, Pizarro, RA & Costa, CS 2018, 'New envelope stress factors involved in σ^E activation and conditional lethality of *rpoE* mutations in *Salmonella enterica'*, *Microbiology*, vol. 164:1293-1307.

Anderson, M & Raetz, CRH 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.

Baba, T, Ara, T, Hasegawa, M, Takai, Y, Okumura, Y, Baba, M, Datsenko, KA, Tomita, M, Wanner, BL & Mori, H 2006, 'Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection', *Mol Syst Biol* 2:2006.0008

Babinski, KJ, Ribeiro, AA & Raetz, CR 2002, 'The *Escherichia coli* gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis', *J Biol Chem* 277:25937-25946.

Bachmann, BJ 1972, 'Pedigrees of some mutant strains of *Escherichia coli* K-12', *Bacteriol Rev* 36:525-557.

Barr, K, Klena, J & Rick, PD 1999, 'The modality of enterobacterial common antigen polysaccharide chain lengths is regulated by o349 of the wec gene cluster of *Escherichia coli* K-12', *J Bacteriol* 181:6564-6568.

Barr, K, Ward, S, Meier-Dieter, U, Mayer, H & Rick, PD 1988, 'Characterization of an *Escherichia coli rff* mutant defective in transfer of N-acetylmannosaminuronic acid (ManNAcA) from UDP-ManNAcA to a lipid-linked intermediate involved in enterobacterial common antigen synthesis', *J Bacteriol* 170:228-233.

Barreteau, H, Kovač, A, Boniface, A, Sova, M, Gobec, S & Blanot, D 2008, 'Cytoplasmic steps of peptidoglycan biosynthesis', *FEMS Microbiol Rev* 32:168-207.

Bennish, MK 2012, 'Shigellosis', *Hunter's Tropical Med and Emerg Infect Dis*, 9:454-461.

Bi, Y, Mann, E, Whitfield, C & Zimmer, J 2018, 'Architecture of a channel-forming O-antigen polysaccharide ABC transporter', *Nature* 553:361.

Bohm, K, Porwollik, S, Chu, W, Dover, JA, Gilcrease, EB, Casjens, SR, McClelland, M & Parent, KN 2018, 'Genes affecting progression of bacteriophage P22 infection in *Salmonella* identified by transposon and single gene deletion screens', *Mol Microbiol* 108:288-305.

Bottoms, CA, Smith, PE & Tanner, JJ 2002, 'A structurally conserved water molecule in Rossmann dinucleotide-binding domains', *Prot Sci* 11:2125-2137.

Boumghar-Bourtchai, L, Mariani-Kurkdjian, P, Bingen, E, Filliol, I, Dhalluin, A, Ifrane, SA, Weill, F-X & Leclercq, R 2008, 'Macrolide-resistant *Shigella sonnei'*, *Emerg Infect Dis* 14:1297-1299.

Bridge, DR, Whitmire, JM, Gilbreath, JJ, Metcalf, ES & Merrell, DS 2015, 'An enterobacterial common antigen mutant of *Salmonella enterica serovar Typhimurium* as a vaccine candidate', *Int J Med Microbiol* 305:511-522.

Bridge, DR, Whitmire, JM, Makobongo, MO & Merrell, DS 2016, 'Heterologous *Pseudomonas aeruginosa* O-antigen delivery using a *Salmonella enterica serovar Typhimurium wecA* mutant strain', *Int J Med Microbiol* 306:529-540.

Brozek, KA & Raetz, CR 1990, 'Biosynthesis of lipid A in *Escherichia coli*. Acyl carrier protein-dependent incorporation of laurate and myristate', *J Biol Chem* 265:15410-15417.

Carter, JA, Jimenez, JC, Zaldivar, M, Alvarez, SA, Marolda, CL, Valvano, MA & Contreras, I 2009, 'The cellular level of O-antigen polymerase Wzy determines chain length regulation by WzzB and WzzpHS-2 in *Shigella flexneri* 2a', *Microbiology* 155:3260-3269.

Castelli, ME, Fedrigo, GV, Clementin, AL, Ielmini, MV, Feldman, MF & Garcia Vescovi, E 2008, 'Enterobacterial common antigen integrity is a checkpoint for flagellar biogenesis in *Serratia marcescens'*, *J Bacteriol* 190:213-220.

Castelli, ME & Vescovi, EG 2011, 'The Rcs signal transduction pathway is triggered by enterobacterial common antigen structure alterations in *Serratia marcescens'*, *J Bacteriol* 193:63-74.

CDC 2019, 'Antibiotic resistance threats in the United States'.

Chilton, P, Embry, C & Mitchell, T 2012, 'Effects of differences in lipid A structure on TLR4 pro-inflammatory signaling and inflammasome activation', *Frontiers in Immuno* 3:154.

Chung The, H, Rabaa, MA, Pham Thanh, D, De Lappe, N, Cormican, M, Valcanis, M, Howden, BP, Wangchuk, S, Bodhidatta, L, Mason, CJ, Nguyen Thi Nguyen, T, Vu Thuy, D, Thompson, CN, Phu Huong Lan, N, Voong Vinh, P, Ha Thanh, T, Turner, P, Sar, P, Thwaites, G, Thomson, NR, Holt, KE & Baker, S 2016, 'South asia as a reservoir for the global spread of ciprofloxacin-resistant *Shigella sonnei*: A cross-sectional study', *PLOS Med* e1002055.

Collins, RF, Kargas, V, Clarke, BR, Siebert, CA, Clare, DK, Bond, PJ, Whitfield, C & Ford, RC 2017, 'Full-length, oligomeric structure of Wzz determined by cryoelectron microscopy reveals insights into membrane-bound states', *Structure* 25:806-815.

Cramer, WA, Sharma, O & Zakharov, SD 2018, 'On mechanisms of colicin import: the outer membrane quandary', *Biochem* 475:3903-3915.

Cremers, CM, Knoefler, D, Vitvitsky, V, Banerjee, R & Jakob, U 2014, 'Bile salts act as effective protein-unfolding agents and instigators of disulfide stress *in vivo*', *Pros Nat Aca Sci* 111:e1610.

Cymer, F, von Heijne, G & White, SH 2015, 'Mechanisms of integral membrane protein insertion and folding', *J Mol Biol* 427:999-1022.

Danese, PN, Oliver, GR, Barr, K, Bowman, GD, Rick, PD & Silhavy, TJ 1998, 'Accumulation of the enterobacterial common antigen lipid II biosynthetic intermediate stimulates *degP* transcription in *Escherichia coli*, *J Bacteriol* 180:5875.

Daniels, C, Vindurampulle, C & Morona, R 1998, 'Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy)', *Mol Microbiol* 28:1211-1222.

Datsenko, KA & Wanner, BL 2000, 'One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products', *Proc Natl Acad Sci USA* 97:6640-6645.

