



MONASH University



The architectural significance of Gram-negative periplasm: Size matters



Mr Eric Joseph Loleh Mandela

MSc

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in February 2020
Faculty of Medicine, Nursing and Health Sciences
Department of Microbiology

Copyright notice

© Eric Mandela (2020).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Abstract

The evolution of the outer membrane (OM) a few billion years ago, introduced a new compartment, the periplasm, between the two membrane bilayers. These diderm bacteria (Gram-negatives) faced challenges to their cell biology. The periplasm now has machineries essential for protein transport, folding and quality control; for biogenesis of a membrane distant from the cellular lipid biosynthetic enzymes; to make and remodel a peptidoglycan sacculus that segregates its two membranes; and to build and energize protein secretion machines and drug-efflux pumps that span the periplasm with components integrated into both the OM and inner membrane (IM).

Working on the hypothesis that these various transport and biogenesis systems evolved to function in a periplasm of a specified size, I artificially manipulated the periplasmic size to understand its architectural significance in Gram-negative bacterial cell biology. To accomplish this, a key player in the cell envelope organisation was investigated; a small, alpha-helical protein called Braun's lipoprotein (Lpp). Lpp provides the only known covalent connection between the OM and the peptidoglycan sacculus (PG). The N-terminus of Lpp is integrated into the OM while the C-terminus is attached to the PG.

To comprehensively investigate periplasmic systems that require strict control of the periplasmic size, a genetic screening approach was used. The approach involved screening for essentiality of *E. coli* genes, to identify those that are only essential in an enlarged periplasm. The successful screen exploited a synthetic lethal interactions principle. Here I report the effect of a widened periplasm on both PG biosynthesis, maturation, turnover, and remodelling, and on biogenesis of the outer leaflet of the OM. The significance of non-covalent OM-PG bridging was also revealed, by the observed essentiality of otherwise non-essential proteins (OmpA and Pal) known to non-covalently assist OM-PG linkage, and identification of two other potentially critical PG-binding OM proteins (TolC and YiaD). This data was interpreted to imply that the cell maintains local areas of close contact between the OM and IM.

Further characterisation of the strain with an enlarged periplasm displayed few distinctive phenotypes relative to the wild-type, suggesting that the strain adapts to growth with an increase in periplasmic size. To address the mechanism of this adaptation, I employed transcriptomic and proteomic profiling approaches to study the response at the transcriptional and post-transcriptional level, respectively. The results revealed a dynamic adaptation response network not only confined in the periplasm but involving the whole cell, observed

predominantly in changes to protein steady-state levels, suggesting post-transcriptional control is critical to adaptation. Consistent with the screen results, proteostasis observed within the periplasm pointed to maintaining optimal PG assembly and LPS biogenesis.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature:

Print Name: Eric Mandela

Date: 29th February 2020

Preface

Publication during enrolment

Chen, L., Wilksch, J.J., Liu, H., Zhang, X., Torres, V.V., Bi, W., **Mandela, E.**, Cao, J., Li, J., Lithgow, T. and Zhou, T., (2020). Investigation of LuxS-mediated quorum sensing in *Klebsiella pneumoniae*. *Journal of Medical Microbiology*, p.jmm 001148.

Experiments conducted by others for this work

- Dr Von Torres (University of Queensland, Australia) analysed the distribution of Lpp length, performed multiple sequence alignment for the Longer Lpp variants in comparison with *E. coli* Lpp, and did the Phyre2 structural predictions of the longer Lpp variants - Chapter 4.0, Figure 4.2.2 and Appendix 3.
- Dr Enzo Cheng Huang (Monash Proteomics and Metabolomics Facility, Monash University, Australia) performed liquid chromatography-tandem mass spectrometry (LC-MS/MS), and the initial analysis to determine the quality of the data set and to identify proteins - Chapter 4.0, Section 5.3 (Appendix 5.0)
- Dr Chaille Webb (Monash University, Australia) performed the conserved domain analysis for OmpA, Pal and YiaD – Chapter 4.0, Figure 4.4.2.2.
- Dr Iain Hay (University of Auckland, New Zealand) performed a comparative analysis to determine relative proteome changes in different cellular compartments in regards to volume changes – Chapter 4.0, Figure 4.3.4.1
- Dr Kirill Tsyganov (Monash Bioinformatics platform, Monash University, Australia) performed the bioinformatics analysis of the RNA-seq sequencing data to produce the raw genes count matrix and performed the differential gene expression analysis between the wild-type and Lpp₊₂₁ *E. coli* – Chapter 5.0, Section 5.2.1

All other experimentation comprises my original work.

Acknowledgements

Foremost, I express my sincere gratitude to my supervisor **Professor Trevor Lithgow** for giving me an opportunity to join one of the best research group in the world. Trevor is the best you can ever get, his sincerity, diligence, dynamism, vision and reliability have deeply inspired me. He offered me the privilege to design my own projects, an opportunity I will forever be grateful for, as it shaped me to an excellent scientist I have become. It has been a great privilege and honour to work and study under his guidance.

To my second supervisor **Dr Iain Hay**, who has a substance of a genius; he always provided a solution to every problem, and when the going got tough he always had an arsenal of new ideas. I admired his calmness and patient even when nothing seemed to work-out; always a source of motivation to keep going. I express my sincere appreciation for his unconditional support and advice he gave throughout my candidature even when he left to start his new research group.

I would like to recognise the invaluable assistance provided by my PhD Committee: **Professor Dena Lyras, Professor Ana Traven**, and **Dr Celeste Donato**. Their helpful scientific discussions and constructive criticism during my milestone meetings were instrumental in shaping my project. I am grateful for their support over the years.

While moving from Kenya to Australia, I was hoping to find a new home in terms of a good working team. Lithgow laboratory members did not disappoint. Including past and present members, I sincerely show my gratitude to **Dr Takuya Shiota, Dr Dilshan Gunasinghe, Dr Jonathan Wilksch, Ms Rebecca Bamert, Ms Kher-Shing Tan, Dr Chaille Webb, Dr Grishma Vadlamani, Dr Rhys Dunstan, Ms Natalia Rosas, Dr Jiawei Wang, Ms Manasa Bharathwaj, Mr Murray White, Dr Matthew Belousoff, and Dr Rhys Grinter**. I am grateful to have worked alongside this team of dynamic scientists with different areas of expertise. I could not have asked for a better team.

My special regards go to **Dr Andrea Rocker, Dr Von Torres** and **Dr Chris Stubenrauch** for taking up a supervisory role that was unsigned. I whole-heartedly appreciate their great advice during those one-on-one side discussions that proved monumental towards the success of this study. Andrea and Von were also very supportive during the initial drafting of this thesis.

I am indebted to the friends I made during my candidature, members from our neighbouring research groups: **Traven**, **Naderer**, **Kwok**, **Li**, and **Boyce** laboratories. I am grateful for their support and for making my PhD journey one of the most memorable experience ever.

Finally, to someone special in my life and to whom, I dedicate this thesis; **Ms Blandina Katumbi Mbevi** (my mom). She has sacrificed everything to educate me, believes in me when I have self-doubt and always finds a way to bring out the best in me. She is an inspiration, my role model and my best friend and I love her for being more than a mother.

The scholarship funding by Monash Biomedicine and Discovery Institute (BDI) is truly appreciated. Without their support and funding, this project could not have reached its goal. In addition, thanks to Lithgow laboratory for funding me during the thesis writing period.

Table of contents

Copyright notice.....	i
Abstract	ii
Declaration.....	iv
Preface.....	v
Acknowledgements.....	vi
Table of contents.....	viii
List of Figures	xiii
List of Tables	xv
Abbreviations.....	xvi
Chapter 1: Cell envelope biogenesis.....	1
1.1 Introduction.....	1
1.2 Cell wall architecture	3
1.2.1 The cell wall.....	3
1.2.2 Models for peptidoglycan architecture	5
1.2.3 Cell wall thickness	6
1.2.4 Cell wall elasticity.....	8
1.2.5 Cell turgor and permeability properties of the cell wall	8
1.3 Peptidoglycan biogenesis, regulation, and function	11
1.3.1 Peptidoglycan core synthesis	11
1.3.1.1 The Cytoplasmic steps: Synthesis of UDP-MurNAc-pentapeptide	11
1.3.1.2 The IM steps: Assembly of lipid I and II.....	12
1.3.1.3 The periplasmic steps: Polymerisation and cross-linking.....	14
1.3.2 Regulation of peptidoglycan synthesis in <i>E. coli</i> by outer membrane lipoproteins	14
1.4 Lipoproteins: structure, function, and biosynthesis	15
1.4.1 Lipoprotein biogenesis.....	17
1.4.2 Lipoprotein export	17
1.4.3 Lpp; Major outer membrane lipoprotein.....	19
1.4.4 Lpp and the peptidoglycan.....	20
1.4.5 Role of Lpp and other proteins in peptidoglycan attachment	22
1.4.6 Peptidoglycan bound Lpp is a distance keeper	25
1.5 Lipopolysaccharide biogenesis	27
1.5.1 Lipopolysaccharide structure and function.....	27

1.5.2 Lipopolysaccharide biosynthesis	29
1.5.3 Molecular mechanism of LPS translocation to the outer membrane	31
1.6 Outer membrane protein biogenesis	33
1.6.1 Outer membrane protein translocation across the inner membrane and transport across the periplasm	33
1.6.2 The β -barrel assembly machinery	34
1.6.3 BAM lipoproteins	38
1.7 Genome-scale genetic manipulation for exploring bacterial biology	39
1.7.1 Gene function investigation by mutation combinations	40
1.7.2 Advances in double recombineering high-throughput methods	42
1.8 Thesis rationale and aims	45
Chapter 2: Material and Methods	47
2.1 Bacterial strains, plasmids, and standard growth conditions	47
2.2 Transformation techniques	50
2.2.1 Transformation and storage solution method.	50
2.2.2 Preparation of CaCl_2 chemically competent cells and transformation protocol	50
2.2.3 Preparing electrocompetent cells and electroporation	51
2.3 Bacterial storage and recovery	51
2.4 DNA/RNA based techniques	52
2.4.1 Plasmid extraction	52
2.4.2 Polymerase chain reaction (PCR)	52
2.4.3 Restriction digestion and alkaline phosphatase treatment	55
2.4.4 DNA ligation	55
2.4.5 DNA sequencing	55
2.4.6 Total RNA isolation	55
2.4.7 RNA data analysis	56
2.5 High-density mutagenesis, screen, and selection	56
2.5.1 Hfr Cavalli gene knock-out	56
2.5.2 Arraying Keio collection for screening	57
2.5.3 High-density generation of double recombinants	58
2.5.3.1 Single mutant arraying	58
2.5.3.2 Conjugation	58
2.5.3.3 Selection of double recombinants.	58
2.5.3.4 Imaging and storage	58

2.5.4 High throughput data analysis	60
2.5.5 Verification of identified synthetic lethal/sick mutants	60
2.6 Protein-based techniques	60
2.6.1 Polyacrylamide gel electrophoresis	60
2.6.2 Coomassie staining and destaining of SDS-PAGE gels.....	61
2.6.3 Western transfer and immunoblotting	61
2.6.4 Membrane isolation from <i>E. coli</i>	63
2.6.5 Quantitative proteomics	63
2.6.5.1 Sample preparation	63
2.6.5.2 Reduction, alkylation	64
2.6.5.3 Chloroform/methanol precipitation.....	64
2.6.5.4 Trypsination and peptide clean-up.....	64
2.6.5.5 Mass spectrometric acquisition.....	65
2.6.5.6 Proteomics data analysis	65
2.7 Cell-based techniques	65
2.7.1 Automated growth analysis.....	65
2.7.2 Antibiotic susceptibility testing and SDS sensitivity analysis.....	66
2.7.3 Construction of Pet fusions	66
2.7.4 Growth in starch; an assay for amylase translocation.....	66
2.7.5 Lugol's iodine starch staining assay	67
2.7.6 Live/Dead cell viability assay	67
2.8 Lipopolysaccharide quantification by silver staining	67
2.9 Computer-based methods.....	68

Chapter 3: A high-throughput screen to investigate autotransporter

biogenesis	70
3.1 Introduction.....	70
3.2. Set-up and optimisation for the high-throughput generation of double-gene knock-outs for synthetic lethal screen application.....	72
3.2.1 Keio collection rearrangement.....	72
3.2.2 High-throughput generation of double-gene deletions – a pilot study	74
3.2.3 Optimised high-throughput generation of double-gene recombinants	77
3.3 Development of a high-throughput autotransporter biogenesis assay	80
3.3.1 Candidate reporter proteins for assessment of Pet biogenesis	81
3.3.2 Engineering <i>E. coli</i> to use starch as the sole carbon source.....	85

3.3.3 Starch growth assay	85
3.3.4 Optimisation of the starch-growth assay for a high-throughput screen	87
3.4 Discussion	89
Chapter 4: The architectural significance of Gram-negative periplasm; an Lpp₊₂₁ synthetic lethal screen.....	92
4.1 Introduction.....	92
4.2 Cloning, expression of Lpp ₊₂₁ and Lpp sequence analysis	92
4.3 Phenotypic characterisation of Lpp ₊₂₁	95
4.3.1 Lpp ₊₂₁ has no growth defect	97
4.3.2 Lpp ₊₂₁ displays defects in outer membrane integrity	99
4.3.3 DTT sensitivity, a protein folding aspect on a widened periplasm.....	103
4.3.4 Changes in protein amounts in the sub-cellular compartments of Lpp ₊₂₁	103
4.3.5. Lipopolysaccharide biogenesis in an enlarged periplasm.....	107
4.4 Lpp ₊₂₁ synthetic lethal screen	109
4.4.1 High-throughput gene replacement.....	109
4.4.2 Results analysis and verification.....	111
4.4.2.1 Synthetic lethality of PBPs and their cognate lipoprotein activators in the Lpp ₊₂₁ strain.....	119
4.4.3 Supplementation of amino acids rescues the synthetic lethality displayed by PG assembly proteins in Lpp ₊₂₁	120
4.5 Discussion	121
Chapter 5: Transcriptomic and proteomic profiling to investigate the adaptive response to an enlarged periplasm	127
5.1 Introduction.....	127
5.2 Gene regulation in response to an enlarged periplasm	128
5.2.1 Transcriptional response to an enlarged periplasm.....	129
5.2.1.1 Upregulation of genes in response to an enlarged periplasm.....	130
5.2.1.2 Genes down-regulated in response to an enlarged periplasm	132
5.3 Proteostasis in an enlarged periplasm	134
5.3.1 Proteomics results; data quality analysis	134
5.3.2 Changes to the proteome in an enlarged periplasm	135
5.3.2.1 Changes to the proteome in an enlarged periplasm in hypo-osmotic medium	137
5.3.2.2 Changes to the proteome in an enlarged periplasm in iso-osmotic medium.....	140

5.4 Rational design to investigate Lpp ₊₂₁ synthetic lethality of over-expressed proteins .	147
5.5 Discussion	149
References	155
Appendices	175
Appendix 1.0.....	175
Appendix 2.0.....	176
Appendix 3.0.....	182
Appendix 4.0.....	183
Appendix 5.0.....	186

List of Figures

Figure 1.1.1: Schematic of a Gram-negative cell envelope.	2
Figure 1.2.1.1: Peptidoglycan cross-linking and variations of the peptide stem.	4
Figure 1.2.3.1: Peptidoglycan thickness.	7
Figure 1.2.4.1: Cell wall elasticity.	9
Figure 1.2.5.1: The role played by osmolytes in periplasmic volume changes.	10
Figure 1.3.1.1 Peptidoglycan biosynthesis.....	13
Figure 1.3.2.1: Activation of penicillin-binding proteins by lipoproteins.	16
Figure 1.4.1.1: Lipoprotein biogenesis and export, exemplified with Lpp.	18
Figure 1.4.4.1: Lpp linkage to the outer membrane and peptidoglycan.....	21
Figure 1.4.5.1: Major and minor peptidoglycan-membrane connections.	23
Figure 1.4.5.2: OmpA and Pal are structural homologues.....	24
Figure 1.4.6.1: Peptidoglycan-bound Lpp is a distance keeper.	26
Figure 1.5.2.1: Lipopolysaccharide biosynthesis in <i>E. coli</i>	30
Figure 1.5.3.1: PEZ Model for lipopolysaccharide translocation across the periplasm.....	32
Figure 1.6.1.1: Outer membrane biogenesis.	35
Figure 1.6.2.1: Structures of the <i>E. coli</i> BAM complex components and interaction map.....	37
Figure: 1.7.1.1: Synthetic lethality and multiplicative principles.	41
Figure 1.7.2.1: Hfr conjugation gene transfer system.	44
Figure 2.5.3.1. Strategy for the construction of “double recombinants”.	59
Figure 3.2.1.1. Robotic-genetic selection screen.	73
Figure 3.2.2.1: High-density generation of double-gene knock-outs; 1536 and 384-density format; a pilot study.	75
Figure 3.2.3.1: Optimised high-throughput synthetic lethal screen.....	78
Figure 3.2.3.2: High-throughput gene knock-out in 384-density format.....	79
Figure 3.3.1.1: A schematic diagram of Pet fusion constructs.....	82
Figure 3.3.1.2: Pet- β -barrel production; an assessment of heterologous protein production.....	84
Figure 3.3.3.1: Starch growth assay in 384-colony density.	86
Figure 3.3.4.1: Lugol’s iodine staining, a starch breakdown assay in <i>E. coli</i> (96-colony density format).	88
Figure 4.2.1: Cloning and expression of Lpp ₊₂₁	94
Figure 4.2.2: Lpp sequence analysis.	96
Figure 4.3.1.1: Comparative growth profiles of <i>E. coli</i> JW5028 and Lpp ₊₂₁ strain.....	98
Figure 4.3.2.1: Comparative SDS sensitivity profiles of JW5028 and Lpp ₊₂₁ strains.	100
Figure 4.3.2.2: Comparative total membrane profile of JW5028 and Lpp ₊₂₁ strains on osmotically stabilised media.....	101

Figure 4.3.2.3: Single-cell assessment of membrane integrity using live/dead staining assay.	102
Figure 4.3.3.1: Sensitivity of Lpp ₊₂₁ to DTT.	104
Figure 4.3.4.1 Relative proteome changes in different cellular compartments.	106
Figure 4.3.5.1: LPS quantification by silver staining of <i>E. coli</i> cell lysate and total membrane extracts.	108
Figure 4.4.1.1: A schematic flowchart illustrating the process for high-throughput replacement of <i>lpp</i> with <i>lpp</i> ₊₂₁	110
Figure 4.4.1.2: 384 high-density <i>lpp</i> : <i>lpp</i> ₊₂₁ gene replacement result.	112
Figure 4.4.2.1: High-throughput verification and colony size quantification of synthetic sick/lethal double recombinants.	114
Figure 4.4.2.2: A predicted role of YiaD as part of the OM-PG bridge.	117
Figure 4.4.2.3: The AcrABZ-TolC complex is predicted to play a structural role as an OM-PG bridge.	118
Figure 4.4.2.1.1: Synthetic lethality of major penicillin-binding proteins 1A and 1B, and their cognate lipoprotein activators to Lpp ₊₂₁	119
Figure 4.5.1: Schematic representation of the implication of periplasmic architecture on the interaction of PBP1A/B with LpoA/B.	123
Figure 5.2.1.1. RNA-seq sample quality analysis.	131
Figure 5.3.1.1: Data quality plots.	136
Figure 5.3.2.1.1: Changes to proteome in an enlarged periplasm; Lpp ₊₂₁ vs wild-type in M9 minimal media.	138
Figure 5.3.2.2.1: Changes to proteome in an enlarged periplasm; Lpp ₊₂₁ vs wild-type in M9+sorbitol condition.	141
Figure 5.5.1: Proteostasis in an enlarged periplasm; quantitative proteomics summary.	152

List of Tables

Table 2.1.1: <i>E. coli</i> strains used in the study.....	48
Table 2.1.2: <i>E. coli</i> plasmids and synthesised DNA fragments (gBlocks™) used in the study.....	48
Table 2.4.2.1: Primers used in the study.	52
Table 2.6.3.1: List of antibodies used in the study.....	62
Table 4.3.2.1: Antimicrobial sensitivity profile of Lpp ₊₂₁ <i>E. coli</i>	100
Table 4.4.2.1: Non-essential genes enhancing the growth of Lpp ₊₂₁ <i>E. coli</i>	115
Table 4.4.3.1: Viability assessment of the Lpp ₊₂₁ double recombinants in M9 minimal media with varying supplements.	120
Table 5.2.1.1.1: Genes up-regulated in Lpp ₊₂₁ strain.	132
Table 5.2.1.2.1: Genes down-regulated in Lpp ₊₂₁ strain.	133
Table 5.3.2.1.1: Over-expressed and diminished proteins in M9 minimal media; Lpp ₊₂₁ vs wild-type.	139
Table 5.3.2.2.1: Over-expressed and diminished proteins in M9 minimal media supplemented with 0.5 M sorbitol; Lpp ₊₂₁ vs wild-type.	143
Table 5.4.1: Synthetic lethality evaluation of over-expressed non-essential proteins in nutrient-limited media conditions (M9 Minimal media 0.2% (w/v) glucose).	148

Abbreviations

ACN	acetonitrile
APS	ammonium persulfate
AT	autotransporter
ATC	anhydrotetracycline
ATP	adenosine triphosphate
BAM	β -barrel Assembly Machinery
BCA	bicinchoninic Acid
CAA	chloroacetamide
CAT	chloramphenicol acetyltransferase
cDNA	complementary DNA
CHS	Cold Spring Harbor
CM	chloramphenicol
Cryo-EM	cryo-electron microscopy
Cryo-TEM	transmission electron cryomicroscopy
CV	coefficient of variation
DTT	dithiothreitol
EAEC	enteroaggregative <i>E. coli</i>
EDTA	ethylenediaminetetraacetic acid
eSGA	<i>E. coli</i> Synthetic Genetic Array
FC	fold change
FDR	false discovery rate
FRT	flanking recombinase target
GFP	green fluorescent protein
GIANT-coli	Genetic Interaction Analysis Technology for <i>E. coli</i>
GT	glycosyltransferase
HDA	high-density array
HRP	horse-radish peroxidase
HT	high-throughput
IM	inner membrane
IMP	inner membrane protein
Kan	kanamycin
kDa	kilo Dalton
LB	lysogeny Broth
LC-MS/MS	liquid chromatography-tandem mass spectrometry

LFQ	label-free quantitative
LOL	localisation of lipoprotein
LOS	lipid-core oligosaccharide
Lpp	Braun's lipoprotein
LPS	lipopolysaccharide
M9 MM	minimal medium (M9 salts)
MCS	multiple Cloning Site
<i>mDAP</i>	<i>meso</i> -diaminopimelate
MIC	minimum inhibitory concentration
MS	mass spectrometry
NAG	UDP- <i>N</i> -acetylglucosamine
NAM	UDP- <i>N</i> -acetylmuramic acid
NMR	nuclear magnetic resonance
OD ₆₀₀	optical density at 600nm
OM	outer membrane
OMP	outer membrane protein
PAGE	polyacrylamide gel electrophoresis
PBP	penicillin-binding protein
PCA	principal component analysis
PCR	polymerase chain reaction
PDB	Protein data bank
PEP	phosphoenolpyruvate
PG	peptidoglycan
PMSF	phenylmethane sulfonyl fluoride
POTRA	polypeptide transport-associated
PP	periplasm
PVDF	polyvinylidene fluoride
RND	resistance-nodulation-cell division
SAXS	small-angle X-ray scattering
SD	standard deviation
SDS	sodium Dodecyl Sulfate
SEM	standard error of the mean
SOB	super optimal broth
SPATE	serine protease ATs of the <i>Enterobacteriaceae</i>
SPI	signal peptidase I

SS	signal sequence
T5SS (a-f)	Type 5 Secretion System (a-f)
TAM	Translocation and Assembly Module
TAT	twin-arginine translocon
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TM	transmembrane
TP	transpeptidase
TPR	tetratricopeptide-repeats
TSS	transformation and storage solution
UDP	uridine diphosphate
WT	wild-type

Chapter 1: Cell envelope biogenesis

1.1 Introduction

Gram-negative bacteria have two membrane bilayers termed the outer membrane (OM) and inner membrane (IM) (Figure 1.1.1), and are therefore referred to as diderms. Current models of evolution suggest that the OM evolved a few billion years ago (Sutcliffe, 2010; Gupta, 2011; Tocheva *et al.*, 2016), resulting in a compartment between the two bilayers termed the periplasm (Miller & Salama, 2018). Discovery of the presence of a periplasm occurred even before its morphological visualisation. In the 1960s, scientists were trying to discern how toxic enzymes such as ribonucleases and phosphatases produced by Gram-negative bacteria were non-toxic to the cell. It later became apparent that these enzymes are confined in a separate compartment from the cytoplasm (Miller & Salama, 2018). Biochemical extraction methods that maintain the integrity of the IM bound cytoplasm were key to this discovery (Neu & Heppel, 1964). Only afterwards, in 1964, was the periplasm visualised through the advent of electron microscopy, where it was seen clearly as a space in between the two membrane bilayers (Bladen & Mergenhagen, 1964).

The OM is an asymmetric bilayer with two distinct leaflets; phospholipids are found in the inner leaflet and lipopolysaccharides in the outer leaflet (Koebnik *et al.*, 2000; Silhavy *et al.*, 2010) (Figure 1.1.1). By contrast, the IM is a symmetrical phospholipid bilayer that encloses the cytoplasm (Ingledew & Poole, 1984). In between the two bilayers and within the periplasm, Gram-negative bacteria have an extra layer in their cell wall termed the peptidoglycan (PG) (Vollmer & Holtje, 2004). PG is a polymer of glycan strands cross-linked by short peptides that provide osmotic protection and also control shape in bacterial cells (Vollmer & Holtje, 2004; Vollmer & Seligman, 2010). Thus, the overall cell envelope is a multilayered structure that serves as an interface to the extracellular milieu, contributes to cellular integrity and modulates permeability (Koebnik *et al.*, 2000; Silhavy *et al.*, 2010). Its significance makes the cell wall indispensable and hence essential for life (Silhavy *et al.*, 2010).

Through their evolution diderm bacteria faced challenges to their cell biology (Miller & Salama, 2018): The assembly of the distant lipid barrier of the OM required novel lipid transport pathways (Sperandeo *et al.*, 2017). The periplasm needed to evolve novel folding pathways as the molecular chaperones in the cytoplasm were isolated (De Geyter *et al.*, 2016).

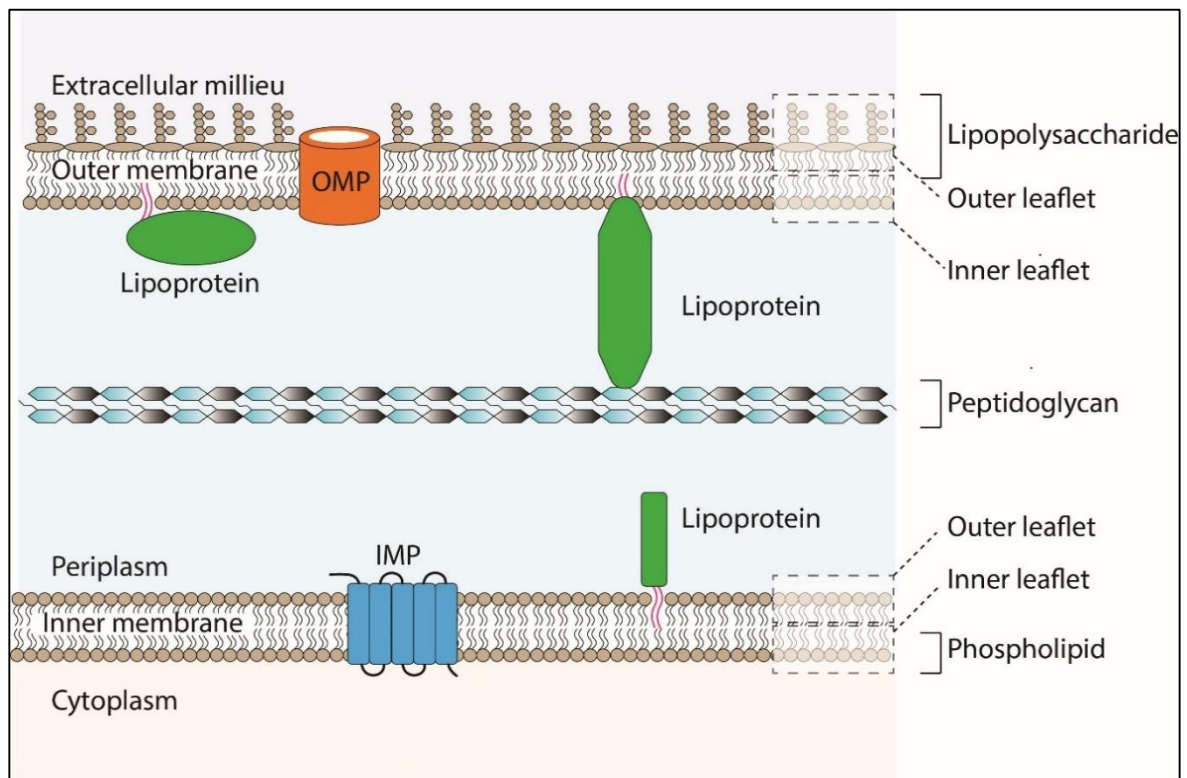


Figure 1.1.1: Schematic of a Gram-negative cell envelope.

As depicted, the cell envelope is composed of two membranes: the outer membrane (OM) and the inner membrane (IM). The OM is asymmetric, where the outer leaflet is composed of lipopolysaccharides, and the inner leaflet is comprised of phospholipids. By contrast, the IM is symmetrical as both leaflets, inner and outer, are made of phospholipids. The aqueous space in between the two membranes is called the periplasm, where a third layer of the cell envelope is located: the peptidoglycan. The cell envelope membranes harbour different types of proteins. Outer membrane proteins (OMPs) shown, are β -barrel proteins embedded within the OM while the inner membrane proteins (IMP) depicted, are α -helical transmembrane proteins in the IM. Lipoproteins have diverse topologies and can be anchored to either membrane (both leaflets of the OM and outer leaflet of the IM).

Membrane protein assembly systems were needed to populate the OM with pores for nutrient influx (Rollauer *et al.*, 2015). The systems that evolved in the periplasm to meet these challenges make this compartment functionally dynamic. It now has machineries essential for transport, folding, and quality control (Miller & Salama, 2018). Some of these machineries span the periplasm with components in both the OM and IM, and sometimes even spanning both membranes (Miller & Salama, 2018).

Some of the diverse functions in the periplasm include protein transport, folding, protein secretion, lipopolysaccharide and lipoprotein secretion, cell division, cell signalling, PG synthesis, osmoregulation and many others (Miller & Salama, 2018). In keeping with the importance of the periplasm in Gram-negative cell biology, this thesis will investigate the architectural significance of the periplasm, including its control over the morphology of the cell wall using an extended OM-PG bridge and the transport of proteins across the wall using an autotransporter reporter system.

1.2 Cell wall architecture

1.2.1 The cell wall

The envelope of most Eubacteria consists of several layers. The PG, also called the murein layer, functions to protect the cell from rupture by its internal turgor pressure (Weidel & Pelzer, 1964) and in bacterial cell shape maintenance (Vollmer *et al.*, 2008). Other functions of the PG include serving as a scaffold for anchoring cell envelope proteins (Dramsı *et al.*, 2008). It is also intimately involved in the processes of cell division and growth (Vollmer & Holtje, 2004).

The PG is a hetero-polymer made of linear glycan strands of alternating, β 1,4-linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues linked together by short peptides (Vollmer *et al.*, 2008), forming a three-dimensional network structure. The peptides are attached by an amide linkage to the lactyl group of the MurNAc and are unusual because they contain rare D-amino acids (Vollmer *et al.*, 2008). In terms of its biogenesis, the initial sequence of the newly synthesised peptide in *Escherichia coli* is L-Ala-D-iGlu-m-DAP-D-Ala-D-Ala, with meso-diaminopimelic acid (m-DAP) at position 3 (Vollmer *et al.*, 2008). However, the sequence characteristics of the pentapeptide vary depending on the bacterial species and growth conditions (Figure 1.2.1.1A). In *E. coli*, measurements of the concentration of pentapeptides show that there is a rapid proteolytic degradation to tetrapeptides, tripeptides and dipeptides (Vollmer *et al.*, 2008). Thus, complex dynamic recycling of these building blocks is at play in the periplasm, ultimately leading to the disaccharide pentapeptide form. In

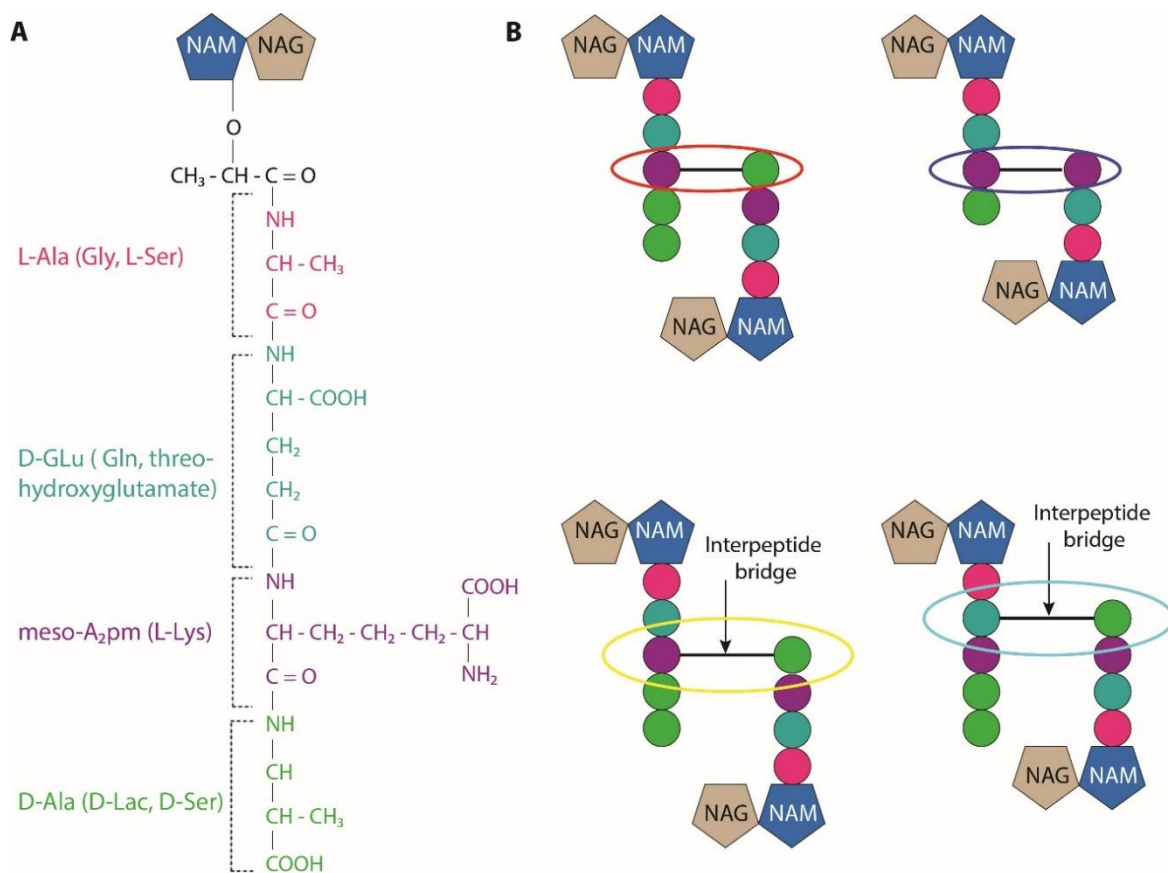


Figure 1.2.1.1: Peptidoglycan cross-linking and variations of the peptide stem.

(A) The chemical structure of the peptidoglycan (PG) building block. The glycan chain is composed of *N*-acetylmuramic acid (MurNAc) and *N*-acetylglucosamine (GlcNAc). The D-lactoyl group of MurNAc is substituted by peptide chains containing D- and L- amino acids. Variations in the chemical composition of the peptide are shown in brackets. (B) Variations in PG cross-linking. Cross-links can either be directly through 4-3 or 3-3 positions (top; red and purple circle, respectively) or in some cases through a short peptide bridge situated between position 3 and 4 or 2 and 4 of the stem-peptide (bottom; yellow circle and cyan circle, respectively). Circles in panel B represent the stem peptides whose colours match with panel A peptide structures. Abbreviations: L-Ala., L-Alanine; Gly., Glycine; L-Ser., L-Serine; D-Glu., D-Glutamic acid; D-Gln., D-Glutamine; L-Lys., L-Lysine; D-Ala., D-Alanine; D-Lac., D-Lactate; D-Ser., D-Serine; NAM, MurNAc; NAG; GlcNAc.

E. coli and other Gram-negative bacteria, two main peptide cross-links form for the extension of the PG net-like structure. The common DD-type formed between the carboxyl group of D-ala (position 4) of one peptide and the amino group at the D-centre of m-DAP (position 3) of another peptide. The second type is the LD-cross-link formed between the L-centre of m-DAP of one peptide and the D-centre of m-DAP of another peptide (Figure 1.2.1.1B, top) (Glauner *et al.*, 1988b; Figueroa-Cuilan & Brown, 2018).

The composition of the PG “sacculus” is not homogenous. This variation is due to several unique aspects of the mucopeptides, where over 50 different types exist (Vollmer *et al.*, 2008). These include the length of the peptide chain (di-, tri-, tetra-, pentapeptide), the state of cross-linkage (monomer, dimer, trimer, tetramer), the presence of either D-ala or Ser at position 4 or 5, the type of cross-linkage (DD or LD), the presence of 1,6-anhydroMurNAc residues (glycan strand termini), and the presence of a L-lysine –L-Arg dipeptide (at position 4) which remains after proteolytic digestion of Braun’s lipoprotein (Lpp) (Glauner, 1988). In what seems to be specific for *E. coli*, the major mucopeptides are the disaccharide tetrapeptide (>30% of the total material) (and the DD-type cross-links (>20% of the total material) (Glauner, 1988). Even within *E. coli*, this composition shows some variation depending on the strain, growth phase, growth medium and temperature. For example, the transition from exponential phase to stationary phase leads to an increased fraction of the L-Lys-L-Arg containing peptides as well as the total cross-linkage of the sacculi (Glauner *et al.*, 1988b). The extra linkage could be significant in the strengthening of the cell envelope as it prepares for stationary phase by increasing the number of cross-links in the PG and the number of contacts between the PG layer and the OM via Lpp.

1.2.2 Models for peptidoglycan architecture

A crucial feature of the PG architecture is the orientation of the glycan strands and peptides relative to the membranes and axis of the cell. Our understanding of this crucial feature is at an early stage, due to the heterogeneous and flexible nature of the PG that has so far made it difficult to determine the structure at high resolution. Two mutually exclusive models have been proposed based on the chemical and biophysical data available. The classical “layered model” and the “scaffold model”.

The layered model was proposed based on the conformations of the PG constituents, the glycans and the peptide side chains as well as the prediction of the three-dimensional structure of the sacculus. In this model, glycans are considered as straight rods that run parallel with the

peptides protruding above or below the glycan plane (Vollmer & Holtje, 2004). The model would only agree with a horizontally layered PG in which glycan strands run parallel to the cytoplasmic membranes. A vertical arrangement of the glycan strands would be too long to be accommodated in the periplasm (Rogers *et al.*, 1980).

The scaffold model proposes that the glycan strands extend perpendicular from the cytoplasmic membrane and are cross-linked by peptides that are parallel to the surface of the membrane (Dmitriev *et al.*, 1999; Dmitriev *et al.*, 2003). According to this model, the PG almost entirely fills the periplasmic space. From analysis of PG amounts using the available data of glycan length distribution, Vollmer *et al.* (2004) concluded that *E. coli* does not contain enough PG for the proposed scaffold architecture. Vollmer *et al.* (2004) further argued that there is a high proportion (25% to 30%) of glycan material consisting of glycan strands that are longer than 30 disaccharide units. The number would mean that many of the glycan strands are longer than the measured distance between the OM and the IM, which is approximately 20 nm (Matias *et al.*, 2003). For these reasons, Vollmer *et al.* (2004) concluded that the scaffold model is less useful than the layered model for understanding PG architecture *in vivo*.

1.2.3 Cell wall thickness

Whether or not it is the most representative species, *E. coli* is the best-studied model for PG architecture. Using Cryo-TEM of frozen-hydrated sections, the PG is seen as a thin layer beneath the OM, with the thickness of 6.35 ± 0.53 nm in *E. coli* versus 2.41 ± 0.54 nm in *Pseudomonas aeruginosa* (Matias *et al.*, 2003) (Figure 1.2.3.1). Does this increased thickness of PG in *E. coli* suggest 2-3 layers compared to *P. aeruginosa*? The standard paradigm suggests that the PG in *E. coli* has a single-layered architecture, but there are several lines of evidence that appear to contradict this. First, the fact that *E. coli* have excess PG as implied by the ability to grow under certain conditions with 50% less PG per cell surface unit (Prats & de Pedro, 1989). Second, the mucopeptides of the PG are trimeric or tetrameric (Glauner *et al.*, 1988a) with three or four connected peptides does not appear to be consistent with an exclusively single-layered sacculus. Third, small-angle scattering experiments show that 20-25% of the surface of sacculi is thicker than a single layer and probably may consist of a triple layer (Labischinski *et al.*, 1991). Finally, that the PG in *P. aeruginosa* is significantly thinner than that of *E. coli* shows that *E. coli* PG does not have the minimum possible thickness of a single layer (Matias *et al.*, 2003). In the absence of high-resolution imaging of the PG, these contentious debates remain unresolved.