Dobson, L, Reményi, I & Tusnády, GE 2015, 'CCTOP: a Consensus Constrained TOPology prediction web server', *Nucleic Acids Research* 43.

Drozdetskiy, A, Cole, C, Procter, J & Barton, GJ 2015, 'JPred4: a protein secondary structure prediction server', *Nucleic Acids Research* 43:389-394.

Eade, CR, Wallen, TW, Gates, CE, Oliverio, CL, Scarbrough, BA, Reid, AJ, Jorgenson, MA, Young, KD & Troutman, JM 2021, 'Making the enterobacterial common antigen glycan and measuring its substrate sequestration', *ACS Chem Biol* 16:691–700.

Ebel, W & Trempy, JE 1999, *'Escherichia coli* RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions to activate its own expression', *J Bacteriol* 181:577-584.

Egan, AJF, Errington, J & Vollmer, W 2020, 'Regulation of peptidoglycan synthesis and remodelling', *Nat Rev Microbiol* 18:446-460.

El Ghachi, M, Howe, N, Huang, CY, Olieric, V, Warshamanage, R, Touze, T, Weichert, D, Stansfeld, PJ, Wang, M, Kerff, F & Caffrey, M 2018, 'Crystal structure of undecaprenyl-pyrophosphate phosphatase and its role in peptidoglycan biosynthesis', *Nat Commun* 9:1078.

Elbein, AD, Gafford, J & Kang, MS 1979, 'Inhibition of lipid-linked saccharide synthesis: comparison of tunicamycin, streptovirudin, and antibiotic 24010', *Arch Biochem Biophys* 196:311-318.

Erbel, PJ, Barr, K, Gao, N, Gerwig, GJ, Rick, PD & Gardner, KH 2003, 'Identification and biosynthesis of cyclic enterobacterial common antigen in *Escherichia coli'*, *J Bacteriol* 185:1995-2004.

Evans, KL, Kannan, S, Li, G, de Pedro, MA & Young, KD 2013, 'Eliminating a set of four penicillin binding proteins triggers the Rcs phosphorelay and Cpx stress responses in *Escherichia coli'*, *J Bacteriol* 195:4415-4424.

Gentle, A, Ashton, PM, Dallman, TJ & Jenkins, C 2016, 'Evaluation of molecular methods for serotyping *Shigella flexneri'*, *J Clin Microbiol* 54:1456-1461.

Gérard, P, Vernet, T & Zapun, A 2002, 'Membrane topology of the *Streptococcus pneumoniae* FtsW division protein', *J Bacteriol* 184:1925-1931.

Gilbreath, JJ, Colvocoresses Dodds, J, Rick, PD, Soloski, MJ, Merrell, DS & Metcalf, ES 2012, 'Enterobacterial common antigen mutants of *Salmonella enterica serovar Typhimurium* establish a persistent infection and provide protection against subsequent lethal challenge', *Infect Immun*, 80:441-450.

Girgis, HS, Hottes, AK & Tavazoie, S 2009, 'Genetic Architecture of intrinsic antibiotic susceptibility', *PLoS ONE* 4:e5629.

Gozdziewicz, TK, Lugowski, C & Lukasiewicz, J 2014, 'First evidence for a covalent linkage between enterobacterial common antigen and lipopolysaccharide in *Shigella sonnei* phase II ECALPS', *J Biol Chem* 289:2745-2754.

Gozdziewicz, TK, Lukasiewicz, J & Lugowski, C 2015, The structure and significance of enterobacterial common antigen (ECA), *Postepy Hig Med Dosw* 69:1003-1012.

Gronow, S, Brabetz, W & Brade, H 2000, 'Comparative functional characterization *in vitro* of heptosyltransferase I (WaaC) and II (WaaF) from *Escherichia coli'*, *Eur J Biochem* 267:6602-6611.

Günther, SD, Fritsch, M, Seeger, JM, Schiffmann, LM, Snipas, SJ, Coutelle, M, Kufer, TA, Higgins, PG, Hornung, V, Bernardini, ML, Höning, S, Krönke, M, Salvesen, GS & Kashkar, H 2019, 'Cytosolic Gram negative bacteria prevent apoptosis by inhibition of effector caspases through lipopolysaccharide', *Nat Microbiol* 5:354-367.

Guzman, LM, Belin, D, Carson, MJ & 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, RW & Elbein, AD 1979, 'Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate GlcNAc-1-phosphate transferase', *Biochemistry* 18:2186-2192.

Hella-Monika Kuhn, UM-D, Hubert Mayer 1988, 'ECA, the enterobacterial common antigen', *FEMS Microbiol Rev* 54:195-222.

Heumann, D & Roger, T 2002, 'Initial responses to endotoxins and Gram negative bacteria', *Clin Chim Acta* 323:59-72.

Hobman, JL, Penn, CW & Pallen, MJ 2007, 'Laboratory strains of *Escherichia coli*: model citizens or deceitful delinquents growing old disgracefully?', *Mol Microbiol* 64:881-885.

Hong, M & Payne, SM 1997, 'Effect of mutations in *Shigella flexneri* chromosomal and plasmid-encoded lipopolysaccharide genes on invasion and serum resistance', *Mol Microbiol* 24:779-791.

Howe, V & Brown, AJ 2017, 'Determining the Topology of Membrane-Bound Proteins Using PEGylation', *Methods Mol Biol* 1583:201-210.

Howe, V, Chua, NK, Stevenson, J & Brown, AJ 2015, 'The regulatory domain of squalene monooxygenase contains a re-entrant loop and senses cholesterol via a conformational change', *J Biol Chem* 290:27533-27544.

Huesa, J, Giner-Lamia, J, Pucciarelli, MG, Paredes-Martínez, F, García-del Portillo, F, Marina, A & Casino, P 2021, 'Structure-based analyses of *Salmonella* RcsB variants unravel new features of the Rcs regulon', *Nucleic Acids Res* 49:2357-2374.

Hvorup, RN, Winnen, B, Chang, AB, Jiang, Y, Zhou, XF & Saier, MH, Jr. 2003, 'The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily', *Eur J Biochem* 270:799-813.

Igura, M, Maita, N, Kamishikiryo, J, Yamada, M, Obita, T, Maenaka, K & Kohda, D 2008, 'Structure-guided identification of a new catalytic motif of oligosaccharyltransferase', *EMBO* 27:234-243.

Inoue, H, Nojima, H & Okayama, H 1990, 'High efficiency transformation of *Escherichia coli* with plasmids', *Gene* 96:23-28.

Islam, ST, Gold, AC, Taylor, VL, Anderson, EM, Ford, RC & Lam, JS 2011, 'Dual conserved periplasmic loops possess essential charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in *Pseudomonas aeruginosa* PAO1', *J Biol Chem* 286:20600-20605.

Islam, ST, Huszczynski, SM, Nugent, T, Gold, AC & Lam, JS 2013, 'Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1', *Sci Rep* 3:3441.