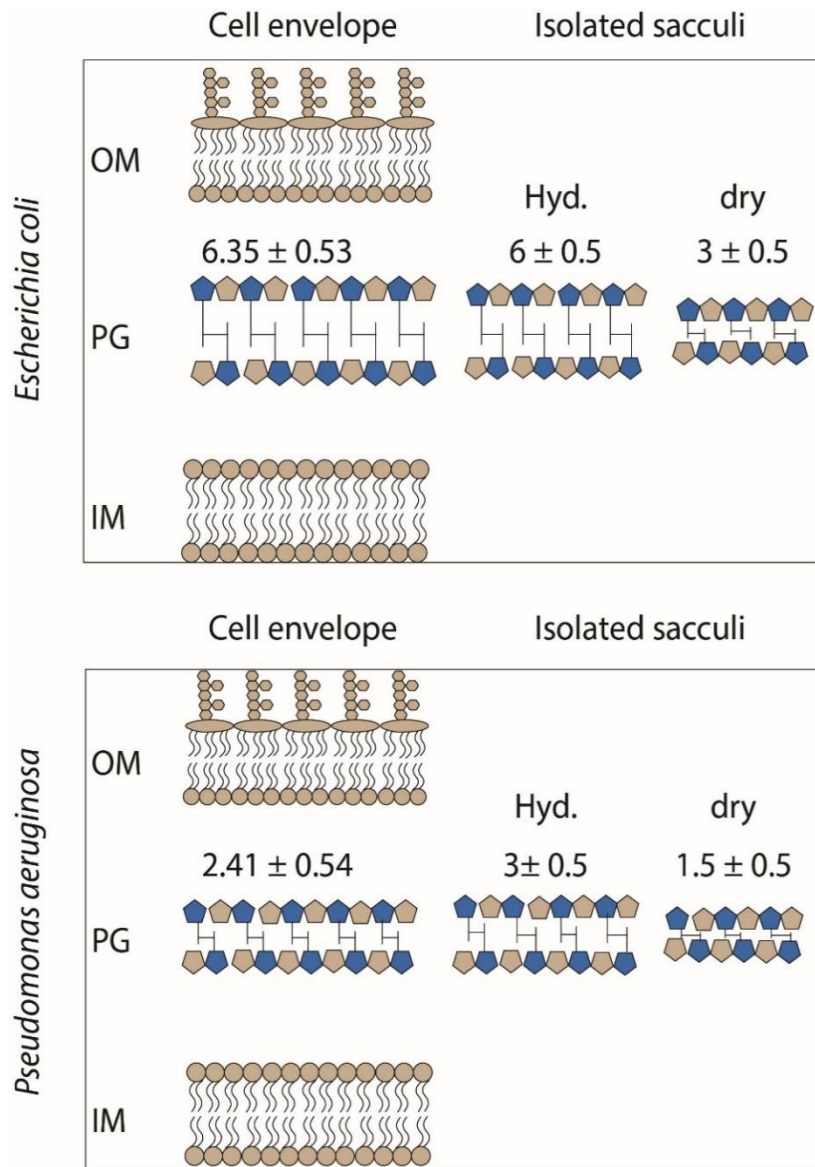


Figure 1.2.3.1: Peptidoglycan thickness.

Dimensions of the cell envelope and isolated peptidoglycan (PG) in Gram-negative: *E. coli* and *P. aeruginosa*. The numbers indicate the thickness of the PG in nm, with \pm showing the standard deviation (Matias *et al.*, 2003). Thickness was measured by atomic force microscopy for both unisolated and isolated PG (Matias *et al.*, 2003). The figure shows that PG is not a solid block of regular thickness. The thickness varies from one organism to another, and even within the same organism. In this case, hydrated isolated PG is thicker than dried PG, and *E. coli* PG is more than double the PG thickness in *P. aeruginosa*. Abbreviations: Hyd., hydrated; OM, outer membrane; IM, inner membrane.

1.2.4 Cell wall elasticity

In *E. coli*, studies using low-angle laser light scattering to determine the change of the mean surface of sacculi, following alteration of either the net charge, pH or chemical modification concluded that the sacculi is elastic and can expand and shrink three-fold without rupture (Koch & Woeste, 1992). Upon detergent solubilisation of the IM, there is a sudden decrease in the cell surface area of about 45% as the PG relaxes (Koch *et al.*, 1987). These results suggest that the PG forms an elastic net which is expanded to some extent in living cells due to the cells turgor pressure.

The observed change in length, but with very little change in diameter seen upon osmotic shock of *E. coli* (van den Bogaart *et al.*, 2007), provides evidence that elasticity occurs in the direction of the long axis (Vollmer & Seligman, 2010). This data is in agreement with a PG model where peptides oriented in the direction of the long axis. By stretching, the sacculus becomes more organised/ordered and more disorganised when relaxed (Vollmer & Seligman, 2010) (Figure 1.2.4.1). There is speculation that the tension of covalent bonds in PG caused by the turgor might be an essential factor for PG growth. This assumption is reinforced by evidence that PG synthesis complexes may recognise and require stretched PG for efficient insertion and attachment of newly synthesised glycan chains (Vollmer *et al.*, 2008).

1.2.5 Cell turgor and permeability properties of the cell wall

Demchick & Koch (1996) used fluorescence-labelled dextrans of different sizes to determine the pore diameter of the PG network in *E. coli*. On average, the pores are relatively homogenous in size, with the mean radius of the pore size in *E. coli* being 2.06 nm (Demchick & Koch, 1996). Due to the elastic nature of the sacculi, the pore size varies based on the state of the sacculi, with these results obtained from relaxed sacculi. A recent study by Turner *et al.*, (2018) confirmed the pore size variability by showing pores with up to 66 nm² area, although most were less than 5 nm².

From these data, it was determined that globular, uncharged proteins with a molecular weight of up to 24 kDa should be able to diffuse across the PG net (Demchick & Koch, 1996). In a living cell, expanded PG due to turgor might even be permeable to proteins up to 100 kDa (Vazquez-Laslop *et al.*, 2001). Considering that proteins of varying sizes are transported across the PG for assembly into the OM, the sacculi can act as a barrier to large protein complexes. These complexes can cross by binding to chaperones that diffuse through the PG. An example

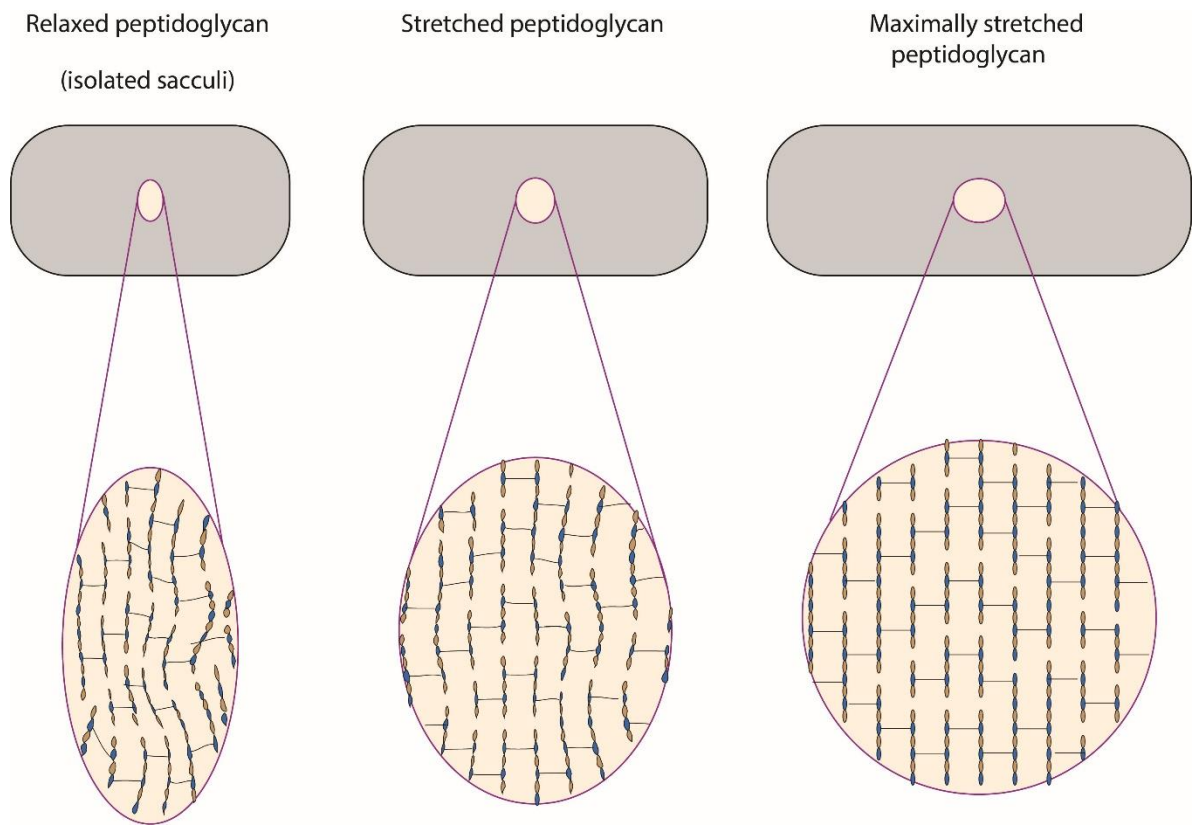


Figure 1.2.4.1: Cell wall elasticity.

Model illustrating the changes in peptidoglycan (PG) organisation of the layered architecture model in *E. coli*. Isolated PG, representing a relaxed form of PG, has been observed and presents a disorganised layered architecture (left side panel) (Gan *et al.*, 2008; Vollmer & Seligman, 2010). In the presence of turgor pressure, the PG is stretched and is presumed to be more organised, as shown in the middle panel. At increased turgor pressure, i.e., in a hypotonic solution, PG is presumed to fully stretch to a very ordered structure as shown (far right panel). The glycan strands are shown as oval-shaped subunits (disaccharides) running mainly in the direction of the short axis of the cell. The glycan strands are connected by peptides (lines) running mainly in the direction of the long axis of the cell. For simplicity, only glycan strands of 3-14 subunits, and cross-linked peptides connecting the glycans are shown, with the omission of unlinked peptides Adapted from Vollmer & Seligman (2010), with minor modifications.

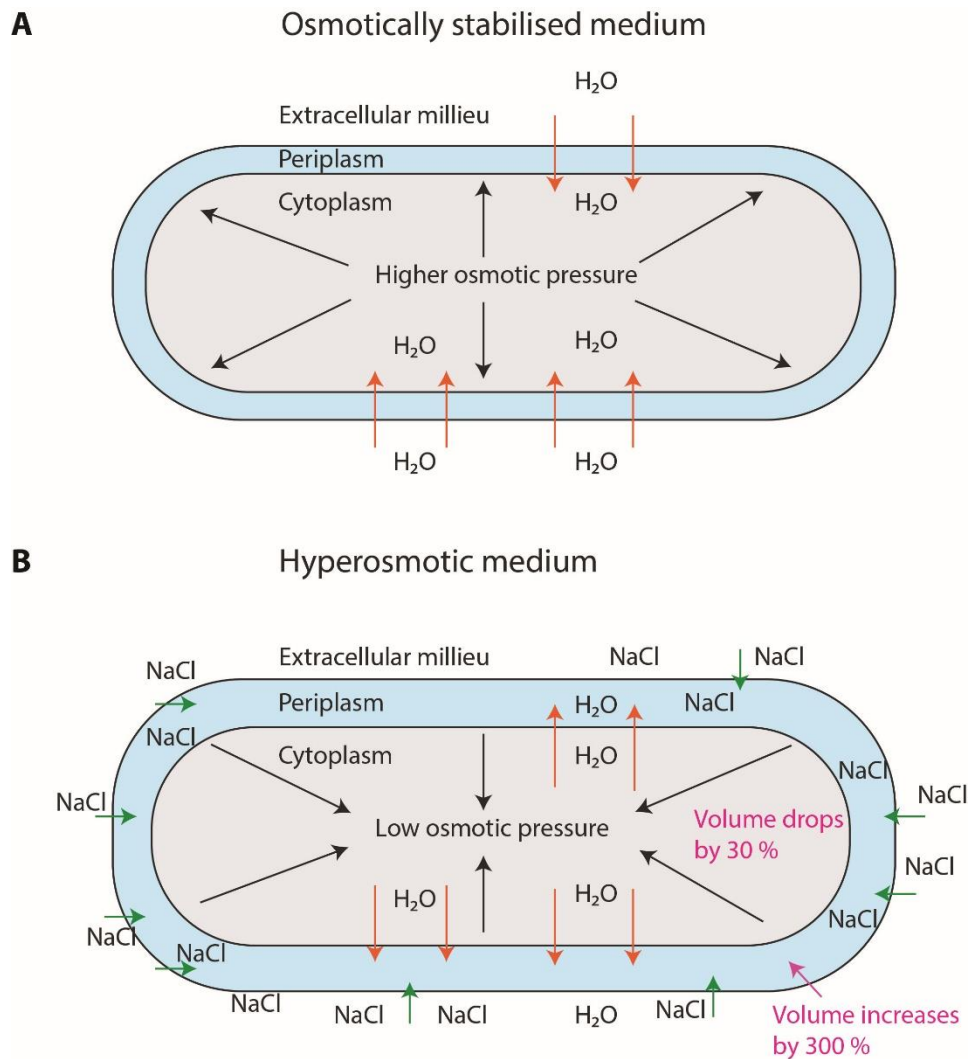


Figure 1.2.5.1: The role played by osmolytes in periplasmic volume changes.

(A) In an osmotically stabilised medium, which is the case for most of laboratory media conditions, the cell is always under turgor pressure of 3-5 atm (Cayley *et al.*, 2000) due to the high concentration of solutes in the cytoplasm. The pressure is contained by the elastic PG, which prevents the cell from bursting. (B) On a hyperosmotic medium, the osmolyte (NaCl) can transverse the OM but not the semipermeable IM. Since the periplasm and the cytoplasm are iso-osmotic (Stock *et al.*, 1977; Cayley *et al.*, 2000; Erickson, 2017), the cell loses water to counteract the pressure, increasing the volume of the periplasm by 300 % and decreasing the cytoplasmic volume by 30% (Cayley *et al.*, 2000).

is the transport of OM lipoproteins which form a soluble complex with the chaperone LolA (Tokuda & Matsuyama, 2004). In order to accommodate other, larger trans-envelope complexes such as secretion systems and flagella, PG hydrolases are involved in order to locally open large spaces in the PG net (Dijkstra & Keck, 1996).

A significant factor to consider in discussions on the size of the periplasm and permeability properties of the PG is how the turgor force of the bacterial cytoplasm affects periplasmic architecture. Many Gram-negative bacteria like *E. coli* are found in diverse environmental conditions with different osmolalities, ranging from freshwater environments to the hyperosmotic environment of the human gut. Experimental findings, therefore, need to take into consideration the osmolality of laboratory growth media (Figure 1.2.5.1). Sodium chloride (NaCl) is often used to supplement growth media (e.g. 85 mM in LB medium), and osmolytes such as sorbitol that are impermeant to the IM can be useful in setting stable iso-osmotic conditions (Weber *et al.*, 2006; van den Bogaart *et al.*, 2007). Studies by Stock *et al.* (1977) found that the periplasmic volume increased rapidly in response to increased osmolyte concentration in the external medium (Figure 1.2.5.1).

Even in hypo-osmotic environments, regulatory pathways act to maintain the periplasm in a relatively high osmolar state, working to match that of the cytoplasm (Stock *et al.*, 1977; Cayley *et al.*, 2000). Due to this, the IM experiences little turgor pressure (Cayley *et al.*, 2000), with the greatest load exerted on the OM-PG layer of the cell wall (Rojas *et al.*, 2018; Hwang *et al.*, 2018; Jefferies *et al.*, 2019).

1.3 Peptidoglycan biogenesis, regulation, and function

1.3.1 Peptidoglycan core synthesis

In a nutshell, PG biogenesis starts in the cytoplasm with the synthesis of its primary building block, GlcNAc-MurNAc pentapeptide attached to a lipid carrier undecaprenyl-phosphate (Und-P), forming what is referred to as lipid II. Lipid II is flipped into the periplasm where biogenesis terminates by the polymerisation and cross-linking of the MurNAc pentapeptides. The overall pathway is divided into three phases based on the localisation of the processes; (i) those occurring in the cytoplasm, (ii) the inner leaflet of the cytoplasmic membrane, or (iii) the periplasm. The pathway is well understood from studies with *E. coli* (Figure 1.3.1.1).

1.3.1.1 The Cytoplasmic steps: Synthesis of UDP-MurNAc-pentapeptide

Initiation of PG biogenesis involves the conversion of the disaccharide, UDP-*N*-acetylglucosamine (UDP-GlcNAc) to UDP-*N*-acetylmuramic acid (UDP-MurNAc) by

MurA and MurB enzymes (van Heijenoort, 2001). MurA transfers enolpyruvyl moiety from phosphoenolpyruvate (PEP) to UDP-GlcNAc generating enolpyruvyl-UDP-GlcNAc. MurB subsequently catalyses the reduction of enolpyruvyl-UDP-GlcNAc to UDP-MurNAc in an NADPH-dependent manner (Barreteau *et al.*, 2008). A series of ATP-dependent amino acid ligases (these proteins are encoded by the *murC-murF* genes) add a pentapeptide chain to UDP-MurNAc forming UDP-MurNAc-pentapeptide (Ruiz, 2016).

1.3.1.2 The IM steps: Assembly of lipid I and II

In the second phase of the pathway, the UDP-MurNAc-pentapeptide is transferred to undecaprenyl-phosphate (Und-P), a lipid carrier, forming Lipid I (Bouhss *et al.*, 2008) (Figueroa-Cuilan & Brown). MraY, an integral membrane protein, catalyses this reversible reaction. Insertional inactivation experiments which result in lethal phenotype demonstrate the essentiality of this enzyme in PG biogenesis (Boyle & Donachie, 1998). Next, the essential membrane-associated glycosyltransferase, MurG, couples GlcNAc to the C-4 hydroxyl moiety of the MurNAc component of Lipid I, generating undecaprenyl-pyrophosphoryl-GlcNAc-MurNAc-pentapeptide known as Lipid II (Figueroa-Cuilan & Brown, 2018). At this stage, Lipid II is anchored in the IM (Figure 1.3.1.1).

There has been an ongoing debate concerning how lipid II is then translocated across the IM. It is a large amphipathic molecule that would require a transporter (van Heijenoort, 2007). One study showed that upon depletion of MurJ, lipid II and its other precursors accumulate in the cytoplasm, suggesting MurJ is the lipid II translocator or “flippase” (Ruiz, 2008). This model was later disputed by another study which presented FtsW as the lipid II flippase based on experiments in reconstituted liposomes (Mohammadi *et al.*, 2011). Both studies presented evidence for their cases, and neither case dismissed the other.

With the advent of new techniques, there is new clarity on how lipid II is translocated into the periplasm. The predicted structure of MurJ includes a V-shaped, solvent-exposed cavity that contains several charged residues required for its function (Butler *et al.*, 2014). Modification of engineered cysteine substitutions in this cavity with sulfhydryl-reacting probes inactivates MurJ in *E. coli* (Sham *et al.*, 2014). Exploiting this technique to rapidly inhibit MurJ, in combination with a new *in vivo* assay to monitor translocation of lipid II, has helped to clear up the MurJ versus FtsW controversy. The ability of colicin M toxin to enter the periplasm and hydrolyse lipid II into its constituents, undecaprenyl and pyrophosphate-disaccharide-pentapeptide forms the basis of the *in vivo* assay (Cherier *et al.*, 2016). Since the toxin cannot

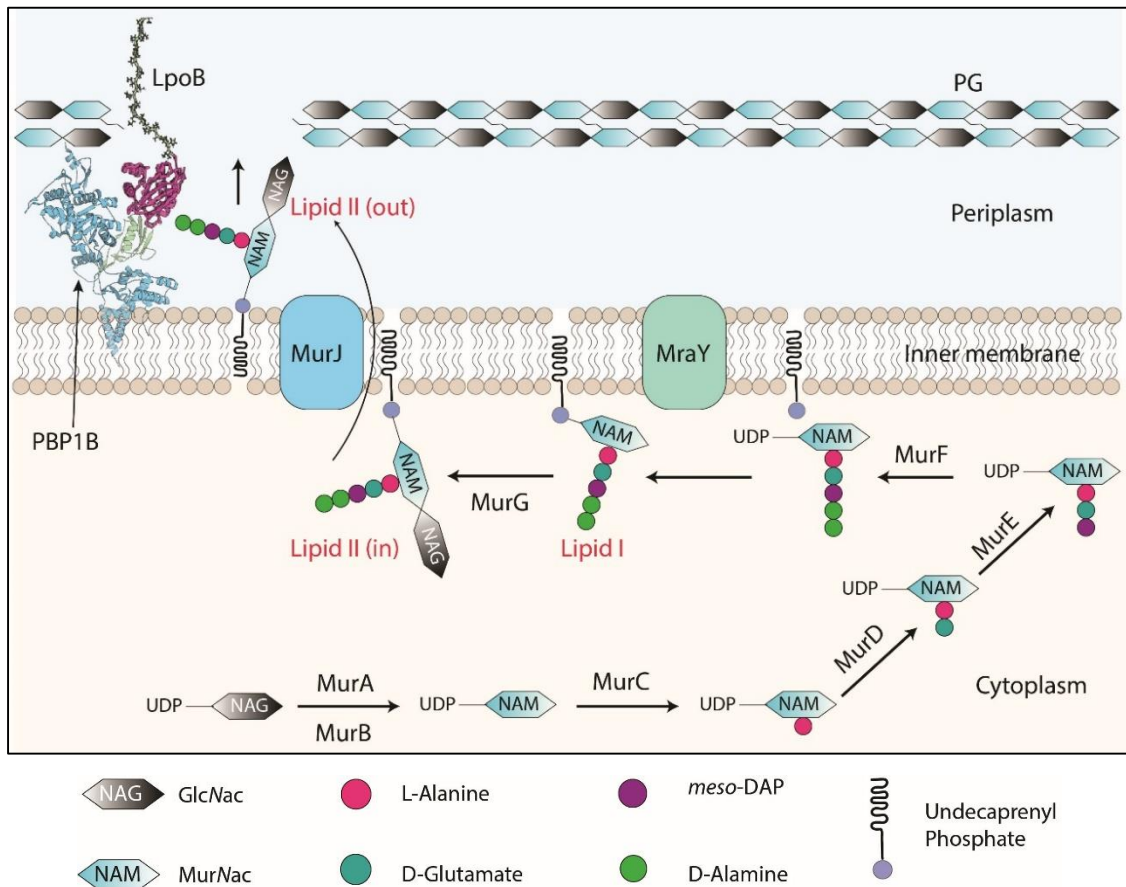


Figure 1.3.1.1 Peptidoglycan biosynthesis.

A cartoon depiction of peptidoglycan (PG) biosynthesis in three cellular locations. Muropeptide synthesis in the cytoplasm, synthesis of lipid I and lipid II in the cytoplasmic leaflet of the IM, and PG chain elongation in the periplasm. First UDP-NAG is converted into UDP-NAM-pentapeptide by a series of Mur enzymes (MurA to MurF), that act in a catalytic cascade. UDP-NAM-pentapeptide is then transferred onto undecaprenyl phosphate to form lipid I, through the action of an integral membrane protein called MraY. UDP-NAG is then used as a substrate to convert lipid I into the muropeptide-containing lipid II, which is subsequently flipped by MurJ to the periplasm where it attaches to a growing muropeptide chain. Growth of the muropeptide chain requires glycosyltransferases (GTase) to polymerise the glycan chains and transpeptidases (TPases) to cross-link the stem peptides. Depicted here is a class A penicillin-binding protein 1B (PBP1B), that can perform both roles.

cross the IM, the cytoplasmic form of Lipid II is not degraded. Thus, the amount of pyrophosphate-disaccharide-pentapeptide produced after addition of the toxin correlates with the activity of lipid II flippase. Combining both techniques showed that all lipid II flippase activity detectable in *E. coli* cells was abolished upon rapid inactivation of MurJ (Sham *et al.*, 2014). In contrast, depletion of FtsW still resulted in lipid II hydrolysed products. These data demonstrate that lipid II translocation requires MurJ, and suggest that FtsW is not required. However, it is still not clear whether MurJ is the flippase itself or a factor required to assist lipid II translocation across the IM.

1.3.1.3 The periplasmic steps: Polymerisation and cross-linking

The final step of PG synthesis involves the incorporation of the GlcNAc-MurNAc pentapeptide from lipid II into the growing sacculus. Two main reactions take place: Transglycosylation to polymerise the disaccharide units, and transpeptidation to cross-link the peptide moieties. (van Heijenoort, 2001; Vollmer, 2008)

1.3.2 Regulation of peptidoglycan synthesis in *E. coli* by outer membrane lipoproteins

Synthesis of the PG requires coordination of glycan synthesis and hydrolysis of the pre-existing sacculus so that new PG material can be incorporated without compromising cell integrity. The factors involved in this process are localised in the OM, IM and the periplasm. Recent findings suggest that the growth of PG is synchronised with that of the OM, which would provide for the two envelope layers to grow in unison.

PG synthases catalyse the glycosyltransferase (GT) and transpeptidase (TP) reactions mentioned in the previous section. The reactions require glycosyltransferases (GTases) that catalyse the transglycosylation reactions and transpeptidases (TPases) that catalyse the transpeptidation reactions. These enzymes have historically been called penicillin-binding proteins (PBPs), based on their ability to bind penicillin covalently (Suginaka *et al.*, 1972), and are the target for penicillin drugs. There are three main types of PBPs: Class A PBPs, which have both GTase and TPase activities (bifunctional), Class B PBPs with only TPase activity (monofunctional), and monofunctional GTases (Vollmer & Bertsche, 2008). *E. coli* has three bifunctional synthases (PBP1A, PBP1B, PBP1C), a GTase (MgtA) and two monofunctional TPases (PBP2 and PBP3). PBP1A and PBP1B are functionally partially redundant as they are synthetic lethal (Yousif *et al.*, 1985), while the function of PBP1C is unknown but may be required in host cells (Budd *et al.*, 2004). PBP2 is required for cell elongation, and PBP3 is required for cell division (Typas *et al.*, 2011).

PBP1A and PBP1B are encoded by *mrcA* and *mrcB*, respectively. Employing genetic approaches that relied on synthetic lethality identified two OM lipoproteins, LpoA and LpoB, as factors that activate PBP1A and PBP1B, respectively (Paradis-Bleau *et al.*, 2010; Typas *et al.*, 2010). The authors showed that *E. coli* depends on at least one of the LpoA-PBP1A or LpoB-PBP1B protein complexes for growth, with deletion of two proteins in both complexes resulting in cell lysis. The Lpo proteins activate their cognate PBPs by direct interaction with a specific docking domain (Paradis-Bleau *et al.*, 2010; Typas *et al.*, 2010), which requires a membrane-spanning topology since the Lpo proteins are located in the OM while the cognate PBPs are IM anchored (Figure 1.3.2.1).

Structural analysis of LpoA and LpoB revealed that they are long enough to span the periplasm and contact their specific docking domains. (Egan *et al.*, 2014; Jean *et al.*, 2014). LpoB interacts with a small non-catalytic UB2H domain in PBP1B (Egan *et al.*, 2014). This interaction increases the GT activity of PBP1B, which indirectly stimulates its TP activity (Lupoli *et al.*, 2014). NMR spectroscopy showed that LpoB has a small globular C-terminal domain, which serves as the interaction site for the UB2H domain of PBP1B (Jean *et al.*, 2015b).

LpoB also has an elongated proline-rich unstructured flexible N-terminal region providing a reach of 14.5 nm across the periplasm, allowing LpoB to reach PBP1B UB2H domain from the OM (Egan *et al.*, 2014) (Figure 1.3.2.1).

Structural studies show that LpoA is larger and more rigid than LpoB. There is no long, flexible region in LpoA to reach PBP1A. Instead, LpoA adapts an elongated shape through two structured domains (Jean *et al.*, 2014). NMR and SAXS data of LpoA suggest that there is no flexibility between the two domains, but that the overall shape of the molecule is elongated giving a length of approximately 14 – 15 nm (Egan *et al.*, 2015). This length is sufficient for LpoA to reach its PBP1A partner (Figure 1.3.2.1). From NMR information derived from the N-terminal domain, LpoA adopts a series of five tetratricopeptide-repeats (TPR) (Jean *et al.*, 2015a).

1.4 Lipoproteins: structure, function, and biosynthesis

Lipoproteins are proteins with a covalently attached lipid moiety added as a post-translational protein modification. A wide range of protein structures can be anchored to a membrane via this common lipid structure. About 2.7% of all proteins encoded in prokaryotic genomes are

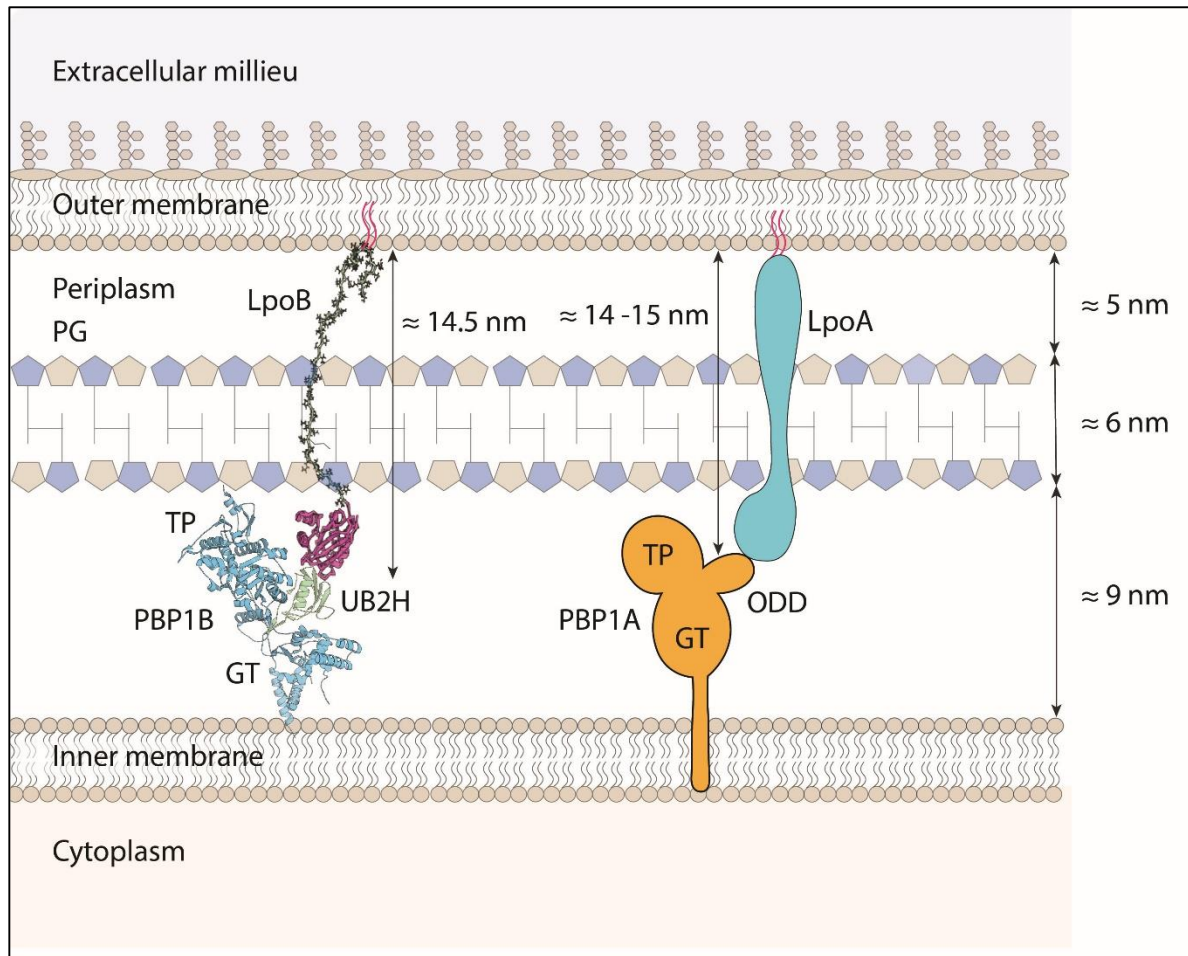


Figure 1.3.2.1: Activation of penicillin-binding proteins by lipoproteins.

Model of stimulation of penicillin-binding proteins (PBPs) by *E. coli* OM Lpo proteins. On the left, LpoB (PDB: 2MII) uses its extended ≈ 14.5 nm long, flexible N-terminal region to span the periplasm and places the globular domain in position to interact with the UB2H domain of PBP1B (PDB: 3VMA). On the right, cartoon representation of PBP1A activation by LpoA. The N-terminal domain of LpoA is anchored to the OM, whereas the C-terminal domain interacts with the ODD domain of PBP1A. The overall shape of the molecule is elongated, giving a total length of $\approx 14 - 15$ nm and a thickness of ≈ 3 nm. These dimensions should enable the protein reach PBP1A through the periplasm and cross the ≈ 6 nm thick peptidoglycan (PG) layer, which has $\approx 4 - 6$ nm wide pores (Demchick & Koch, 1996). The thickness of the PG layer and the distances to the OM and IM are according to Matias *et al.* (2003). Abbreviations: TP., transpeptidase domain; GT., glycosyltransferase domain.

predicted to be lipoproteins (Braun & Hantke, 2019). Lipoproteins adopt a diverse range of topologies in the cell envelope: they can be monomers, homo-oligomers or exist in hetero-oligomeric complexes with β -barrel proteins (in the OM) or α -helical membrane proteins (in the IM). The lipid modification is site-specific, corresponding to the N-terminal Cys residue of a mature lipoprotein (Tokunaga *et al.*, 1982). This modification is carried out by a unique multi-step pathway, is irreversible, and occurs on the periplasmic side of the IM. Some of the acylated lipoproteins are then transferred from the IM across to the OM by a unique protein transport system called the Lol pathway.

1.4.1 Lipoprotein biogenesis

Lipoprotein biosynthesis starts in the cytoplasm, when a protein is translated with an N-terminal signal sequence designating it for translocation via the Sec or Tat secretion machineries (Kudva *et al.*, 2013). The signal sequence of these prolipoproteins has a conserved sequence motif called the lipobox, where the critical residue in the lipobox is a conserved cysteine (+1 position of a mature protein) that serves as a lipid acceptor in a specific acylation reaction. This acylation involves three sequential steps catalysed by three membrane-bound enzymes. The first step involves the addition of diacyl glyceryl from phosphatidyl-glycerol to the sulfhydryl group of cysteine by phosphatidyl-glycerol site-specific transferase (Lgt) (Sankaran & Wu, 1994). The second step involves the cleavage of the signal peptide by prolipoprotein signal peptidase (LspA) liberating the α -amino group of diacylated cysteine. The final step involves the acylation of the free cysteine amino group by phospholipid transacylase (Lnt), generating mature triacylated lipoprotein (Buddelmeijer, 2015) (Figure 1.4.1.1).

1.4.2 Lipoprotein export

Mature lipoproteins in *E. coli* are targeted to the OM by the Lol pathway unless they contain a Lol avoidance signal, also known as the +2 rule (Konovalova & Silhavy, 2015). The identity of the amino acid after the conserved cysteine residue is what constitutes the lipobox sorting signal (Yamaguchi *et al.*, 1988). According to this rule, an aspartic acid (D) residue at position +2 causes retention of lipoproteins in the IM, acting as a sorting signal. The residue at position +3 also plays an additional role in IM localisation with either aspartic acid, glutamic acid or glutamine (Terada *et al.*, 2001). This rule is conserved in enterobacteria (Lewenza *et al.*, 2006).

The translocation of lipoproteins to the OM requires five Lol proteins. LolA, LolB, LolC, LolD, and LolE (Figure 1.4.1.1). A LolCDE complex includes ATP-hydrolysis activity and belongs

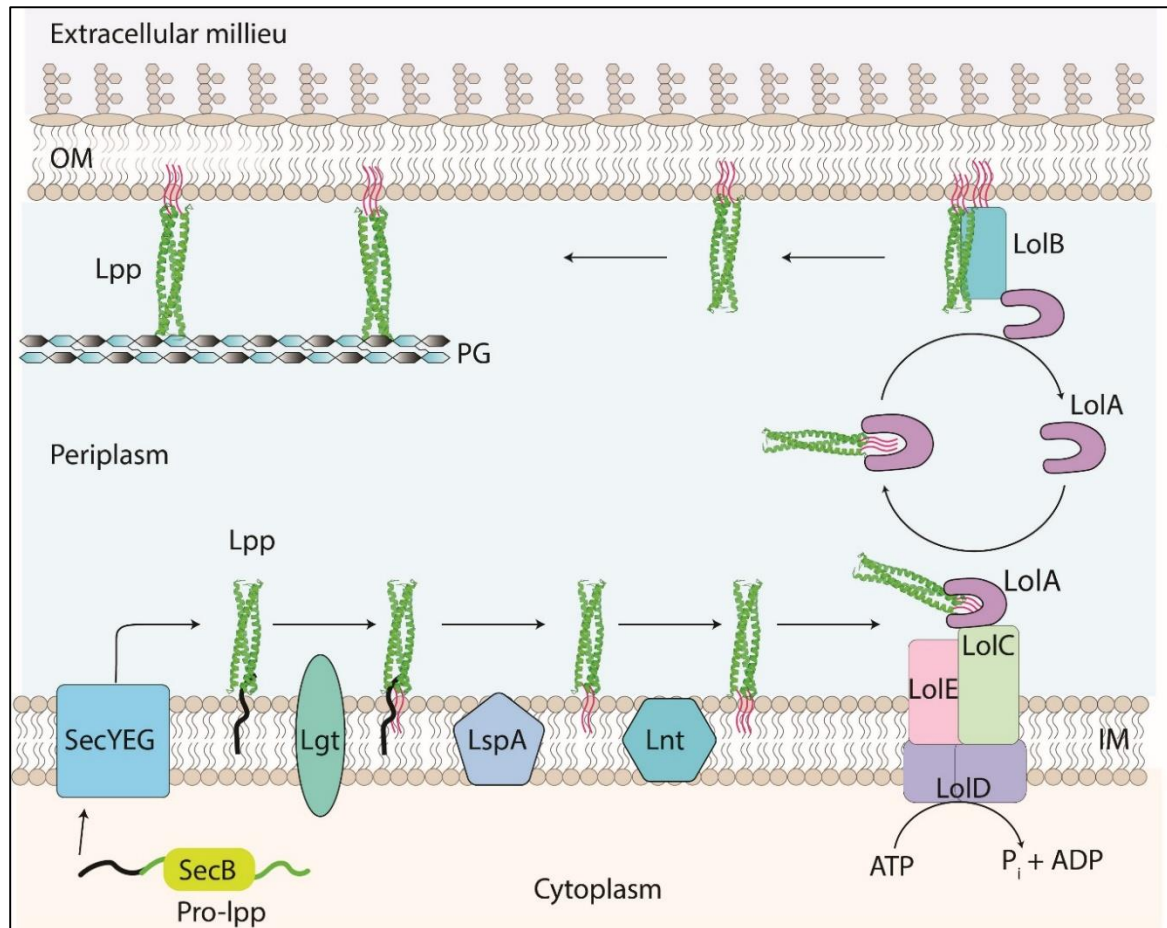


Figure 1.4.1.1: Lipoprotein biogenesis and export, exemplified with Lpp.

Lpp (green) (PDB: 1EQ7) is a major lipoprotein in *E. coli*. Pro-Lpp is synthesised in the cytoplasm with an N-terminal signal sequence (SS) (black), which targets it for translocation across the IM by SecYEG. Pro-Lpp remains anchored in the IM by its SS, and a diacylglyceryl transferase (Lgt) adds a diacylglycerol moiety to the cysteine residue. A signal peptidase (LspA) cleaves the SS, and a phospholipid acyltransferase (Lnt) adds another acyl chain to the newly formed N-terminus. Mature Lpp still anchored in the IM is recognised by the IM LolCDE complex, which powers its extraction from the IM using the energy of ATP. Lpp is then released to the periplasm in a complex with chaperone LolA. LolA delivers Lpp to the OM acceptor protein LolB which inserts it in the inner leaflet of the OM. Empty LolA returns to LolCDE and is recycled. Lpp is then covalently attached to the peptidoglycan (PG). Abbreviations: IM, inner membrane; OM, outer membrane.

to the ABC transporter superfamily (Yakushi *et al.*, 2000; Braun & Hantke, 2019). LolE recognises and binds lipoprotein substrates, while LolC recruits LolA. ATP hydrolysis by LolD then releases the lipoprotein from LolCDE to the molecular chaperone LolA. LolA delivers the lipoprotein to LolB, located in the OM (Konovalova & Silhavy, 2015). The affinity of LolB for substrate lipoproteins is higher than that of LolA, which promotes a unidirectional flow of the substrates from LolA to LolB (Konovalova & Silhavy, 2015). LolB distorts the packing of the lipid bilayer. Braun and Hantke (2019) speculate that this distortion may facilitate entry of the triacyl group of substrate lipoproteins into the inner leaflet of the OM (Figure 1.4.1.1).

In addition to the Lol pathway, some lipoproteins are transported across the periplasm by the Type 2 secretion system (T2SS) (Rondelet & Condemine, 2013). Unlike the Lol system, in which lipoproteins insertion takes place in the inner leaflet of the OM, the T2SS is thought to enable lipoproteins to insert into the outer leaflet of the OM (Zückert, 2014). Well-characterised substrates for T2SS dependent export include the lipoprotein Pula from *Klebsiella* species (d'Enfert *et al.*, 1987) and the lipoprotein SslE from *E. coli* (Baldi *et al.*, 2012b).

1.4.3 Lpp; Major outer membrane lipoprotein

Also, known as Braun's lipoprotein, Lpp is a 5.8 kDa α -helical lipoprotein, the first bacterial protein to have been identified with this now canonical lipid modification (Braun & Wolff, 1970). As far as known, Lpp is unique among lipoproteins in that it is covalently bound to the PG (Braun, 1975). It is the most abundant protein in *E. coli* being present at approximately 1 million copies per cell (Li *et al.*, 2014). Several features contribute to the extraordinary expression levels achieved for Lpp. An AT-rich promoter in the *lpp* gene facilitates strand unwinding by RNA polymerase for highly efficient transcription initiation (Nakamura & Inouye, 1979), and the mRNA encoding Lpp is very stable, having a half-life of 12 minutes, ten-fold more than other *E. coli* mRNAs (Hirashima & Inouye, 1973).

As a mature protein, Lpp consists of 58 residues, including the N-terminal modified cysteine. Even before its structure was available, the Lpp sequence suggested 14 repeats composed of seven amino acids (termed heptad repeats) of which every third or fourth residue is hydrophobic (Braun & Hantke, 2019). Such a sequence recurrence is common in coiled-coils of α -helices, where the seven residues form two turns of a helix and the side chains of parallel helices interlock systematically (Asmar & Collet, 2018). The X-ray crystal structure solved by Shu *et al.* (2000) confirmed these predictions. The protein has three domains: an N-terminal

capping motif, a long coiled-coil domain containing three parallel α -helices, and a hydrophobic helix-termination motif (Shu *et al.*, 2000).

1.4.4 Lpp and the peptidoglycan

Lpp exists in two forms: the PG linked form (bound form) and the free form (unbound form). The bound form is monotopic, facing the periplasm for PG attachment, while at least some of the free form is surface exposed (Cowles *et al.*, 2011).

In *E. coli*, one-third of Lpp is covalently bound to the PG (Inouye *et al.*, 1972) via a C-terminal lysine residue (Figure 1.4.4.1). This residue is highly conserved among Lpp homologues and often present in a tyrosine-arginine-lysine motif. Mutating this residue results in loss of PG cross-linking (Zhang & Wu, 1992). The linkage occurs between the ϵ -amino group of the C-terminal lysine residue of Lpp and the carboxyl group of the meso-diaminopimelate (mDAP) of the PG tetrapeptide (Braun & Hantke, 2019) catalysed by three L,D-transpeptidases, LdtA, LdtB and LdtC (Magnet *et al.*, 2007).