Islam, ST & Lam, JS 2014, 'Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway', *Can J Microbiol* 11:697-716.

Islam, ST, Taylor, VL, Qi, M & Lam, JS 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:e00189-10.

Jiang, XE, Shrivastava, R, Tan, WB, San Seow, DC, Chen, SL, Guan, XL & Chng, S-S 2018, 'Biosynthetic intermediates of the enterobacterial common antigen overcome outer membrane lipid dyshomeostasis in *Escherichia coli*', *BioRxiv* 480533.

Jiang, XE, Tan, WB, Shrivastava, R, Seow, DCS, Chen, SL, Guan, XL & Chng, S-S 2020, 'Mutations in enterobacterial common antigen biosynthesis restore outer membrane barrier function in *Escherichia coli tol-pal*mutants', *Mol Microbiol* 114:991-1005.

Jorgenson, MA, Kannan, S, Laubacher, ME & Young, KD 2016, 'Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate', *Mol Microbio* 100:1-14.

Jorgenson, MA & Young, KD 2016, 'Interrupting biosynthesis of O-antigen or the lipopolysaccharide core produces morphological defects in *Escherichia coli* by sequestering undecaprenyl phosphate', *J Bacteriol* 198:3070-3079.

Jukič, M, Rožman, K, Sova, M, Barreteau, H & Gobec, S 2019, 'Anthranilic acid inhibitors of undecaprenyl pyrophosphate synthase (UppS), an essential enzyme for bacterial cell wall biosynthesis', *Frontiers in Microbiol* 9:3322.

Jumper, J, Evans, R, Pritzel, A, Green, T, Figurnov, M, Ronneberger, O, Tunyasuvunakool, K, Bates, R, Žídek, A, Potapenko, A, Bridgland, A, Meyer, C, Kohl, SAA, Ballard, AJ, Cowie, A, Romera-Paredes, B, Nikolov, S, Jain, R, Adler, J, Back, T, Petersen, S, Reiman, D, Clancy, E, Zielinski, M, Steinegger, M, Pacholska, M, Berghammer, T, Bodenstein, S, Silver, D, Vinyals, O, Senior, AW, Kavukcuoglu, K, Kohli, P & Hassabis, D 2021, 'Highly accurate protein structure prediction with AlphaFold', *Nature* 596:583-589.

Junkermeier, EH & Hengge, R 2021, 'A novel locally c-di-GMP-controlled exopolysaccharide synthase required for bacteriophage N4 infection of *Escherichia coli*', *MBio* 12:e0324921.

Kajimura, J, Rahman, A, Hsu, J, Evans, MR, Gardner, KH & Rick, PD 2006, 'O acetylation of the enterobacterial common antigen polysaccharide is catalyzed by the product of the *yiaH* gene of *Escherichia coli* K-12', *J Bacteriol* 188:7542-7550.

Kajimura, J, Rahman, A & Rick, PD 2005, 'Assembly of cyclic enterobacterial common antigen in *Escherichia coli* K-12', *J Bacteriol* 187:6917-6927.

Källberg, M, Wang, H, Wang, S, Peng, J, Wang, Z, Lu, H & Xu, J 2012, 'Template-based protein structure modeling using the RaptorX web server', *Nature Protocols* 7:1511.

Kalynych, S, Cherney, M, Bostina, M, Rouiller, I & Cygler, M 2015, 'Quaternary structure of WzzB and WzzE polysaccharide copolymerases', *Protein Sci* 24:58-69.

Kalynych, S, Morona, R & Cygler, M 2014, 'Progress in understanding the assembly process of bacterial O-antigen', *FEMS Microbiol Rev* 38:1048-1065.

Karimova, G & Ladant, D 2017, 'Defining membrane protein topology using *pho-lac* reporter fusions', in L Journet & E Cascales, *Bac Protein Secretion Systems: Methods and Protocols*, Springer 129-142.

Kattke, MD, Gosschalk, JE, Martinez, OE, Kumar, G, Gale, RT, Cascio, D, Sawaya, MR, Philips, M, Brown, ED & Clubb, RT 2019, 'Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria', *PLoS Pathog* 15:e1007723.

Kelly, JA, Moews, PC, Knox, JR, Frère, JM & Ghuysen, JM 1982, 'Penicillin target enzyme and the antibiotic binding site', *Science* 218:479-481.

Kiino, DR, Licudine, R, Wilt, K, Yang, DH & Rothman-Denes, LB 1993, 'A cytoplasmic protein, NfrC, is required for bacteriophage N4 adsorption', *J Bacteriol* 175:7074-7080.

Kim, TH, Sebastian, S, Pinkham, JT, Ross, RA, Blalock, LT & Kasper, DL 2010, 'Characterization of the O-antigen polymerase (Wzy) of *Francisella tularensis*', *J Biol Chem* 285:27839-27849.

Kim, YC, Tarr, AW & Penfold, CN 2014, 'Colicin import into *E. coli* cells: a model system for insights into the import mechanisms of bacteriocins', *Biochim Biophys Acta* 1843:1717-1731.

Klein, G, Müller-Loennies, S, Lindner, B, Kobylak, N, Brade, H & Raina, S 2013, 'Molecular and structural basis of inner core lipopolysaccharide alterations in *Escherichia coli*: Incorporation of glucuronic acid and phosphoethanolamine in the heptose region', *J Biol Chem* 288:8111-8127.

Klein, G & Raina, S 2015, 'Regulated control of the assembly and diversity of LPS by noncoding sRNAs', *Biomed Res Int* 2015:153561.

Klein, G & Raina, S 2019, 'Regulated assembly of LPS, its structural alterations and cellular response to LPS defects', *Int J Mol Sci* 20:356.

Klein, KA, Fukuto, HS, Pelletier, M, Romanov, G, Grabenstein, JP, Palmer, LE, Ernst, R & Bliska, JB 2012, 'A transposon site hybridization screen identifies *galU* and *wecBC* as important for survival of *Yersinia pestis* in murine macrophages', *J Bacteriol* 194:653-662.

Knirel, Y, Kondakova, A, Vinogradov, E, Lindner, B, Perepelov, A & Shashkov, A 2011, 'Lipopolysaccharide core structures and their correlation with genetic groupings of *Shigella* strains. A novel core variant in *Shigella boydii* type 16', *Glycobiology* 21:1362-1372.

Kotloff, KL, Riddle, MS, Platts-Mills, JA, Pavlinac, P & Zaidi, AKM 2018, 'Shigellosis', *Lancet* 391:801-812.

Kozlowski, LP 2016, 'IPC – Isoelectric Point Calculator', *Biology Direct* 11:55.

Krogh, A, Larsson, B, von Heijne, G & Sonnhammer, EL 2001, 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes', *J Mol Biol* 305:567-580.

Kuhn, HM, Basu, S & Mayer, H 1987, 'Comparison of enterobacterial common antigen from different species by serological techniques', *Eur J Biochem* 162:69-74.

Kuk, AC, Hao, A, Guan, Z & Lee, SY 2019, 'Visualizing conformation transitions of the Lipid II flippase MurJ', *Nat Commun* 10:1736.

Kuk, AC, Mashalidis, EH & Lee, SY 2017, 'Crystal structure of the MOP flippase MurJ in an inward-facing conformation', *Nat Struct Mol Biol* 24:171-176.

Kumar, S, Rubino, FA, Mendoza, AG & Ruiz, N 2019, 'The bacterial lipid II flippase MurJ functions by an alternating-access mechanism', *J Biol Chem* 294:981-990.

Lamothe, Sp 2013, 'PEGylation as a novel tool to investigate the topology of *Escherichia coli* WecA, a membrane enzyme involved in lipopolysaccharide O antigen initiation'.

Laubacher, ME & Ades, SE 2008, 'The Rcs phosphorelay is a cell envelope stress response activated by peptidoglycan stress and contributes to intrinsic antibiotic resistance', *J Bacteriol* 190:2065-2074.

Le Brun, AP, Clifton, LA, Halbert, CE, Lin, B, Meron, M, Holden, PJ, Lakey, JH & Holt, SA 2013, 'Structural characterization of a model Gram negative bacterial surface using lipopolysaccharides from rough strains of *Escherichia coli'*, *Biomacromolecules* 14:2014-2022.

Lee, S, Cho, E & Jung, S 2009, 'Periplasmic glucans isolated from Proteobacteria', *BMB Rep* 42:769-775.

Leo, V 2020, 'Polysaccharide co-polymerase WzzB/WzzE chimeras reveal transmembrane 2 region of WzzB is important for interaction with WzyB', *J Bacteriol* 203:e00598-20.

Leo, V, Teh, MY, Tran, ENH & Morona, R 2021, 'Identification of a region in *Shigella flexneri* WzyB disrupting the interaction with Wzz(pHS2)', *J Bacteriol* 203:e0041321.

Li, F, Egea, PF, Vecchio, AJ, Asial, I, Gupta, M, Paulino, J, Bajaj, R, Dickinson, MS, Ferguson-Miller, S, Monk, BC & Stroud, RM 2021, 'Highlighting membrane protein structure and function: A celebration of the Protein Data Bank', *J Biol Chem* 296:100557.

Liang, DM, Liu, JH, Wu, H, Wang, BB, Zhu, HJ & Qiao, JJ 2015, 'Glycosyltransferases: mechanisms and applications in natural product development', *Chem Soc Rev* 44:8350-8374.

Little, K, Tipping, MJ & Gibbs, KA 2018, 'Swarmer cell development of the bacterium *Proteus mirabilis* requires the conserved ECA biosynthesis gene, *rffG'*, *J Bacteriol* 200:e00230-18.

Liu, B, Knirel, YA, Feng, L, Perepelov, AV, Senchenkova, SN, Reeves, PR & Wang, L 2014, 'Structural diversity in *Salmonella* O antigens and its genetic basis', *FEMS Microbiol Rev* 38:56-89.

Liu, B, Knirel, YA, Feng, L, Perepelov, AV, Senchenkova, SyN, Wang, Q, Reeves, PR & Wang, L 2008, 'Structure and genetics of *Shigella* O-antigens', *FEMS Microbiol Rev* 32:627-653.

Liu, D, Cole, RA & Reeves, PR 1996, 'An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase', *J Bacteriol* 178:2102-2107.

Liu, D & Reeves, PR 1994, *'Escherichia coli* K12 regains its O antigen', *Microbiology* 140:49-57.

Liu, L, Zha, J, DiGiandomenico, A, McAllister, D, Stover, CK, Wang, Q & Boons, GJ 2015, 'Synthetic enterobacterial common antigen (ECA) for the development of a universal immunotherapy for drug-resistant *Enterobacteriaceae'*, *Angew Chem Int Ed Engl* 54:10953-10957.

Liu, MA, Morris, P & Reeves, PR 2019, 'Wzx flippases exhibiting complex O-unit preferences require a new model for Wzx–substrate interactions', *Microbiology* 8:e00655.

Liu, Q, Shen, X, Bian, X & Kong, Q 2020, 'Effect of deletion of gene cluster involved in synthesis of Enterobacterial common antigen on virulence and immunogenicity of live attenuated *Salmonella* vaccine when delivering heterologous *Streptococcus pneumoniae* antigen PspA', *BMC Microbiol* 20:150.

Lugtenberg, B, Meijers, J, Peters, R, van der Hoek, P & van Alphen, L 1975, 'Electrophoretic resolution of the ‘major outer membrane protein’ of *Escherichia coli* K12 into four bands', *FEBS Lett* 58:254-258.

Luo, Q, Yang, X, Yu, S, Shi, H, Wang, K, Xiao, L, Zhu, G, Sun, C, Li, T, Li, D, Zhang, X, Zhou, M & Huang, Y 2017, 'Structural basis for lipopolysaccharide extraction by ABC transporter LptB(2)FG', *Nat Struct Mol Biol* 24:469-474.

Ma´ire Begley, CGMG, Colin Hill 2004, 'The interaction between bacteria and bile', *ELSEVIER* 29:625-651.

Maciejewska, A, Kaszowska, M, Jachymek, W, Lugowski, C & Lukasiewicz, J 2020, 'Lipopolysaccharide-linked enterobacterial common antigen (ECA(LPS)) occurs in rough strains of *Escherichia coli* R1, R2, and R4', *Int J Mol Sci* 21.

Macpherson, DF, Manning, PA & Morona, R 1994, 'Characterization of the dTDP-rhamnose biosynthetic genes encoded in the *rfb* locus of *Shigella flexneri*', *Mol Microbiol* 11:281-292.

Macpherson, DF, Morona, R, Beger, DW, Cheah, KC & Manning, PA 1991, 'Genetic analysis of the *rfb* region of *Shigella flexneri* encoding the Y serotype O-antigen specificity', *Mol Microbiol*, 5:1491-1499.

Maczuga, N, Tran, ENH, Qin, J & Morona, R 'Interdependence *of Shigella flexneri* O antigen and enterobacterial common antigen biosynthetic pathways', *J Bacteriol* e00546-00521.

Maffei, E, Shaidullina, A, Burkolter, M, Druelle, V, Willi, L, Estermann, F, Michaelis, S, Hilbi, H, Thaler, DS & Harms, A 2021, 'Systematic exploration of *Escherichia coli* phage-host interactions with the BASEL phage collection', *BioRxiv* 434280.

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

Marolda, CL, Tatar, LD, Alaimo, C, Aebi, M & Valvano, MA 2006, 'Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide O-antigen', *J Bacteriol* 188:5124-5135.