It is not clear why cells express three different L, D-transpeptidases. Potentially, these enzymes may have specific roles under different conditions. The three enzymes are homologous with a conserved catalytic site (Magnet *et al.*, 2007). Among the three, LdtB is the most abundant in exponential phase, with estimates of 5000 molecules of LdtB per cell compared to 500 and 70 molecules per cell for LdtC and LdtA, respectively (Li *et al.*, 2014). Magnet *et al.* (2007) also showed that the formation of Lpp-PG cross-links is almost abolished in exponentially growing *E. coli* cells lacking LdtB.

In contrast to the bound form of Lpp, the role of the free form of Lpp remains a mystery. Cowles *et al.* (2011) using the selective surface biotinylation technique showed that for at least some of these free molecules, the C-terminus is surface exposed (Cowles *et al.*, 2011). How Lpp could cross the OM to reach the surface remains completely unknown. In terms of quantities, the total amount of free Lpp is twice that of PG-bound Lpp (Braun & Hantke, 2019). Asmar *et al.* (2018), hypothesised that free Lpp acts as Lpp reserve, maintaining a pool of free Lpp molecules, available for cross-linking, enabling modulation of PG-OM attachment levels during the cell cycle or along the cell axis (Asmar & Collet, 2018). In support of this hypothesis, bacterial cells growing under normal conditions, increase the number of covalent connections upon entry into stationary phase by 70% (Glauner *et al.*, 1988b). However, only a pool of free Lpp located in the periplasm would be available for this important function.

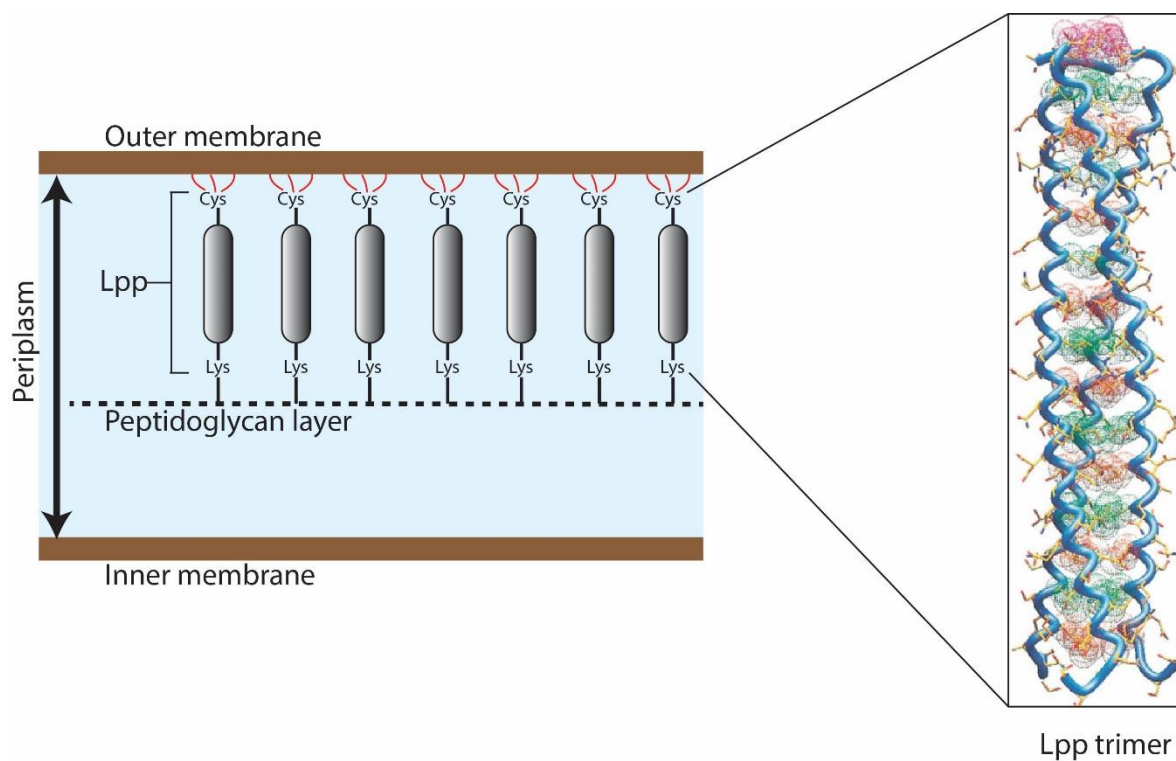


Figure 1.4.4.1: Lpp linkage to the outer membrane and peptidoglycan.

On the left, a cartoon depiction of Lpp attachment to both the outer membrane (OM) and the peptidoglycan (PG) layer. The attachment of a triacylglycerol moiety (red lines) to the N-terminus anchors Lpp in the OM. Lpp is also covalently bound to the mDAP residue in the peptide stems of PG via its invariant C-terminal lysine residue. On the right, a high-resolution crystal structure of the Lpp trimer (PDB: 1EQ7) (residues 2-56) comprising a parallel three-stranded coiled-coil. Red and green colours depict van der Waals surfaces of residues at *a* and *d* heptad positions, respectively, superimposed on the helix backbone.

1.4.5 Role of Lpp and other proteins in peptidoglycan attachment

While the high abundance of Lpp would seem to suggest that it is vital for the cell to attach the OM to the PG covalently, an *lpp* deletion mutant grows and divides like wild-type under standard laboratory conditions (Hirota *et al.*, 1977). It does, however, exhibit several phenotypes suggesting defects in OM integrity, including sensitivity to hydrophobic antibiotics, detergents like SDS, chelating agents such as EDTA, and relatively large antibiotics, i.e. vancomycin, as well as exhibiting enzyme leakage from the periplasm (Cascales *et al.*, 2002; Nichols *et al.*, 2011). The *lpp* mutant also forms OM blebs leading to hyper-vesiculation (Cohen *et al.*, 2017), suggesting that Lpp functions to stabilise the OM (Cohen *et al.*, 2017). Cells expressing an Lpp variant without the C-terminal lysine, or mutants with all three L, D-transpeptidases deleted, share most of the phenotypes of the *lpp* deletion mutant (Asmar *et al.*, 2017). This highlights the functional significance of the covalent attachment of Lpp to the PG layer.

The integral outer membrane protein A (OmpA), and peptidoglycan associated lipoprotein, Pal, are also crucial for anchoring the OM to the PG. However, for these two proteins, the interaction is non-covalent involving a strong hydrogen-bonded network with mDAP (Park *et al.*, 2012) (Figure 1.4.5.1). OmpA adopts a two-domain structure with an N-terminal eight-stranded β -barrel inserted in the OM and a C-terminal, globular domain present in the periplasm (Park *et al.*, 2012). A sub-population of Pal is engaged in the Tol-Pal trans-envelope complex that is involved in membrane constriction during cell division (Cascales *et al.*, 2002). Despite the periplasmic domain of OmpA and Pal having a low sequence similarity, they are structural homologs (Figure 1.4.5.2) that can bind to PG non-covalently (Egan, 2018). These non-covalent interactions are significant: Cascales *et al.* (2002), showed that Pal overexpression complements the sensitivity of *lpp* deletion mutant to SDS and vancomycin.

Other proteins also anchor to the PG layer non-covalently (Scheurwater & Burrows, 2011) (Figure 1.4.5.1B). Apart from the OmpA-type PG-binding domain displayed by OmpA and Pal, four other types of PG-binding domains exist. These are LytM-like PG-binding domain found in EnvC and NlpD (Uehara *et al.*, 2010; Tsang *et al.*, 2017), LysM-like PG-binding domain found in MltD (Bateman & Bycroft, 2000), SPOR PG-binding domain found in DamX (Gerding *et al.*, 2009), and PF01471 PG-binding domain displayed by EpsAB (Martynowski *et al.*, 2013). Despite these five well-characterised PG binding domains, proteins like InvH,

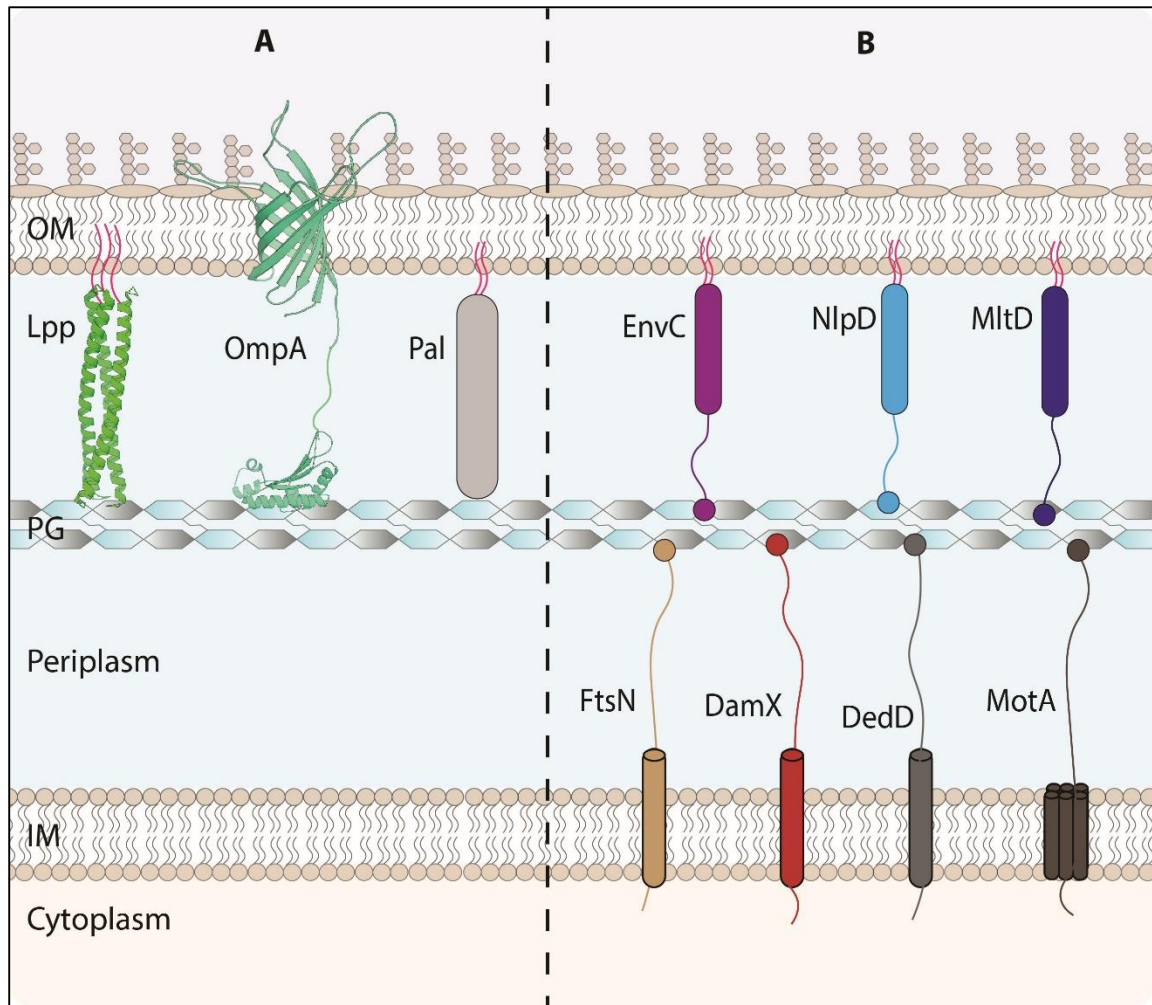


Figure 1.4.5.1: Major and minor peptidoglycan-membrane connections.

(A) The major connections of the outer membrane (OM) to the peptidoglycan (PG). OM is tethered to the PG by Lpp, OmpA and Pal. Lpp provides the only covalent connection to the PG layer (green trimer) (PDB: 1EQ7). Next to Lpp is a ribbon diagram of OmpA (teal) (PDB: 2GE4) and a cartoon depiction of Pal (lilac). Both proteins are tethered to the PG non-covalently via a similar OmpA-type PG-binding domain, depicted as a ribbon diagram (PDB: 4G4Z) for OmpA. (B) Other minor connections between the PG to either the OM or the inner membrane (IM). These proteins harbour unique PG binding domains. As an example, EnvC contains LytM-like PG-binding domain, MltD has a LysM-like PG-binding domain, and DamX has a SPOR PG-binding domain

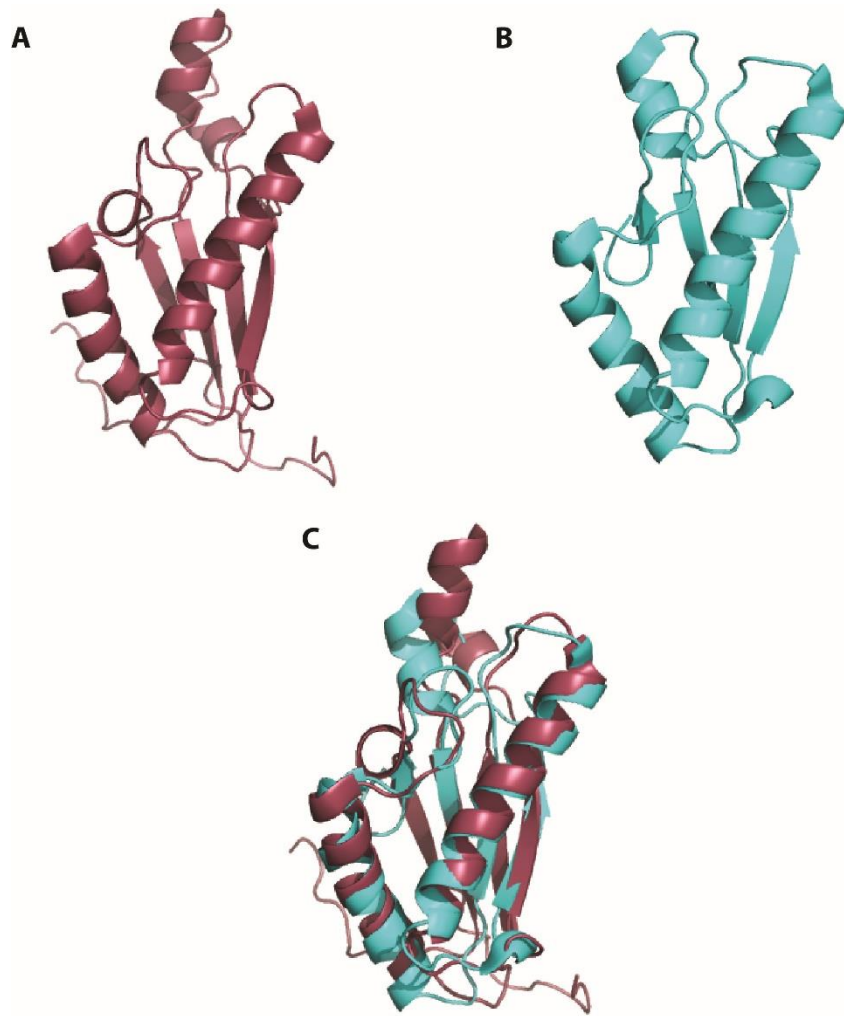


Figure 1.4.5.2: OmpA and Pal are structural homologues.

(A) Crystal structure of the PG-binding domain of the β -barrel protein OmpA from *E. coli* (PDB 2MQE residues 180 - 325). (B) Crystal structure of full-length Pal from *E. coli* (PDB 10AP). (C) A superimposition of the two crystal structures, OmpA (maroon) and Pal (turquoise) show that they are structural homologues, despite having a low sequence similarity.

PrgH and PrgK have been reported to harbour as yet uncharacterised PG binding domains (Pucciarelli & Garcia-del Portillo, 2003). These non-covalent PG interactions are presumed to couple PG biogenesis to other cellular processes, e.g. EnvC and NlpD are divisome associated proteins shown to activate PG hydrolases (Uehara *et al.*, 2010; Tsang *et al.*, 2017) and thus couple cell division to PG recycling and turnover. Others are implicated in structural roles for mobility and secretion systems (Scheurwater & Burrows, 2011).

1.4.6 Peptidoglycan bound Lpp is a distance keeper

Recently, studies performed in *Salmonella enterica* (Cohen *et al.*, 2017) and *E. coli* (Asmar *et al.*, 2017) showed that the distance between the OM and the IM was proportional to the length of Lpp variants: insertions of 14 amino acids (Lpp₊₁₄) or 21 amino acids (Lpp₊₂₁) in Lpp were observed to increase the distance between the OM and IM by 3 and 4 nm, respectively (Figure 1.4.6.1). Both studies used cryoelectron microscopy (cryo-EM), a technique in which cells are rapidly frozen and hence preserved as close as possible to their near-native state to measure these nanoscale changes in distance.

In the first study, Cohen *et al.* (2017) identified LppA as a suppressor mutant of FlgG non-motile mutants. FlgG is a component of the rod section of the basal-body assembly of the flagellar motor (Jones *et al.*, 1990). In *S. enterica*, the length of a wild-type flagellum rod is 25 nm (Cohen *et al.*, 2017). Mutations mapping to the *flgG* gene results in longer rods of about 60 nm, and the mutants with longer rods are non-motile. The authors speculated that the placement of the rod tip was no longer perpendicular to the OM, compromising PL-ring assembly and resulting in a non-motile phenotype. Selection for motile suppressor mutants of the FlgG non-motile mutants identified a deletion of *lppA*, which resulted in a significant increase in swim diameter on agar. The authors hypothesised that the deletion of LppA would allow the longer rods of the FlgG mutant to be positioned perpendicular to the OM, which would allow PL-ring pore formation. The authors confirmed the hypothesis when purified rods of LppA lacking 21 amino acids were shorter than those of the LppA wild-type. The same applied for purified mutants of Lpp increasing length variants (+21, +42, +63) which had longer rods. The authors concluded that, during flagellar biogenesis, the flagellum complex senses when it hits the OM and stops growing and LppA determines the distance between the OM and the IM (Cohen *et al.*, 2017).

In the second study, Asmar *et al.* (2017) used the same extended forms of Lpp to show that increasing the length of Lpp resulted in the inability of the OM stress sensor lipoprotein RcsF

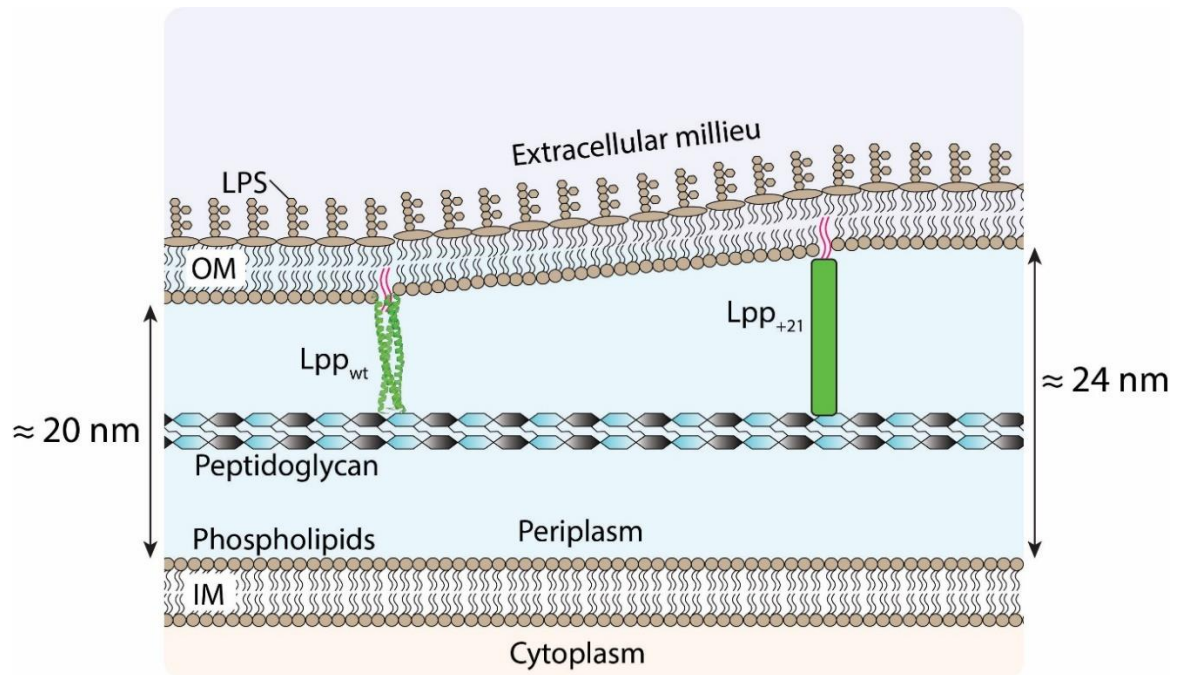


Figure 1.4.6.1: Peptidoglycan-bound Lpp is a distance keeper.

PG-bound lpp determines the outer membrane (OM) to inner membrane (IM) distance. Lpp (green trimer) (PDB: 1EQ7) maintains the periplasmic width at ≈ 20 nm (left side). Upon increase in Lpp length (Lpp₊₂₁), depicted by the green rectangle, the periplasmic width increases by ≈ 4 nm (right side) (Cohen *et al.*, 2017) (Asmar *et al.*, 2017). The authors showed that the increase in distance was proportional to the length of Lpp using Lpp₊₁₄, not shown, which increased the OM-IM distance by ≈ 3 nm.

to transfer stress signals to IgaA, its downstream partner in the IM. The implication is that the increase in OM to IM distance prevents RcsF from reaching its partner. The result is the inactivation of the regulation of the capsule synthesis stress response system when PG damage occurs. Remarkably, when the authors increased the length of RcsF, to compensate for the longer OM to IM distance, efficient signalling was restored (Asmar *et al.*, 2017). These two studies demonstrated the importance of Lpp as a crucial determinant of the OM to IM distance.

1.5 Lipopolysaccharide biogenesis

1.5.1 Lipopolysaccharide structure and function

Rietschel *et al.* (1994) suggested that 1-3 million molecules of LPS must be assembled on the bacterial surface per generation (Rietschel *et al.*, 1994). This reported amount makes LPS the major surface component of the OM, occupying three-quarters of the bacterial surface area with the remaining surface area occupied by proteins (Nikaido, 1976). LPS molecules have an amphiphilic architecture with a lipophilic moiety, termed lipid A, and a hydrophilic moiety composed of core oligosaccharide and the O-antigen components (Sperandeo *et al.*, 2017) (Figure 1.5.1.1).

While not all bacterial lineages have LPS, lipid A is conserved in all Gram-negative bacteria and is the primary lipid component of the OM outer leaflet (Sweeney & Lowary, 2019). It is also the component responsible for activating the host immune response (Park & Lee, 2013). The core oligosaccharide connects lipid A to the O-antigen and shows more variability across Gram-negative bacteria (Sweeney & Lowary, 2019). The O-antigen is hyper-variable and an essential virulence determinant (Clarke *et al.*, 1995) responsible for resistance to complement-mediated serum killing (Joiner, 1988).

LPS forms a rigid lipid monolayer. The rigidity is in part due to the dense packing of the acyl groups attached to lipid A and buried in the hydrophobic bilayer (Wu *et al.*, 2013), and enhanced by the chelation of divalent cations by the negatively charged groups in both lipid A and the oligosaccharide. The tight packing of LPS in the OM makes it a formidable permeability barrier, protecting the cell from the entry of many toxic compounds, including many of the clinically useful antibiotics (Nikaido, 2003). Beyond these well-established roles, Rojas *et al.* (2018) showed that this LPS layer contributes to the overall cell stiffness, such that when LPS biosynthesis is compromised, the cell is more susceptible to mechanical stress such as turgor pressure (Rojas *et al.*, 2018). This study was the first to show that the OM also protects the cell from mechanical stress and not just chemical stresses.

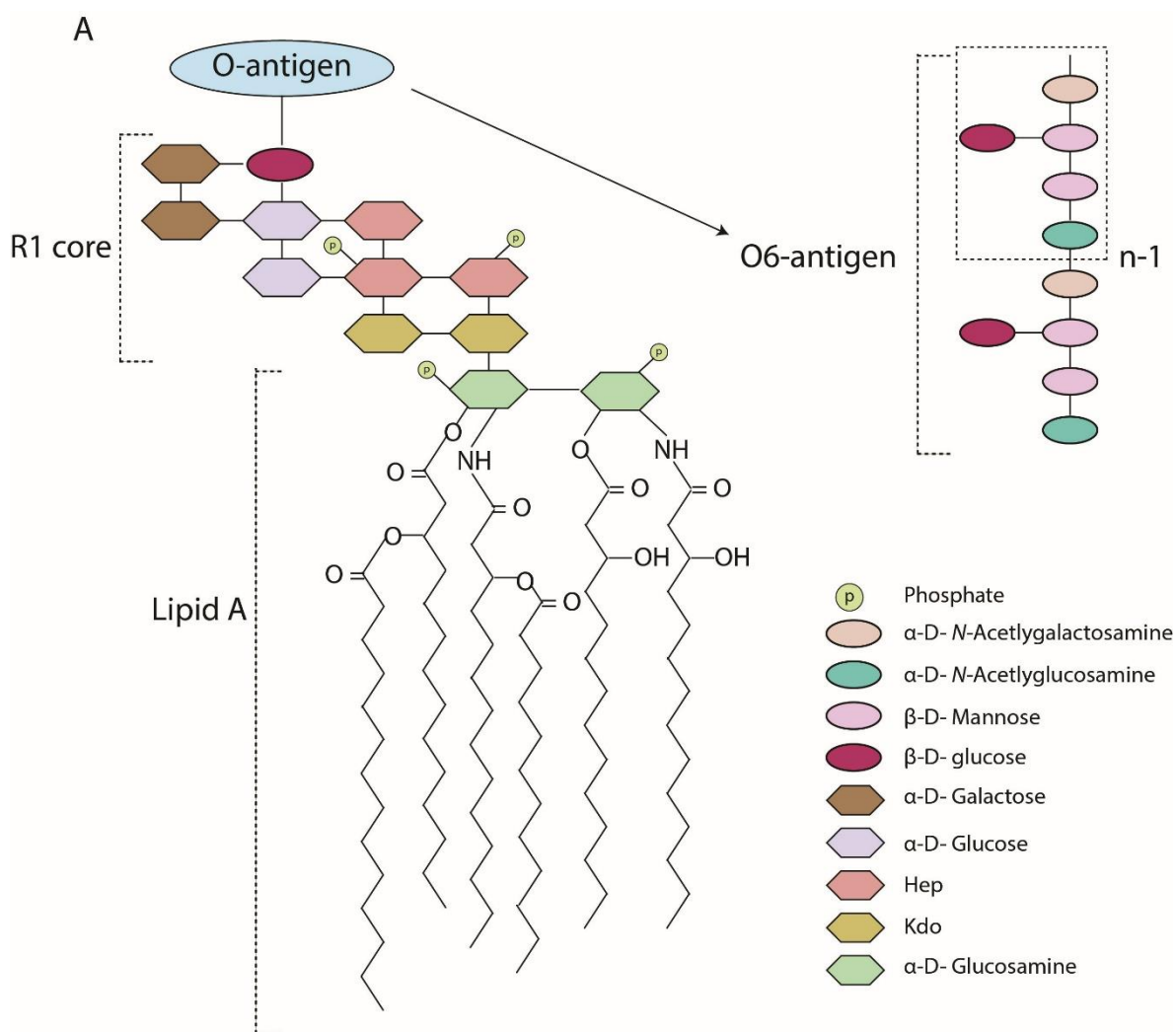


Figure 1.5.1.1: Chemical structure of *E. coli* lipopolysaccharide. Smooth-type LPS consists of three distinct domains: lipid A, inner and outer core oligosaccharides, and the highly variable O-antigen repeat. Of the five core oligosaccharides and almost 170 O-antigen variations, the R1 core and the O6 antigen depicted here are highly prevalent within pathogenic *E. coli* (Gibbs *et al.*, 2004; Stenutz *et al.*, 2006). The number of O-antigen repeating units ('n') varies between 10 and 25 (Stenutz *et al.*, 2006). Cartoon depictions of the various subunits that make up the core and the O-antigen are shown in the legend. Abbreviations: Kdo, 3-deoxy-D-manno-oct-2-ulonic acid; Hep, L-glycero-D-manno-heptose.

Bacteria with a complete LPS are said to have a ‘smooth LPS’ (Figure 1.5.1.1). However, several exceptions exist in LPS structure. Some bacteria lack the O-antigen and are said to have a ‘rough LPS’ while others only have the lipid A component, said to have a ‘deep rough LPS’ (Friedrich & Whitfield, 2005). Most laboratory strains of *E. coli* do not synthesise the O-antigen, due to a mutation in the *rfb* locus, cluster of genes responsible for the O-antigen biosynthesis, and hence have a rough LPS (Chart *et al.*, 2000). Lipid A had long been regarded as an essential structural component necessary to build the OM, but the viability of LPS-null mutants in *Neisseria meningitidis* (Steehgs *et al.*, 1998) and *Moraxella catarrhalis* (Peng *et al.*, 2005) with mutations in the *lpxA* gene challenged this view. Colistin-resistant *Acinetobacter baumannii* mutants unable to produce LPS due to a mutation in LPS biosynthetic genes are also viable (Moffatt *et al.*, 2010). These mutants can survive with an OM comprised entirely of phospholipids, in which the cell upregulates a range of cell envelope and membrane biogenesis machineries to compensate for the absence of LPS.

1.5.2 Lipopolysaccharide biosynthesis

LPS biosynthesis takes place in three distinct cellular compartments: the cytoplasm, the IM and the periplasm. It involves several independent pathways that converge in an ordered assembly to give a mature molecule. Lipid A and the inner core, are synthesised in the cytoplasm via the Raetz pathway (Sperandeo *et al.*, 2017), involving a series of enzymes, some of which are essential targets of drugs that compromise LPS biosynthesis, e.g., LpxC. The core-lipid A complex, anchored to the IM with its hydrophilic moiety exposed to the cytoplasm is then flipped across the IM by the ABC transporter MsbA and thus becomes exposed to the periplasm (Polissi & Georgopoulos, 1996).

In parallel, O-antigens are synthesised in the cytoplasm and flipped to the periplasmic side of the IM (Polissi & Sperandeo, 2014). The extension of the O-polysaccharide chain may occur in the periplasm via Wzy-dependent, ATP-dependent or synthase-dependent pathways (Greenfield & Whitfield, 2012). The synthase-dependent pathway is not ubiquitous, with no known *E. coli* O-antigens following this pathway. The final step in the LPS biosynthesis pathway is the convergence of both pathways by the ligation of O-antigen to the core-lipid A at the periplasmic leaflet of the IM, catalysed by WaaL ligase to form the mature LPS (Han *et al.*, 2012).

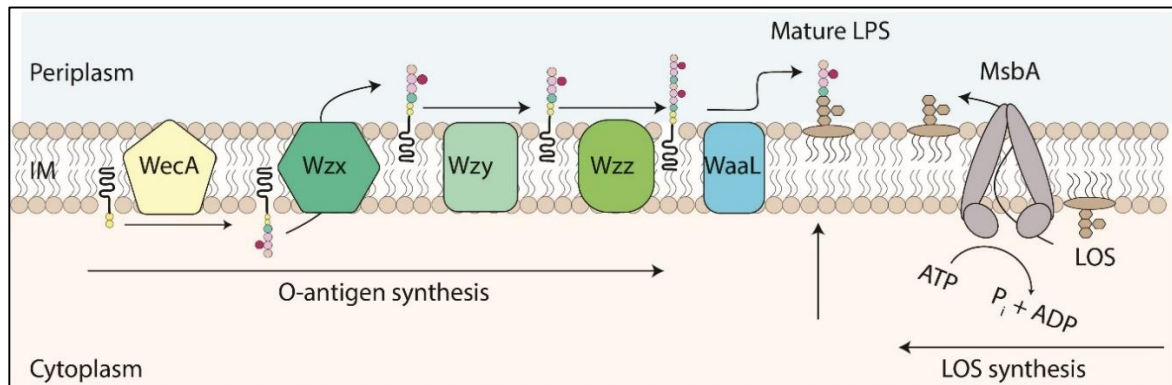


Figure 1.5.2.1: Lipopolysaccharide biosynthesis in *E. coli*.

Arrows pointing towards the right show the steps involved in the synthesis of the O-antigen via the Wzy-dependent pathway. Wec initiates O-antigen synthesis by transferring an *N*-acetylhexosamine sugar onto undecaprenyl diphosphate (UND-PP), a universal glycan lipid carrier (Stenutz *et al.*, 2006). Flippase, Wzx, flips the precursor O-antigen (UND-PP-oligosaccharide) to the outer leaflet of the inner membrane (IM) (Marolda *et al.*, 2010). Wzy then polymerises oligosaccharide subunits to a length determined by Wzz (Marolda *et al.*, 2010). Arrow pointing towards the left depict a nine-step enzymatic process for the assembly of lipid A (the Raetz pathway) (Whitfield & Trent, 2014a). The core oligosaccharide which connects lipid A to the O-antigen is also synthesised in the cytoplasm. After the lipid-core oligosaccharide (LOS) is assembled, it is flipped by the ABC transporter MsbA (Zhou *et al.*, 1998) to the outer leaflet of the IM. Upward facing arrow shows the final biosynthesis step involving the transfer of the entire O-antigen by WaaL to LOS producing a mature ‘smooth’ LPS molecule, which is subsequently transferred to the Lpt machinery for OM translocation.

1.5.3 Molecular mechanism of LPS translocation to the outer membrane

The translocation of LPS, either ‘smooth’ or ‘rough’, across the periplasm to the OM, is carried out by the lipopolysaccharide transport (Lpt) machinery composed of seven Lpt components (LptB₂FGCADE) (Whitfield & Trent, 2014b). With species-dependent exceptions, all seven of the Lpt proteins are essential for Gram-negative bacterial survival (Hicks & Jia, 2018). The Lpt machinery is divided into three subassemblies: LptB₂FGC, LptA, and LptDE, located in the IM, the periplasm, and at the OM, respectively (Figure 1.5.3.1). LptB₂FG constitutes the ABC transporter that powers the transport of LPS. The LPS is extracted from the IM by LptB₂FG and passed on to LptC, an auxiliary IM protein that strongly associates with LptB₂FG and interacts with LptA. LptA acts as a bridge by connecting the IM subcomplex to the OM subcomplex. The LptDE subcomplex is composed of the β -barrel protein LptD and the lipoprotein LptE, and this subcomplex receives LPS from LptA and is responsible for the final stage of assembly into the cell surface (Polissi & Sperandio, 2014).

Okunda *et al.* (2016) likened the Lpt system to a “PEZ” candy dispenser, where a spring at the bottom of the candy-filled chamber, in this case, the LptB₂FGC subcomplex, slides the contents (LPS) up as candies are removed from the top one at a time. The model means that there is a continuous passage of LPS from the previous subunit to the next until it reaches the OM

(Okuda *et al.*, 2016) (Figure 1.5.3.1). However, this model does not fully account for how the individual Lpt subunits bind and release LPS.

LptA function and relative placement in the Lpt system is understood, but its stoichiometry *in vivo* remains uncertain. The crystal structure of LptA by Suits *et al.* (2008) showed an end-to-end fibrous tetramer arrangement forming a continuous hydrophobic groove between the LptA monomers (Suits *et al.*, 2008). Mass spectrometry analysis confirmed the observations, showing that LptA forms a range of oligomers, from dimers to pentamers, in a concentration-dependent manner when purified and that LPS stabilises the resultant complexes (Santambrogio *et al.*, 2013).

Hicks & Jia (2018) speculate that LptA oligomerisation allows the Lpt system to tolerate changes in the intermembrane space that result from osmotic stress. The authors speculate that a single LptA complex can bridge a typical *E. coli* intermembrane distance, but the size of the periplasm could vary between species and in response to osmotic stress (Hicks & Jia, 2018).

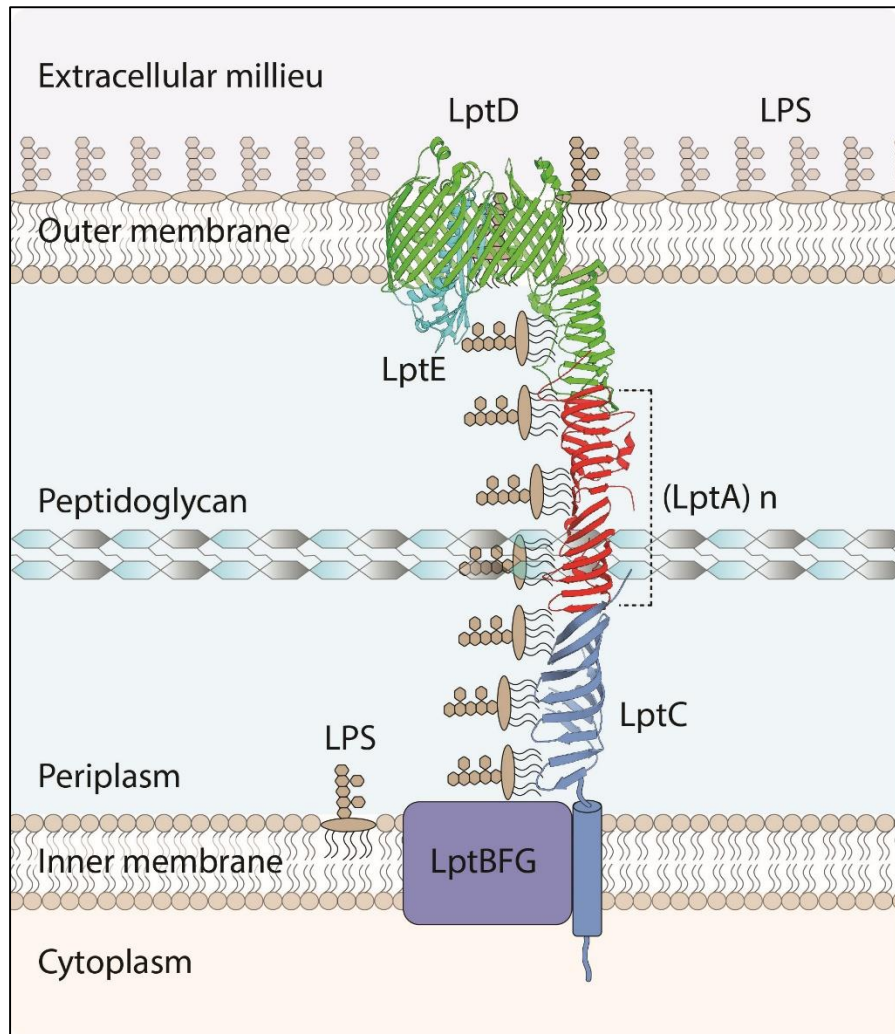


Figure 1.5.3.1: PEZ Model for lipopolysaccharide translocation across the periplasm.

The LPS translocation system forms a continuous bridge across the inner membrane (IM), periplasm and outer membrane (OM). The LptB₂GF complex, depicted as a purple rectangle, extracts LPS from the outer leaflet of the IM through an ATPase-depended mechanism. LptC (PDB: 3MY2), LptA (PDB: 2R19), and LptD form a continuous hydrophobic groove along which LPS transverses. LptDE (PDB: 4Q35) sorts LPS to the outer leaflet of the OM. Although LptA (red) is depicted as a dimer, the exact number of subunits *in vivo* remains uncertain. Hicks & Jia (2018) hypothesised that LptA might vary in terms of subunit number in response to increased LptA availability or changes in periplasmic size.

Functional analysis of the promoters associated with LptA expression shows additional responsiveness to changes in LPS biogenesis (Martorana *et al.*, 2011), unlike the other Lpt components. The authors concluded that LptA might vary in terms of subunit number in response to increased LptA availability, or by the disruption of a single-subunit complex upon the expansion of the periplasm (Hicks & Jia, 2018).

1.6 Outer membrane protein biogenesis

The protein components of the OM are also essential to membrane function. As much as 3% of a Gram-negative bacterial genome may encode OM proteins (Wimley, 2003). Outer membrane protein (OMP) synthesis takes place in the cytoplasm, followed by a multistage translocation process that minimises OMP aggregation maintaining a competent state for insertion into the OM (Weiner & Li, 2008; Selkrig *et al.*, 2014; Rollauer *et al.*, 2015). The process presents a significant challenge to the cell: OMPs must transverse the hydrophobic IM, but also the aqueous crowded periplasmic space. The cell overcomes these challenges through the use of soluble chaperones and membrane-embedded machines, e.g., the SecYEG and the β -barrel assembly machinery (BAM) (Selkrig *et al.*, 2014; Rollauer *et al.*, 2015; Leyton *et al.*, 2015; Plummer & Fleming, 2016) (Figure 1.6.1.1).

OMPs are integral membrane proteins which adopt a β -barrel conformation with antiparallel β -strands connected via short “turns” facing the periplasmic side and large “loops” on the extracellular side (Koebnik *et al.*, 2000; Fairman *et al.*, 2011). The β -strands interact via extensive hydrogen bonding in which the first and the last β -strand interact to form a closed cylindrical structure called a β -barrel, with a hydrophilic lumen and a hydrophobic exterior. This architecture presents a challenge in the folding of nascent OMPs before insertion into the membrane. To date, all solved crystal structures of OMPs from Gram-negative bacteria contain an even number of antiparallel β -strands (Fairman *et al.*, 2011). For monomeric β -barrels, the strand number ranges from 8 β -strands, e.g., OmpX (Vogt & Schulz, 1999) to 26 β -strands, e.g., LptD (Wu *et al.*, 2006).

1.6.1 Outer membrane protein translocation across the inner membrane and transport across the periplasm

OMP precursors (pre-OMPs) contain an intrinsic N-terminal targeting sequence required for translocation across the IM. This signal sequence is a stretch of amino acids with a tripartite structure containing a positively charged N-terminal region, a hydrophobic stretch (typically 5-16 amino acids) and a hydrophilic segment that contains a protease cleavage site (Hegde &

Bernstein, 2006; Driessen & Nouwen, 2008). The pre-OMPs are delivered to the SecYEG by one of two well-documented mechanisms: either co-translationally during their synthesis on the ribosome (Matlack *et al.*, 1998) or post-translationally by the use of the SecA and SecB chaperones which maintain the protein in an unfolded conformation (Economou & Wickner, 1994) (Figure 1.6.1.1). By either mechanism, the pre-OMPs are targeted to the SecYEG, which mediates translocation across the IM (Driessen & Nouwen, 2008). The translocon is ATP-dependent, powered by the ATPase domains of SecA (Kudva *et al.*, 2013). During translocation, the signal sequence is cleaved from the pre-OMPs by signal peptidase I (SPI) liberating mature OMP polypeptide (Paetzel *et al.*, 2002; Hegde & Bernstein, 2006).