Marolda, CL & Valvano, MA 1995, 'Genetic analysis of the dTDP-rhamnose biosynthesis region of the *Escherichia coli* VW187 (O7:K1) *rfb* gene cluster: identification of functional homologs of *rfbB* and *rfbA* in the *rff* cluster and correct location of the *rffE* gene', *J Bacteriol* 177:5539-5546.

Martinić, M, Hoare, A, Contreras, I & Alvarez, SA 2011, 'Contribution of the lipopolysaccharide to resistance of *Shigella flexneri* 2a to extreme acidity', *PLoS ONE* 6:e25557.

Mavris, M, Manning, PA & Morona, R 1997, 'Mechanism of bacteriophage SfII-mediated serotype conversion in *Shigella flexneri*', *Mol Microbiol* 26:939-950.

May, JF & Groisman, EA 2013, 'Conflicting roles for a cell surface modification in *Salmonella*', *Mol Microbiol* 88:970-983.

Mazur, A, Krol, JE, Marczak, M & Skorupska, A 2003, 'Membrane topology of PssT, the transmembrane protein component of the type I exopolysaccharide transport system in *Rhizobium leguminosarum* bv. *trifolii* strain TA1', *J Bacteriol* 185:2503-2511.

Meeske, AJ, Riley, EP, Robins, WP, Uehara, T, Mekalanos, JJ, Kahne, D, Walker, S, Kruse, AC, Bernhardt, TG & Rudner, DZ 2016, 'SEDS proteins are a widespread family of bacterial cell wall polymerases', *Nature* 537:634-638.

Meier-Dieter, U, Starman, R, Barr, K, Mayer, H & Rick, PD 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.

Meng, J, Xu, J, Huang, C & Chen, J 2020, 'Rcs phosphorelay responses to truncated lipopolysaccharide-induced cell envelope stress in *Yersinia enterocolitica'*, *Molecules* 25:5718.

Meng, J, Young, G & Chen, J 2021, 'The Rcs system in *Enterobacteriaceae*: Envelope stress responses and virulence regulation', *Frontiers in Microbiol* 12:627104.

Mestrom, L, Przypis, M, Kowalczykiewicz, Pollender, A, Kumpf, Marsden, Bento, I, Jarzębski, Szymańska, Chruściel, A, Tischler, D, Schoevaart, Hanefeld & Hagedoorn, P 2019, 'Leloir glycosyltransferases in applied biocatalysis: A multidisciplinary approach', *Int J Mol Sci* 20:5263.

Mitchell, AM, Srikumar, T & Silhavy, TJ 2018, 'Cyclic enterobacterial common antigen maintains the outer membrane permeability barrier of *Escherichia coli* in a manner controlled by YhdP', *MBio* 9:e02846-21.

Morgenstein, RM, Clemmer, KM & Rather, PN 2010, 'Loss of the *waaL* O-antigen ligase prevents surface activation of the flagellar gene cascade in *Proteus mirabilis'*, *J Bacteriol* 192:3213-3221.

Morgenstein, RM & Rather, PN 2012, 'Role of the Umo proteins and the Rcs phosphorelay in the swarming motility of the wild type and an O-antigen (*waaL*) mutant of *Proteus mirabilis'*, *J Bacteriol* 194:669-676.

Morona, R, Mavris, M, Fallarino, A & Manning, PA 1994, 'Characterization of the *rfc* region of *Shigella flexneri'*, *J Bacteriol* 176:733.

Morona, R & Van Den Bosch, L 2003, 'Lipopolysaccharide O antigen chains mask IcsA (VirG) in *Shigella flexneri'*, *FEMS Microbiol Lett* 221:173-180.

Morten Kallberg, GM, Sheng Wang, Jianzhu Ma, Jinbo Xu 2014, 'RaptorX server: A resource for template-basedn protein structure modeling', *Springer Sci* 1137.

Murray, GL, Attridge, SR & Morona, R 2003, 'Regulation of *Salmonella typhimurium* lipopolysaccharide O-antigen chain length is required for virulence; identification of FepE as a second Wzz', *Mol Microbiol* 47:1395-1406.

Muszynski, A, Rabsztyn, K, Knapska, K, Duda, KA, Duda-Grychtol, K, Kasperkiewicz, K, Radziejewska-Lebrecht, J, Holst, O & Skurnik, M 2013, 'Enterobacterial common antigen and O-specific polysaccharide coexist in the lipopolysaccharide of *Yersinia enterocolitica* serotype O:3', *Microbiology* 159:1782-1793.

Muthuirulandi Sethuvel, DP, Devanga Ragupathi, NK, Anandan, S & Veeraraghavan, B 2017, 'Update on: *Shigella* new serogroups/serotypes and their antimicrobial resistance', *Lett Appl Microbiol* 64:8-18.

Nath, P & Morona, R 2015a, 'Detection of Wzy/Wzz interaction in *Shigella flexneri'*, *Microbiology* 161:1797-1805.

Nath, P & Morona, R 2015b, 'Mutational analysis of the major periplasmic loops of *Shigella flexneri* Wzy: identification of the residues affecting O-antigen modal chain length control, and Wzz-dependent polymerization activity', *Microbiology* 161:774-785.

Nath, P, Tran, EN & Morona, R 2015, 'Mutational analysis of the *Shigella flexneri* O-antigen polymerase Wzy: identification of Wzz-dependent Wzy mutants', *J Bacteriol* 197:108-119.

Neelamegham, S, Aoki-Kinoshita, K, Bolton, E, Frank, M, Lisacek, F, Lütteke, T, O'Boyle, N, Packer, NH, Stanley, P, Toukach, P, Varki, A & Woods, RJ 2019, 'Updates to the symbol nomenclature for glycans guidelines', *Glycobiology* 29:620-624.

Nikolaidis, I, Favini-Stabile, S & Dessen, A 2014, 'Resistance to antibiotics targeted to the bacterial cell wall', *Protein Sci* 23:243-259.

Noszczynska, M, Kasperkiewicz, K, Duda, KA, Podhorodecka, J, Rabsztyn, K, Gwizdala, K, Swierzko, AS, Radziejewska-Lebrecht, J, Holst, O & Skurnik, M 2015, 'Serological characterization of the enterobacterial common antigen substitution of the lipopolysaccharide of *Yersinia enterocolitica* O : 3', *Microbiology* 161:219-227.

O'Toole, KH, Imperiali, B & Allen, KN 2021, 'Glycoconjugate pathway connections revealed by sequence similarity network analysis of the monotopic phosphoglycosyl transferases', *Pro Nat Aca Sci* 118:e2018289118.

Papadopoulos, M & Morona, R 2010, 'Mutagenesis and chemical cross-linking suggest that Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal chain length control', *J Bacteriol* 192:3385-3393.

Papadopoulos, M, Tran, ENH, Murray, GL & Morona, R 2016, 'Conserved transmembrane glycine residues in the *Shigella flexneri* polysaccharide co-polymerase protein WzzB influence protein-protein interactions', *Microbiology* 162:921-929.

Paradis-Bleau, C, Kritikos, G, Orlova, K, Typas, A & Bernhardt, TG 2014, 'A genome-wide screen for bacterial envelope biogenesis mutants identifies a novel factor involved in cell wall precursor metabolism', *PLoS Genetics* 10:e1004056-e1004056.

Park, JS, Lee, WC, Yeo, KJ, Ryu, KS, Kumarasiri, M, Hesek, D, Lee, M, Mobashery, S, Song, JH, Kim, SI, Lee, JC, Cheong, C, Jeon, YH & Kim, HY 2012, 'Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the Gram negative bacterial outer membrane', *Faseb j*, 26:219-228.

Pastoret, S, Fraipont, C, den Blaauwen, T, Wolf, B, Aarsman, MEG, Piette, A, Thomas, A, Brasseur, R & Nguyen-Distèche, M 2004, 'Functional analysis of the cell division protein FtsW of *Escherichia coli', J Bacteriol* 186:8370-8379.

Paunova-Krasteva, TS, Pavlova, VA, De Castro, C, Ivanova, RM, Molinaro, A, Nikolova, EB & Stoitsova, SR 2014, 'Cyclic enterobacterial common antigens from *Escherichia coli* O157 as microbe-associated molecular patterns', *Can J Microbiol* 60:173-176.

Perepelov, AV, Shekht, ME, Liu, B, Shevelev, SD, Ledov, VA, Senchenkova, SN, L'Vov V, L, Shashkov, AS, Feng, L, Aparin, PG, Wang, L & Knirel, YA 2012, '*Shigella flexneri* O-antigens revisited: final elucidation of the O-acetylation profiles and a survey of the O-antigen structure diversity', *FEMS Immunol Med Microbiol* 66:201-210.

Perez, C, Köhler, M, Janser, D, Pardon, E, Steyaert, J, Zenobi, R & Locher, KP 2017, 'Structural basis of inhibition of lipid-linked oligosaccharide flippase PglK by a conformational nanobody', *Scientific Reports* 7:46641.

Phan, M-D, Peters, KM, Sarkar, S, Lukowski, SW, Allsopp, LP, Moriel, DG, Achard, MES, Totsika, M, Marshall, VM, Upton, M, Beatson, SA & Schembri, MA 2013, 'The serum resistome of a globally disseminated multidrug resistant uropathogenic *Escherichia coli* clone', *PLoS Genetics* 9:e1003834.

Pluschke, G, Mayden, J, Achtman, M & Levine, RP 1983, 'Role of the capsule and the O-antigen in resistance of O18:K1 *Escherichia coli* to complement-mediated killing', *Infect Immun* 42:907-913.

Purins, L, Van Den Bosch, L, Richardson, V & Morona, R 2008, 'Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2', *Microbiology*, 154:1104-1116.

Qiao, J, Tan, X, Huang, D, Li, H, Wang, Z, Ren, H, Hu, X & Wang, X 2021, 'Construction and application of an *Escherichia coli* strain lacking 62 genes responsible for the biosynthesis of enterobacterial common antigen and flagella', *J Agric Food Chem* 87:e0038121.

Qiao, S, Luo, Q, Zhao, Y, Zhang, XC & Huang, Y 2014, 'Structural basis for lipopolysaccharide insertion in the bacterial outer membrane', *Nature* 511:108-111.

Raetz, C & Whitfield, C 2002, 'Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. Annu Rev Biochem 71: 635-700', *Annual Rev Biochem* 71:635-700.

Rai, AK, Carr, JF, Bautista, DE, Wang, W & Mitchell, AM 2021, 'ElyC and cyclic enterobacterial common antigen regulate synthesis of phosphoglyceride-linked enterobacterial common antigen', *MBio* 12:e0284621.

Rai, AK & Mitchell, AM 2020, 'Enterobacterial common antigen: Synthesis and function of an enigmatic molecule', *MBio* 11:e01914-20.

Raivio, TL, Popkin, DL & Silhavy, TJ 1999, 'The Cpx envelope stress response is controlled by amplification and feedback inhibition', *J Bacteriol* 181:5263-5272.

Ram, PK, Crump, JA, Gupta, SK, Miller, MA & Mintz, ED 2008, 'Part II. Analysis of data gaps pertaining to *Shigella* infections in low and medium human development index countries, 1984-2005', *Epidemiol Infect* 136:577-603.

Ramos-Morales, F, Prieto, AI, Beuzon, CR, Holden, DW & Casadesus, J 2003, 'Role for *Salmonella enterica* enterobacterial common antigen in bile resistance and virulence', *J Bacteriol*, 185:5328-5332.

Rapp, M, Drew, D, Daley, DO, Nilsson, J, Carvalho, T, Melén, K, De Gier, JW & Von Heijne, G 2004, 'Experimentally based topology models for *E. coli* inner membrane proteins', *Protein Sci*, 13:937-945.

Reeves, P 1995, 'Role of O-antigen variation in the immune response', *Trends Microbiol* 3:381-386.

Rick, PD, Barr, K, Sankaran, K, Kajimura, J, Rush, JS & Waechter, CJ 2003, 'Evidence that the *wzxE* gene of *Escherichia coli* K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen', *J Biol Chem* 278:16534-16542.

Rick, PD, Wolski, S, Barr, K, Ward, S & Ramsay-Sharer, L 1988, 'Accumulation of a lipid-linked intermediate involved in enterobacterial common antigen synthesis in *Salmonella typhimurium* mutants lacking dTDP-glucose pyrophosphorylase', *J Bacteriol* 170:4008-4014.

Rogers, HJ 1980, *Microbial cell walls and membranes / H.J. Rogers, H.R. Perkins, J.B. Ward*, eds HR Perkins & JB Ward.

Ruan, X, Loyola, DE, Marolda, CL, Perez-Donoso, JM & Valvano, MA 2012, 'The WaaL O-antigen lipopolysaccharide ligase has features in common with metal ion-independent inverting glycosyltransferases', *Glycobiology* 22:288-299.

Saidijam, M, Azizpour, S & Patching, SG 2018, 'Comprehensive analysis of the numbers, lengths and amino acid compositions of transmembrane helices in prokaryotic, eukaryotic and viral integral membrane proteins of high-resolution structure', *J Biomol Struct Dyn* 36:443-464.

Samsudin, F, Ortiz-Suarez, Maite L, Piggot, Thomas J, Bond, Peter J & Khalid, S 2016, 'OmpA: A flexible clamp for bacterial cell wall attachment', *Structure* 24:2227-2235.

Sarkar, P, Yarlagadda, V, Ghosh, C & Haldar, J 2017, 'A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics', *Med Chem Comm* 8:516-533.