Upon emerging from the SecYEG, the unfolded OMPs encounter the periplasm, an aqueous environment in which they would be prone to aggregation. To avert this, periplasmic chaperones bind nascent OMPs until their eventual folding and insertion into the OM (Figure 1.6.1.1). The three best characterised and important periplasmic chaperones present in *E. coli* are Survival Protein A (SurA), Seventeen kilodalton protein (Skp) and DegP (Rollauer *et al.*, 2015; Plummer & Fleming, 2016). Single deletions of any of these three chaperones are viable, but double deletions of *surA* and either *skp* or *degP* show a synthetic lethal phenotype (Rizzitello *et al.*, 2001). Taken together, several studies propose two distinct pathways for nascent OMP passage across the periplasm, the SurA-mediated pathways and the Skp/DegP-mediated pathway, in which at least one pathway must be functional for bacterial cell viability (Rizzitello *et al.*, 2001; Plummer & Fleming, 2016) (Sklar *et al.*, 2007b). According to this model, SurA plays the major role in OMP biogenesis, while Skp/DegP act together in a separate secondary pathway (Rizzitello *et al.*, 2001). Apart from chaperone activity, DegP also exhibits protease activity. Spiess *et al.* (1999) showed that DegP could switch between protease and chaperone activity during the transition from low to high temperatures (Spiess *et al.*, 1999). However, recent data (Ge *et al.*, 2014) suggest that DegP functions primarily as a protease and has very little chaperone activity either at high and low temperatures. DegP, therefore, remains somewhat enigmatic in biogenesis.

1.6.2 The β -barrel assembly machinery

Nascent OMPs are assembled into the OM by conserved multicomponent complex termed the β -barrel assembly machinery (BAM) (Knowles *et al.*, 2009; Selkrig *et al.*, 2012; Selkrig *et al.*, 2014; Noinaj *et al.*, 2015; Noinaj *et al.*, 2017). It is generally accepted that the BAM complex accepts substrates from periplasmic chaperones and catalyses their insertion into the OM. In *E. coli*, the BAM complex consists of BamA, an OMP itself, and four lipoproteins

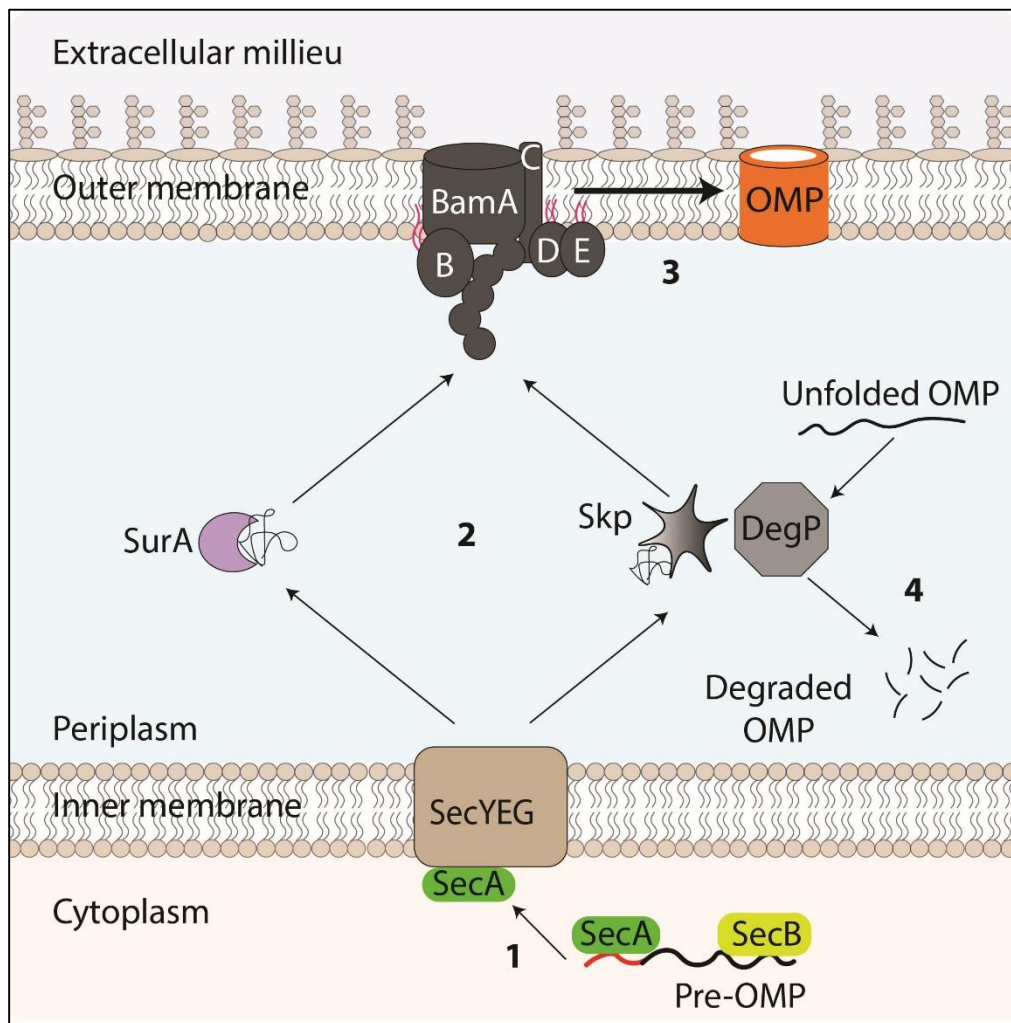


Figure 1.6.1.1: Outer membrane biogenesis.

The figure portrays the journey of outer membrane proteins (OMPs) from the cytoplasm to the outer membrane (OM). **(1)** Nascent OMP (pre-OMP) (black) is synthesised in the bacterial cytoplasm and delivered by SecA and SecB chaperones to the SecYEG inner membrane (IM) translocation system. At the IM, the pre-OMP is processed where its N-terminal signal sequence (red) is cleaved then released into the periplasm. **(2)** The nascent OMP then takes either SurA or the Skp/DegP chaperone pathway for transit to the β -barrel assembly machinery (BAM) complex in the OM. **(3)** The BAM complex then catalyses the OMP folding into the OM. **(4)** If the OMP misfolds and cannot be rescued by Skp, it is degraded by protease DegP to prevent toxicity to the cell.

named BamB, BamC, BamD and BamE, all anchored to the periplasmic leaflet of the OM (Hagan *et al.*, 2011; Webb *et al.*, 2012) (Figure 1.6.2.1). Of the five components, BamA and BamD are essential for cell viability, while single deletions of the other components are non-lethal (Wu *et al.*, 2005). However, the single deletions are reported to negatively impact OM integrity and efficiency of OMP assembly due to an overall decreased catalytic activity (Charlson *et al.*, 2006; Hagan *et al.*, 2010). BamB and BamD interact directly with BamA, while BamC and BamE interact directly with BamD (Malinverni *et al.*, 2006; Kim *et al.*, 2007). This interaction map (Figure 1.6.2.1) is based on crystal structures (Albrecht & Zeth, 2011; Kim *et al.*, 2011a; Gu *et al.*, 2016) and data for detergent solubilisation of the native complex from the OM of *E. coli*, which yields a BamA:B module and a distinct BamC:D:E module (Anwari *et al.*, 2010; Hagan *et al.*, 2010).

From various crystal structures, the N-terminal domain of BamA is seen to have a large periplasmic domain consisting of five polypeptide transport-associated (POTRA) segments followed by a C-terminal membrane-inserted β -barrel with 16 antiparallel β -strands (Noinaj *et al.*, 2013). The POTRA domains form a scaffold for interacting with BamB and BamD and are also thought to play a role in the initial recognition of substrates (Kim *et al.*, 2007). BamA has an unusual β -barrel, displaying a reduced barrel height on the side where strands $\beta 1$ and $\beta 16$ join (Noinaj *et al.*, 2013). This feature has been suggested to enable the opening and closing of the BamA lateral gate (Noinaj *et al.*, 2013; Noinaj *et al.*, 2014). Following the gate opening, BamA is thought to use its terminal β -strand as a template for substrate OMP β -strand assembly, a process termed β -augmentation (Noinaj *et al.*, 2014; Estrada Mallarino *et al.*, 2015). One current model posits that BamA distorts the OM leading to local thinning of the lipid bilayer, it then recognises the substrate via the β -signal, which enters the lumen of BamA. The substrate's extracellular domains may translocate across the OM through the hydrophilic pore, whereas the transmembrane portion laterally enters the disordered regions of the lipid bilayer and assembles via β -augmentation (Noinaj *et al.*, 2014). There remains considerable debate over whether this model best represents the mechanism of OMP assembly in the OM.

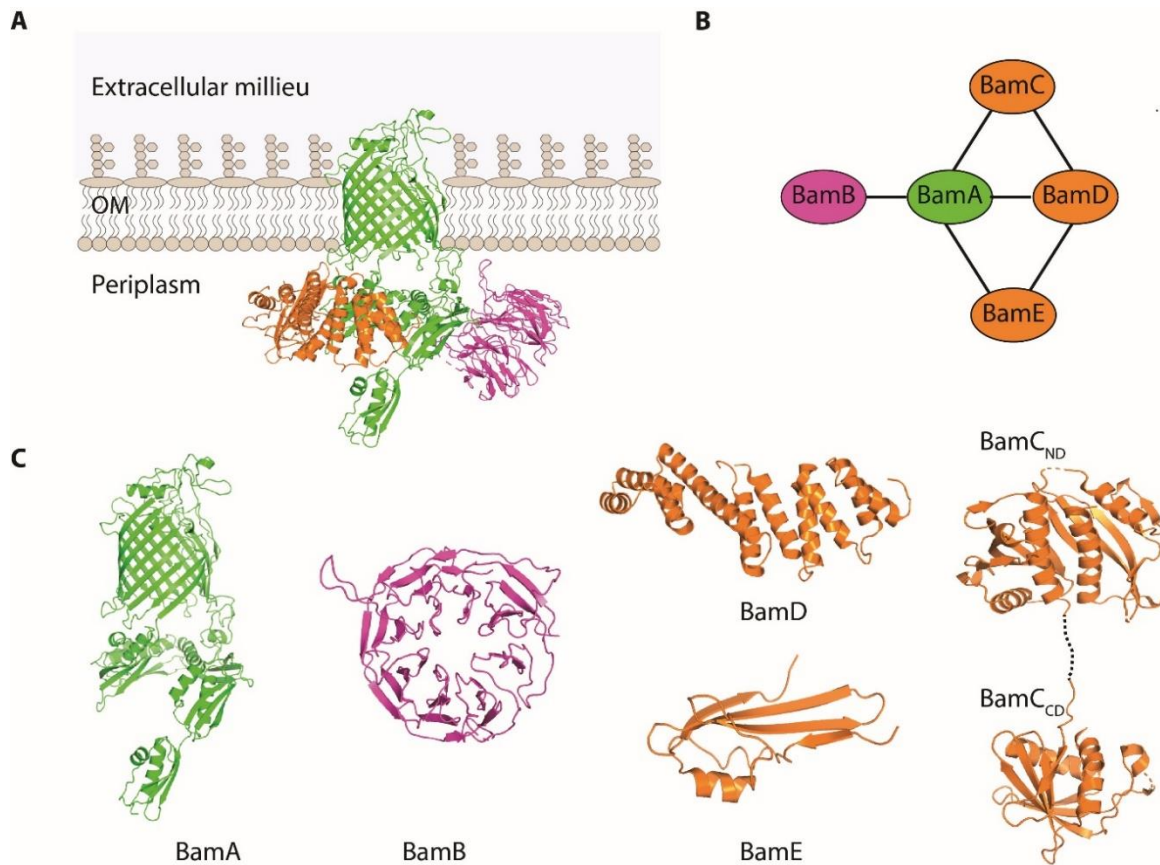


Figure 1.6.2.1: Structures of the *E. coli* BAM complex components and interaction map.

(A) Side view of the ribbon diagram of the BAM complex (PDB: 5D0O) with all five components modelled on the Gram-negative outer membrane (OM). (B) The interaction map showing the direct binding between BAM subunits determined through structural and experimental studies. (C) Ribbon diagrams of individual BAM components where the colours match the ribbon diagram of BAM complex in panel A (PDB: 5D0O), except BamC (PDB: 2YH6 and 3SNS for its N- and C- terminals, respectively).

1.6.3 BAM lipoproteins

Of the four BAM complex lipoproteins, BamD is the only essential component. Its essentiality can be attributed to its function in maintaining a stable interaction of BamA with the BamC:D:E subcomplex (Malinverni *et al.*, 2006; Sklar *et al.*, 2007a). BamD significance is further shown using depletion studies, where its depletion results in lower steady-state levels of OMPs as judged by immunoblotting studies of bacterial membrane extracts (Wu *et al.*, 2005; Malinverni *et al.*, 2006). Structural studies show that BamD is comprised of five tetratricopeptide (TPR) repeats (Figure 1.6.2.1C): 3 TPR repeats in the N-terminus and 2 TPR repeats in the C-terminus (Dong *et al.*, 2012a). The N-terminal TPR domains have been suggested to interact with substrates and, interestingly, with BamC (Malinverni *et al.*, 2006; Albrecht & Zeth, 2011; Kim *et al.*, 2011a) while the C-terminal TPR domains have been shown to interact with BamA, and BamE (Gu *et al.*, 2016; Han *et al.*, 2016).

Of the three nonessential BAM lipoproteins, the absence of BamB produces the most significant OMP biogenesis defect (Charlson *et al.*, 2006; Sklar *et al.*, 2007a; Ureta *et al.*, 2007). The defect is further exacerbated when BamC or BamE are also simultaneously absent (Sklar *et al.*, 2007a) with the double deletion mutant $\Delta bamB \Delta bamE$ displaying the most drastic growth and OMP biogenesis defects (Tellez & Misra, 2012). BamB is a WD40 protein and adopts an eight-bladed β -propeller structure with four β -strands per blade (Dong *et al.*, 2012b) (see Figure 1.6.2.1C). Structural and biochemical studies show that BamB interacts directly with POTRA domain 3 of BamA (Vuong *et al.*, 2008; Heuck *et al.*, 2011). Using super-resolution imaging and in situ cross-linking, Gunasinghe *et al.*, (2018) showed this interaction was required in mediating interactions between neighbouring BAM complexes, forming OMP assembly precincts across the bacterial cell surface.

BamC is intriguing due to its exposure on the surface of the bacterial cell (Webb *et al.*, 2012; Gunasinghe *et al.*, 2018). However, its role as part of the BAM complex is still unclear. Unlike *bamB*, deletion of *bamC* does not affect OMP levels but still displays OM permeability defects. (Onufryk *et al.*, 2005; Wu *et al.*, 2005). However, double deletion of *bamC* in combination with *bamB* or *bamE* results in lower levels of BamD (Wu *et al.*, 2005; Malinverni *et al.*, 2006; Rigel *et al.*, 2012). BamC is composed of three distinct domains. The first 50 amino acids make up the unstructured N-terminus of the mature protein, and the remainder of the polypeptide folds into two approximately 10 kDa globular domains (Kim *et al.*, 2011a) (see Figure 1.6.2.1C). Kim *et al.* (2011) solved the crystal structure of the BamC:D complex and showed that the two proteins interact via the N-terminal unstructured domain. The interaction of BamC

with BamD in the same region that BamD interacts with substrate OMP has prompted a suggestion that BamC may play a role in regulating BamD activity (Kim *et al.*, 2011a; Misra *et al.*, 2015).

BamE, with a molecular weight of 10.4 kDa, is the smallest of the four BAM complex lipoproteins. It is monomeric, consisting of two N-terminal α -helices and a C-terminal β -sheet containing three β -strands (Endo *et al.*, 2011) (Figure 1.6.2.1C). BamE plays a significant role in stabilising the interaction of BamD and BamA (Sklar *et al.*, 2007a; Rigel *et al.*, 2012). Rigel *et al.* (2012) further showed that BamE either directly or indirectly modulates the conformation of BamA during its functional cycle. The authors showed this using protease sensitivity assays, where BamA was degraded beyond detection in $\Delta bamE$ mutant cells treated with proteinase K.

1.7 Genome-scale genetic manipulation for exploring bacterial biology

The influx of whole genome sequences is staggering: as of February 1st, 2020, the National Centre for Biotechnology Information database (NCBI) records a total of 16,658 complete bacterial genome sequences ("Database resources of the National Center for Biotechnology Information," 2018). The data from genome sequencing has contributed immensely to our understanding of bacterial cell biology. However, this information provides a new challenging task on how to think about biological problems. Bacterial databases based on the available information are full of entries such as 'hypothetical gene', leaving hundreds of genes without assigned functions. For example, 12 years ago, roughly 35% of protein-coding genes of *E. coli* lacked experimental annotations (Karp *et al.*, 2007). 12 years later, the figure is still 35% (Ghatak *et al.*, 2019). The situation is worse for bacterial species less-studied than *E. coli*.

Recent advances in high-throughput genetic manipulation and screening approaches can address this gap in functional information. Historically, forward genetics, i.e., screening a mutation from a pool of mutants based on assayed phenotypes then mapping the mutants alleles (Shuman & Silhavy, 2003), has been very successful. However, with the advancement of gene modification and sequencing technologies, reverse genetics has proven more useful than previously imagined. Reverse genetics entails modification of specific gene segments, e.g., gene deletion, then assaying the effect of each perturbation using fitness or functional assays to determine the roles of query genes or even operons (Shuman & Silhavy, 2003).

Advancement of the λ "Red" recombineering technique (Datsenko & Wanner, 2000) that exploits the use of the bacteriophage λ "Red" proteins has significantly improved the efficiency of targeted mutagenesis. Baba *et al.* (2006) used the technique to construct the *E. coli* Keio

single-gene deletion library by replacing each target open reading frame with a selectable marker. The collection, now commercially available, is widely used and has been instrumental in many studies of gene functions in *E. coli*.

1.7.1 Gene function investigation by mutation combinations

Baba *et al.*, (2006) reported that only 303 of the 4453 protein-coding genes in *E. coli* are essential for growth under standard laboratory conditions. The low percentage of essential genes is striking and suggests both a high degree of gene redundancy and that standard laboratory growth is not taxing to *E. coli*. Genetic buffering (Figure 1.7.1.1A) complicates the study of gene functions where deletion of a redundant gene gives no phenotype due to the buffering effect of another gene on a different pathway that contributes to the same essential process (Babu *et al.*, 2011). In this case, mutations on both genes may be necessary to reveal the participation of each gene in a biological process that is subject to extensive functional redundancy. Knowledge of gene interactions can also reveal genetic interaction networks between genes that function within different pathways, providing an opportunity to investigate pathway inter-connections (Gagarinova & Emili, 2012) (Figure 1.7.1.1A).

Phenotypic comparison between single and double mutants is used to assess genetic interactions (Figure 1.7.1.1B). The commonly used phenotype is fitness, based on the simplicity of acquiring comparative measurements. Microbial fitness can be measured spectrophotometrically in case of liquid cultures or by colony size quantification on solid media (Butland *et al.*, 2008; Gagarinova & Emili, 2012). When mutations in separate genes are combined, and the resulting mutant grows better, the two genes are said to be ‘alleviating’. In contrast, if the double mutant grows worse, the two genes are said to be ‘synthetic sick’ or ‘synthetic lethal’, i.e., an ‘aggravating interaction’ (Butland *et al.*, 2008; Typas *et al.*, 2008). Large-scale genetic interaction studies use a multiplicative model (Butland *et al.*, 2008; Typas *et al.*, 2008; Babu *et al.*, 2014) based on this principle. According to the model (Figure 1.7.1.1B), most genes are likely to be functionally unrelated; the fitness of the double mutants is expected to be equal to the product of each corresponding single mutant fitness. An aggravating interaction would be less than the product of every single mutant fitness, and an alleviating interaction would be the same as the single mutant fitness (Figure 1.7.1.1B). Gene pairs that function within the same pathway exhibit alleviating interactions while gene pairs

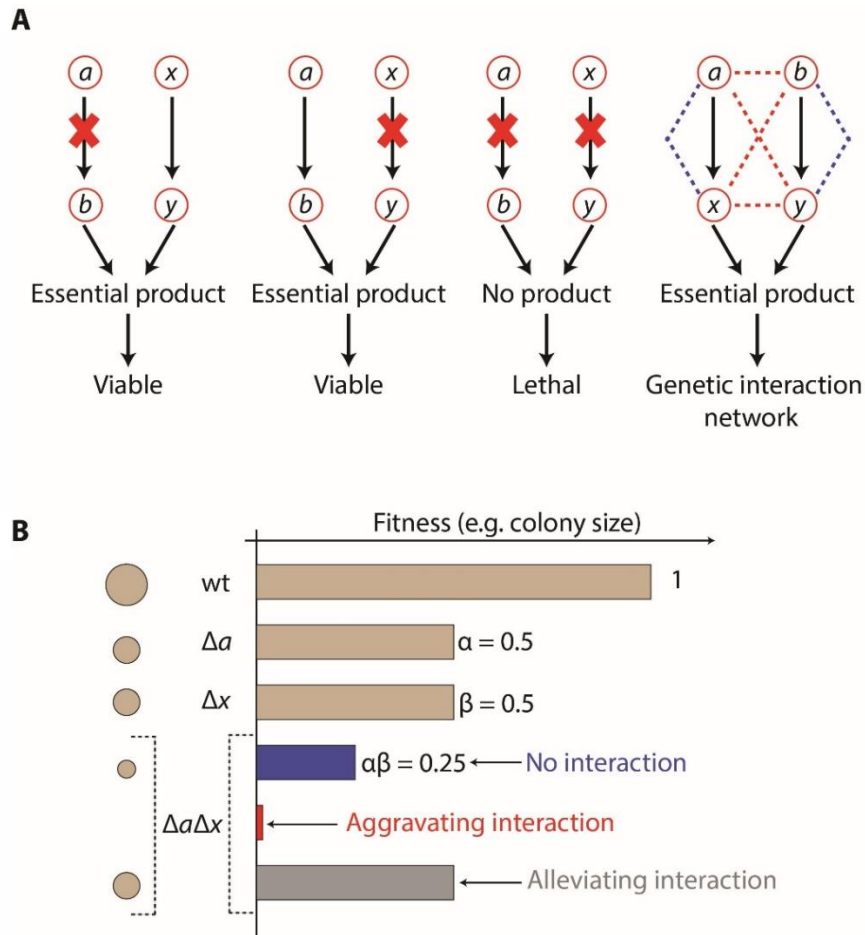


Figure: 1.7.1.1: Synthetic lethality and multiplicative principles.

(A) Genes ‘a’ and ‘b’ are components of the same pathway, which is buffered by a second, parallel pathway, composed of gene ‘x’ and ‘y’. Both pathways contribute to an essential product/process. Inactivating any single gene in the same pathways may slow but not abolish viability due to the buffering effect of the parallel pathway. However, when two genes in both pathways are knocked-out, the essential product cannot be formed, and the cell can no longer survive. The two genes are said to be synthetic lethal and show an aggravating interaction. Panel A, far right, a genetic interaction network can be determined when all combinations of pair-wise mutations for two pathways are interrogated. Blue dotted, and red dotted lines show alleviating and aggravating interactions, respectively. (B) Multiplicative principle for scoring genetic interactions. The model takes into account the extent to which mutations in two genes impact the phenotype by evaluating the single and double mutant fitness. The figure shows how fitness is determined using the measurements of colony sizes of mutants arrayed on plates. The wild-type strain is assigned an arbitrary fitness of one unit. The single mutants have half (0.5) the fitness relative to the wild-type, assigned as α and β for mutant ‘ Δa ’ and ‘ Δx ’, respectively. When the two mutants are combined, the double mutant $\Delta a \Delta x$ is expected to have the $\alpha\beta$ fitness of 0.25, if the two genes do not interact. However, if they do a fitness of less than 0.25 (aggravating interaction) or more than 0.25 is expected (alleviating interaction).

whose products function within buffering pathways that converge on the same essential process exhibit aggravating interactions (Gagarinova & Emili, 2012).

Data obtained from genetic interaction can be used to assign annotations of genes of unknown functions based on a principle of “guilt-by-association” where genes are assigned biological labels based on the functions of their interacting partners (Tian *et al.*, 2008). Genes performing similar functions are expected to have a similar pattern of genetic interaction with all other genes in a given system.

1.7.2 Advances in double recombineering high-throughput methods

Two similar high-throughput screening methods were developed independently to investigate the genetic interaction networks in *E. coli*. The Genetic Interaction Analysis Technology for *E. coli* (GIANT-*coli*) (Typas *et al.*, 2008) and *E. coli* Synthetic Genetic Array (eSGA) (Butland *et al.*, 2008). Both techniques systematically combine pairs of loss-of-function mutations using a unique bacterial conjugation system that requires a robotic platform for high-density mating.

The conjugation system in *E. coli* exploiting a self-transmissible plasmid, i.e., fertility factor (F^+), and high-frequency recombination (Hfr) strains have long been used to study plasmid transfer, genetic control of recombination, and mapping genes (Prozorov, 2003). Cells lacking the fertility factor are known as F^- strains. Hfr strains arise when the F plasmid integrates into the chromosome of the host cell via recombination between insertion elements (Chumley *et al.*, 1979). The Hfr strain can revert to F^+ via the reverse mechanism and, in some cases, a small portion of the chromosome may be captured in the F plasmid resulting in F' plasmids (Deonier & Mirels, 1977). This section will highlight the Hfr conjugation system (Figure 1.7.2.1) as it was central to this study.

The transfer of the chromosome begins at the *oriT* as a single-stranded DNA and then continues in a unidirectional manner into the recipient strain until mating is interrupted. The transfer takes place via a conjugative bridge known as the sex pilus (Lawley *et al.*, 2003). The newly transferred DNA becomes a substrate for DNA replication: however, the resultant linear double-strand is unstable and must be integrated into the host chromosome via homologous recombination in order to be maintained (Babu *et al.*, 2009).

Compared to plasmid transfer, chromosomal transfer takes longer but is nonetheless complete after about 100 minutes (Prozorov, 2003). This efficiency of the Hfr gene transfer system was critical to the development of large scale gene transfer technologies.

GIANT-coli and eSGA technologies have further exploited the natural genetic transfer capacity of Hfr strains and the available Keio deletion collection as a basis of a high-throughput method for examining genetic interactions in *E. coli* (Butland *et al.*, 2008; Typas *et al.*, 2008). The Keio collection consists of 3985 single-gene deletion mutants covering all non-essential *E. coli* genes (Baba *et al.*, 2006). The library background strain (BW25113) is an F⁻ strains, i.e., a recipient, allowing conjugative gene transfer. The two techniques, though developed independently, are very similar with minor differences. In summary, the ‘query’ gene is deleted using an antibiotic resistance cassette in the Hfr donor strain. The deletion exploits the λ red-mediated system for homologous recombination. The donor strain is pinned robotically onto a plate of ordered F⁻ colonies, each harbouring a single antibiotic marked deletion mutation and allowed to conjugate. The double recombinants produced by conjugation are then selected, first on single antibiotic plates then on double antibiotics plates. The plates are imaged, and the relative fitness scored by comparing the colony sizes between the double recombinants and the recipients. A comparison of the expected fitness and the double recombinant observed fitness is done for identification of significantly interacting gene pairs. The authors in both studies proved the concept by investigating well known parallel pathways with experimental confirmation of the anticipated synthetic lethal phenotypes (Butland *et al.*, 2008; Typas *et al.*, 2008).

While efficient, the technique does have some downsides. Closely linked genes exhibit aggravating interactions due to lower recombination frequencies and thus tend to be missed in large-scale screens. This effect was evident in both studies and was inversely proportional to genetic distance (Butland *et al.*, 2008; Typas *et al.*, 2008). The GIANT-coli study reported that gene pairs separated by less than 60 kb in M9 and 30 kb in LB displayed negative interactions (Typas *et al.*, 2008). The authors also reported merodiploidy; a partial duplication of the chromosome in haploid microorganisms (Johnston *et al.*, 2013) that leads to false-negative results.

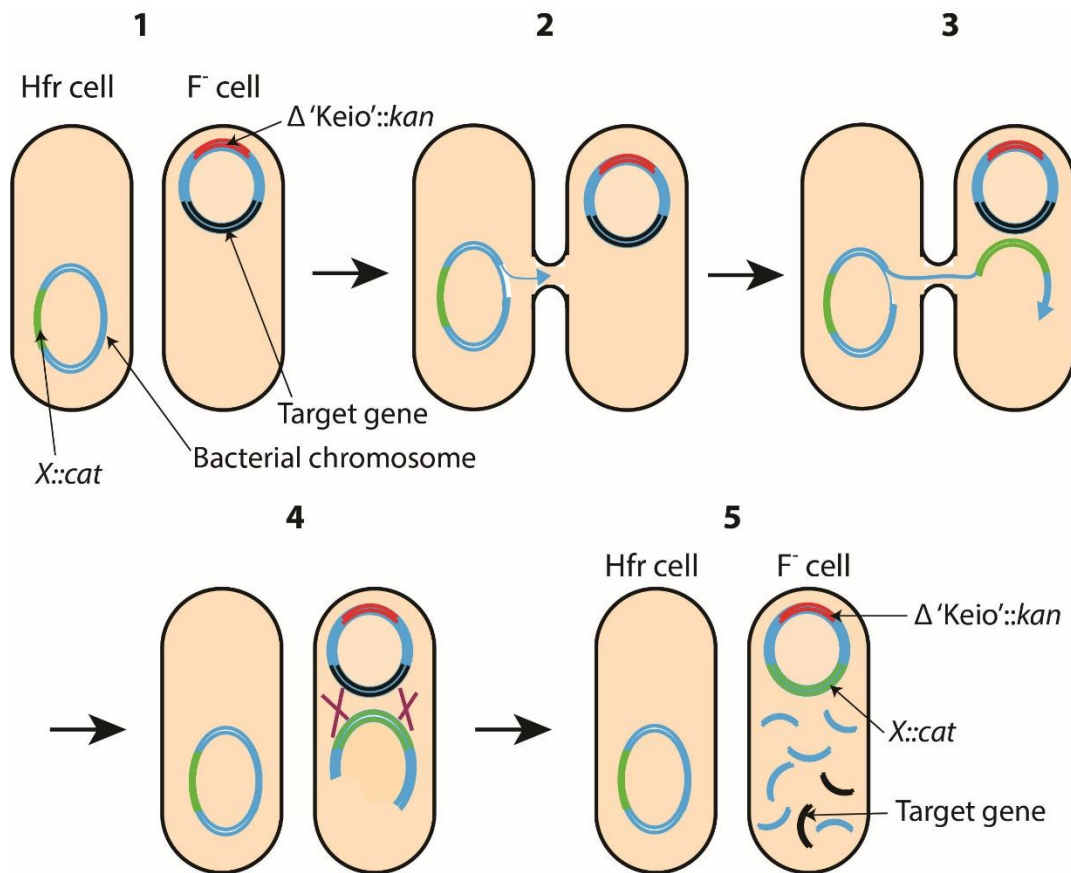


Figure 1.7.2.1: Hfr conjugation gene transfer system.

A step by step illustration of gene knock-out exploiting the Hfr conjugation system. **(1)** The target gene for deletion (black) in the recipient cell (F^+) is first knocked-out in the donor (Hfr cell) using an antibiotic marker, i.e., chloramphenicol (green gene). **(2)** During conjugation, transfer of the donor chromosome initiates at the *oriT* site (blue arrow) and is transferred to the recipient through the conjugative bridge formed (pilus). **(3)** The single-stranded DNA (ssDNA) transferred the recipient replicates to form a partially duplicated chromosome in one cell. **(4)** Homologous recombination may occur due to the homologous regions, and by doing so, the target gene (black) is replaced by the antibiotic marker (green). **(5)** Selection in double antibiotics selects cells that have undergone successful recombination and also keep the original deletion (red), resulting in a double deletion strain. The linear chromosome is then degraded.

Typas *et al.* (2008) developed two strategies to avoid the merodiploids: a single antibiotic intermediate selection step and use of M9 minimal media. The single antibiotic intermediate selection step selects-out duplicated regions that confer resistance to both antibiotics by competition from the growth of single antibiotic resistant parent and by spontaneous resolution of the unstable state. Additionally, only one parent was transferred to the subsequent double antibiotic selection plate, eliminating new rounds of mating and further generation of rare duplications (Typas *et al.*, 2008).

1.8 Thesis rationale and aims

Despite the impressive advances in our knowledge of cell envelope biogenesis as detailed in this chapter, we still have gaps in information regarding the architectural aspect of the envelope. The work presented here describes efforts to address these gaps in knowledge using GIANT-coli and eSGA techniques.

1. The first aim was to optimise the GIANT-coli and eSGA techniques to be better suited for addressing specific questions in envelope architecture and to increase the overall efficiency of the screening process. (Chapter 3).
2. The second aim was to develop an assay to investigate the transport of a model protein across the wall, using assembly of a model autotransporter (Pet β -barrel) domain expressed in the OM of *E. coli*. (Chapter 3)
3. In the third aim, I sought to address what happens when the periplasmic architecture is artificially changed. Specifically, which periplasmic processes rely on strict control of the periplasmic width? Also, why did the cell evolve to have that specific width? (Chapter 4).
4. The final aim was to understand how the cell copes with an enlarged periplasmic volume. What are the adaptive measures taken by the cell to cope with a larger periplasm? How is proteostasis monitored and regulated? (Chapter 5).

Chapter 2: Material and Methods

2.1 Bacterial strains, plasmids, and standard growth conditions

E. coli strains used in this study are listed in Table 2.1.1, sourced either internally from the Lithgow laboratory collection, or sourced externally from other laboratories or depository centres. Much of the work in this project relied on two strains: *E. coli* K-12 BW25113, the Keio collection parental strain, and the *E. coli* Hfr Cavalli conjugation donor strain, obtained from *E. coli* genetic stock centre, (Yale, #4410). The Keio collection library developed by Baba *et al.* (2006) was part of the Lithgow laboratory collection. Some of the strains in Table 2.1.1 were transformed to harbour a plasmid or were genetically manipulated to introduce a chromosomal mutation. A list of all the plasmids used in the study is shown in Table 2.1.2.

Unless otherwise stated, *E. coli* strains were grown in lysogeny broth (LB) media (CSH Protocols, 2016) containing (1% tryptone, 0.5% yeast extract and 0.5% NaCl) for rich media or M9 minimal media (standard) (CSH Protocols, 2010) supplemented with 1 mM MgSO₄, 0.1 mM CaCl₂, 1.12 mM thiamine and 0.2% (w/v) glucose for a defined minimal media.

Reagents for media supplementation were filter sterilised using Stericups® with 0.22 µm Express TM Plus membranes (Merck Millipore). In case of low volumes (less than 50 ml), e.g., antibiotics, filtration was done using an Acrodisc® syringe filter with a 0.2 µm Supor® membrane (PALL Life Sciences). In all the other cases, sterilisation was by autoclaving at 121 °C for 20 minutes at 15 psi.

Culture media was supplemented with antibiotics for plasmid selection and maintenance at the following concentrations: 100 µg/ml ampicillin, 30 µg/ml kanamycin, 34 µg/ml chloramphenicol unless otherwise stated. 15 g/l agar (Merck Millipore, 12177) was added to media before autoclaving when solid media was required.

Unless otherwise indicated, liquid cultures of *E. coli* were incubated at 37 °C on an orbital platform rotating 200 strokes per minute (25 mm orbit). In the case of solid cultures, incubation was for 16-24 hours (rich media) or 36-48 hours (minimal media) at 37 °C. Unless under unique circumstances, new media (3-5 ml liquid) was inoculated with single colonies for an overnight incubation (12-16 hours). The saturated cultures obtained were used in subculturing new media or preparation of glycerol stocks. Growth of *E. coli* was monitored by measuring the optical density at 600 nm (OD₆₀₀) with an Ultrospec 10 cell density meter (GE healthcare life sciences) or the Spark™ 10M multimode plate reader (Tecan).

Table 2.1.1: *E. coli* strains used in the study.

<i>E. coli</i> strains	Genotype / description	Purpose	Source
Hfr Cavalli	F+ <i>relA1</i> , <i>spoT1</i> , <i>metB1</i> , <i>rrnB</i> -2, <i>mcrB1</i> , <i>creC510</i>	Hfr conjugation donor strain	<i>E. coli</i> genetic stock centre (Yale)
DH5 α	F- <i>endA1</i> <i>hsdR17</i> (r_k2^- , m_k^+) <i>supE44</i> <i>thi</i> -1 λ^- <i>recA1</i> <i>gyrA96</i> <i>relA1</i> <i>deoR</i> Δ (<i>lacZYA-argF</i>)-U169 Φ 80 <i>dlacZ</i> Δ M15; NaI ^R	Cloning and vector storage	Invitrogen
S17-1 λ pir	F+ <i>recA</i> , <i>thi</i> , <i>pro</i> , <i>hsdR</i> -M ⁺ RP4::2-Tc::Mu::Km Tn7 λ pir, Tp ^R Sm ^R	Conjugation donor (Plasmid transfer)	Richard A. Strugnell, University of Melbourne
K-12 strain MG1655	F- λ^- <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1	PCR template	(Bachmann, 1996)
K-12 strain BW25113	F- Δ (<i>araD</i> - <i>araB</i>)567, Δ <i>lacZ</i> 4787(:: <i>rrnB</i> -3), λ^- , <i>rph</i> -1, Δ (<i>rhaD</i> - <i>rhaB</i>)568, <i>hsdR</i> 514	PCR template and Keio collection neat control	(Datsenko & Wanner, 2000)
Keio collection	A single gene knock-out library of all non-essential genes of <i>E. coli</i> replaced by a Kanamycin cassette.	Synthetic lethal screen library	(Baba <i>et al.</i> , 2006)
JW5028	BW25113 Kan ^R	Synthetic lethal screen control	Keio collection

Table 2.1.2: *E. coli* plasmids and synthesised DNA fragments (gBlocksTM) used in the study.

Plasmid or gBlock	Genotype or description	Purpose	Source or reference
Plasmids			
pBADpet	pBADHisA derivative expressing <i>de novo</i> synthesised Pet, amp ^R	<i>pet</i> PCR template	(Leyton <i>et al.</i> , 2010)
pKD3	FRT- <i>cat</i> -FRT <i>oriR6K</i> Amp ^r	CAT cassette template	(Datsenko & Wanner, 2000)
pKD4	FRT- <i>kan</i> -FRT <i>oriR6K</i> Amp ^r	Kan cassette PCR template	(Baba <i>et al.</i> , 2006)

pKD46	<i>repA101(ts) oriR101 bla P_{araB}-(gam bet exo)</i>	Gene knock-out	(Baba <i>et al.</i> , 2006)
pCP20	<i>cI857 λPR flp pSC101 oriTS Amp^r Cam^r</i>	Antibiotic resistance cassette flipping/removal	(Cherepanov & Wackernagel, 1995)
pJP168	P _{BAD} promoter and <i>araC</i> replaced by P _{tetA} and <i>tetR</i> gene from Tn10 in pBAD24, Amp ^r	Pet fusions expression vector	(Baldi <i>et al.</i> , 2012a)
pELM1	pJP168 (<i>BglII/PstI</i>) <i>amyA</i> gene within Pet passenger domain	Pet-AmyA expression vector	This study
pELM2	pJP168 (<i>BglII/PstI</i>) <i>sfGFP</i> gene within Pet passenger domain	sfGFP expression vector	This study
pELM3	pJP168 (<i>BglII/PstI</i>) <i>phoA</i> gene within Pet passenger domain	PhoA expression vector	This study
pELM4	pJP168 (<i>BglII/PstI</i>) <i>mCherry</i> gene within Pet passenger domain	mCherry expression vector	This study
Gene blocks			
Lpp ₊₂₁	A 406 bp gene block containing 234 bps of <i>lpp</i> CDS and extra bps for Gibson assembly	<i>lpp₊₂₁</i> cloning	Integrated DNA Technologies (IDT)
mCherry	A 754 bp gene block containing 705 bps of mCherry CDS and <i>pet</i> overhangs	<i>pet-mCherry</i> cloning	Integrated DNA Technologies (IDT)
Amylase	A 1531 bp gene block containing 1482 bps of <i>amyA</i> CDS and <i>pet</i> overhangs	<i>pet-amylase</i> cloning	Integrated DNA Technologies (IDT)
sfGFP	A 760 bp gene block containing bps of sfGFP CDS and <i>pet</i> overhangs	<i>pet-sfGFP</i> cloning	Integrated DNA Technologies (IDT)
phoA	A 1399 bp gene block containing 1350 711 bps <i>phoA</i> CDS and <i>pet</i> overhangs	<i>pet-phoA</i> cloning	Integrated DNA Technologies (IDT)

2.2 Transformation techniques

Three transformation techniques were employed to transform a plasmid or a PCR fragment in *E. coli* strains: the transformation and storage solution (TSS) method, the CaCl_2 chemical method, and the electroporation method. TSS method was employed for typical plasmid transformations, where cells do not have to be highly competent. CaCl_2 chemical method was used while cloning, i.e. the transformation of a ligation mixture. The electroporation method was used when cells were required to be highly competent, i.e. for gene knock-out experiments. The three techniques are outlined below in detail.

2.2.1 Transformation and storage solution method.

In this technique, developed by Chung and co-workers (1989), cells are made competent by mixing with TSS (10% (w/v) polyethene glycol, 5% (v/v) dimethyl sulfoxide, and 50 mM Mg^{2+} in LB and pH adjusted to 6.5) (Chung *et al.*, 1989). To prepare TSS competent cells, saturated overnight cultures were diluted 1:100 into fresh 10 ml LB. The cultures were grown to the logarithmic phase ($\text{OD}_{600} = 0.4 - 0.5$ AU), collected by centrifugation ($5000 \times g$, 10 minutes, 4°C) and resuspended in 150 μl LB. An equal volume of ice-cold 2x TSS was added and sample well mixed.

For transformation, 100 ng of plasmid DNA was added to 50 μl of the mixture and incubated on ice for 30 minutes; no heat shock is required. Transformed cells were transferred to 950 μl of non-selective SOB media supplemented with 20mM MgCl_2 and 0.2% (w/v) glucose for a 45-minute incubation. Cells were concentrated by centrifuging ($5000 \times g$, 5 minutes, 4°C) and resuspending in 100 μl LB. The concentrated transformation mixture was spread plated on selective solid media and incubated overnight (12 – 16 hours) at 37°C . Well-isolated single colonies of transformants were purified by streaking in the same media conditions twice and cultured in fresh LB (liquid) media for subsequent experiments or storage.

2.2.2 Preparation of CaCl_2 chemically competent cells and transformation protocol

A subculture was prepared by diluting 1:100 of the saturated overnight culture into fresh 200 ml LB. The culture was grown to an OD_{600} of 0.4 - 0.5 AU. The cells were chilled for 30 minutes on ice, collected by centrifugation ($5000 \times g$, 10 minutes, 4°C), and resuspended in 40 ml ice-cold 0.1 M CaCl_2 . Chilling was repeated for a further 15 minutes, and cells collected by centrifugation as described previously. The pellet was resuspended in 1 ml ice-cold 0.1 M CaCl_2 and incubated on ice for 2 hours. 500 μl ice-cold LB media containing 30% v/v glycerol

was added to the suspension, and 50 µl aliquots were snap-frozen in liquid nitrogen. The aliquots were stored at -80 °C freezer.