Sellner, B, Prakapaitė, R, van Berkum, M, Heinemann, M, Harms, A & Jenal, U 2021, 'A new sugar for an old phage: a c-di-GMP-dependent polysaccharide pathway sensitizes *Escherichia coli* for bacteriophage infection', *MBio* 12:e0324621.

Sham, LT, Butler, EK, Lebar, MD, Kahne, D, Bernhardt, TG & Ruiz, N 2014, 'Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis', *Science* 345:220-222.

Sham, LT, Zheng, S, Yakhnina, AA, Kruse, AC & Bernhardt, TG 2018, 'Loss of specificity variants of WzxC suggest that substrate recognition is coupled with transporter opening in MOP-family flippases', *Mol Microbiol* 109:633-641.

Sherman, DJ, Lazarus, MB, Murphy, L, Liu, C, Walker, S, Ruiz, N & Kahne, D 2014, 'Decoupling catalytic activity from biological function of the ATPase that powers lipopolysaccharide transport', *Proc Natl Acad Sci USA* 111:4982-4987.

Sievers, F, Wilm, A, Dineen, D, Gibson, TJ, Karplus, K, Li, W, Lopez, R, McWilliam, H, Remmert, M, Söding, J, Thompson, JD & Higgins, DG 2011, 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Mol Syst Biol* 7:539.

Silhavy, TJ, Kahne, D & Walker, S 2010, 'The bacterial cell envelope', *Cold Spring Harb Perspect Biol* 2:a000414.

Sivapalasingam, S, Nelson, JM, Joyce, K, Hoekstra, M, Angulo, FJ & Mintz, ED 2006, 'High prevalence of antimicrobial resistance among *Shigella* isolates in the United States tested by the National Antimicrobial Resistance Monitoring System from 1999 to 2002', *Antimicrobial agents and chemotherapy* 50:49-54.

Sjodt, M, Brock, K, Dobihal, G, Rohs, PDA, Green, AG, Hopf, TA, Meeske, AJ, Srisuknimit, V, Kahne, D, Walker, S, Marks, DS, Bernhardt, TG, Rudner, DZ & Kruse, AC 2018, 'Structure of the peptidoglycan polymerase RodA resolved by evolutionary coupling analysis', *Nature* 556:118.

Smith, SM 2017, 'Strategies for the purification of membrane proteins', in D Walls & ST Loughran, *Methods and Protocols*, Springer 389-400.

Steimle, A, Autenrieth, IB & Frick, J-S 2016, 'Structure and function: Lipid A modifications in commensals and pathogens', *Int J Med Microbiol* 306:290-301.

Stenutz, R, Weintraub, A & Widmalm, G 2006, 'The structures of *Escherichia coli* O-polysaccharide antigens', *FEMS Microbiol Rev* 30:382-403.

Stevenson, G, Neal, B, Liu, D, Hobbs, M, Packer, NH, Batley, M, Redmond, JW, Lindquist, L & Reeves, P 1994, 'Structure of the O-antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster', *J Bacteriol* 176:4144-4156.

Suits, MD, Sperandeo, P, Dehò, G, Polissi, A & Jia, Z 2008, 'Novel structure of the conserved Gram negative lipopolysaccharide transport protein A and mutagenesis analysis', *J Mol Biol* 380:476-488.

Sun, Q, Knirel, YA, Wang, J, Luo, X, Senchenkova, SN, Lan, R, Shashkov, AS & Xu, J 2014, 'Serotype-converting bacteriophage SfII encodes an acyltransferase protein that mediates 6-O-acetylation of GlcNAc in *Shigella flexneri* O-antigens, conferring on the host a novel O-antigen epitope', *J Bacteriol* 196:3656-3666.

Swoboda, JG, Campbell, J, Meredith, TC & Walker, S 2010, 'Wall teichoic acid function, biosynthesis, and inhibition', *Chem Biochem* 11:35-45.

Taguchi, A, Welsh, MA, Marmont, LS, Lee, W, Sjodt, M, Kruse, AC, Kahne, D, Bernhardt, TG & Walker, S 2019, 'FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein', *Nat Microbiol* 4:587-594.

Talaat, KR, Alaimo, C, Martin, P, Bourgeois, AL, Dreyer, AM, Kaminski, RW, Porter, CK, Chakraborty, S, Clarkson, KA, Brubaker, J, Elwood, D, Frölich, R, DeNearing, B, Weerts, H, Feijoo, BL, Halpern, J, Sack, D, Riddle, MS & Fonck, VG 2021, 'Human challenge study with a *Shigella* bioconjugate vaccine: Analyses of clinical efficacy and correlate of protection', *EBio Med* 66:103310.

Tamae, C, Liu, A, Kim, K, Sitz, D, Hong, J, Becket, E, Bui, A, Solaimani, P, Tran Katherine, P, Yang, H & Miller Jeffrey, H 2008, 'Determination of antibiotic hypersensitivity among 4,000 single-gene-knockout mutants of *Escherichia coli*', *J Bacteriol* 190:5981-5988.

Tao, K, Narita, S & Tokuda, H 2012, 'Defective lipoprotein sorting induces lolA expression through the Rcs stress response phosphorelay system', *J Bacteriol* 194:3643-3650.

Tatar, LD, Marolda, CL, Polischuk, AN, van Leeuwen, D & Valvano, MA 2007, 'An *Escherichia coli* undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling', *Microbiology* 153:2518-2529.

Tavares-Carreon, F, Ruan, X, Ford, A & Valvano, MA 2019, 'Sulfhydryl labeling as a tool to investigate the topology of membrane proteins involved in lipopolysaccharide biosynthesis', *Methods Mol Biol* 1954: 203-213.

Taylor, VL, Hoage, JF, Thrane, SW, Huszczynski, SM, Jelsbak, L & Lam, JS 2016, 'A bacteriophage-acquired O-antigen polymerase (Wzyβ) from *P. aeruginosa* serotype O16 performs a varied mechanism compared to its cognate Wzyα, *Frontiers Microbiol* 7:393.

Teh, MY, Furevi, A, Widmalm, G & Morona, R 2020, 'Influence of *Shigella flexneri* 2a O-antigen acetylation on its bacteriophage sf6 receptor activity and bacterial interaction with human cells', *J Bacteriol* 202:e00363-20.

Tocilj, A, Munger, C, Proteau, A, Morona, R, Purins, L, Ajamian, E, Wagner, J, Papadopoulos, M, Van Den Bosch, L, Rubinstein, JL, Fethiere, J, Matte, A & Cygler, M 2008, 'Bacterial polysaccharide co-polymerases share a common framework for control of polymer length', *Nat Struct Mol Biol* 15:130-138.

TouzÉ, T & Mengin-Lecreulx, D 2008, 'Undecaprenyl Phosphate Synthesis', *EcoSal Plus* 3:1.