To transform CaCl₂-competent *E. coli* strains, the cells were first thawed on ice for about 10 minutes. Purified plasmid (10 ng) was added, and after gentle mixing, the sample was incubated on ice for 5 minutes. The sample was then heat-shocked for 45 seconds at 42 °C, incubated on ice for 5 minutes, and transferred to 950 µl of non-selective SOB media supplemented with 20mM MgCl₂ and 0.2% (w/v) glucose for a 45-minute incubation. 100 – 200 µl of the transformation mixture was then spread plated on selective solid media and incubated overnight (12 – 16 hours) at 37 °C. Subsequent steps were the same as the TSS method detailed in Section 2.2.1.

2.2.3 Preparing electrocompetent cells and electroporation

Chromosomal gene modification by the transformation of a PCR fragment occurs at relatively low efficiency, and therefore success is dependent on highly competent cells. To achieve high competency, electrocompetent *E. coli* strains were prepared. Saturated overnight cultures were diluted 1:100 with 300 ml of fresh LB and incubated until the OD₆₀₀ was between 0.4-0.5 AU. Series of centrifugation (5000 × g 10 minutes, 4 °C) and washing followed, where the cells were resuspended in increasingly smaller volumes of ice-cold glycerol (10% v/v) or sterile water, e.g. 20 ml, then 10 ml, then 5 ml, then 0.3 ml. Aliquots (50 µl) were either used for transformation immediately or snap-frozen in liquid nitrogen and stored at -80 °C. When required, frozen aliquots were thawed on ice before subsequent electroporation steps.

For *E. coli* gene knock-out transformation, 50 µl electrocompetent cells were thawed on ice for 5 minutes. 0.5 – 1 µg of the PCR product was added, and the mixture transferred to pre-chilled ice-cold electroporation cuvettes (1 mm gap). Cells were electroporated (1800 V, 200 Ω and 100 µF) and immediately transferred to pre-warmed 950 µl of non-selective SOB media, supplemented with 20mM MgCl₂ and 0.2% (w/v) glucose. The transformants were recovered for 1 hour at 37 °C. Subsequent steps were the same as the TSS method detailed in Section 2.2.1.

2.3 Bacterial storage and recovery

To store new *E. coli* strains after routine procedures, 800 µl aliquot of fresh, saturated culture was transferred to a cryogenic vial and mixed with an equal volume of 50% (v/v) glycerol (25% v/v final concentration). The sample was snap-frozen in liquid nitrogen and transferred to a -80 °C freezer for long term storage. New plasmids stocks were transformed in *E. coli* DH5α,

a $\Delta recA$ mutant for long-term storage. For bacterial recovery, a glycerol stock was retrieved, avoiding full thawing, a small portion of the culture was scooped and inoculated in solid LB media. The strain was recovered by incubating at 37 °C for subsequent inoculation of starter cultures.

2.4 DNA/RNA based techniques

The plasmids and synthesised DNA fragments (gBlocks™) used in this study are listed in Table 2.1.2. Oligonucleotides used in the study are listed in Table 2.4.2.1.

2.4.1 Plasmid extraction

Plasmids were isolated from 5 ml overnight cultures using the Wizard® Plus SV Miniprep DNA purification system (Promega) kit as per the manufacturer's instructions.

2.4.2 Polymerase chain reaction (PCR)

Primers were designed using the SnapGene® viewer (GSL biotech) and synthesised by Integrated DNA Technologies (USA). Genes of interest were PCR amplified using, Taq DNA polymerase (Roche) kits or Phusion® High-Fidelity DNA polymerase (New England Biolabs) kits as per the manufacturer's instructions. For diagnostic PCR, Go Taq® Green Master Mix (Promega) was used as per the manufacturer's instructions. Primers used in PCR are listed in Table 2.4.2.1. Template DNA used for PCR was provided either as a water resuspended single colony from an overnight incubation, synthesised DNA fragments (gBlocks) or as a purified plasmid. In cases where whole cells were used, the colony suspension was first boiled for 5 minutes at 95 °C. Following PCR amplification, DNA fragments were stained using SYBR® Safe (Thermo Fisher Scientific) as per the manufacturer's instructions and analysed by gel electrophoresis using 0.5-2.0% (w/v) molecular biology grade agarose (Scientifix, 9010E) in TAE buffer (CHS protocols, 2013) at 80 – 110 volts for 45 – 60 minutes. Separated DNA fragments were visualised using a LAS-300 Imaging system (Fuji), and fragments of the correct size were purified using Wizard® SV Gel and PCR Clean-up System kit (Promega).

Table 2.4.2.1: Primers used in the study.

Primer number	Primer name	Sequence (5'→3')	Primer type/purpose
EP7	CATF	catatgaatatcctccttagttcc	CAT cloning
EP8	CATR	tgtgtaggctggagctgcttc	CAT cloning
EP11	CATSeqF	gagaatatgttttcgtctcagcc	CAT sequencing

EP12	CATseqR	cgtaatatccagctgaacggctctg	CAT sequencing
EP13	BamCKoF	cagagcggcgcttaagcatgccggttgctgtaaag tttagggagatttg catatgaatatcctccttagttcc	<i>bamC</i> knock-out
EP14	BamCKoR	gtatcatatcagaaaaaggccggatgattccagc cctgtattt tgtgtaggctggagctgcttc	<i>bamC</i> knock-out
EP15	BamC-95F	gcgcctgccaatgacaccaatcacc	<i>bamC</i> sequencing
EP16	BamC+95R	cgtgcgaattaccgctaaaatcgcg	<i>bamC</i> sequencing
EP17	PetF	gaattaacatgggcatgaacaaaatctactc	Pet cloning
EP18	PetR	cagccaagctttcagaaagagtaacggaag	Pet cloning
EP19	PetNcoISeqF	ccactcctatcagtgatagagaaaag	Pet sequencing
EP20	PetPstISeqR	ctttgatagaaccctggtaagaggtg	Pet sequencing
EP21	PetHindIIISeqR	gccaggcaaattctgtttatcagaccg	pJP168 Sequencing
EP31	F1PetAsn1018Gly	gcgttcctggcggaagtggcaacctgaacaaacg tatggg	Pet cloning
EP32	R2PetAsn1018Gly	cccatacgtttgttcaggttgccaactccgccagg aacgc	Pet cloning
EP33	F2PetAsp1115Gly	gtttgaatctgggtgcgtacatcgccctgatcggtaa atacgttc	Pet cloning
EP34	R1PetAsp1115Gly	gaacgtatttaccgatcaggccgatgtacgcaccag attcaaac	Pet cloning
EP35	PETF3forward	cgttatccctgattctgtggataac	Pet cloning
EP36	PETR3reverse	gttatccacagaatcaggggataacg	Pet cloning
EP37	PJP168FHindIII	cgtaaagctt ggctgtttggcgatgagagaag	pJP168 cloning
EP38	PJP168RHindIII_ MCS	ctgatgcaagcttgcctgcaggtcgactcta gaggatccccgggtaccatggtgaattcctcctgctag	pJP168 cloning
EP39	FN1018GSeq	ctacaaaaccgttgcgaacgcg	Pet sequencing
EP40	RD1115GSeq	gagtgagaagagtagtcacggg	Pet sequencing
EP41	FMCSHindIIIEP	cctaattttgttgacactctatcattg	pJP168 sequencing
EP42	HindIIIMCS2	ctgatgc aagctt ccatggtgaattcctcctgcta gcattcatttc	pJP168 sequencing
EP43	BamCFSeq	cttcgctacttgagccactatcgactac	<i>bamC</i> sequencing
EP44	BamCRSeq	ggaaggagctgactgggtgaaggctctc	<i>bamC</i> sequencing
EP51	LppKOF	gtgtaatacttgtaacgctacatggagattaactcaat ctagagggtattaata catatgaatatcctccttagttcc	<i>lpp</i> knock-out
EP52	LppKOR	cagacaaaaaaatggcgcacaatgtgcgcca	<i>lpp</i> knock-out

		ttttcacttcacaggtactatgtgttagctggagctgcttc	
EP53	Lpp-95F	gtgccttcccatcaaaaaatattctcaac	<i>lpp</i> sequencing
EP54	Lpp+95R	ccagagcaagggaatatgttacgcg	<i>lpp</i> sequencing
EP64	Lpp-95FseqV2	gtgtaatacttgaacgctacatggag	<i>lpp</i> sequencing
EP65	CatseqRV2	gccattgggatatacaacggtgg	CAT sequencing
EP66	MrcAUF	gtcactctgacgggtatatcaatgcgtc	<i>mrcA</i> PCR
EP67	MrcADR	cgattttgtagccggataacgcgttcac	<i>mrcA</i> PCR
EP68	MrcBUF	ggaccgatcgagcacaattttgagag	<i>mrcB</i> PCR
EP69	MrcBDR	gggttaataaacaaccagatgaaaagaagg	<i>mrcB</i> PCR
EP70	LpoAUF	cctgagaattatccgccgattggtgttg	<i>lpoA</i> PCR
EP71	LpoADR	gccacgctcggtcacgttagcggcg	<i>lpoA</i> PCR
EP72	LpoBUF	cccgacattcgccggtaacgctgtac	<i>lpoB</i> PCR
EP73	LpoBDR	caatgagaaaactcccgccactaagccc	<i>lpoB</i> PCR
EP74	KanR	Cagtcatagccgaatagcct	Kan cloning
EP75	KanF	Cggtgccctgaatgaactgc	Kan cloning
EP76	LppR	ttacttgcggtatttagtagccatgtgtcc	<i>lpp</i> cloning
EP77	OmpAUF	catggcggtatttggatgataacg	<i>ompA</i> PCR
EP78	OmpAUR	gcagcgggggttttctaccagacgag	<i>ompA</i> PCR
EP79	PalUF	cgccgtatctgtgataataattaattgaatag	<i>pal</i> PCR
EP80	PalUR	gttgatgtctgaagtactgctcatgc	<i>pal</i> PCR
EP83	AcrBUF	gtccaagtcttaactaaacaggagccg	<i>AcrB</i> PCR
EP84	AcrBUR	ccgcttacgcggccttagtgattacacg	<i>AcrB</i> PCR
EP87	NlpDUF	gcacggtaattatgtcactggttattaacc	<i>NlpD</i> PCR
EP88	NlpDUR	ggaacattcaagcaaaagcctggttc	<i>NlpD</i> PCR
EP89	YcaLUF	gtcgcttgtaatgtaagttgttgatac	<i>ycaL</i> PCR
EP90	YcaLUR	gaagaccagcccgaagatgacaatg	<i>ycaL</i> PCR
EP97	YraNUF	ctaaccgggattgcgtgattaacaggaac	<i>yraN</i> PCR
EP98	YraNUR	ctttcttgcacgctaataccttaaac	<i>yraN</i> PCR
EP99	TolCUF2	cgtttgactgccgttgagcagtcatgtg	<i>tolC</i> PCR
EP100	TolCDR2	gaggatggctggtcgaaattgaagcga	<i>tolC</i> PCR
EP109	MdtLUF	gtatgctttgcgtcaccttactatcagg	<i>mdtL</i> PCR
EP110	MdtLDR	ctgcacagcagttgcagacaaagcagcag	<i>mdtL</i> PCR
EP111	MdtKUF	cctcgctacaatagactgaatttccctg	<i>mdtK</i> PCR
EP112	MdtKDR	gcaacccgtgaaggtgcgccttgc	<i>mdtK</i> PCR

EP141	YiaDKOF	Acgcggcataatctcccgccacggaacccgtggcaaag aataaaaaggttattaaggattaacaatgattccggggatcc gtcgacc	<i>yiaD</i> knock-out
EP142	YiaDKOR	gatttctgcatcacctggcatgaaagggattacagcgggct taaggtaattgtaggctggagctgcttcg	<i>yiaD</i> knock-out
EP143	YiaDseQF	ctcccgccacggaacccgtggcaaag	<i>yiaD</i> PCR
EP144	YiaDseqR	gatttctgcatcacctggcatgaaaggg	<i>YiaD</i> PCR

2.4.3 Restriction digestion and alkaline phosphatase treatment

Where possible, high fidelity restriction enzymes (New England Biolabs) were used for digestion of plasmids or purified PCR amplicons as per manufacturer's instructions, but that the digestion time was increased to one hour. Plasmids linearized by a single restriction digest were treated with alkaline phosphatase (Promega, M182A) as per the manufactures instructions. Digested DNA samples were stained, analysed by gel electrophoresis, visualised and purified as described in PCR Section 2.4.2 above.

2.4.4 DNA ligation

The concentrations of purified insert DNA and plasmid DNA were determined using the NanoDrop 1000 Spectrophotometer (Thermo Scientific) as per the manufacturer's instructions. The insert DNA and the plasmid DNA were mixed at a ratio of 3:1, and a T4 DNA Ligase kit (New England Biolabs) used for ligation as per the manufacturer's instructions. 10 µl of the ligation mixture was used to transform chemically competent *E. coli* DH5α.

2.4.5 DNA sequencing

Ligated plasmids or amplified PCR fragments were sequenced by Macrogen, Inc. (South Korea) using appropriate sequencing primers (Table 2.4.2.1). The sequence data obtained was then analysed by Geneious® software version 8.1.9 (Biomatters) to determine if desired deletions, insertions or mutations were present.

2.4.6 Total RNA isolation

A subculture was prepared by diluting 1:100 of a saturated overnight culture into 10 ml media. For minimal media subculturing, the overnight culture was washed twice in 1x M9 salts. The culture was grown to an OD₆₀₀ of 0.5 AU and 2 ml of the culture added to 10 ml of RNAprotect Bacteria Reagent (Qiagen). Samples were immediately vortexed for 5 seconds and incubated for an additional 5 minutes at room temperature. The samples were then centrifuged (15000 ×

g, 10 minutes, room temperature) and the supernatant discarded (The pellet was stored at -80 °C, in cases where RNA extraction was not done immediately).

Total RNA was then isolated from the cell pellets using an RNAeasy Plus Mini kit (Qiagen) according to the manufacturer's instructions. Extracted RNA samples were quantified using a NanoDrop 100 Spectrophotometer, which was also used to check the quality of the isolated RNA by measuring the sample's A₂₆₀: A₂₈₀ ratio, expected to be above 1.8. The quality of isolated RNA was also determined by visualisation on agarose gels.

2.4.7 RNA data analysis

An RNA sequencing library was prepared and sequenced at Micromon (Monash University), and the generated fastq files were analysed with RNAsik pipeline (Tsyganov *et al.*, 2018) to produce raw genes count matrix and various quality control metrics, all summarised in MultiQC report (Ewels *et al.*, 2016). For this analysis RNAsik pipeline (Tsyganov *et al.*, 2018) was run with BWA-MEM aligner (Li, 2013) option and reads were quantified with feature counts (Liao *et al.*, 2014). The reference GFF and FASTA files were downloaded from the RefSeq database. The gene counts matrix were then analysed with Degust (Powell, 2019) web tool, including differential expression analysis and several quality plots such as classical multidimensional scaling (MDS) and MA plots. Limma voom (Law *et al.*, 2014) was used in this analysis for model fitting. Degust by Powell, 2019 largely follows limma voom workflow, including the trimmed mean of M values (TMM) normalisation (Robinson & Oshlack, 2010) for RNA composition. The analysis was performed at Monash Bioinformatics platform (Dr Kirill Tsyganov).

2.5 High-density mutagenesis, screen, and selection

2.5.1 Hfr Cavalli gene knock-out

Knock-in and knock-out mutants, in which target genes were deleted or replaced with an antibiotic cassette, or with an antibiotic cassette including a gene of interest, were constructed in *E. coli* Hfr Cavalli donor strain using the λ -red homologous recombination technique (Datsenko & Wanner, 2000). All oligonucleotides for amplification of PCR knock-in fragments and introduction of flanking homology regions are listed in Table 2.4.2.1.

Donor strain carrying the desired mutation was constructed as follows. A chloramphenicol cassette containing flanking recombinase target (FRT) sites was amplified from a linearised pKD3 plasmid (Table 2.1.2). The resulting PCR product contained chloramphenicol acetyltransferase (CAT) gene flanked by long (> 50) base pairs arms homologous to the target

gene. The linear PCR fragments were purified, quantified and stored at -20 °C awaiting electroporation.

The donor was prepared for gene knock-out by transforming pKD46 plasmid, which contains the genes for λ -red homologous recombination. TSS transformation technique was used as described in Section 2.2.1, but the recovery of the transformants was done at 30 °C as the plasmid is heat sensitive. Electrocompetent cells were prepared as described in Section 2.2.3 with minor modifications. Growth was strictly at 30 °C, and pKD46 was induced by addition of 10 mM L-arabinose (Sigma-Aldrich). Induction was done at log phase ($OD_{600} = 0.4$ AU) 30 minutes before cells harvesting, and recovery of transformants was done at 30 °C for 2 hours. Successful gene knock-out strains were selected on LB solid media supplemented with 34 μ g/ml chloramphenicol at 37 °C for 24 hours.

To screen for correct knock-out mutants, colony PCR was performed as described in Section 2.4.2. Using a colony as a template, the chromosomal DNA was amplified using two sets of knock-out confirmation primers (Table 2.1.2). The first set consisted of flanking primers, 200 base pairs upstream and downstream of the deleted region. This primer set screened out any rare merodiploid strains.

The second set of primers included either one of the upstream or downstream primers paired with one internal primer complementary to the CAT cassette. This set confirmed the position and location of the cassette and target genes, respectively. The confirmed deletion mutants were stored, as described in Section 2.3.

2.5.2 Arraying Keio collection for screening

The kanamycin resistance, recipient (F⁻) KEIO collection *E. coli* strains, initially arranged in 96-well plates were robotically rearranged to a 384-well format using a Singer RoTor high-density array station (HDA) (Singer instruments). Strains were pinned from 96-well plates to 384-well plates containing 100 μ l of LB medium supplemented with 30 μ g/ml kanamycin and grown overnight at 37 °C to an OD_{600} of 0.5 AU with gentle shaking at 190 rpm. The cultures were supplemented with 15% glycerol and stored at -80 °C before pinning. A comprehensive list of KEIO *E. coli* deletion mutants used in the screen is available online (Baba *et al.*, 2006). In each plate, empty wells were included as internal quality controls and negative conjugation controls. JW5028 strain, a Keio collection strain containing a kanamycin resistance gene in place of a pseudogene, was added to the side border wells as a positive control for conjugation.

2.5.3 High-density generation of double recombinants

Briefly, the process was broken down into four separate sections, mutant array preparation, conjugation, selection and imaging. The high throughput generation of double recombinants was achieved by replica plating employing the Singer RoToR HDA robotics platform. The process was carried out in both LB and M9 minimal media (0.2% glucose). In this section, the technique is highlighted in LB media. The overview of the procedure is shown in Figure 2.5.3.1.

2.5.3.1 *Single mutant arraying*

The frozen Hfr Cavalli donor strain, the query deletion mutant, with the target gene replaced with CAT, was grown overnight at 37 °C in LB liquid media supplemented with 34 µg/ml of chloramphenicol. On the second day, the culture was dispensed in 384-well microtiter plates (50 µl in each well) and pinned onto LB solid media containing 34 µg/ml chloramphenicol at 384-density (24 columns x 16 rows). Keio collection recipient strains prepared in Section 2.5.2, were thawed and replica pinned in LB solid media containing 30 µg/ml kanamycin at a similar density as the donor strain. The cultures were incubated overnight at 37 °C.

2.5.3.2 *Conjugation*

On the third day of the protocol, conjugation was carried out on LB (solid) media without antibiotics by replica pinning the overnight donor and recipient deletion mutants on the same plate and mixing at 384-density.

2.5.3.3 *Selection of double recombinants.*

“Double recombinants,” i.e. *E. coli* strains with both genes deleted (*X::cat*, *Y::kan*) (Figure 2.5.3.1), were subjected to two rounds of selection. The first round employed a single antibiotic (kanamycin 30 µg/ml), while the second round employed two antibiotics, kanamycin and chloramphenicol at 30 µg/ml and 34 µg/ml respectively. For both rounds of selection, mutants were replica pinned from one plate to another at the same density and incubated overnight at 37 °C. In the case of double selection, the mutants were incubated for longer (\approx 24 hours) to best allow for uniform colony growth.

2.5.3.4 *Imaging and storage*

The colonies from each selection plate were photographed using a special camera system, the PhenoBooth (Singer instruments). The captured images were saved as PNG files and the

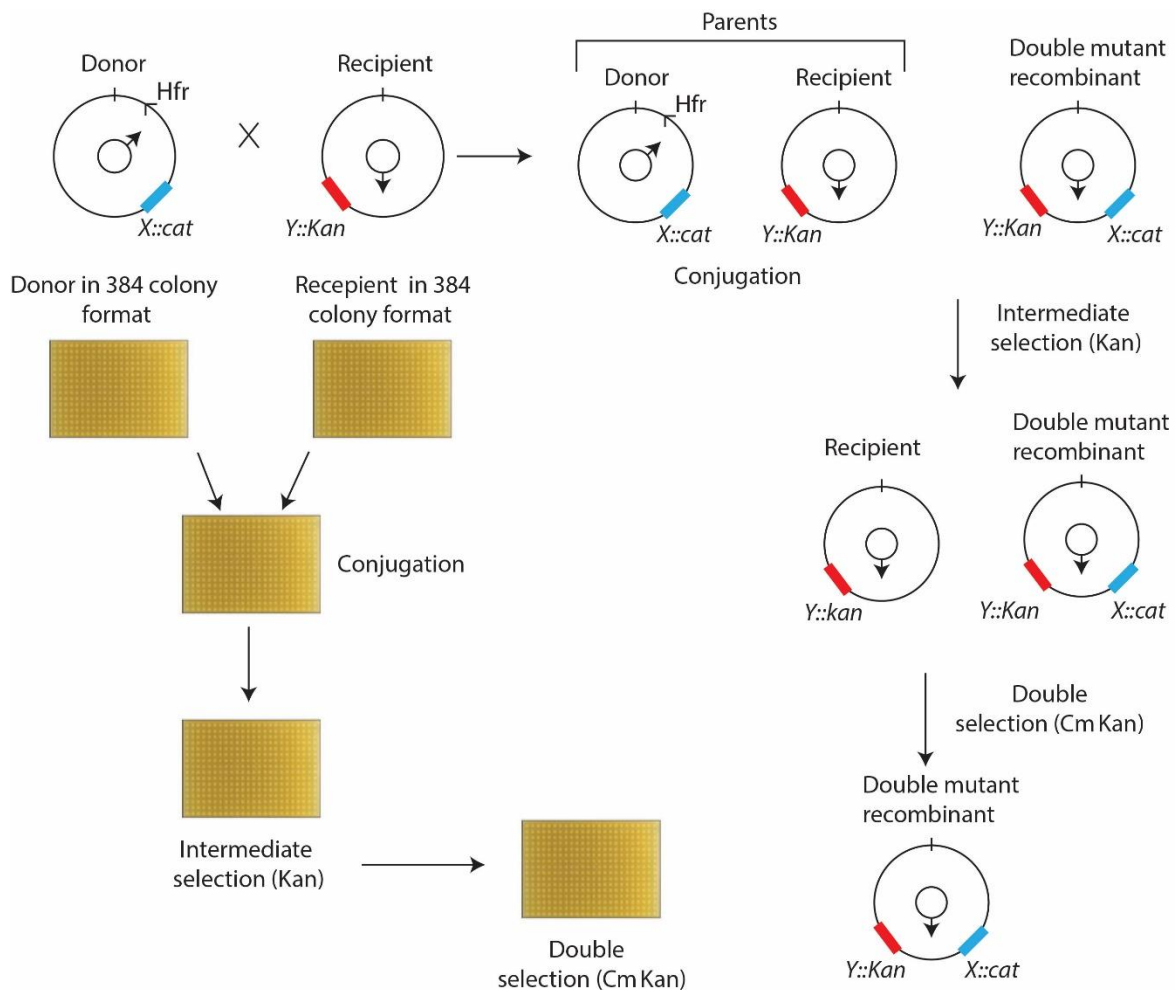


Figure 2.5.3.1. Strategy for the construction of “double recombinants”.

The Hfr Cavalli donor strain has the query gene (“X”) deleted using the gene encoding CAT ($X::cat$). Using a robotic platform and 384-well plates, conjugations reactions are set up between the donor and hundreds of recipient strains, each one carrying a gene of interest (“Y”) deleted with a kanamycin cassette ($Y::kan$). The donor is arrayed in the same format, 384-density format, as the recipients. Conjugation is performed by robotic mixing of both the donor and the recipient on LB (solid) media without antibiotics, and plates incubated for 16 hours at 37 °C. The conjugants then undergo the first round of selection, (intermediate selection) on kanamycin (Kan) containing plates to select out the donor parent. The mutants are further subjected to a second round of selection using both antibiotics, Kan and chloramphenicol (CM), selecting-out the recipient parent and generating a plate with only the “double recombinants” (Typas *et al.*, 2008). Images of representative plates generated in each step of the procedure are shown on the left of the strain cartoons.

growth phenotype of the double recombinants was assayed using an automated image processing software detailed in Section 2.5.4 below.

The double recombinants were replica pinned into 2x LB liquid culture, supplemented with antibiotics at the appropriate concentration, and incubated overnight at 37 °C with gentle shaking at 190 rpm. Each overnight culture was supplemented with 15% glycerol and transferred to -80 °C freezer for long term storage.

2.5.4 High throughput data analysis

The contrast of high-resolution images obtained by the PhenoBooth (Section 2.5.3.4) was enhanced through batch processing level adjustment in ImageJ software (NIH). Raw colony size data for each location on the plate array was obtained by automated image analysis software, HT Colony Grid Analyser (Collins *et al.*, 2006), available online as open-source software. Generated data were processed manually for the 384-density using Excel spreadsheets for each of the media condition used. Relative colony size plots were generated using GraphPad prism® version 8.0 (GraphPad Software, Inc.)

2.5.5 Verification of identified synthetic lethal/sick mutants

Synthetic lethal or growth-compromised mutants selected from the high-throughput screen above were subjected to another round of screen in the same conditions as previously identified. The process was termed as a mini-screen. For increased accuracy, each KEIO biological mutant replicate was arrayed in 16 technical replicates. Verified mutants from the mini-screen were further screened by colony PCR as described in Section 2.4.2, to screen-out merodiploids and confirm both the presence and chromosomal location of both gene modifications.

2.6 Protein-based techniques

2.6.1 Polyacrylamide gel electrophoresis

Proteins were analysed by polyacrylamide gel electrophoresis (PAGE) using Mini-PROTEAN® Tetra vertical electrophoresis cell (Bio-Rad) equipment as per the manufacturer's instructions. For gradient gels, the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) was used as per the manufacturer's instructions.

Sodium dodecyl sulphate (SDS)-PAGE analysis was performed as previously described by (Simpson, 2006), with several modifications. The stocks of the two gel buffers, 4× stacking gel buffer (1.0 M Tris pH 6.8, 2 mM EDTA, and 0.08% w/v SDS) and 4× resolving gel buffer (1.5 M Tris pH 8.8, 2 mM EDTA, and 0.4% w/v SDS) were adjusted to 1x before use. For different

resolving specifications, the 40% (w/v) acrylamide solution (Bio-Rad, #1610146) was adjusted accordingly to make gradients gels, e.g., 4-16% or uniform gels, e.g., 8%, 10%, 12%. The stacking gel (4%) was also prepared from the same 40% (w/v) acrylamide stock solution. The polymerisation of the gel solutions was initiated by the addition of 50 µl of APS (20% w/v) and 10 µl of TEMED, per 10 ml of gel solution, immediately before casting. Gels cast using 0.75 mm combs and spacers. For resuspension or dilution of protein samples, the final concentration for SDS sample buffer was adjusted to 10% (v/v) glycerol, 100 mM Tris-HCl pH 6.8, 100 mM dithiothreitol (DTT), 1% (w/v) SDS, and 0.01% (w/v) bromophenol blue.

Dual colour Precision Plus Protein™ (Bio-Rad), pre-stained standards stock, was diluted 1:10 with SDS sample buffer, of which 5 µl were loaded into polyacrylamide gels as an SDS-PAGE molecular weight marker. SDS running buffer (25 mM Tris, 0.1% SDS, and 192 mM glycine) was used to separate proteins by electrophoresis. Protein samples were normalised by growth (OD₆₀₀) or concentration (µg/ml) and appropriate volumes loaded on the gels. Typically, 1-20 µl or 0.5-50 µg.

2.6.2 Coomassie staining and destaining of SDS-PAGE gels

To visualise proteins following SDS-PAGE analysis gels were stained with Coomassie staining solutions (10% v/v acetic acid, 50% v/v methanol, and 0.05% w/v Coomassie Brilliant Blue R-250). Gels were incubated in the staining solution for at least 1 hour with gentle shaking (40-50 rpm, 32 mm orbit), and the stain was then gradually removed with sequential incubations in a destaining solution comprising 7% (v/v) acetic acid 5% (v/v) methanol. Gels were sometimes destained for up to 24 hours at 40-50 rpm (32 mm orbit) until the background staining was minimal for clear visualisation of protein bands. Gel images were acquired by scanning using CanoScan 8600F with software version 5.0.1.2 (Canon) and subsequently dried using the Slab Gel Dryer model 1125 (Bio-Rad) as per the manufacturer's instructions.

2.6.3 Western transfer and immunoblotting

Following SDS-PAGE analysis, proteins were transferred from the gels to 0.45 µm Protran nitrocellulose hybridisation transfer membranes (PerkinElmer), or 0.45 µm hydrophobic Immobilon-P PVDF transfer membranes (Merck Millipore). The western transfer was performed using the semi-dry Trans-Blot® Turbo transfer system (Bio-Rad) as per the manufacturer's instructions. PVDF membranes were activated in methanol before transfer.

For immunoblotting, antibodies and conjugates used as detection reagents are listed in Table 2.6.3.1. Primary and secondary antibodies were diluted in blocking solution, 5% (w/v) instant

skim milk powder (Coles) in Western washing buffer. Western washing buffer was prepared as per CSH Protocols (2013) except the final concentration of Tween-20 was readjusted to 0.2 % (v/v). Incubations were done at either room temperature or 4 °C at 40-50 rpm (25 mm orbit). Primary and secondary antibodies were stored at -80 °C and -20 °C, respectively, and when required, aliquots were thawed on ice and subsequently stored at 4 °C. Antibodies (either concentrated or diluted in blocking solution) were supplemented with 0.025 % (w/v) sodium azide for prolonged storage at 4 °C.

Table 2.6.3.1: List of antibodies used in the study.

Antibody	Type	Dilution	Source
Mouse α Lpp	Primary	1:400,000	Thomas Silhavy, Princeton University, USA.
Mouse α BamA	Primary	1: 20,000	Susan Buchanan, NIH, USA
Rabbit α BamB	Primary	1:20,000	In-house, Lithgow laboratory
Rabbit α BamC	Primary	1: 20,000	In-house, Lithgow laboratory
Rabbit α BamD	Primary	1:20,000	In-house, Lithgow laboratory
Rabbit α BamE	Primary	1:20,000	In-house, Lithgow laboratory
Mouse α BamE	Primary	1:500	Susan Buchanan, NIH, USA
Rabbit α Pet	Primary	1;5000	In-house, Lithgow laboratory
Goat α Mouse-HRP	Secondary	1:20,000	Sigma (A4416)
Goat α Rabbit-HRP	Secondary	1:20,000	Sigma (A6154)

Following a Western transfer, nitrocellulose membranes were incubated for 1 hour in blocking solution at room temperature or overnight at 4 °C with gentle rocking. The blocking solution was removed, and the nitrocellulose membranes were incubated for one hour in a diluted primary antibody. The nitrocellulose membranes were sometimes incubated with the primary antibody overnight at 4 °C. The primary antibody solution was removed and stored at 4 °C for reuse, and the nitrocellulose membranes were washed 5 times for 8 minutes in Western washing buffer. The nitrocellulose membranes were incubated for 1 hour in a diluted secondary antibody solution and washed as before.

Proteins were detected using the enhanced chemiluminescence method with the ECL Prime Western blotting detection reagent (GE Healthcare Life Sciences) as per the manufacturer's instructions. Chemiluminescent membranes were then exposed to Super RX-N film (Fujifilm) in an Amersham Hypercasesette™ (GE Healthcare Life Sciences) for up to 1 hour, depending

on the signal strength. Usually, a good signal was detected within one minute. Immunoblot images were scanned using CanoScan 8600F with software version 5.0.1.2 (Canon).

2.6.4 Membrane isolation from *E. coli*

Saturated overnight *E. coli* cultures were first washed twice in 1x M9 salts. The cultures were then diluted into 200 ml of fresh M9 minimal media supplemented with 0.2% (w/v) glucose to a starting OD₆₀₀ of 0.05 AU. Cultures were grown to the logarithmic phase (OD₆₀₀ ≈ 0.5 AU) and cells collected by centrifugation (5000 × g, 10 min, 4 °C). The cell pellets were then resuspended in TS buffer (0.75 M sucrose, 10 mM Tris-HCl pH 7.5). To aid in cell lysis and to minimize proteolysis mediated by endogenous proteases, lysozyme (50 µg/ml), PMSF (2 mM) and two volumes of 1.65 mM EDTA (pH 7.5) were added. The cells were then lysed with an EmulsiFlex (Avestin Inc.) at 15,000 psi in 2-3 passes. Cell debris was removed by centrifugation (15000 × g, 10 minutes, 4 °C). From the supernatant, the membrane fraction was collected by ultracentrifugation (100000 × g, 10mM Tris-HCl, 1 hour, 4 °C). The membranes were then washed by resuspending the membrane pellet in TES buffer (1.1 mM EDTA, 3.3 mM Tris-HCl, 0.25 M sucrose pH 7.5), and re-isolating the membranes by a second round of ultracentrifugation (100000 × g, 1 hour, 4 °C). Membranes were resuspended in a final volume of 400 µl, 25% (w/v) sucrose in 5 mM EDTA (pH 7.5), and stored in aliquots at -80 °C.

2.6.5 Quantitative proteomics

2.6.5.1 Sample preparation

Saturated overnight *E. coli* cultures were washed twice in 1x M9 salts and diluted 1:100 in 10 ml of M9 minimal media supplemented with 0.2% (w/v) glucose and M9 minimal media supplemented with both 0.2% (w/v) glucose and 0.5 M D-sorbitol (Sigma-Aldrich). Cultures were grown to the logarithmic phase (OD₆₀₀ ≈ 0.5 AU) and collected by centrifugation (5000 × g, 10 min, 4 °C). Cells were then washed using PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄) and collected by centrifugation, same as before. The pellet was stored at -80 °C.

The cell pellet was homogenised in 1 ml of 4 % SDS, 100 mM Tris, pH 8.1 and boiled at 95 °C for 10 minutes. The lysate was then sonicated with a Bioruptor® Pico (Diagenode) for 20 cycles with the setting of 30 seconds on and 30 seconds off. The lysate was again boiled as described previously, and the insoluble pellets were removed by centrifugation (16000 x g 5 minutes). The protein concentration was determined with Bicinchoninic Acid assay (BCA, Thermo Fisher) according to the manual.

2.6.5.2 Reduction, alkylation

To prepare protein samples for mass spectrometry, reduction and alkylation reactions were carried out, using equal amounts of protein from each the replicates. Proteins were reduced with 10 mM DTT by incubating at 50-65 °C for 30 minutes and alkylated with the addition chloroacetamide (CAA) to a final concentration of 40mM followed by incubation in the dark for 20 minutes.

2.6.5.3 Chloroform/methanol precipitation

Samples were kept on ice in all the following steps. First, the sample volume was adjusted to 80 µl with lysis buffer. 400 µl ice-cold methanol was added to the samples, mixed well and then 200 µl of ice-cold chloroform was added. Samples were mixed and chilled on ice for 5 minutes. 300 µl ice-cold water was then added, mixed well, and incubated on ice for another 5 minutes. Centrifugation was followed (20000 x g, 4 °C, 5 minutes) to separate the aqueous and organic phases. The upper aqueous phase was carefully removed, avoiding disturbance of the proteins in the interphase. 200 µl of ice-cold methanol was added to the organic phase and the denatured protein was precipitated by centrifugation (20000 x g, 4 °C, 5 minutes). The supernatant was removed, and the protein pellet was washed by resuspending in 500 µl ice-cold methanol and centrifugation as described previously. The pellet was then air-dried for 20 minutes in a fume hood.

2.6.5.4 Trypsination and peptide clean-up

The protein pellet was resuspended in 100 µl Triethylammonium Bicarbonate buffer (TAEB), (Sigma-Aldrich T7408) and mixed vigorously by vortexing to ensure increased trypsin access to the protein substrates. Trypsin (Promega, V528X) was added to the samples at a 1:100 ratio; 1 µg trypsin to 100 µg of protein substrate, and digestion continued overnight at 37 °C on a ThermoMixer® (Eppendorf) with shaking. Trypsin digestion was stopped by addition of 1 µl of neat formic acid and the samples centrifuged (16000 x g, 5 minutes) to remove any insoluble material.

Tryptic peptides were further purified using ZipTips® (Agilent) and dried in the SpeedVac (Labconco). The samples were reconstituted in 20 µl of 0.1% formic acid and sonicated for 10 minutes in a sonicator water bath. Samples were centrifuged at 16000 x g for 5 minutes and transferred to mass spectrometry vials.

2.6.5.5 Mass spectrometric acquisition

Mass spectrometry made use of a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75 μm x 50 cm, nanoViper, C18, 2 μm , 100Å; Thermo Scientific) and an Acclaim PepMap 100 trap column (100 μm x 2 cm, nanoViper, C18, 5 μm , 100Å; Thermo Scientific). The tryptic peptides were separated by increasing concentrations of 80% ACN / 0.1% formic acid at a flow of 250 nl/min for 120 minutes and analysed with a QExactive Plus mass spectrometer (Thermo Scientific). The instrument was operated in the data-dependent acquisition mode to switch between full-scan MS and MS/MS acquisition automatically. Each full survey scan (m/z 375-1575) was acquired in the Orbitrap with 60,000 resolution (at m/z 200) after an accumulation of ions to a 3×10^6 target value with a maximum injection time of 54 ms. Dynamic exclusion was set to 30 seconds. The 20 most intense multiply charged ions ($z \geq 2$) were sequentially isolated and fragmented in the collision cell by higher-energy collisional dissociation (HCD) with a fixed injection time of 54 ms, 15000 resolution and automatic gain control (AGC) target of 2×10^5 . MS acquisition was done at Monash Proteomics and Metabolomics Facility (Dr Enzo Cheng Huang).

2.6.5.6 Proteomics data analysis

The raw data files were analysed using MaxQuant software suite v1.6.5.0 (Tyanova *et al.*, 2016) against Andromeda search engine (Cox *et al.*, 2011) for protein identification and to obtain their respective label-free quantification (LFQ) values using in-house standard parameters.

Statistical analysis was performed using Perseus software. The LFQ data was converted to \log_2 scale, samples were grouped by conditions, and missing values were imputed. Two sample t-test was performed, and the p-value was adjusted with Benjamini-Hochberg algorithm (Benjamini & Hochberg, 1995).

2.7 Cell-based techniques

2.7.1 Automated growth analysis

Single colonies were inoculated for overnight culture in LB media. The cultures were diluted 1:100 in LB media and grown to $\text{OD}_{600} \approx 0.5$ AU. Dilution was repeated to a final OD_{600} of 0.001 AU in round-bottom 96-well plates using LB or minimal media to a final volume of 200 μl . For minimal media growth analysis, the log phase cultures were washed twice in 1x M9 salts before the dilution. Growth was monitored by measuring the absorbance at 600 nm every

30-60 minutes in a SparkTM 10M microplate reader (Tecan, Australia) with constant orbital shaking at 37 °C. Graphs were prepared using GraphPad prism® version 8.0 (GraphPad Software, Inc.)

2.7.2 Antibiotic susceptibility testing and SDS sensitivity analysis

Minimum inhibitory concentrations (MICs) of antibiotics were determined by broth microdilution (BMD) reference method according to Clinical Laboratory Standards Institute (CLSI) 2015 guidelines (Jorgensen & Turnidge, 2015). The tests were performed in triplicate, repeated thrice, and median MIC values were used for analysis.

SDS sensitivity was determined by the same method as MICs in LB media with minor modifications. 100 µL (2.0×10^6 CFU/ml) of each bacterial strain was added to 100 µL of varying concentrations of SDS (10%-0.05%) diluted in LB media. After 24-hour incubation at 37 °C, the minimum SDS concentration inhibiting growth was obtained by checking the turbidity of the bacterial growth. The minimum inhibitory value corresponded to the concentration that inhibited 99% of bacterial growth. SDS sensitivity was also determined by streaking strains, in duplicate, in LB media (solid) with increasing concentration of SDS (0.05%-5%). LB media without SDS was used as a baseline growth control.

2.7.3 Construction of Pet fusions

As detailed in Chapter 3 Introduction, Pet is an autotransporter. The open-reading frame encoding Pet was amplified from pBADpet and cloned into pJP168 expression vector (Table 2.1.2), to create plasmid template onto which the genes encoding heterologous proteins were cloned for further Pet β -barrel assembly assays. The proteins for Pet fusion were chosen based on initial considerations of assays for detection upon translocation to the extracellular milieu. The genes encoding these proteins were synthesised *de novo* and cloned into pJP168, replacing the region encoding the Pet passenger domain. Four proteins were chosen: (1) α -Amylase (AmyA), (2) mCherry, a red fluorescent protein and a derivative of *Discosoma sp* DsRed, (3) alkaline phosphatase (PhoA), a periplasmic *E. coli* dephosphorylating enzyme, and (4) super fold green fluorescent protein (sfGFP) a GFP variant that folds robustly even when fused to poorly folded proteins. The Pet fusion constructs were made using techniques described in Section 2.4.

2.7.4 Growth in starch; an assay for amylase translocation

Starch is a large polysaccharide that cannot cross the outer membrane but can be hydrolysed to glucose by extracellular amylase. The strain expressing Pet-amylase chimera was inoculated

in M9 minimal media (solid) supplemented with 5 g/l soluble starch (Sigma-Aldrich) as the sole carbon source and 50 ng/ml anhydrotetracycline (ATC) to induce the expression of the Pet⁻amylase fusion construct. The plates were incubated at 37 °C for 5 days for growth assessment. Control strains containing empty pJP168 were also inoculated.

The assay was also performed in high-throughput by robotic pinning of strains using the Singer RoToR HDA robotics platform (Singer instruments). Strains, including controls, were first arranged manually in 96-well plates containing M9 minimal media supplemented with 0.2% (w/v) D-maltose (Sigma-Aldrich). After 48-hour incubation at 37 °C, colonies were replica-pinned to M9 minimal media (solid) supplemented with 5 g/l soluble starch and 50 ng/ml anhydrotetracycline (ATC) in both 96 and 384-colony density formats. The plates were incubated at 37 °C for 5 days for growth assessment.

2.7.5 Lugol's iodine starch staining assay

Plates from starch amylase assays were stained using Lugol's iodine solution (10% potassium iodide, 5% iodine crystals). The solution was further diluted to 0.25% (v/v) using water and stored in the dark. Plates were flooded with the solution for 1 minute at room temperature to stain starch. Following staining, the solution was removed and the plates incubated for 30 – 40 minutes at room temperature for the destaining of the simple sugars. The plates were then imaged using the PhenoBooth (Singer instruments) for analysis.

2.7.6 Live/Dead cell viability assay

Single colonies were inoculated for overnight culture in LB media. The cultures were diluted 1:100 in LB + (0.4 M NaCl) media and grown to OD₆₀₀ ≈ 0.5 AU. Dilution was repeated to adjust the cell concentration to a final OD₆₀₀ of 0.3 AU. 20 µl aliquots were obtained for staining. Staining was performed using the LIVE/DEAD® BacLight™ bacterial viability kit (Thermo Scientific) as per the manufacturer's instructions.

2.8 Lipopolysaccharide quantification by silver staining

Membranes were isolated from *E. coli* strains as described in Section 2.6.4 and the amount of membrane material normalised according to protein content before loading in SDS gels. SDS-PAGE analysis was performed as described in Section 2.6.1, and the gels were silver stained for outer membrane lipopolysaccharide (LPS) quantification as detailed below.

The gels were washed twice in ultrapure water and transferred to a fixer solution (30% ethanol, 10% acetic acid) for 2 – 16 hours (overnight) at room temperature, the solution was replaced

once during the period. The gel was oxidised by placing in a solution made up of 700 µl of 50% (w/v) periodic acid added to 100 ml of fixer solution for 5 minutes at room temperature with gentle shaking. The oxidising solution was removed by washing the gel thrice for 15 minutes in ultrapure water. Staining proceeded using the Pierce Silver Stain kit (Thermo Scientific) as per the manufacturer's instructions.

2.9 Computer-based methods

In analyses of proteins sequences to determine the presence or absence of a signal sequence for protein translocation, and also for determining the primary sequence of a mature protein following signal peptide cleavage, SignalP version 5.0 was used (Armenteros *et al.*, 2019). TMHMM server version 2.0 was used to predict the presence and orientation of transmembrane α -helices in protein sequences (Krogh *et al.*, 2001). ExPASy ProtParam was used to estimate the molecular weight and to calculate the distribution of amino acids or atoms within a protein (Gasteiger *et al.*, 2003), while multiple sequence alignments were performed using the MUSCLE algorithm (Edgar, 2004).

EcoCyc, a database for *E. coli* K-12 MG1655 (Keseler *et al.*, 2016), was used as an electronic reference source for analysis of data retrieved from high-throughput screens. Some of the information retrieved from the database includes information on gene sequence and function, operon information, protein sequence features, and protein localisation. The database also includes information on *E. coli* gene essentiality on nutrient conditions that do or do not support the growth of *E. coli*.

Protein structures were predicted using Phyre2 (Kelley *et al.*, 2015) and the Dali server (Holm, 2019) was used to compared protein structures in three-dimensional space (3D) using Dali method (Holms & Sander, 1993). The Dali server was particularly crucial in searching for proteins with similar domains across the protein data bank (PDB) database. Rendering of protein images to PNG files, including superimpositions, of models and crystal structures was done by use of PyMolTM Open-source version 2.3.4 (Schrödinger).

Figures were prepared using Adobe Illustrator, version CC 2020 unless otherwise stated.

Chapter 3: A high-throughput screen to investigate autotransporter biogenesis

3.1 Introduction

Bacteria possess nine secretion systems, numbered Type I through to Type IX. Type VII is found only in Gram-positive bacteria while the rest are found only in Gram-negative bacteria. These secretion systems facilitate the navigation of secreted proteins through the IM, periplasm, and the OM. The Type V secretion system (T5SS) is widespread in Gram-negative bacteria allowing the traffic of proteins across the OM for secretion or display of surface-associated structures (Fan *et al.*, 2016). A sub-class of this system were referred to as autotransporters (ATs), as they were previously thought to translocate through both membranes independent of accessory factors (Halter *et al.*, 1984), an idea that has been revised (Desvaux *et al.*, 2004). Thus, the T5SS is now categorised into six sub-classes: T5SSa (classical ATs), T5SSb (two-partner secretion ATs), T5SSc (trimeric ATs), T5SSd (secreted phospholipases), T5SSe (inverse ATs) and T5SSf (two-partner inverse ATs). This chapter focuses on T5SSa sub-class, which contains the model protein used in this study.

Unlike the other AT sub-classes, classical ATs (T5SSa) are characterised by their monomeric nature. The structure of a T5SSa protein is composed of three major domains that contribute to its secretion; a cleavable N-terminal signal sequence that directs the protein via Sec-dependent translocation to the periplasm, the passenger domain that contains the effector function of the protein, and the C-terminal domain that forms a β -barrel and which acts a translocation unit for the export of the passenger domain through the OM (Pohlner *et al.*, 1987).

The molecular mechanisms of AT biogenesis are not entirely understood. However, it is established that a variety of accessory factors are necessary for the release of the AT passenger domain into the extracellular milieu (Jain & Goldberg, 2007; Rossiter *et al.*, 2011; Leyton *et al.*, 2014) (Ruiz-Perez *et al.*, 2009; Selkrig *et al.*, 2012). The BAM complex catalyses the steps in the assembly of ATs in the OM as it does for other types of β -barrel proteins. BamA and BamD, the essential components of the complex, have been shown to work in concert during the insertion of AT β -barrel into the OM (Jain & Goldberg, 2007; Sauri *et al.*, 2009; Rossiter *et al.*, 2011). Further analysis of the other non-essential components of the BAM complex (BamB, BamC, and BamE) established they are not required in AT biogenesis (Rossiter *et al.*, 2011). Apart from BAM, a variety of other accessory factors are necessary for the delivery of

a functional AT into the OM for β -barrel assembly. Periplasmic chaperones such as SurA and DegP also participate in the folding and insertion stages for AT (Ruiz-Perez *et al.*, 2009; Wagner *et al.*, 2009), as does the translocation and assembly module (TAM) that spans the outer and inner membrane (Selkrig *et al.*, 2012). Other implicated periplasmic chaperones include Skp, FkpA and VirK (Ruiz-Perez *et al.*, 2009; Ruiz-Perez *et al.*, 2010; Tapia-Pastrana *et al.*, 2012). It is thought that some of these factors might facilitate or retard aspects of protein folding, to assist overall coordination of the process.

This study aimed to address the details of the mechanism determining how ATs transverse the periplasm to reach the OM for surface display. To study this question, an archetypal AT called the plasmid-encoded toxin (Pet) was used. Pet is a functionally well-characterised cytotoxin encoded on a 65-MDa adherence related plasmid of Enteroaggregative *E. coli* (EAEC) strain 042 (Eslava *et al.*, 1998). Analysis of the toxins nucleotide sequence suggested that it is a member of the autotransporter class of proteins designated as serine protease ATs of the *Enterobacteriaceae* (SPATEs) (Eslava *et al.*, 1998). Pet is characterised by an N-terminal signal sequence terminating at a cleavage site between residues A52 and A53, a passenger domain that contains the toxin, and a conserved C-terminal domain that folds into a β -barrel in the bacterial OM and through which the functional domain (toxin) is translocated (Eslava *et al.*, 1998). Upon export of the passenger domain, Pet undergoes a posttranslational processing step where the cleavage of the passenger domain from the β -domain takes place, releasing the toxin into the extracellular milieu (Eslava *et al.*, 1998; Domingo Meza-Aguilar *et al.*, 2014)

T5SSs have been adapted for successful extracellular accumulation of a variety of heterologous target molecules in the bacterial cell surface in a process termed autodisplay (Kjaergaard *et al.*, 2002; Jose & Meyer, 2007). The passenger domain of autotransporters displays structural and functional heterogeneity and can be substituted with heterologous proteins (Henderson *et al.*, 2004). Sevastsyanovich *et al.*, (2012), demonstrated the ability of this system to secrete heterologous proteins into the culture media using SPATE protein Pet. The authors cloned proteins of distinctive size, structural and functional properties in-frame of the *pet* gene, replacing a relatively significant portion of the passenger domain and giving rise to chimeric proteins. All the chimeric proteins were secreted to the culture medium at levels stoichiometrically, similar to the wild-type Pet (Sevastsyanovich *et al.*, 2012). The SPATEs class of ATs are preferred as they do not require additional proteases to release the heterologous passenger protein (Sevastsyanovich *et al.*, 2012).

The prevailing model for AT secretion comprises entry to the periplasm via the Sec translocation apparatus, followed by an unclear series of steps in which the C-terminus of the periplasmic AT inserts into the OM as a β -barrel protein followed by the translocation of the passenger domain to the bacterial surface. The information regarding the fate of ATs in the periplasm, including the accessory proteins involved in Pet's translocation across the periplasm is still insufficient despite the identification of critical periplasmic chaperones involved in the process (Ruiz-Perez *et al.*, 2009; Ruiz-Perez *et al.*, 2010; Tapia-Pastrana *et al.*, 2012). Apart from the well-known chaperones, I was interested in investigating whether there are any other accessory proteins that aid in AT periplasmic translocation, either working in concert with the known chaperones or independently. To answer these questions, I aimed to develop a high-throughput Pet biogenesis assay. The plan was to use the assay to screen the Keio collection, a single gene knock-out library of all non-essential genes in *E. coli*, to determine single-gene deletion clones that impair AT periplasmic translocation or screen a double-gene knock-out library, where a known Pet chaperone is combined with an extra 'Keio' single-gene knock-out.

3.2. Set-up and optimisation for the high-throughput generation of double-gene knock-outs for synthetic lethal screen application

3.2.1 Keio collection rearrangement

The source of mutants for the Pet biogenesis screen was the Keio collection (Baba *et al.*, 2006). This collection is sourced in 96-well format covering 3,985 non-essential *E. coli* single-gene deletion clones, marked with a kanamycin-resistance cassette (*kan*). The library contains two independent biological knock-outs per gene for a total of 7,970 clones distributed in 90 plates. This number of plates are laborious and cumbersome to work on, and for this reason, the 96-well plates were robotically rearranged to a 384-well format using a Singer RoTor robotics platform. Four-96 well plates were combined into one-384 well plate, significantly reducing the number of plates. Figure 3.2.1.1 summarises the operations needed in order to re-plate these clones to a 384-well format. See section 2.5.2 for a detailed procedure.

The rearrangement was not only to make the screen convenient by improving the efficiency of handling the library but to also provide means to include controls. In each plate, empty wells were included as internal quality controls and negative conjugation controls. JW5028 strain, a Keio collection clone containing a kanamycin resistance gene in place of a pseudogene, was added to the vertical border wells as a positive control for conjugation required for the high-throughput generation of double-gene knock-outs described in the next section. The JW5028

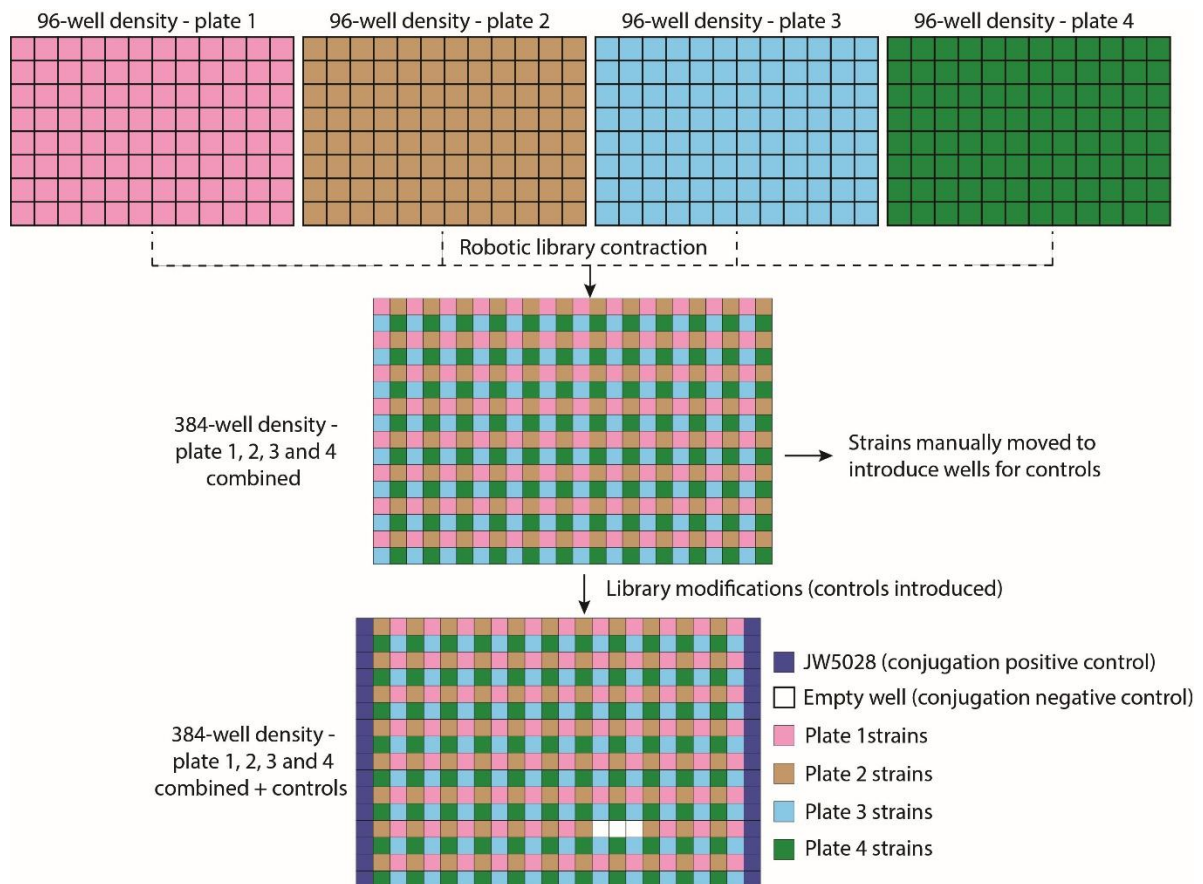


Figure 3.2.1.1. Robotic-genetic selection screen.

The source for all screens in this thesis was the Keio collection. This collection of *E. coli* mutants is sourced in 96-well formats across 90 microtiter plates (the collection contains additional unannotated strains). The schematic summarises the operations needed in order to re-plate these clones for a robotic-genetic screen of the collection. The screening format of tailored versions of the Keio collection made use of a final format in 384-well plates, as shown in the final panel. The placement of positive- and negative-controls is shown as per the colour-coding in the legend.

strain has been previously validated to confirm it displays no genetic interactions with any individual *E. coli* genes (Gagarinova *et al.*, 2012).

3.2.2 High-throughput generation of double-gene deletions – a pilot study

To generate double-gene knock-out clones, I utilised the already developed eSGA (Butland *et al.*, 2008) and GIANT-coli (Typas *et al.*, 2008) techniques discussed in detail in Section 1.7.2. The techniques exploit the Hfr conjugation system and take advantage of the already available Keio collection (Baba *et al.*, 2006) as the recipient library. The techniques, though based on the same principle have minor variation in methods. To get the best out of these techniques, I designed a pilot study to assess the best conditions.

The query gene, the second gene deleted on the single-gene knock-out recipient clone, was first replaced with a chloramphenicol resistance gene (*cat*) in the Hfr donor (F^+) strain using a λ red-mediated system for homologous recombination (see Section 2.5.1 for detailed methods) (Datsenko & Wanner, 2000). The donor strain, with the query gene deleted, was then mated with a sample plate of single-gene deletion clones from the Keio collection, which are all recipients (F^-). To mate the donor with all 384 sample clones, a Singer RoTor HDA robotics platform was used. The platform uses A2 plus plates that can accommodate a variety of desired colony densities (96, 384, 1536 or 6144). Post mating the conjugants were subjected to two rounds of selection on relevant antibiotics to generate the double-gene deletions. All the steps to this stage were automated. The double recombinant clones were then imaged for further analysis. The primary reason for the development of these techniques (eSGA and Giant-coli) was to investigate genetic interactions. This was accomplished by determining the relative fitness score of each clone achieved by comparing the colony sizes between the double-gene knock-outs and the single-gene knock-out (see Section 2.5.4 or 4.4.2 for more details).

As an optimisation study, I trialled several variations: growth medium, length of incubation time in each step, use of one or two antibiotics in the intermediate selection step, the incubation temperature, robotic pinning parameters, and density of colony pinning. For the media used, tests were conducted on both rich (LB) media and nutrient-limited (M9) media. The incubation time was varied depending on the media used and the step of the technique, e.g. the conjugation step was extended to 16 hours in LB and 24 hours in M9 minimal media to avert the problem of poor transfer of markers distal to *oriT*, as the Hfr chromosome transfer to the recipient is time-dependent. Another significant test was the number of antibiotics to use in the first selection step as it is a crucial step in determining the frequency of false negatives.

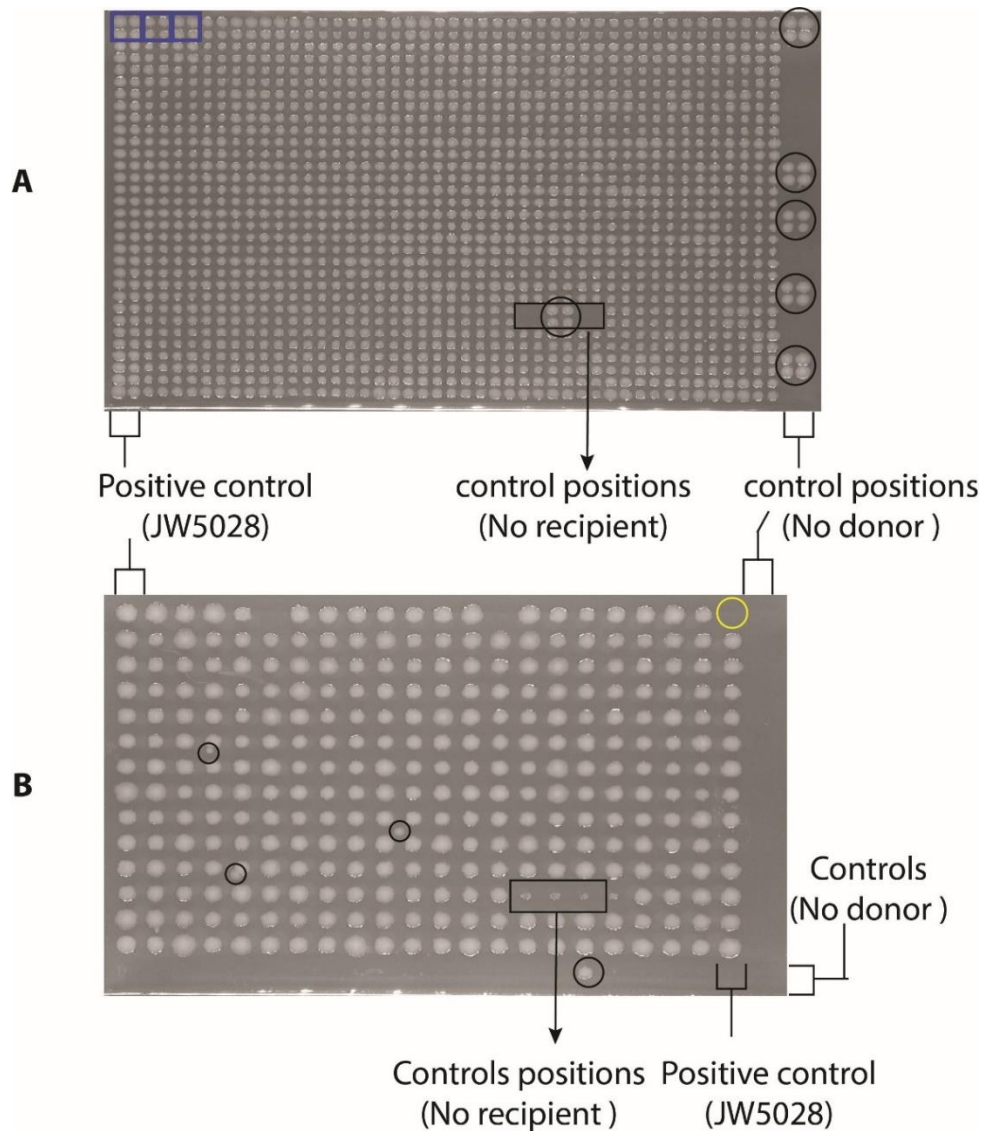


Figure 3.2.2.1: High-density generation of double-gene knock-outs; 1536 and 384-density format; a pilot study.

The panels present results for high-density gene knock-out on a background of single-gene deletion genes (Keio collection). Panel A shows the results in 1536-density format, where each clone is replicated four times (blue squares) while panel B shows the results in 384-density format, where each spot is a unique clone, but all with a common query gene deletion (*cat* marker transfer from the donor). Black rectangles indicate internal controls where recipients were excluded. Border controls where the donor is excluded are indicated. Black circles show false-negative results and contamination while the yellow circle shows a false-positive result (False synthetic lethal interactions). Plates from both panels show second selection results, pinned from first selection LB media plates. Both first and second selections of the double-gene knock-outs contained Kan (30 $\mu\text{g/ml}$) and Cm (34 $\mu\text{g/ml}$). Panel A plate was incubated at 30 $^{\circ}\text{C}$ for 14 hours while panel B plate was incubated at 37 $^{\circ}\text{C}$ for 16 hours.

I also considered which among two colony densities, 384 or 1536, was better for the screen. The 1536-density shown in Figure 3.2.2.1A was, at first, the preferred choice due to the higher number of replicates. At this density, the number of technical replicates on one plate is four. However, due to the close proximity between the colonies, higher rates of false-negatives (unexpected growth) were observed likely due to mixing of neighbouring clones. In an attempt to avert this problem, I shortened the incubation time and lowered the temperature to 30 °C to avoid colony mixing as the clones grew, which resulted in unreproducible growth results from one plate to another. The 384-density result shown in Figure 3.2.2.1B, though with less number of technical replicates, was superior due to reproducibility of colony size from one plate to another. Also noted at 384-colony density, there were no false-negatives on the control positions were recipients were excluded unlike at 1536-colony density. However, as Figure 3.2.2.1B shows, other problems were noticed at 384-colony density that also needed optimisation.

Among the issues noted in the analysis of Figure 3.2.2.1B was contamination or splashing, seen as small colonies next to the main pin colony (small black circles). The problem was fixed by adjusting the robot pinning parameters for the target plate. These included trial-and-error adjustments to: the pin pressure, the pinning speed, the diameter of pinning while mixing, and the overshoot of the pin (how deep the pin goes in the agar). The contamination/splashing problem was eliminated when the above parameters were adjusted as follows: pin pressure was set at 7%, the pinning speed was set at 19 mm/sec, the overshoot was set at 2 mm, and the pinning diameter was set at 1 mm. Another problem noted was the unexpected absence of growth, yellow circled positions (Figure 3.2.2.1B). This likely occurred when either the donor or recipients were not transferred to the conjugation plate. In this case, the problem presents a false-positive result for the recipient gene in that location. The problem was fixed by increasing the number of repeats while pinning from the source plates, which ensured that enough cells were transferred to the target plates. These parameters were kept constant for the rest of the screen for reproducibility and consistency of results.

The use of either one or two antibiotics for the first selection was evaluated in both LB and M9 minimal media. The eSGA method recommended the use of both antibiotics, while the Giant-colli method recommended the use of one antibiotic. Using ‘self-mating’ pairs, when the query and ‘Keio’ genes are the same, this inconsistency was tested. Since *E. coli* is haploid, self-mating pairs are synthetic lethal. Using this notion, the frequency of recovery of false-positives due to duplication of the chromosome was evaluated. The recovery of false-positives was lower

when the transconjugants were first selected on a single antibiotic, kanamycin, which also selects the recipient parent. The frequency was even lower in M9 minimal media compared to LB media. For this reason, the Giant-coli recommendation to use a single antibiotic for the first selection was adapted.

3.2.3 Optimised high-throughput generation of double-gene recombinants

Following the optimisation-pilot studies, a detailed procedure shown in Figure 3.2.3.1 was developed, incorporating the best options chosen in all the variable aspects investigated. As the strategy was to perform the screens in both the LB and M9 minimal medium, separate procedures were developed for each media condition.

The donor and the recipient clones were pinned in 384-colony density and incubated at 37 °C for 16 hours in LB and 20 hours in M9 minimal media supplemented with 0.2% (w/v) glucose (M9-glucose). Mating between the donor and the recipients was performed in either LB or M9-glucose without antibiotics, and the plates were incubated at 37 °C for 16 hours in LB or 24 hours in M9-glucose. The intermediate (first) selection was done in one antibiotic (Kan 30 µg/ml) and the plates were incubated at 37 °C for 16 hours in LB or 24 hours in M9-glucose. The double-gene knock-outs were then transferred to the final selection plates, containing double antibiotics (Kan 30 µg/ml and CM 34 µg/ml) and incubated at 37 °C for 14 hours in LB or 25-30 hours in M9-glucose. The double-gene deletion clones were then imaged. All the robotic pinning parameters determined in the previous section were kept constant in all pinning steps in both conditions. For reproducibility purposes, the query gene was screened independently twice against the recipient array collection to test each combination of mutations in duplicate.

Figures 3.2.3.2 shows the results using the optimised procedure. A total of 768 double-gene knock-outs (one plate of 384 clones in duplicate) were generated in six days. Since the entire Keio collection can be pinned in 2 days, including duplicates, it is possible to generate over 15,000 double-gene knock-outs in 12 days, exhibiting the power of using conjugation as a tool for high-density gene knock-out. In addition to the generation of double-gene deletions, lethal or alleviating genetic interactions were analysed. Figure 3.2.3.2A shows the single-gene knock-out recipient plate, imaged before the start of the procedure and Figure 3.2.3.2B shows the image of the final plate after the gene knock-out procedure. The plates were compared side by side for both duplicates, and as shown a lethal genetic interaction was observed. The clone was viable as a single-gene deletion mutant (red square in Figure 3.2.3.2A) but failed to grow when

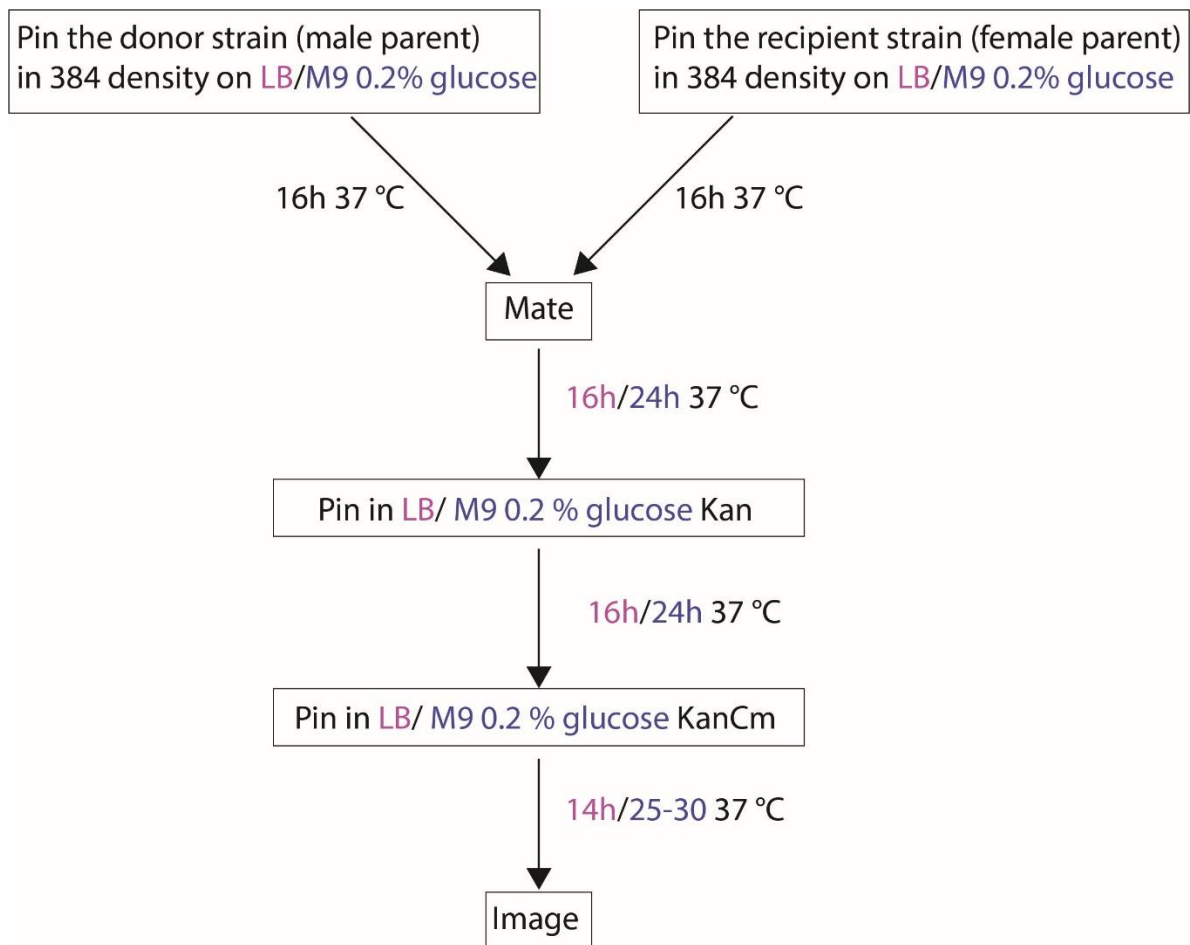


Figure 3.2.3.1: Optimised high-throughput synthetic lethal screen

The figure depicts a flow diagram of a detailed high-throughput gene knock-out protocol developed by optimisation of the eSGA and the GIANT-coli protocols. Magenta and blue italics represent specific steps in the rich and minimal media protocol, respectively.

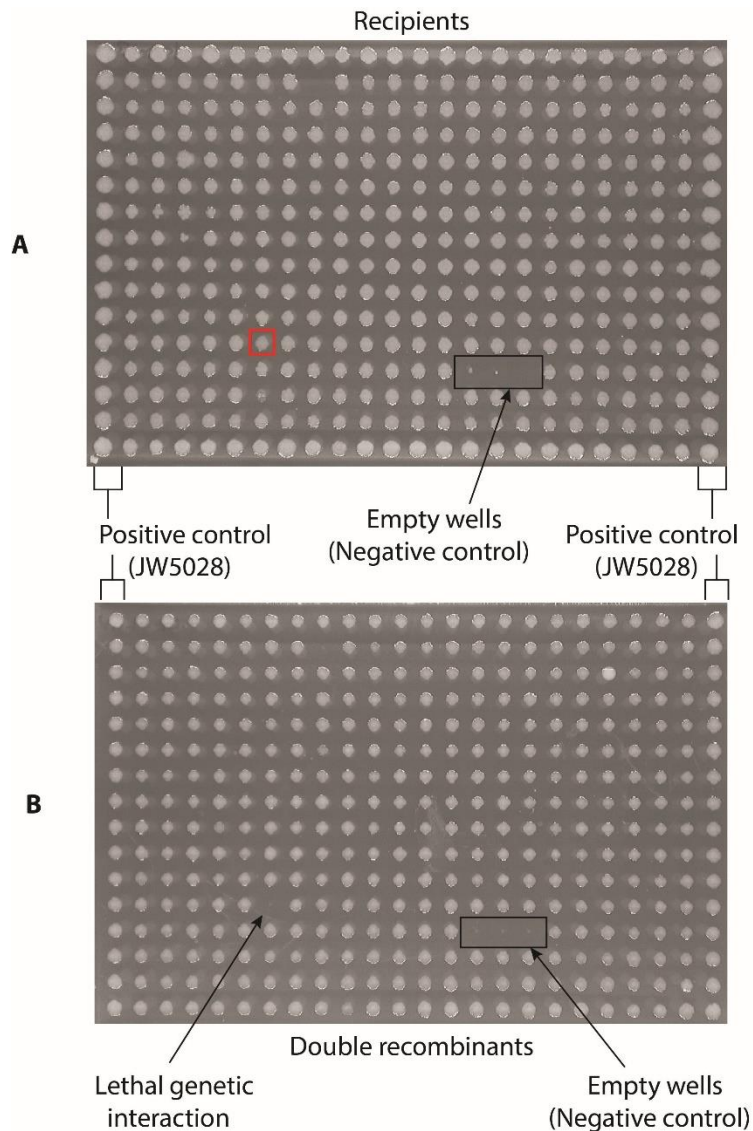


Figure 3.2.3.2: High-throughput gene knock-out in 384-density format

(A) The image presents a sample recipient's plate, captured before the start of the procedure. Each spot (colony) is a unique single-gene knock-out clone in the Keio collection library. (B). The image presents the double-gene knock-outs, captured at the end of the second selection process. Each spot is a double-gene knock-out clone composed of a unique 'Keio' gene knock-out and the query gene deletion, which is common in all the clones. For both panels, the first and the last columns (1 and 24) show the positive control, strain JW5028, which has not been shown to exhibit any genetic interactions with any other gene in *E. coli*. Position M15, M16, and M17 indicated as "empty well", act as negative controls for qualitative analysis of the pinning process and conjugation. The recipients were excluded in these specific wells. No growth in position L7 is an example of a synthetic lethal interaction between the query gene and the 'Keio' gene at that specific position. The single-gene knock-out clone at position L7 is viable (panel A red square). The plates are representative data shown from experiments performed in biological duplicate (n=2).

the second gene (query) was deleted. In this case, the two genes exhibit a lethal genetic interaction and can be subjected to further analysis.

Since the process is highly automated, controls were included as shown. JW5028 is one of the Keio single-gene deletion strains with the *kan* gene replacing a pseudogene. The strain has been shown not to exhibit any genetic interactions in previous studies (Gagarinova *et al.*, 2012); hence, it was used as a positive control for conjugation. Negative controls were included as empty wells on the recipient plates, which were expected to remain blank at the end of the screening processes. The empty wells also served as internal quality controls scanning for contamination, splashing or other pinning associated anomalies that would exhibit false-positive results.

Post high-throughput double-gene knock-out, the double-gene recombinants generated were verified by colony PCR. Clones were randomly selected from different positions on the plate for representative analysis. This verification step was necessary to screen out gene duplications that may arise and also confirm both the presence and chromosomal location of both gene modifications targeting *kan* and *cat* genes as markers. Colony PCR verification screen was performed as described in Section 2.4.2. The results confirmed a successful generation of double recombinants and also showed a very low rate of gene duplication (false positives). A generated double-gene knock-out library can be stored for later use (see Methods section 2.5.3.2 regarding storage).

3.3 Development of a high-throughput autotransporter biogenesis assay

Following the set-up of the single-gene knock-out library in a format suitable for high-throughput screening, and the optimisation of the high-throughput generation of double recombinants, I aimed to develop a high-throughput assay on autotransporter biogenesis that would make use of these available resources.

Using Pet as a model AT, I projected to screen for non-essential *E. coli* genes that code for proteins involved in AT biogenesis precisely during translocation across the periplasm. The assay could also be applied to screen a double-gene knock-out library, where a known Pet chaperone is combined with an extra ‘Keio’ single-gene knock-out to identify genes that code for accessory proteins that work in concert with the known chaperones. To develop the assay, I needed a “cargo” which I could easily monitor the surface exposure and folding (i.e. being transported via a functional T5SS). Noting that the T5SS (and Pet in particular) has successfully been used to transport various foreign proteins previously (Henderson *et al.*, 2004)

(Sevast'syanovich *et al.*, 2012), was a precedent for using T5SS to display reporter proteins. For this assay, I exploited the heterologous nature of the Pet passenger domain, where a reporter protein was cloned for autodisplay. The amount of reporter protein secreted was used as an indirect measure of the role of the protein in question on Pet biogenesis. The aim was to identify new accessory proteins that work independently or in concert with known chaperones involved the biogenesis of Pet.

3.3.1 Candidate reporter proteins for assessment of Pet biogenesis

The success of an assay such as this would rely greatly on the ability to detect and quantify the reporter protein in a high-throughput style and a format similar to the rearranged clone library. Four potential reporter proteins were assessed: (a) cytoplasmic α -amylase (56.6 kDa), (b) mCherry (26.6 kDa), (c) alkaline phosphatase (48.2 kDa), and (d) super fold green fluorescent protein (26.7 kDa).

The four proteins were selected based on the potential ease of detection, quantification and the possibility of high-throughput application. The α -amylase AmyA is a starch degrading enzyme, expressed in *E. coli* but normally restricted to the cytoplasm. Export of this protein would facilitate the use of starch as a sole carbon source and thus growth of the cells correctly exporting the enzyme on starch. The alkaline phosphatase PhoA is a periplasmic *E. coli* dephosphorylating enzyme that can be colourimetrically detected using various alkaline phosphatase substrates such as 5-Bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro blue tetrazolium (NBT). mCherry is a red fluorescent protein, a derivative of *Discosoma sp* DsRed, preferred in many biotechnological applications due to its monomeric form that translates to faster folding and hence unlikely to perturb the system. mCherry absorbs light between 540-590 nm and emits light in the range of 550-650 nm (Shu *et al.*, 2006). Super fold green fluorescent protein (sfGFP) was previously engineered from GFP to achieve a faster maturation and higher stability and was shown to fold robustly even when fused to poorly folded proteins (Pedelacq *et al.*, 2006). sfGFP presents a maximum peak of excitation at 485 nm and a maximum peak of emission at 510 nm (Pedelacq *et al.*, 2006). Both mCherry and sfGFP can be detected by special imaging systems using filters specific to each protein's emission wavelengths for high-throughput application.

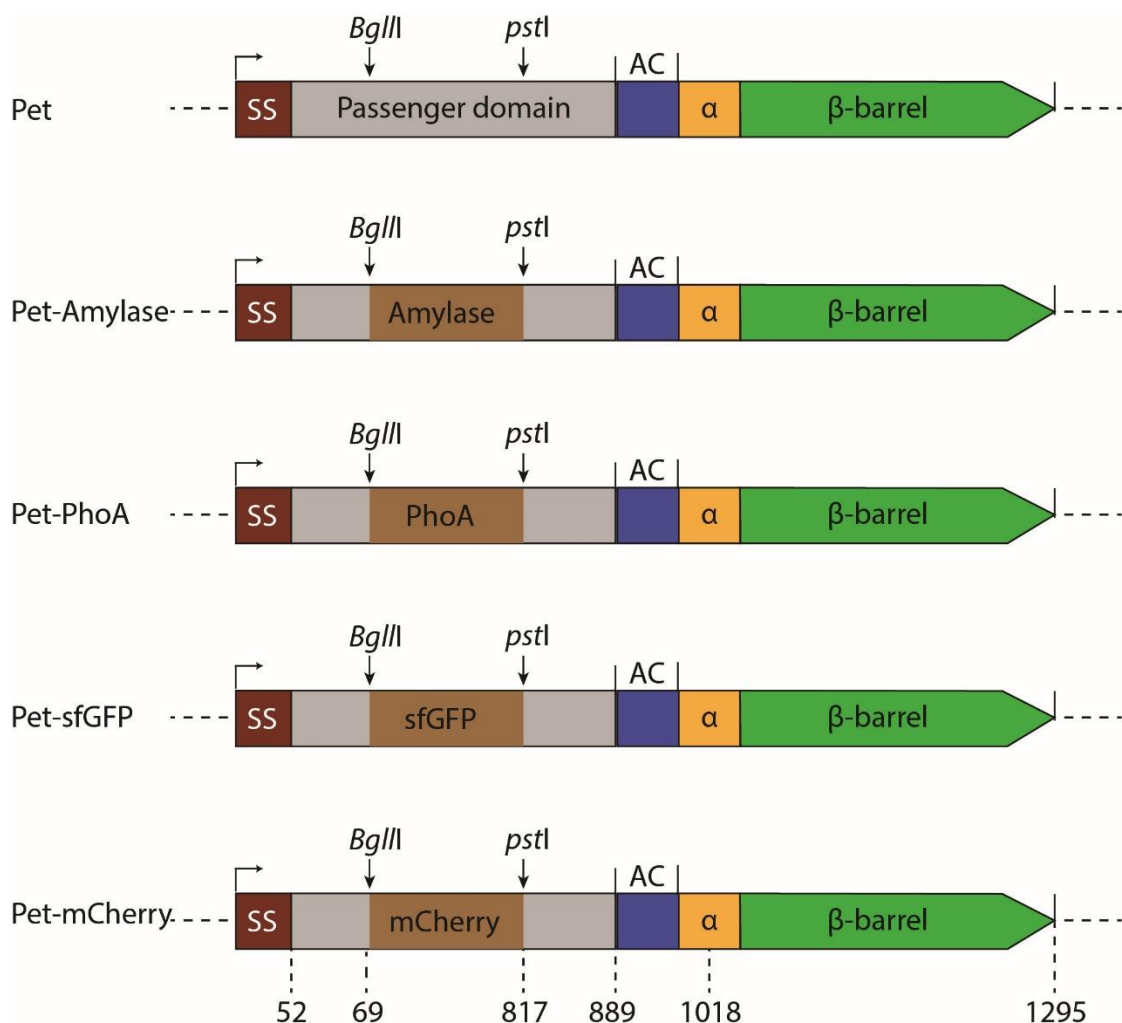


Figure 3.3.1.1: A schematic diagram of Pet fusion constructs.

The figure depicts the strategy used to design Pet chimeras. The names of the heterologous protein insertions are shown within the Pet passenger domain. The genes were cloned using the restriction sites *BglII* and *PstI* indicated by the arrows. The coordinates below the figure are given for the amino acids derived from *de novo* synthesised *pet* gene. Position 1018 denotes the cleavage site in the α -helix for the release of the passenger domain into the culture medium. Modification of this site results in surface display of molecules (Sevastyanovich *et al.*, 2012). The names of the resulting chimeras are shown on the left; abbreviations SS, signal sequence; AC, autochaperone domain; α , α -helix.

The gene encoding Pet was amplified from pBADPet (Leyton *et al.*, 2011) and cloned into A pJP168 expression vector, to create plasmid template onto which the genes encoding heterologous passenger proteins (synthesised *de novo*) were cloned. The Pet-chimeras were assembled as described in Section 2.7.4 and visually presented in Figure 3.3.1.1 above. Correct plasmid constructs were verified by PCR and sequencing before transformation into *E. coli* for assessment of Pet β -barrel production.

Production of the Pet β -barrel was monitored by immunoblotting using an anti-Pet antibody and indicated whether the new Pet-chimeras were functional. Since the heterologous protein is cleaved and released into the extracellular milieu, the molecular weight expected was that of the Pet β -barrel, ≈ 30 kDa determined by ExPASy-ProtParam tool from the point of cleavage.