Tran, Day, CJ, McCartney, E, Poole, J, Tse, E, Jennings, MP & Morona, R 2020, *'Shigella flexneri* targets human colonic goblet cells by O-antigen binding to sialyl-Tn and Tn antigens via glycan-glycan interactions', *ACS Infect Dis* 6:2604-2615.

Tran, AX, Dong, C & Whitfield, C 2010, 'Structure and functional analysis of LptC, a conserved membrane protein involved in the lipopolysaccharide export pathway in *Escherichia coli', J Biol Chem* 285:33529-33539.

Tran, EN, Papadopoulos, M & Morona, R 2014, 'Relationship between O-antigen chain length and resistance to colicin E2 in *Shigella flexneri', Microbiology* 160:589-601.

Tran, ENH, Doyle, MT & Morona, R 2013, 'LPS Unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles', *PLoS ONE* 8:e70508.

Trent, MS, Ribeiro, AA, Lin, S, Cotter, RJ & Raetz, CR 2001, 'An inner membrane enzyme in *Salmonella* and *Escherichia coli* that transfers 4-amino-4-deoxy-L-arabinose to lipid A: induction on polymyxin-resistant mutants and role of a novel lipid-linked donor', *J Biol Chem* 276:43122-43131.

Tsai, CM & Frasch, CE 1982, 'A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels', *Anal Biochem* 119:115-119.

Typas, A, Banzhaf, M, Gross, CA & Vollmer, W 2011, 'From the regulation of peptidoglycan synthesis to bacterial growth and morphology', *Nat Rev Microbiol* 10:123-136.

Valtonen, MV, Larinkari, UM, Plosila, M, Valtonen, VV & Mäkelä, PH 1976, 'Effect of enterobacterial common antigen on mouse virulence of *Salmonella typhimurium', Infec Immun* 13:1601-1605.

Varki, A, Cummings, RD, Aebi, M, Packer, NH, Seeberger, PH, Esko, JD, Stanley, P, Hart, G, Darvill, A, Kinoshita, T, Prestegard, JJ, Schnaar, RL, Freeze, HH, Marth, JD, Bertozzi, CR, Etzler, ME, Frank, M, Vliegenthart, JF, Lütteke, T, Perez, S, Bolton, E, Rudd, P, Paulson, J, Kanehisa, M, Toukach, P, Aoki-Kinoshita, KF, Dell, A, Narimatsu, H, York, W, Taniguchi, N & Kornfeld, S 2015, 'Symbol nomenclature for graphical representations of glycans', *Glycobiology* 25:1323-1324.

Vollmer, W, Blanot, D & de Pedro, MA 2008, 'Peptidoglycan structure and architecture', *FEMS Microbiol Rev* 32:149-167.

Wall, E, Majdalani, N & Gottesman, S 2018, 'The Complex Rcs Regulatory Cascade', *Annu Rev Microbiol* 72:111-139.

Wang, L, Wang, Q & Reeves, PR 2010, 'The variation of O-antigens in gram-negative bacteria', *Subcell Biochem* 53:123-152.

Wang, RF & Kushner, SR 1991, 'Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*', *Gene* 100:195-199.

Wang, Z, Fan, F, Wang, J, Wang, L, Hu, H, Wang, C & Wang, X 2021, 'Engineering *Escherichia coli* to produce *Bordetella pertussis* oligosaccharide with multiple trisaccharide units', *Metab Eng* 69:147-162.

Waterhouse, AM, Procter, JB, Martin, DM, Clamp, M & Barton, GJ 2009, 'Jalview Version 2--a multiple sequence alignment editor and analysis workbench', *Bioinformatics* 25:1189-1191.

Whitfield, C, Kaniuk, N & Frirdich, 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 & Trent, MS 2014, 'Biosynthesis and export of bacterial lipopolysaccharides', *Annu Rev Biochem* 83:99-128.

Whitfield, C, Wear, SS & Sande, C 2020, 'Assembly of bacterial capsular polysaccharides and exopolysaccharides', *Annu Rev Microbiol* 74:521-543.

Wiseman, B, Nitharwal, RG, Widmalm, G & Högbom, M 2021, 'Structure of a full-length bacterial polysaccharide co-polymerase', *Nat Commun* 12:369.

Woodward, R, Yi, W, Li, L, Zhao, G, Eguchi, H, Sridhar, PR, Guo, H, Song, JK, Motari, E, Cai, L, Kelleher, P, Liu, X, Han, W, Zhang, W, Ding, Y, Li, M & Wang, PG 2010, '*In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz', *Nat Chem Biol* 6:418-423.

Workman, SD & Strynadka, NCJ 2020, 'A slippery scaffold: synthesis and recycling of the bacterial cell wall carrier lipid', *J Mol Biol* 432:4964-4982.

Yang, J, Yan, R, Roy, A, Xu, D, Poisson, J & Zhang, Y 2015, 'The I-TASSER Suite: protein structure and function prediction', *Nat Methods* 12:7-8.

Yates, LE, Natarajan, A, Li, M, Hale, M, Mills, DC & DeLisa, MP 2019, 'Glyco-recoded *Escherichia coli*: Recombineering-based genome editing of native polysaccharide biosynthesis gene clusters', *Metab Eng* 53:59-68.

Yethon, JA, Heinrichs, DE, Monteiro, MA, Perry, MB & Whitfield, C 1998, 'Involvement of *waaY*, *waaQ*, and *waaP* in the modification of *Escherichia coli* lipopolysaccharide and their role in the formation of a stable outer membrane', *J Biol Chem* 273:26310-26316.

York, A, Lloyd, AJ, del Genio, CI, Shearer, J, Hinxman, KJ, Fritz, K, Fulop, V, Dowson, CG, Khalid, S & Roper, DI 2021, 'Structure-based modeling and dynamics of MurM, a *Streptococcus pneumoniae* penicillin resistance determinant present at the cytoplasmic membrane', *Structure* 29:731-742.

Yuasa, R, Levinthal, M & Nikaido, H 1969, 'Biosynthesis of cell wall lipopolysaccharide in mutants of *Salmonella*. V. A mutant of *Salmonella typhimurium* defective in the synthesis of cytidine diphosphoabequose', *J Bacteriol* 100:433-444.

Zhang, H, Zhu, F, Yang, T, Ding, L, Zhou, M, Li, J, Haslam, SM, Dell, A, Erlandsen, H & Wu, H 2014, 'The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold', *Nat Commun* 5:4339.

Zheng, S, Sham, L-T, Rubino, FA, Brock, KP, Robins, WP, Mekalanos, JJ, Marks, DS, Bernhardt, TG & Kruse, AC 2018, 'Structure and mutagenic analysis of the lipid II flippase MurJ from *Escherichia coli*, *Pro Nat Aca Sci* 115:6709-6714.

Zhou, Z, White, KA, Polissi, A, Georgopoulos, C & Raetz, CR 1998, 'Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis', *J Biol Chem* 273:12466-12475.