In the maturation of native Pet, cleavage to release the passenger domain occurs after translocation and at positions Asn-1018 and Asp-1115 at the autocatalytic (AC) domain, located next to the β -barrel domain (Leyton *et al.*, 2011). As a measure of cleavage, the production of the Pet β -barrel was compared between induced and uninduced cells. The results presented in Figure 3.3.1.2 show AmyA-Pet β -barrel as the best-expressed construct. sfGFP-Pet β -barrel and PhoA-Pet β -barrel were also successfully expressed though at lower levels compared to AmyA-Pet β -barrel. mCherry-Pet β -barrel was barely detectable. Also noted was that the plasmid exhibited some leaky expression determined by faint bands at wells with uninduced samples. The β -barrel detected was of the correct size for all constructs, and an unspecific band of approximately 50 kDa was also detected in all samples in both conditions.

For protein induction of chimera β -barrel production, 1 ng/ml anhydrotetracycline (ATC) was used. Higher levels (>5 ng/ml) of ATC were investigated, but these resulted in toxicity of the chimeras in *E. coli* cells; cells stopped growing immediately upon induction with any of the four constructs. Slight toxicity was observed even at 1 ng/ml ATC, as the growth rate of *E. coli* slowed down significantly, though reached the desired cell density for analysis ($OD_{600} \approx 0.5$)

Based on the results, the presence of significant amounts of AmyA-Pet β -barrel indicated a stable chimera, and for this reason, I selected the AmyA-Pet chimera for further analysis and considered sfGFP-Pet chimera as the second choice. Apart from high expression levels, the AmyA-Pet chimera had a significant advantage in that it was easier to optimise the AmyA autodisplay capability for a high-throughput screen discussed in detail in the next section.

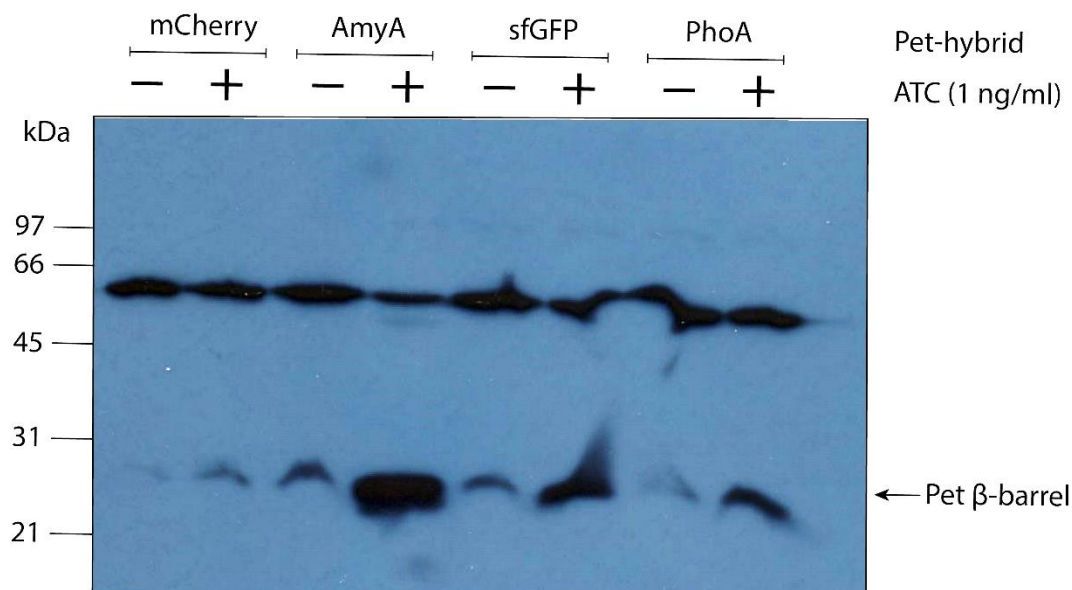


Figure 3.3.1.2: Pet- β -barrel production; an assessment of heterologous protein production.

The image depicts an immunoblot for detection of Pet β -barrel production as an assessment of functional production of Pet-chimeric proteins. Pet-chimera constructs were cloned in a pJP168 expression vector and the plasmid transformed in *E. coli* BW25113 cells. Whole-cell lysates were prepared using mid-exponential phase cells ($OD_{600} \approx 0.5$) of induced and uninduced cultures. For induction of cultures, anhydrotetracycline (ATC) was added to a final concentration of 1 ng/ml at $OD_{600} \approx 0.1$. Prior to SDS-PAGE analysis, whole-cell lysates were standardised based on cell concentration (OD_{600} readings). A 15% (w/v) acrylamide SDS-PAGE gel was used. + and – signs indicate induction and no induction, respectively. The expected molecular weight of the Pet β -barrel is ≈ 30 kDa determined by ExPASy-ProtParam tool (Gasteiger *et al.*, 2003). Also noted was an unspecific band of approximately 50 kDa detected in all samples in both conditions.

3.3.2 Engineering *E. coli* to use starch as the sole carbon source

α -Amylases are glycoside hydrolases that cleave α -1-4-glycosidic linkages in starch and maltodextrins. *E. coli* expresses two α -amylases, AmyA located in the cytoplasm (Raha *et al.*, 1992) and MalS located in the periplasm (Freundlieb & Boos, 1986). Raha *et al.* (1992) showed that Amylose, a linear α -glucan is the most effective substrate *in vitro* for AmyA. Starch and amylopectin, a lightly branched α -glucan, and linear oligomeric α -glucans of six or more glucose moieties also serve as good substrates (Raha *et al.*, 1992). Glycogen is a poor substrate, and since it is the only polysaccharide present in high amounts in the cytoplasm (Preiss, 1984), the physiological role of AmyA remains uncertain. Raha *et al.* (1992) hypothesised that in the absence of exogenous oligosaccharides that could act as primers for glycogen synthesis, the cytoplasmic AmyA could catabolise existing cellular glycogen, of which the oligosaccharides generated could act as a source of primers for the synthesis of more molecules of glycogen. Since AmyA expression is restricted in the cytoplasm, *E. coli* cannot use extracellular starch as the sole carbon source

3.3.3 Starch growth assay

In this study, the capability for AmyA autodisplay provided an excellent opportunity to engineer *E. coli* to use starch as the sole carbon source. In this setting, the AmyA-Pet chimera would export AmyA to the extracellular medium in which AmyA cleaves α -1-4-glycosidic linkages in starch thereby releasing maltotriose, maltotetraose, and maltose, which are imported into the cell and further catabolised for energy production.

M9 Minimal media supplemented with soluble starch was used to investigate growth of *E. coli* expressing AmyA-Pet chimera using starch as the sole carbon source. The assay was performed in 384-density format to assess the prospect of using the assay in a high-throughput manner. The experiment was split into two conditions; induced and uninduced. Controls, *E. coli* containing an empty plasmid, were included in the positions shown (Figure 3.3.3.1). The results for the uninduced experiment show no growth for the AmyA-Pet chimera (Figure 3.3.3.1A), an affirmation that *E. coli* cannot use starch as a carbon source. The control, containing the empty plasmid, also displayed no growth.

Contrastingly, the induced plate showed growth of both strains, indicating successful engineering of *E. coli* able to utilise starch. However, it was noted that the recombinants grew very slow compared to growth in glucose, attaining the optimal size for growth quantification

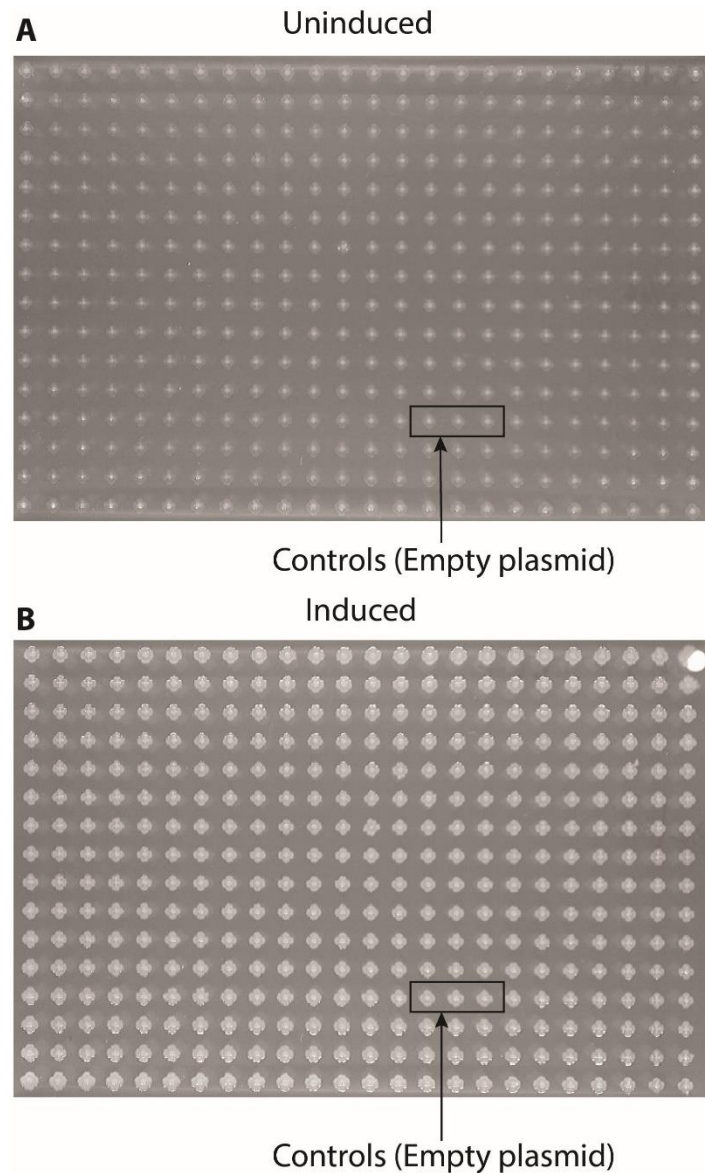


Figure 3.3.3.1: Starch growth assay in 384-colony density.

The panels present results showing successful engineering of *E. coli* to utilise starch as the sole carbon source. Each grey spot, apart from the three spots surrounded by a black rectangle, show growth of wild-type *E. coli* (BW25113) transformed with pJP168 plasmid expressing the AmyA-Pet chimera protein. The indicated controls show the same strain of *E. coli* but with an empty pJP168 plasmid. Grey spots in Panel A are a result of pinning and not bacterial growth. The strains were replica pinned from M9 minimal media plates (0.2 % w/v glucose) to new M9 minimal media plates supplemented with soluble starch 5 g/l (w/v) instead of glucose. Panel A and B show results without and with induction, respectively. To induce the production of the chimera, M9 minimal media was supplemented with 1 ng/ml of ATC. Both plates in panel A and B were incubated at 37 °C for five days.

after five days at 37 °C. Of concern was the unanticipated growth of controls (*E. coli* with empty plasmid) (Figure 3.3.3.1B). The likely reason was the diffusion of AmyA through the solid media from the neighbouring colonies which had successfully exported the protein; leading to starch degradation throughout the plate. Indeed, Lugol's staining (described in detail in the next section) confirmed starch breakdown in the entire assay plate, which was visualised as a complete loss of starch and thus the stain in the plate. To use the assay for high-throughput screening, the diffusion of AmyA had to be restricted to a limited area around the colonies. Two ideas were explored, increasing the distance between the clones and modifying the AmyA-Pet chimera to retain AmyA on the extracellular surface by abolishing its auto cleavage and release. The two strategies are explored in detail in the next section.

3.3.4 Optimisation of the starch-growth assay for a high-throughput screen

Leyton *et al.* (2011) described a Pet derivative containing mutations in amino acid residues Asn-1018 and Asp-1115 that are responsible for autocatalytic cleavage of the passenger domain from the translocator. The authors replaced the two amino acids with glycine, and the resultant secretion-blocked variant translocated the passenger domain to the cell surface but was not processed or released from the cell. A similar strategy was employed in this study; I made a new construct of Pet-AmyA with two point mutations, Asn1018Gly and Asp1115Gly. The new construct, denoted as AmyA-Pet**, would translocate AmyA to the cell surface but would not release it in the medium, avoiding AmyA diffusion to neighbouring strains. The expression of the new construct was confirmed in a similar manner as described in Section 3.3.1

As a further possible solution, I increased the distance between the strains by switching from 384-colony pinning density to 96-colony pinning density. Following these two adjustments, the starch growth assay was performed as described in Section 3.3.3 with minor alterations in the arrangement to include more controls. Assessment of growth in starch showed growth for only the strains expressing AmyA-Pet** and no growth for strains containing the empty plasmid as anticipated, though I did note some inconsistency in the growth of the clones.

To get a clear picture of these results, a colourimetric detection assay was used to analyse whether AmyA was diffusing away from the colonies. Lugol's iodine staining assay was employed. Lugol's iodine, also known as Lugol's solution, is a solution mix of elemental iodine and potassium iodide in water. The solution is an indicator of starch in organic compounds,

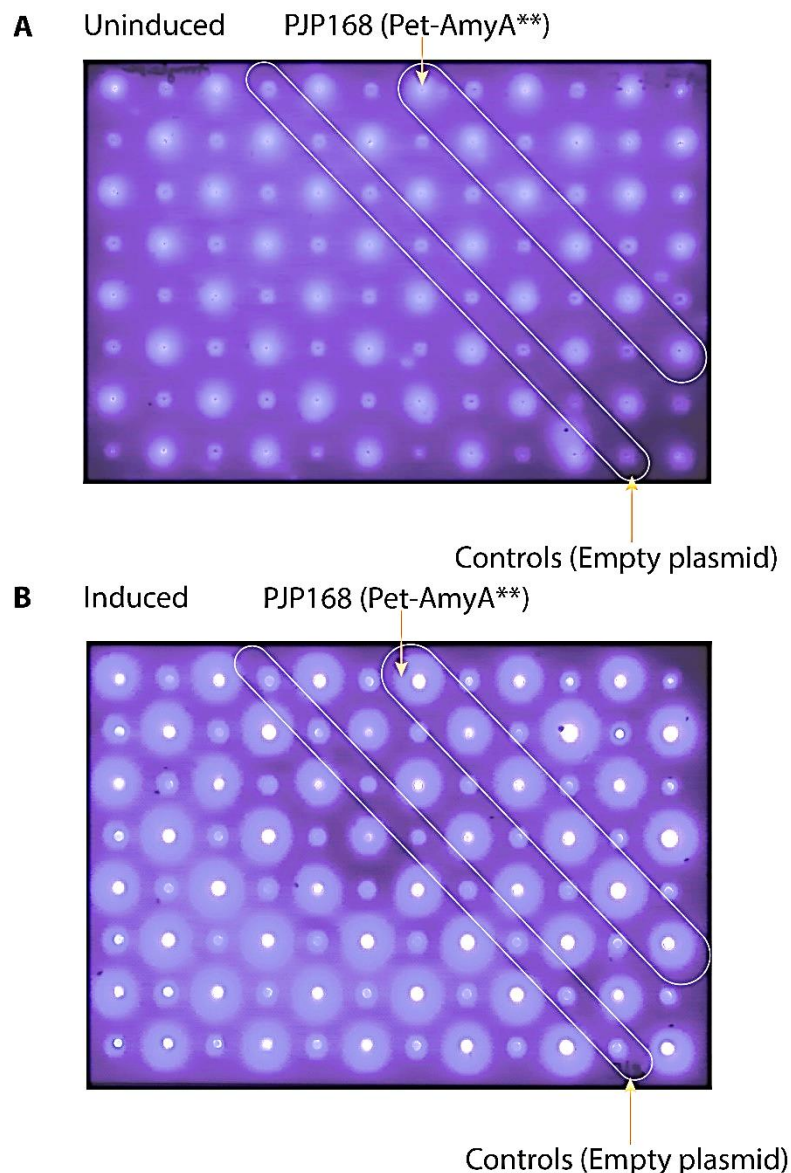


Figure 3.3.4.1: Lugol's iodine staining, a starch breakdown assay in *E. coli* (96-colony density format).

The panels show results of Lugol's iodine staining after a starch growth assay on M9 minimal media supplemented with soluble starch 5 g/l (w/v) as the sole carbon source. The spots show the location of strains under investigation, where the two clones were arranged in an alternating manner resulting in a diagonal orientation of the same clone. Starting from the top-left most position is the assay strain, wild-type *E. coli* (BW25113) transformed with pJP168 plasmid expressing the Pet-AmyA^{**} chimera alternating with the control strain, *E. coli* transformed with an empty pJP168 plasmid. Panel A and B show results without induction and with induction (1 ng/ml ATC), respectively. Prior to staining, plates in both panels were incubated at 37 °C for five days to allow growth on starch. Plates were then flooded with Lugol's iodine, 0.25% (v/v), for 1 min. The stain was removed, and plates were incubated at room temperature for about 30 minutes for destaining before the depicted images were captured.

with which it reacts to displaying a ‘black-blue’ colour, but does not stain simple sugars. The ability of Lugol’s iodine to stain starch, excluding simple sugars, formed the basis of the assay. I exploited the assay to assess the breakdown of starch by evaluating the decolourisation of the ‘black-blue’ colour. Analysis of the staining results presented in Figure 3.3.4.1A (uninduced condition) showed slight discolouration around the AmyA-Pet** expressing colonies, though the strains do not exhibit growth. These results were not surprising as I had previously observed that the plasmid displayed slightly leaky expression (Figure 3.3.1.2). Results presented on Figure 3.3.4.1B show a light blue halo surrounding a clear colony pin, only for the strains expressing AmyA-Pet** chimera and not for the strains with an empty plasmid.

While the results looked promising, AmyA diffusion through the media was still evident, as indicated by the light blue halo surrounding the AmyA-Pet** expressing clone. However, at this pinning density (96), and arrangement, the AmyA cross-strain contamination was not evident as the distance between the strains was sufficiently big. Further scrutiny of the destained halos showed inconsistent sizes despite using the same clone.

To further investigate the inconsistency in growth and staining, I performed pinning replica experiments to assess the reproducibility of the results across multiple plates. Analysis of growth in starch from one plate to another showed a high degree of variability, despite processing all the samples in similar conditions. Upon Lugol’s staining of the replicate plates, the variability was more evident. Due to the disparity in growth and also in colour detection, the technique was deemed unsuitable for a high-throughput application, and a decision was made to cease further efforts regarding development of Pet biogenesis high-throughput screening assays.

3.4 Discussion

The development of eSGA and GIANT-coli techniques provided an opportunity to investigate the genetic interaction networks in bacteria in a high-throughput manner. In this chapter, pilot studies were set-up to evaluate and reconcile some of the variations between the two techniques to develop a standardised protocol that was reliable and reproducible for consistency of results. Pilot studies were first conducted (Section 3.2.2), and several readjustments were made according to the results (Section 3.2.3). In line with this aim, the Keio collection library was rearranged to a format suitable for a high-throughput screen (384-colony density format).

Following optimisation, the analysis of both the positive and negative controls showed consistency with the same plate and also between replicate plates (Figure 3.2.3.2). Apart from

the controls, the results were further validated using colony PCR, where randomly selected clones were confirmed to have both of the expected gene deletions and in the correct chromosomal positions. Based on these results, the optimised protocol was adapted for future screens.

In this chapter, I also describe the development of a starch growth assay to investigate Pet biogenesis (Section 3.3). Exploiting the autodisplay capability of Pet, I successfully engineered *E. coli* to use starch as the sole carbon source by making Pet chimeras, AmyA-Pet and AmyA-Pet** that excrete and secrete AmyA, respectively, for the breakdown of starch. However, despite the enormous potential of the assay, attempts to optimise the assay to be robust enough for a high-throughput screen proved futile. The growth of the same strain in starch as the sole carbon source, upon induction of AmyA-Pet** chimera, was inconsistent and could not be reproduced in duplicate assay plates. Analysis of starch breakdown by Lugol's iodine staining showed diffusion of AmyA from the pin colonies (Figure 3.5.2.1), despite modifying Pet to abolish the autocatalytic cleavage of the AmyA containing domain.

AmyA contains seven cysteines in its amino acid sequence, accumulation of this protein in the oxidising periplasm is likely to result in misfolding due to the formation of unnatural disulphide bonds. I speculate that AmyA misfolds in the periplasm resulting in its degradation, which further overwhelms the periplasmic quality control machinery, ultimately leading to cell lysis. Lysis in turn, releases the cytoplasmic AmyA to the extracellular milieu further breaking down more starch. This hypothesis is backed-up by an earlier observation I made, where the induction of the cells with high levels of ATC (> 1 ng/ml) resulted in toxicity exhibited by the reduced growth rate. Despite the unsuccessful attempt to develop a Pet biogenesis screen, the Giant-col and eSGA techniques that I optimised were adapted to an optimised protocol for a high-throughput Lpp₊₂₁ synthetic lethal screen as described in the next chapter.

Chapter 4: The architectural significance of Gram-negative periplasm; an Lpp₊₂₁ synthetic lethal screen

4.1 Introduction

Gram-negative bacteria have evolved mechanisms to attach the OM to the PG, with Lpp being the most significant as it provides the only covalent connection between the two layers (Braun, 1975). Recently, Lpp was shown to be the primary determinant of OM to IM distance (Asmar *et al.*, 2017; Cohen *et al.*, 2017). The importance of Lpp is further illustrated as being the most abundant protein in *E. coli* (Li *et al.*, 2014). Despite this seemingly important role of Lpp, an *lpp* deletion mutant grows and divides normally under standard laboratory conditions (Hirota *et al.*, 1977; Braun & Hantke, 2019).

This chapter focuses on a lengthened functional isoform of Lpp (Lpp₊₂₁) to study the role of the periplasmic architecture on Gram-negative envelope biology. An increased OM to IM distance must also mean a substantial increase in periplasmic volume, and I hypothesised that this could impact on many aspects of the envelope biology. To address this, I screened for non-essential *E. coli* genes that become essential in an Lpp₊₂₁ background (larger periplasm). By doing so, I identified periplasmic systems that rely on strict control of the periplasmic dimensions and also further realised the redundancy of OM-PG attachment (i.e. systems that become essential for providing the physical OM-PG attachment). The chapter will further focus on proteins that become essential, by characterising a series of mutants to better understand their role in regards to changes in periplasmic size.

4.2 Cloning, expression of Lpp₊₂₁ and Lpp sequence analysis

The periplasmic architecture was modified by creating *E. coli* strains that expressed a long form of Lpp (Lpp₊₂₁). A gene block containing an insertion of 21 amino acids (3 heptad repeats) was used to construct the variant (Figure 4.2.1A). The gene block also had 50 bp DNA flanking each end of the *lpp* gene (Appendix 1). One of these extensions was homologous to DNA sequence upstream of *lpp*, while the other DNA sequence was homologous to the *cat* gene. The gene block was first introduced in the donor strain, *E. coli* Hfr Cavalli, using the lambda (λ) red recombination technique as described in Section 2.3.1.

The Lpp structure is based on the heptad repeats of interacting monomers in the trimer (Shu *et al.*, 2000). The structural and functional aspects of the protein were maintained by designing the length variant to retain the interacting heptad repeats (Asmar *et al.*, 2017; Cohen *et al.*,

2017). Attempts by Cohen *et al.* (2017) to increase the Lpp length in *S. enterica* by fusing LppA with LppB resulted in cells severely defective in cell shape and division. Since *S. enterica* LppA and LppB share high sequence degree of homology, the authors predicted this might have been due to the chimera being too long, the heptad chimera being out of register or the fact that the N-termini of LppA and LppB have diverged considerably, or all three issues combined.

Colony PCR was used to confirm the replacement of *lpp* with *lpp*₊₂₁. Figure 4.2.1B shows the variant (300 bp) and the wild-type *lpp* (234 bp). The PCR assays also ruled out gene duplication events that would introduce the wild-type *lpp* in a different chromosomal location, as no extra bands were observed for any of the strains analysed.

To determine the expression levels of the Lpp₊₂₁ variant in these *E. coli* strains relative to wild-type, immunoblotting with a monoclonal Lpp antibody was performed (Section 2.6.3). Figure 4.2.1C below confirmed the expression of the Lpp₊₂₁ isoform, though at relatively lower levels than the wild-type by visual comparison. An *lpp* null mutant confirmed the specificity of the antibody, and antibodies recognising OmpA were used to ensure equal loading of extracts.

Lpp exists in two forms, the free and the PG-bound form (Cowles *et al.*, 2011), and SDS-solubilisation of *E. coli* will liberate only the free form. The PG-bound form can be liberated only after the digestion of the PG with lysozyme (Cowles *et al.*, 2011). The results depicted in Figure 4.2.1C show the free form Lpp as the cell lysates were generated without lysozyme treatment.

However, even when the bound PG was degraded using lysozyme, the wild-type Lpp and Lpp₊₂₁ levels remained the same as the cell lysate results. These results showed that for both conditions and both strains, with and without lysozyme treatment resulted in equivalent amounts of the Lpp bound and free forms.

Given the significant role of Lpp as a primary determinant of periplasmic size (Asmar *et al.*, 2017; Cohen *et al.*, 2017), I was curious to see that the length of Lpp in *E. coli* (78 residues) is conserved in other species as judged by molecular phylogeny evaluation of Lpp length (analysis was performed by Dr Von Torres, University of Queensland, Queensland). Non-redundant Lpp sequences were identified in 757 publicly available bacterial sequences, and

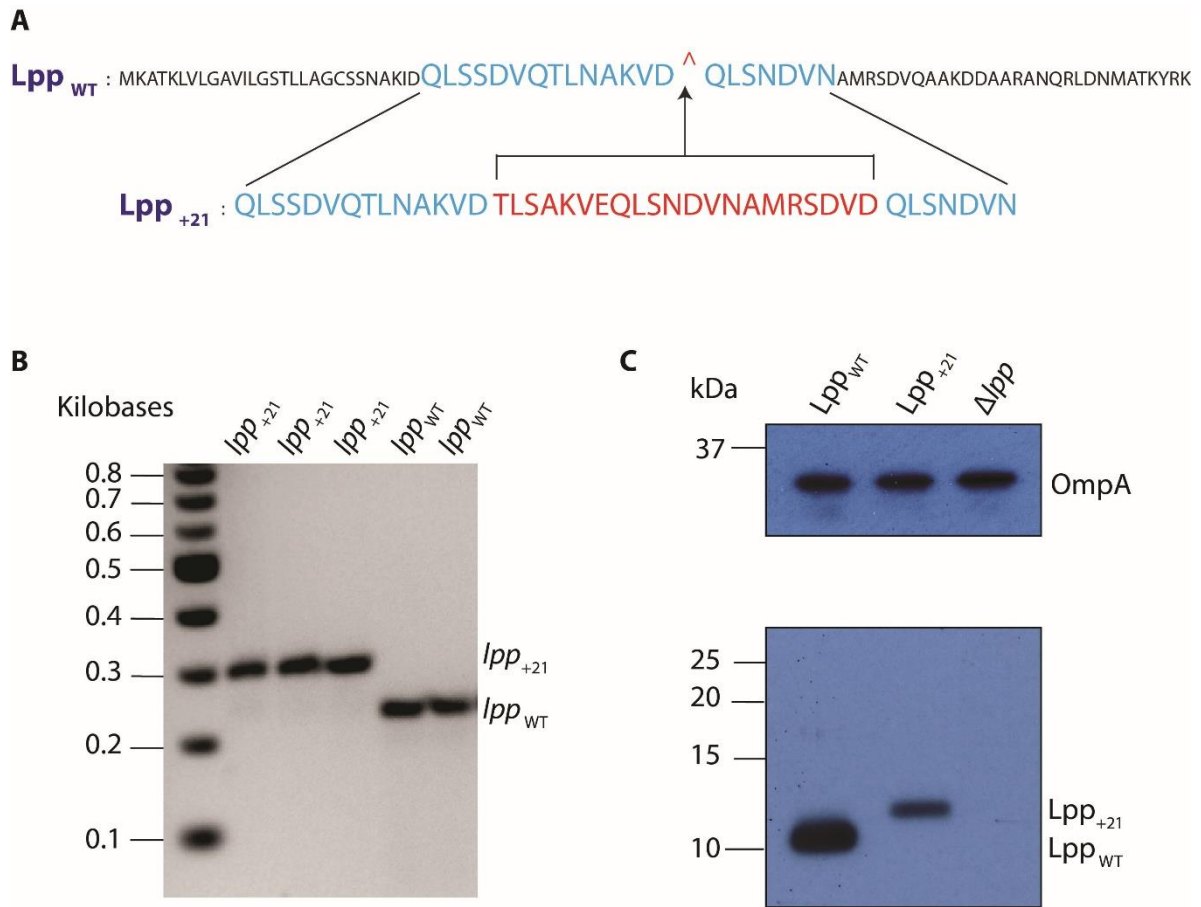


Figure 4.2.1: Cloning and expression of Lpp₊₂₁.

(A) Lpp₊₂₁ isoform was constructed by inserting a sequence of three heptad amino acid repeats (red lettering) between residue D42 and Q43 (^) of the wild-type Lpp. Blue letters show the amino acid sequence of the original three heptad repeats. (B) 2% (w/v) agarose gel electrophoresis results of an amplified DNA segment using primers amplifying the *lpp* gene for both the isoform and the wild-type demonstrate successful cloning of the isoform. (C) Western blot probing for Lpp in whole-cell lysates demonstrated that the length variant of Lpp (Lpp₊₂₁) was present at lower levels compared to the wild-type. Δ *lpp* strain was included to confirm antibody specificity, and OmpA were used to ensure equal loading of extracts. A single percentage of 15% (w/v) acrylamide SDS-PAGE gel, was used. Representative data are shown from experiments performed in biological triplicate.

protein length was plotted against the number of non-redundant sequences showing that length (Figure 4.2.2A). The results showed a narrow window of amino acid length, with Lpp being 78 residues in *E. coli* and most other species of bacteria. The 99 residues of Lpp₊₂₁ place it at the upper end of the range for this protein analysis. In this context, it was interesting to know that the length of Lpp is highly conserved with small variation ranging from 1 to 8 amino acids (Figure 4.2.2A).

Intriguingly, several exceptions were noted with longer variants. The most represented genera were *Geobacter* and *Vibrio*, with 43 additional amino acid residues (*Geobacter bemidjensis*). It should be noted that the number of Lpp entries for *Geobacter* spp. and *Vibrio* spp. is comparatively low compared to more representative sequences (1 unique representative sequence versus 15-58 representative sequences) (Appendix 2). Using these longer variants, a multiple sequence alignment was performed comparing them to *E. coli* Lpp (Figure 4.2.2B). As anticipated, all the sequences contain the lipobox for lipoprotein sorting and a terminal lysine for PG linkage, suggesting an active role in OM-PG linkage. Notable sequence differences were insertions, mostly in the 4'3 heptad repeat structural domain. Also noted was that *Geobacter* species contain extra C-terminal residues adding an extra alternative lysine residue. Phyre2 predicts all these putative Lpp proteins to adopt the α -helical of Lpp with high confidence (>90%) (Appendix 3).

4.3 Phenotypic characterisation of Lpp₊₂₁

Gram-negative bacteria are found in various environments with varying osmotic conditions. The diderms have evolved mechanisms to not only cope but thrive in these conditions (Lucht & Bremer, 1994). A good example is *E. coli*, which can thrive in hyperosmotic environments, e.g. the gut. One of the ways the diderms cope is by maintaining the periplasm and the cytoplasm in an iso-osmotic state (Stock *et al.*, 1977; Cayley *et al.*, 2000). This is achieved by adjusting the solute concentration in the cell compartments by influx or efflux of water. By doing so, osmolality contributes immensely to the architectural aspects of cellular compartments (Stock *et al.*, 1977; Cayley *et al.*, 2000).

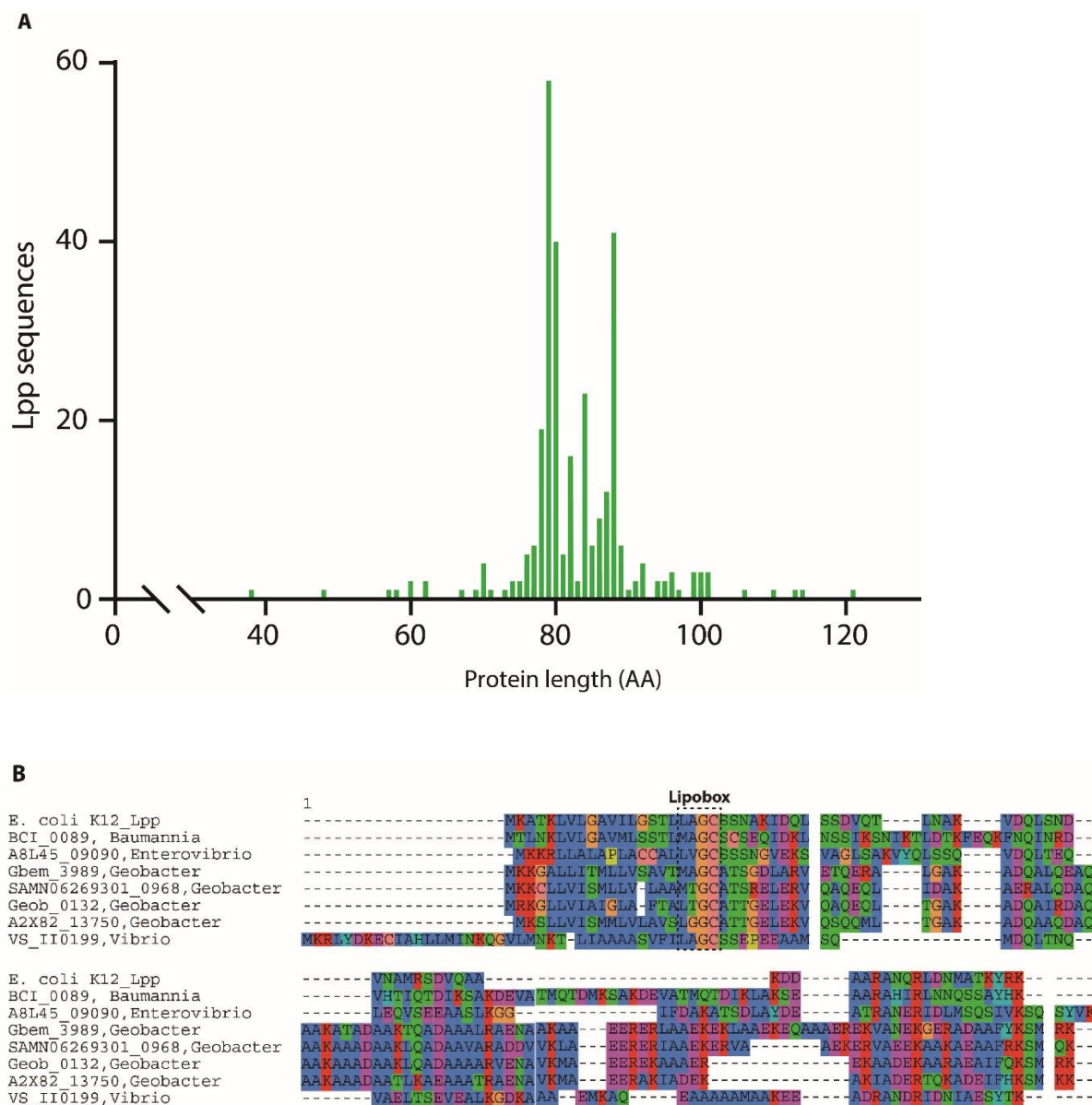


Figure 4.2.2: Lpp sequence analysis.

(A) Chart showing the distribution of Lpp length. Non-redundant Lpp sequences were identified in 757 bacterial genomes and the protein length charted on the x-axis. The number of non-redundant sequences showing that length is shown on the y-axis. (B) Multiple sequence alignment of the Longer Lpp variants in comparison with typical 78 residue *E. coli* Lpp. The analysis was done by Dr Von Torres, University of Queensland, Queensland.

In this study, sorbitol was used to mimic physiological osmotic conditions in phenotypic characterisation of Lpp₊₂₁ strain. The filtration barrier of the OM-PG layer is primarily based on the size of the osmolyte (Demchick & Koch, 1996; Sochacki *et al.*, 2011), with sorbitol, at 182.17 Da, permeating freely through the OM-PG barrier into the periplasm. However, the IM is impermeable to sorbitol (Pilizota & Shaevitz, 2013). Thus, sorbitol in the growth medium diffuses across the OM-PG layer into the periplasm but does not cross the IM into the cytoplasm.

The accumulation of sorbitol into the periplasm increases the osmolality of this compartment, creating an osmotic imbalance. To restore the iso-osmotic state, water is drawn out of the cytoplasm, increasing the osmotic pressure until it matches with that of the periplasm (Figure 1.2.5.1). Studies by Stock *et al.* (1977), found that the periplasmic volume increased rapidly in response to an increase in sucrose in the external medium. This phenomenon can reduce the cytoplasmic volume by about 30%, pulling the IM inward, and substantially increasing the periplasmic volume by around 300% (Cayley *et al.*, 2000).

4.3.1 Lpp₊₂₁ has no growth defect

I assessed growth of Lpp₊₂₁ strain in both rich media and minimal media conditions. In addition to this, I osmotically adjusted both of the media conditions to mimic physiological osmotic conditions, using sorbitol titrated in increasing concentrations from 0.5 – 1 M (Figure 4.3.1.1).

As a non-glucose sugar, sorbitol was the preferred osmolyte in M9 minimal media as it is a poor carbon source (Aidelberg *et al.*, 2014). Unlike the disaccharide sucrose, that can readily be used as a carbon source, the sugar alcohol sorbitol was chosen to avoid carbon source switch likely to take place during long growth assessment due to glucose depletion.

The Lpp₊₂₁ strain grew at wild-type (JW5028) rates in all media conditions tested (Figure 4.3.1.1). Increases in osmolality affected the growth rate of Lpp₊₂₁ strain and JW5028 equally. The growth profiles suggest that this strain of *E. coli* has adapted to whatever impact Lpp₊₂₁ expression has on the periplasm.

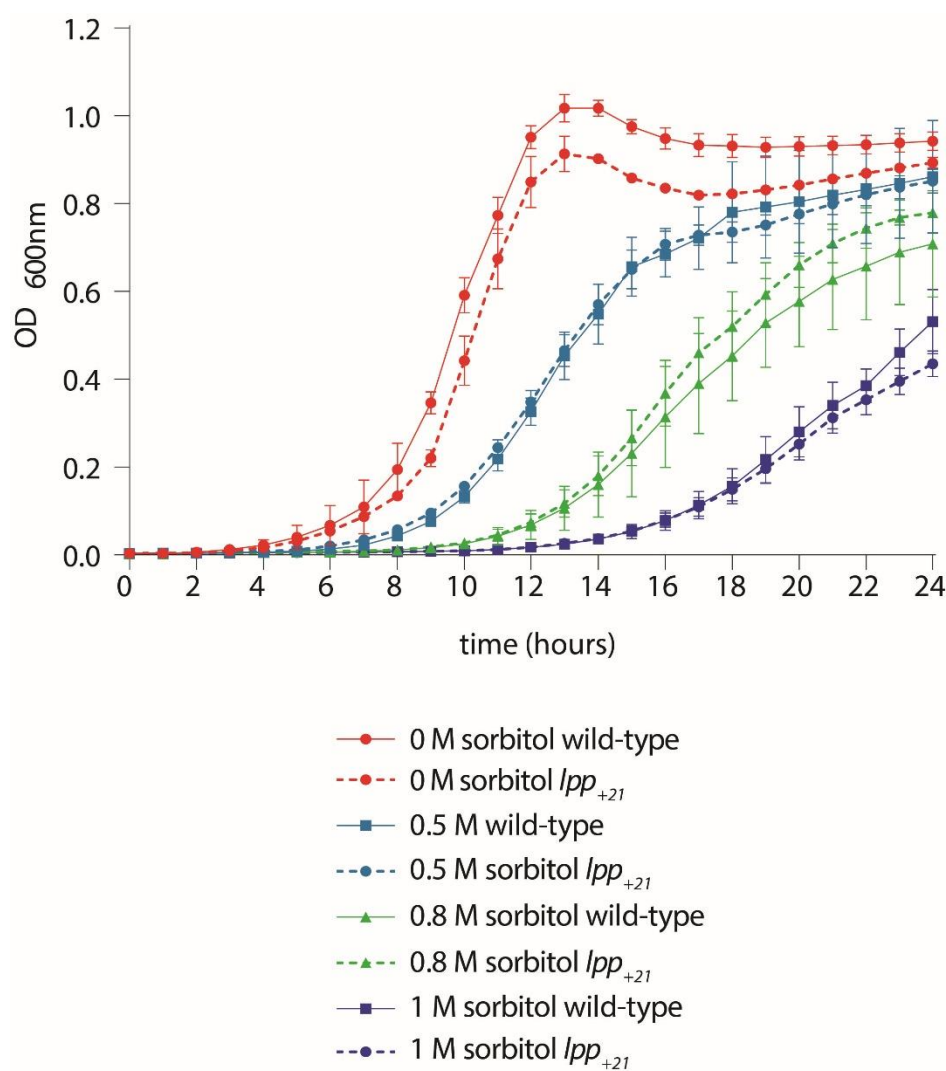


Figure 4.3.1.1: Comparative growth profiles of *E. coli* JW5028 and *Lpp*₊₂₁ strain.

Growth assessment was done in M9 minimal media 0.2% (w/v) glucose supplemented with increasing concentration of sorbitol (0.5 – 1 M). Cells expressing *Lpp*₊₂₁ from the chromosome exhibited a similar growth profile to the Wild-type (JW5028) at 37 °C, with constant orbital shaking. Increase in osmolality, i.e. increasing sorbitol concentration, affected both strains equally. Values are averaged from independent clones (n = 3); error bars depict standard deviation.

4.3.2 Lpp₊₂₁ displays defects in outer membrane integrity

Previous studies have suggested that loss of Lpp leads to membrane integrity defects, as measured by protein release into the medium and by increased sensitivity to hydrophobic antibiotics, chelating agents and toxic compounds such as EDTA, SDS, dibucaine, deoxycholate, and vancomycin (Hirota *et al.*, 1977; Cascales *et al.*, 2002; Nichols *et al.*, 2011). To assess membrane integrity for the Lpp₊₂₁ background, three distinct assays were employed: growth in the presence of the detergent SDS, growth in the presence of hydrophobic antibiotics, and single-cell imaging using a membrane impermeant stain.

SDS can potentially solubilise membranes by disrupting the hydrophobic associations between the lipids. An intact membrane with an LPS interface tends to resist SDS (Falchi *et al.*, 2018), while a strain with a compromised OM is more sensitive to SDS (Benedet *et al.*, 2016).

The wild-type strain (JW5028) grows on SDS concentrations up to 5% (w/v). SDS sensitivity analysis of the Lpp₊₂₁ strain showed an increased sensitivity, failing to grow at 5% (w/v) SDS, and growing only poorly at 2% (w/v) (Figure 4.3.2.1).

As a further measure of membrane integrity, the minimum inhibitory concentrations (MICs) were determined for tetracycline, ciprofloxacin, and nalidixic acid. These relatively large, hydrophobic antibiotics are largely impermeant to an intact LPS barrier (Cascales *et al.*, 2002; Nichols *et al.*, 2011). This measure of membrane integrity suggested that the differences were minor. The Lpp₊₂₁ strain showed only a two-fold increase in sensitivity relative to JW5028 (Table 4.3.2.1). Consistent with minor defects in OM integrity, there is no obvious defect in membrane protein profile either (Figure 4.3.2.2).

As a more sensitive, single-cell based assay for membrane integrity, a live/dead staining assay on Lpp₊₂₁ strain and a control mutant was performed. The assay employed the use of two dyes, SYTO 9 and Propidium iodine. Differential staining is achieved as propidium iodide penetrates cells with damaged membranes, causing a reduction in SYTO 9-staining when both dyes are present (Berney *et al.*, 2007). Intact bacterial cells are impermeant to propidium iodide and therefore fluoresce green from SYTO 9. This assay is a more sensitive means to investigate membrane integrity defects than growth on SDS as it gives single-cell readouts.

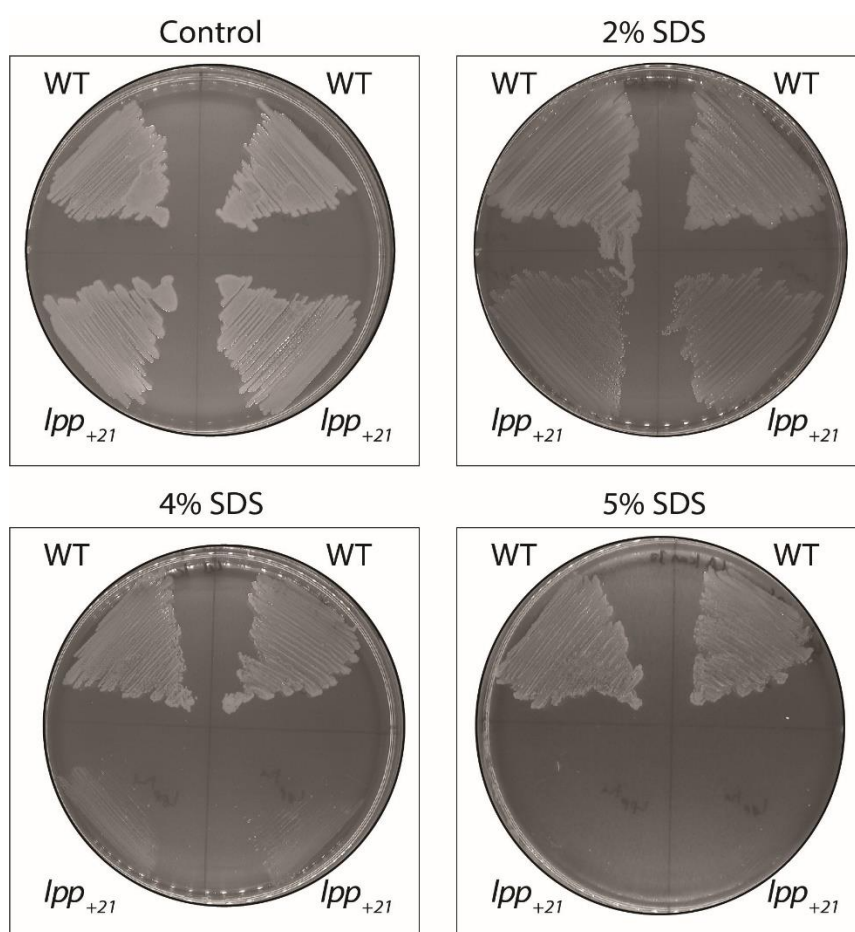


Figure 4.3.2.1: Comparative SDS sensitivity profiles of JW5028 and *Lpp*₊₂₁ strains.

In rich media (LB) cells expressing *Lpp*₊₂₁ from the chromosome exhibit increased sensitivity to SDS compared to the wild-type (JW5028). At 5% (w/v) SDS, *Lpp*₊₂₁ strain exhibits no growth and displays poor growth at 2% (w/v) SDS. The wild-type is able to resist SDS concentration up to 5% (w/v). A control plate without SDS is used to confirm viability of both strains. Plates were incubated at 37 °C for 16 hours. Representative data are shown from experiments performed in biological triplicate.

Table 4.3.2.1: Antimicrobial sensitivity profile of *Lpp*₊₂₁ *E. coli*.

Strain	Tetracycline (µg/mL)	Ciprofloxacin (ng/mL)	Nalidixic acid (µg/mL)
Wild-type (JW5028)	2	16	4
<i>Lpp</i> ₊₂₁	1	8	2

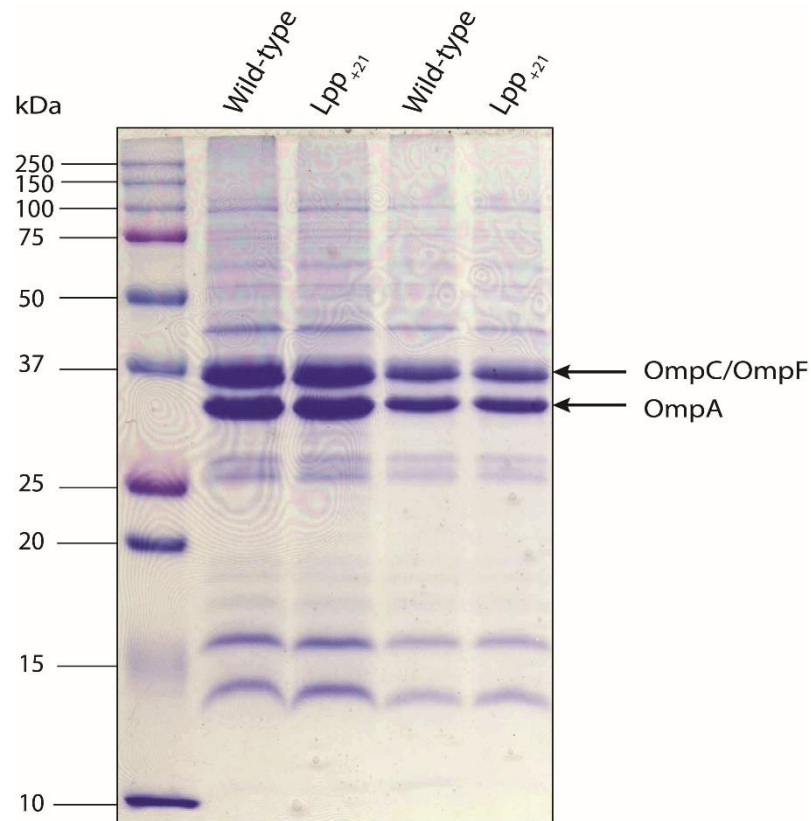


Figure 4.3.2.2: Comparative total membrane profile of JW5028 and Lpp_{+21} strains on osmotically stabilised media.

Total membrane extracts from cells grown in M9 minimal media supplemented with both glucose (0.2% w/v), and sorbitol (0.5 M) were compared for both the wild-type (JW5028) and the Lpp_{+21} strains. The concentration of the membranes is halved for the second set of results (lane 3 and 5) compared to the first set on the left (lane 2 and 3). Both strains show similar total membrane profiles. A gradient percentage of 4-16% (w/v) acrylamide SDS-PAGE gel, was used. Representative data are shown from experiments performed in biological triplicate.

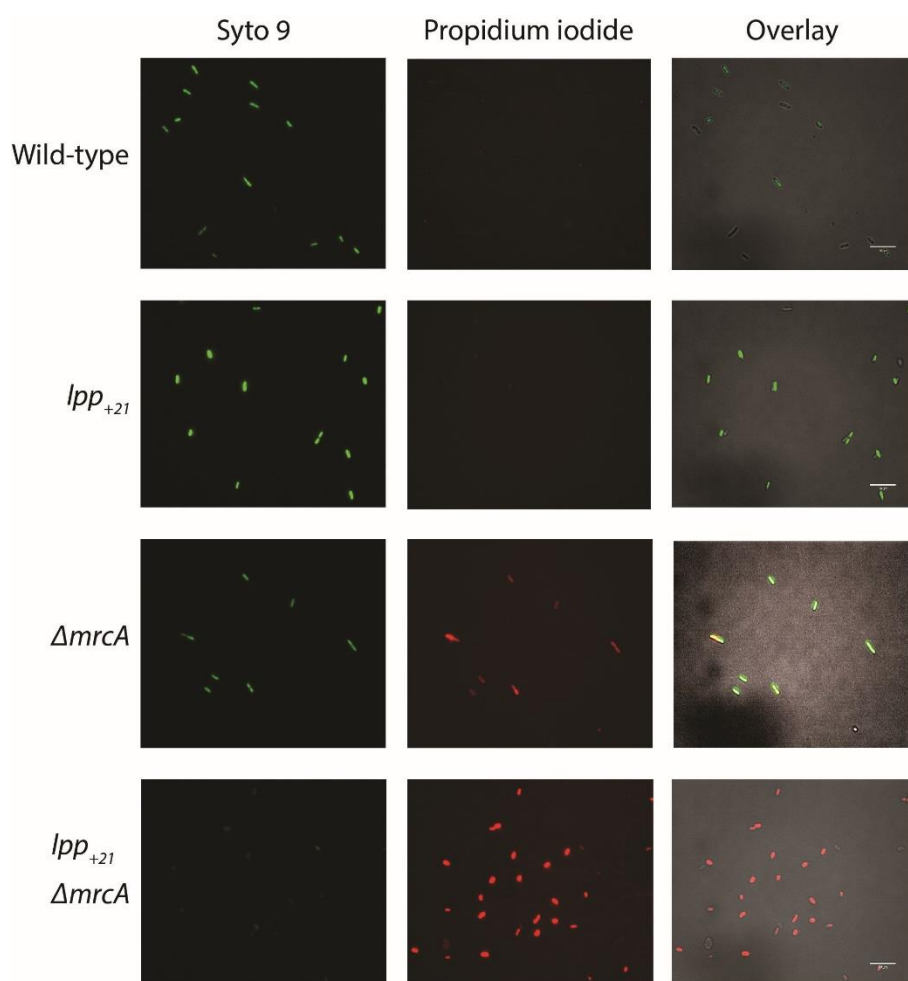


Figure 4.3.2.3: Single-cell assessment of membrane integrity using live/dead staining assay.

Fluorescence microscopy images of *E. coli* cells: wild-type (JW5028), *lpp*₊₂₁, $\Delta mrcA$, and double recombinant (*lpp*₊₂₁ $\Delta mrcA$). SYTO 9 fluorescence (green) indicates live cells with an intact membrane, while dead, membrane-permeable cells exhibited propidium iodide fluorescence (red). The four strains were grown to log phase ($OD_{600} \approx 0.5$) in LB stabilised osmolality (0.4 M NaCl) (a condition used in the early stages of the project). The concentration of the cells was standardised using the OD_{600} readings to a final concentration of $OD_{600} = 0.1$. 20 μ l of the cells were then stained for analysis. The rightmost panels show the overlay of the first two images. The double recombinant strain (*lpp*₊₂₁ $\Delta mrcA$) was used as a positive control for membrane integrity defect. Representative data are shown from experiments performed in biological triplicate; $n=3$.

As shown in Figure 4.3.2.3, across the population of the Lpp₊₂₁ strain, the cells fluoresce green. The control mutant strain (*lpp*₊₂₁, *ΔmrcA*) exhibited the red fluorescence typical of propidium iodide staining, consistent with a severe membrane integrity defect. Neither the *lpp*₊₂₁ strain nor the *ΔmrcA* mutant in wild-type *lpp* background showed this defect (Figure 4.3.2.3). The membrane integrity defect seen in the *lpp*₊₂₁, *ΔmrcA* was only evident in rich, stabilised osmolality (LB + 0.4 M NaCl), growth medium. *lpp*₊₂₁, *ΔmrcA* mutant was identified from the synthetic lethal screen in Section 4.4.2 below and was solely used here as a positive control.

4.3.3 DTT sensitivity, a protein folding aspect on a widened periplasm

Unlike the cytoplasm, the periplasm is separated from the external medium by only a permeable OM, and thus, more vulnerable to environmental fluctuations such as pH, temperature and osmolality. To assist proteostasis in this compartment, the periplasm has at least two crucial folding catalysts; the peptidyl-prolyl *cis-trans* isomerases and protein disulphide oxidases, reductases and isomerases (Dsbs) (Raina & Missiakas, 1997). Many periplasmic proteins acquire disulphide bonds after export to the periplasm (Dutton *et al.*, 2008), which contribute to their stability and catalytic activity (Missiakas & Raina, 1997).

Sensitivity to the reducing agent DTT can be a proxy for measuring redox proteostasis (Missiakas *et al.*, 1993). This was evaluated by comparing DTT sensitivity of JW5028 and the Lpp₊₂₁ strain. From the results shown in Figure 4.3.3.1, the Lpp₊₂₁ strain exhibits increased sensitivity to DTT in M9 minimal media. The sensitivity is higher in liquid media (broth) at 5 mM DTT compared to solid media at 7.5 mM DTT. From the results, one might speculate that an increase in periplasmic volume demands by some unknown mechanism an increase in protein amounts of periplasmic proteins to maintain the concentration. If this was the case, there could be a higher demand for the periplasmic folding catalysts, or some compromises in the mechanism of action of the protein folding catalysts within a larger periplasm. To address this point, a proteomics-based comparison was made of JW5028 and Lpp₊₂₁ strains. A protein-by-protein deep analysis of this data is described in Chapter 5. Here, the data was simply used to assess sub-cellular changes in protein amounts.

4.3.4 Changes in protein amounts in the sub-cellular compartments of Lpp₊₂₁

The JW5028 and Lpp₊₂₁ strains were grown on M9 minimal media condition on stabilised osmolality (+sorbitol) and samples prepared for mass spectrometry as described in Chapter 2 (Section 2.6.5). Quality control assessments and in-depth analysis of the data is presented in Chapter 5 (Section 5.3). Individual protein identities were binned according to localisation

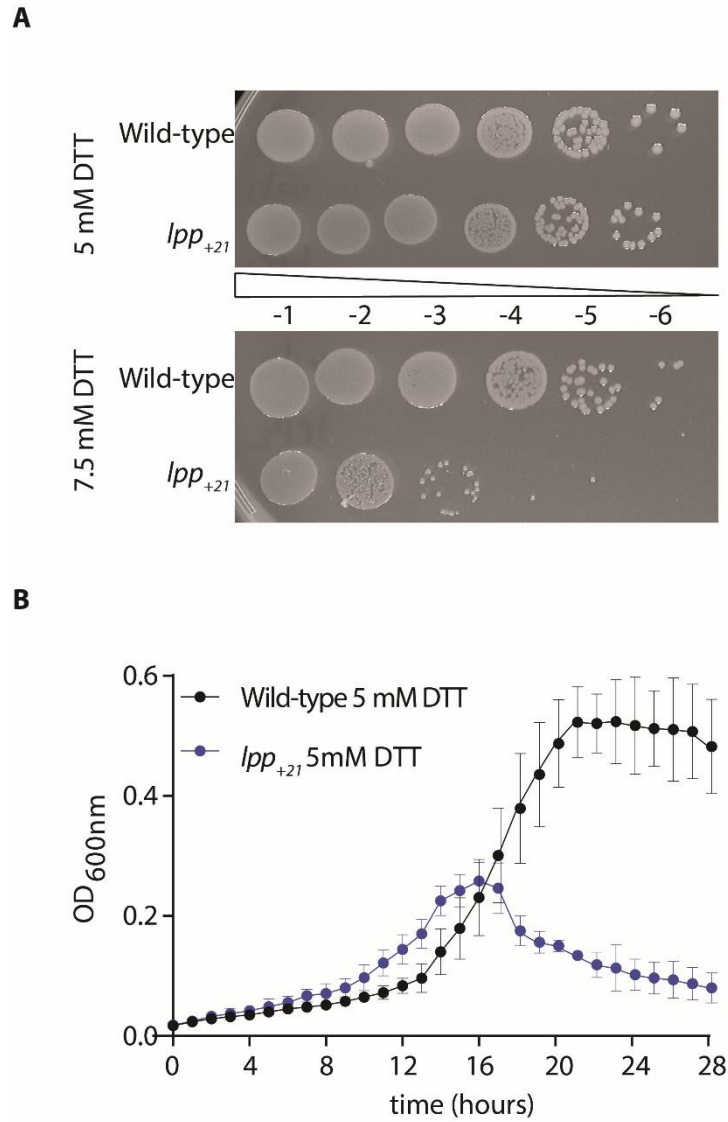


Figure 4.3.3.1: Sensitivity of *Lpp₊₂₁* to DTT.

Cells expressing *Lpp₊₂₁* from the chromosome exhibited increased sensitivity to DTT compared to the wild-type (JW5028) when analysed in (A) M9 minimal media 0.2% (w/v) glucose in a spotting experiment (37 °C), and (B) using a simple growth curve in liquid media under the same media condition but with constant orbital shaking. In (A), representative data are shown from experiments performed in biological triplicate while in (B) values are averaged from independent clones (n = 3); error bars depict standard deviation.

determined using STREpdb, a database of the sub-subcellular topology and localisation of *E. coli* proteins (Loos *et al.*, 2019). The various protein types are indicated diagrammatically in Figure 4.3.4.1A.

With an increase in linear periplasmic dimension by approximately 4 nm (Asmar *et al.*, 2017), and according to the parameters established by Prochnow *et al.* (2019), I calculated that this would correspond to an increase in periplasmic volume from 0.32 μm^3 to 0.39 μm^3 (calculations shown in Appendix 4). The raw protein intensity scores of each protein were extracted, and proteins grouped based on sub-cellular localisation (Figure 4.3.4.1B). The total intensity score of all proteins in each sub-cellular localisation was determined for the two strains and presented as a percentage of the total intensity of all detected proteins in the cell.

Cytoplasm: the proteins relative total intensity is immensely high in the cytoplasm compared to other cellular compartments. This is anticipated given the cytoplasm accounts for the biggest total cell volume percentage ($\approx 88\%$) (Volkmer & Heinemann, 2011; Prochnow *et al.*, 2019) and has an extraordinary concentration of proteins. There was a small, but insignificant, increase in cytoplasmic protein amounts in the Lpp₊₂₁ strain (Fig. 4.3.4.1B).

Membrane proteins: Analysis of the total protein intensity changes in the membranes, IM (B, E) and the OM (H, I, X), did not show any significant ratio changes, comparing JW5028 and Lpp₊₂₁ strain (Figure 4.3.4.1B). The increase in periplasmic volume would presumably require an increase in OM surface area. No change in percentage intensity suggests that the increase in membrane in the Lpp₊₂₁ strain is provided by the increased amount of lipid. The concomitant decrease in protein:lipid ratio in the OM might provide some explanation for the mild membrane integrity defects seen in the Lpp₊₂₁ strain.

Periplasm: the amount of periplasmic proteins decreased in the Lpp₊₂₁ strain. Statistical analysis showed that the decrease in protein ratio in the periplasm of Lpp₊₂₁ was significant ($P < 0.05$). The increase in periplasmic volume with a decrease in the amounts of proteins in Lpp₊₂₁ strain suggests dilution of proteins in this compartment.

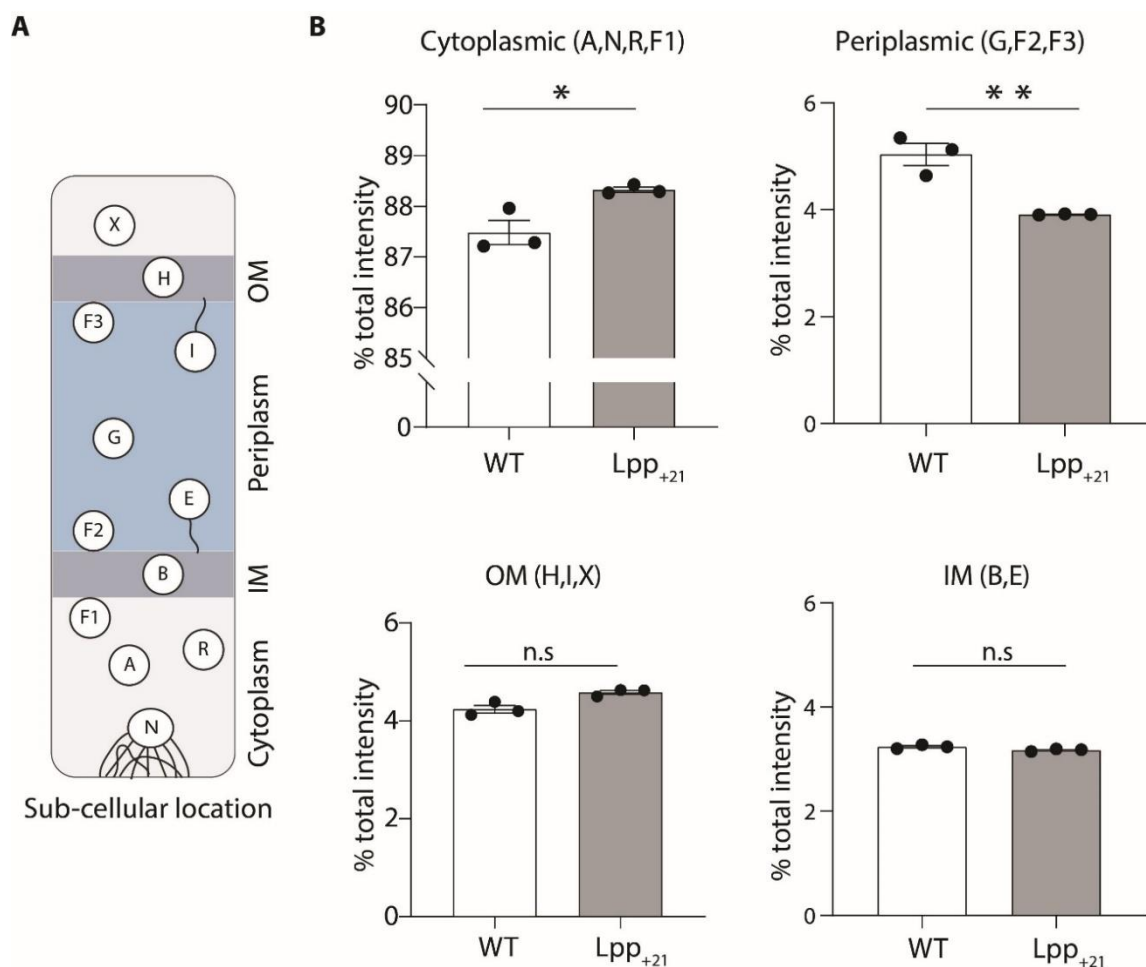


Figure 4.3.4.1 Relative proteome changes in different cellular compartments.

(A) A diagrammatic representation of various sub-cellular locations of proteins. The letters represent groups of proteins localised in the indicated cellular compartment. (B) The figures represent the raw intensity scores of proteins versus the sub-cellular location of all proteins identified from the quantitative proteomics data set. The y-axis shows the ratio of the total protein intensity in that compartment relative to the total cell proteome. The x-axis shows the comparison of the protein intensity between the wild-type (JW5028) and Lpp₊₂₁ strain in M9 sorbitol (0.5 M) condition. Error bars correspond to SEM of fit (n=3). An unpaired two-tailed student's t-test determined the statistical significance ($P < 0.05$). Abbreviations. X, secreted proteins; H, outer membrane (OM) β -barrel proteins; F3, peripheral OM proteins facing the periplasm; I, OM lipoproteins; G, periplasmic proteins; E, peripheral inner membrane (IM) proteins facing the periplasm; B, integral IM proteins; F1, peripheral IM proteins facing the cytoplasm; R, ribosomal proteins; A, cytoplasmic proteins; N, nucleoid proteins; n.s., non-significant. (Analysis was performed with the help of Dr Iain Hay, University of Auckland, New Zealand)

4.3.5. Lipopolysaccharide biogenesis in an enlarged periplasm

The proteomics analysis implied that the level of LPS and/or phospholipid is decreased in the Lpp₊₂₁ strain. As the signature lipid for the OM, I attempted to evaluate whether any decrease in LPS steady-state levels was able to be measured in the Lpp₊₂₁ strain. The strain background for this study is BW25113, is a derivative of *E. coli* K-12 which has a rough LPS phenotype. Thus, the lipid extracted from its OM is lipooligosaccharide (LOS), lacking an O-antigen (Chart *et al.*, 2000). For this reason, silver staining identifies lipid A and the inner and outer core oligosaccharides (LOS) as a single species, unlike the ladder of LPS species observed for strains with smooth LPS (Datta & Basu, 1999). LOS has an approximate molecular weight of 2.3 kDa but migrates on SDS-PAGE gels as if it were a 10 kD molecule relative to protein markers (Grover *et al.*, 2012). LOS levels for the Lpp₊₂₁ strain were measured relative to JW5028 strain using the carbohydrate silver staining technique in the cell lysate and total membrane extracts. Sample loading for the cell lysate was controlled using the same number of cells (OD₆₀₀ readings) while for the total membrane extracts, the same amount of protein concentration was used.

The results showed relatively similar levels of LOS in both JW5028 and Lpp₊₂₁, strains (Figure 4.3.5.1). This is consistent with the mild membrane integrity phenotype of Lpp₊₂₁ the strain and suggests that any change in LPS amounts is too small to be measured and therefore unlikely to be a major consideration in interpreting further analyses of the strain and its biology.

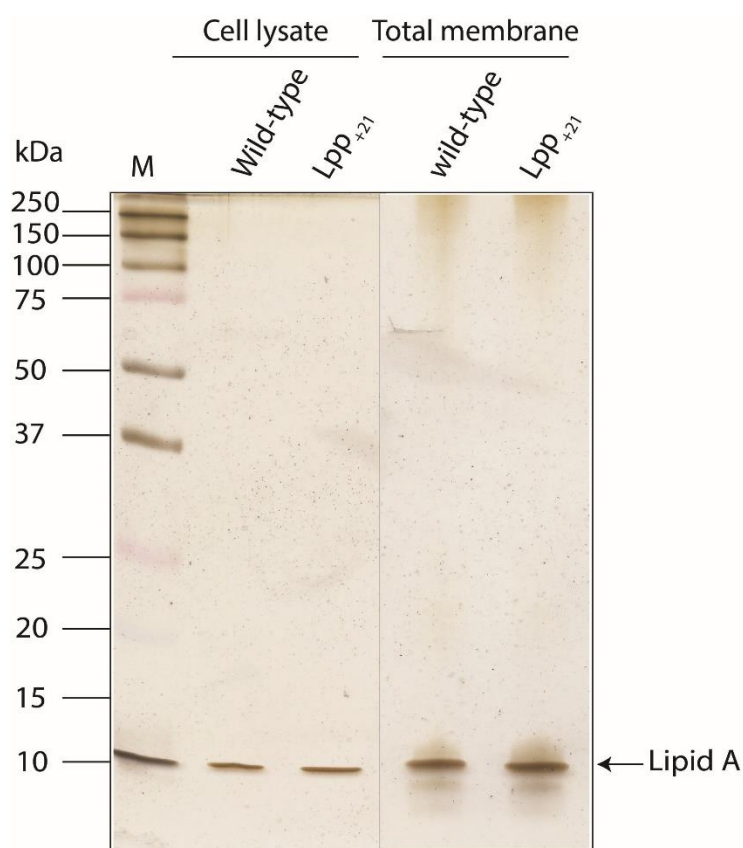


Figure 4.3.5.1: LPS quantification by silver staining of *E. coli* cell lysate and total membrane extracts.

LPS levels were analysed in both whole-cell lysate and isolated total membranes. Whole-cell lysates were prepared from wild-type (JW5028) *E. coli* and *Lpp₊₂₁* strains using mid-exponential phase cells ($OD_{600} \approx 0.5$) grown in M9 minimal (0.2% w/v) glucose plus 0.5 M sorbitol. Total membranes were extracted from the same whole-cell lysates analysed in this experiment. Prior to SDS-PAGE analysis, total membranes were standardised by proteins concentration while cell lysates were standardised by cell concentration (OD_{600} readings). A single percentage of 15% (w/v), acrylamide SDS-PAGE gel, was used. 2.5 μ l and 0.1 μ g of whole-cell lysate and total membranes were analysed, respectively. The experimental molecular weight of the Lipid A core is ≈ 10 kDa (Grover *et al.*, 2012). Representative data are shown from experiments performed in biological triplicate.

4.4 Lpp₊₂₁ synthetic lethal screen

Typically, bacteria exhibit redundancy of essential cellular processes, and *E. coli* is no exception. This redundancy has complicated efforts to elucidate gene functions, since there are many cases where knock-out of one gene does not result in any distinguishable phenotype, as a redundant gene in a substitute pathway masks the resultant phenotype. Both genes have to be removed, or one of the genes can be modified, to identify the redundant genes/pathways.

Quantitative screening techniques, such as eSGA (Butland *et al.*, 2008) and Giant-colli (Typas *et al.*, 2008) could be applied to identify critical periplasmic processes that depend on strict control of the periplasmic architecture by performing a screen for synthetic phenotype using an Lpp₊₂₁ strain background. For example, if gene X is an essential factor of a specific periplasmic process that is strictly dependent on periplasmic architecture, one could hypothesise a stronger (and perhaps lethal) phenotype in an Lpp₊₂₁ strain background.

4.4.1 High-throughput gene replacement

A high-density screen was performed to establish the fitness of all possible digenic mutant combinations in the Lpp₊₂₁ strain background. A schematic flowchart illustrating the process is shown in Figure 4.4.1.1. In this study, the Giant-colli technique was adapted, where the Hfr Cavalli donor cell was used to replace the target gene (*lpp*) with an antibiotic marker and an artificial gene (i.e., *lpp*₊₂₁-*cat*).

The recipient strains arrayed in twenty-seven 384-well plates included the Keio single-gene deletions mutants strain collection (Baba *et al.*, 2006) covering 3,985 non-essential single-gene replacements marked with a Kanamycin-resistance cassette (*kan*). The library contains duplicate independent biological knock-outs for a total of 7,970 clones, and since the screen was done in duplicate, the target single-gene deletions totalled to 15,940 clones. Figure 4.4.1.2 displays the power of using conjugation as a tool for high-density gene knock-out. The plate in panel A shows an array of single-gene knock-outs, while plate B displays knock-out strains of the same genes but in this case in combination with *lpp*₊₂₁ as an additional gene (double recombinants).

Controls were included, as shown in Figure 4.4.1.2 JW5028 is one of the Keio single-gene deletion strains with the *kan* gene replacing a pseudogene. The strain has been shown not to exhibit any genetic interactions from previous studies (Gagarinova *et al.*, 2012); hence, its use as a positive control for conjugation. Negative controls were included as empty wells on the recipient plate. The wells were expected to remain empty by the end of the screening processes

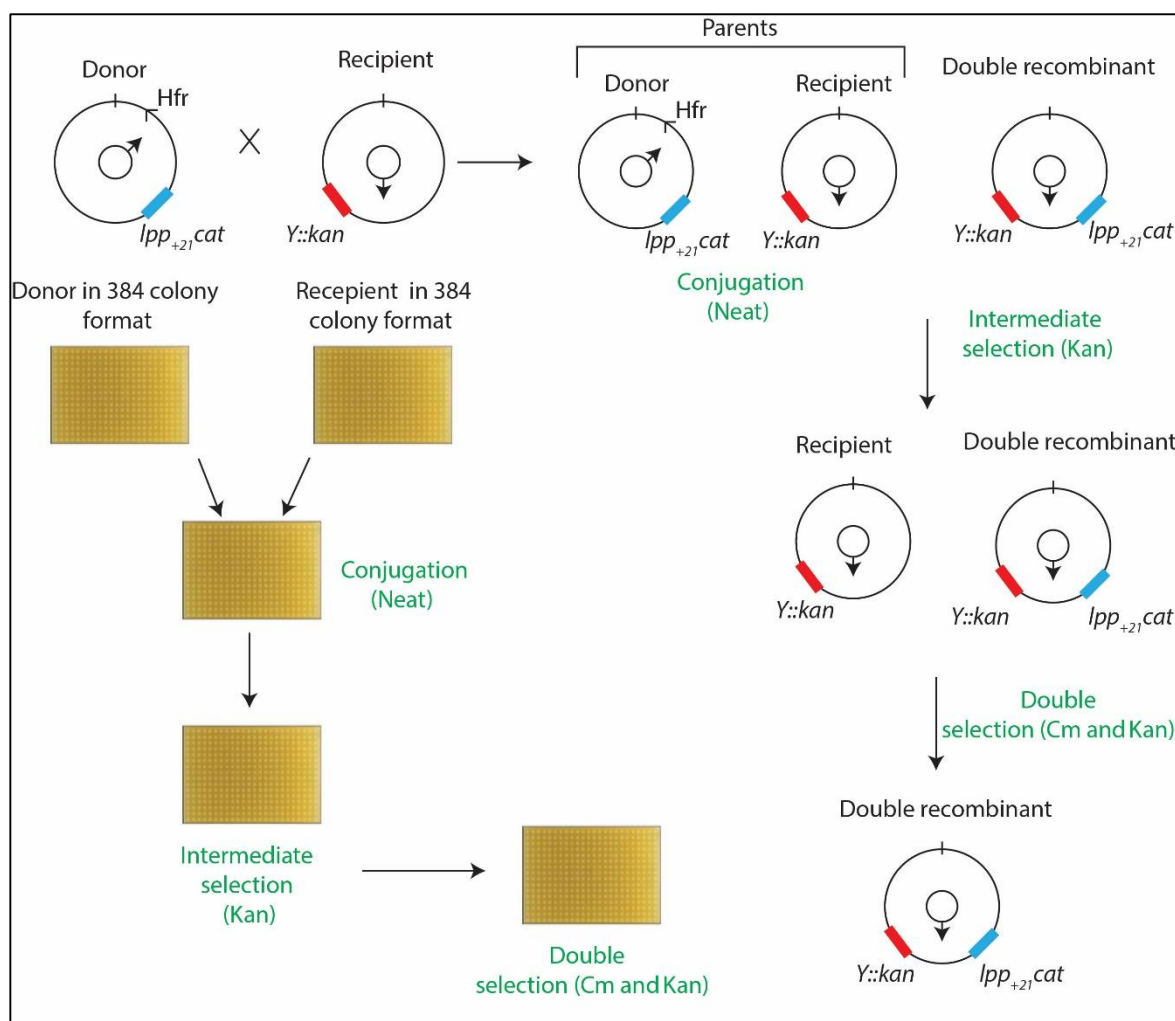


Figure 4.4.1.1: A schematic flowchart illustrating the process for high-throughput replacement of *lpp* with *lpp*₊₂₁.

First, the Hfr Cavalli donor strain was constructed to have *lpp*₊₂₁ gene (encoding Lpp₊₂₁) alongside the gene encoding CAT (*cat*). Using a robotic platform and 384-well plates, conjugations reactions were set up between the donor and hundreds of recipient strains (Keio collection), each one carrying a gene of interest (“Y”) deleted with a kanamycin cassette (*Y::kan*). The donor was arrayed in the same format, 384-density format, as the recipients. Conjugation was performed by robotic mixing of both the donor and the recipient on LB (solid) media without antibiotics, and plates incubated for 16 hours at 37 °C. The conjugants then underwent the first round of selection, (intermediate selection) on kanamycin (Kan) containing plates to select out the donor parent. The mutants were further subjected to a second round of selection using both antibiotics, Kan and chloramphenicol (CM), selecting-out the recipient parent and generating a plate with only the “double recombinants”. Images of representative plates generated in each step of the procedure are shown on the left of the strain cartoons.

as conjugation requires both the donor and recipient strains. The empty wells also served as internal quality controls to check the automated process for contamination or pinning errors. The recipient plates and double recombinant plates were compared side by side to identify genes that showed growth alteration on an *lpp₊21* background (Figure 4.4.1.2). The phenotype could either have been growing worse or synthetic lethal (aggravating interaction) or growing better (alleviating interaction).

4.4.2 Results analysis and verification

The high-resolution images obtained from the synthetic lethal screen described in the previous section were analysed manually by visual comparison of both plates (recipient and double recombinants) and also analysed with an automated image analysis software, the HT colony grid analyser (Collins *et al.*, 2006). Although the synthetic lethal combinations were straight forward to score visually (Figure 4.4.1.2), the automated quantitative system was necessary to gauge subtle changes in growth.

Data normalisation was performed to improve data quality and accuracy of interpretation. It involved correcting for typically larger colony sizes observed at the outermost edges of plates (due to more available nutrients) relative to the centre of the plates (where neighbouring colonies compete more for nutrients). Data was also normalised to eliminate potential sources of systemic error that arise due to pinning defects, unusual noisiness or unreliable results. Results were considered unreliable when a substantial fraction of colonies were missing or when the colony sizes among the duplicate experiments were not correlated. Normalisation was also performed to correct for differences in growth conditions from one plate to another (Collins *et al.*, 2006).

Manual and automated analysis of the data generated from the synthetic lethal screen identified a total of 40 genes that were subjected to a verification process. The verification process involved two aspects; (i) a custom mini-screen as a second round of synthetic lethal screen, and (ii) a confirmation of the final double-gene recombinants by colony PCR using a combination of locus-specific and kanamycin-specific primers.

The secondary mini-screen was a repeat of the large-scale screening process but in this case, using only a few identified genes instead of the whole library. Genes for the mini-screen were pinned in a higher number of biological replicates (n=8) for each independent gene of interest. In all other parameters, the mini-screen was a replica of the main screen. During the early stages

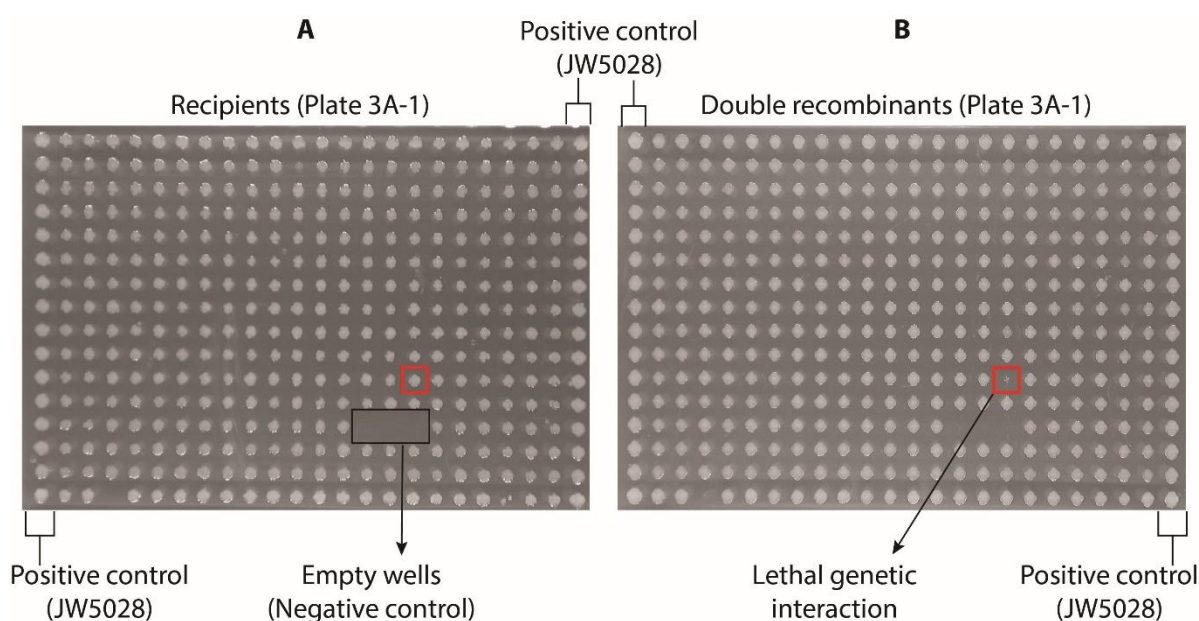


Figure 4.4.1.2: 384 high-density *lpp*: *lpp*_{+2l} gene replacement result.

For both panels, the first and the last columns (1 and 24) show the positive control, strain JW5028, which has not been shown to exhibit any genetic interactions with any other gene in *E. coli* (Gagarinova *et al.*, 2012). In panel A, position M15, M16, and M17 indicated as “empty wells”, act as negative controls for qualitative analysis of the pinning process and conjugation. The recipients were not included in these specific wells. Each spot (colony) in panel A represent a unique single-gene knock-out strain in the Keio collection library while each colony is panel B represents a double-gene recombinant at the end of the conjugation and selection process. No growth in position L17 (red square panel B) is an example of a synthetic lethal interaction between *lpp*_{+2l} and the ‘Keio gene’ at that specific position. Representative data are shown from experiments performed in biological quadruplicate.

of the project, the mini-screens were replicated on multiple conditions. Figure 4.4.2.1A depicts a mini-screen plate with 20 mutants selected from the first screen in osmotically stabilised (+NaCl) rich (LB) media. Other conditions were also screened for the same set of genes.

Colony sizes from the mini-screen were quantified using the data from the HT colony grid analyser. The data on colony sizes were plotted relative to the normalised JW5028 strain (Figure 4.4.2.1B). Plots were generated for each mini-screen in different media conditions.

From the results, Δlpp served as a reliable internal control exhibiting very low false-positive rates. Given that *E. coli* is a haploid organism, the *lpp₊₂₁-cat* antibiotic cassette has no target in the Δlpp strain, and since the final selection was on double antibiotics, the Δlpp strain exhibits a lethal phenotype due to chloramphenicol sensitivity. In addition, genes of close genetic linkage, in this case, *sufE* and *ldtE* located immediately next to the *lpp* gene in the chromosome, also exhibits a false-positive aggravating interaction due to the low frequency of recombination (Typas *et al.*, 2008).

More than half (21) of the 40 genes subjected to the second round of screen (mini-screen) were confirmed as synthetic lethal or impaired in growth on an *lpp₊₂₁* background (Table 4.4.2.1). These genes are otherwise non-essential, but become essential upon the expansion of the periplasm. A majority of the genes in Table 4.4.2.1 were isolated from the M9 minimal media condition on stabilised osmolality, and this selection proved to yield highly reproducible results.

Before further analysis of the double recombinants, the two gene replacements were independently confirmed using PCR. Locus-specific and kanamycin-specific PCR primers were used to check both for correct location of the gene on the chromosome, and also to rule out merodiploidy (Baba *et al.*, 2006; Johnston *et al.*, 2013). PCR verification was performed for both the Keio gene knock-out and the *lpp₊₂₁* knock-in. All mutants listed in Table 4.4.2.1 were verified to have both gene modifications.

The selected synthetic lethal genes were grouped into five categories, as shown in Table 4.4.2.1. In a JW5028 background, most of the components involved in PG biosynthesis exhibit a high degree of redundancy (Vollmer & Bertsche, 2008). In contrast to the JW5028 scenario, the presence of *Lpp₊₂₁* changed note-worthy genes to become essential for efficient growth. Among these, the genes coding for the major penicillin-binding proteins, *mrcA* and *mrcB*, and the genes coding for their regulatory lipoproteins, *lpoA* and *lpoB* become essential in M9 minimal media. These hits are further investigated in Section 4.4.2.1 below. Likewise, other

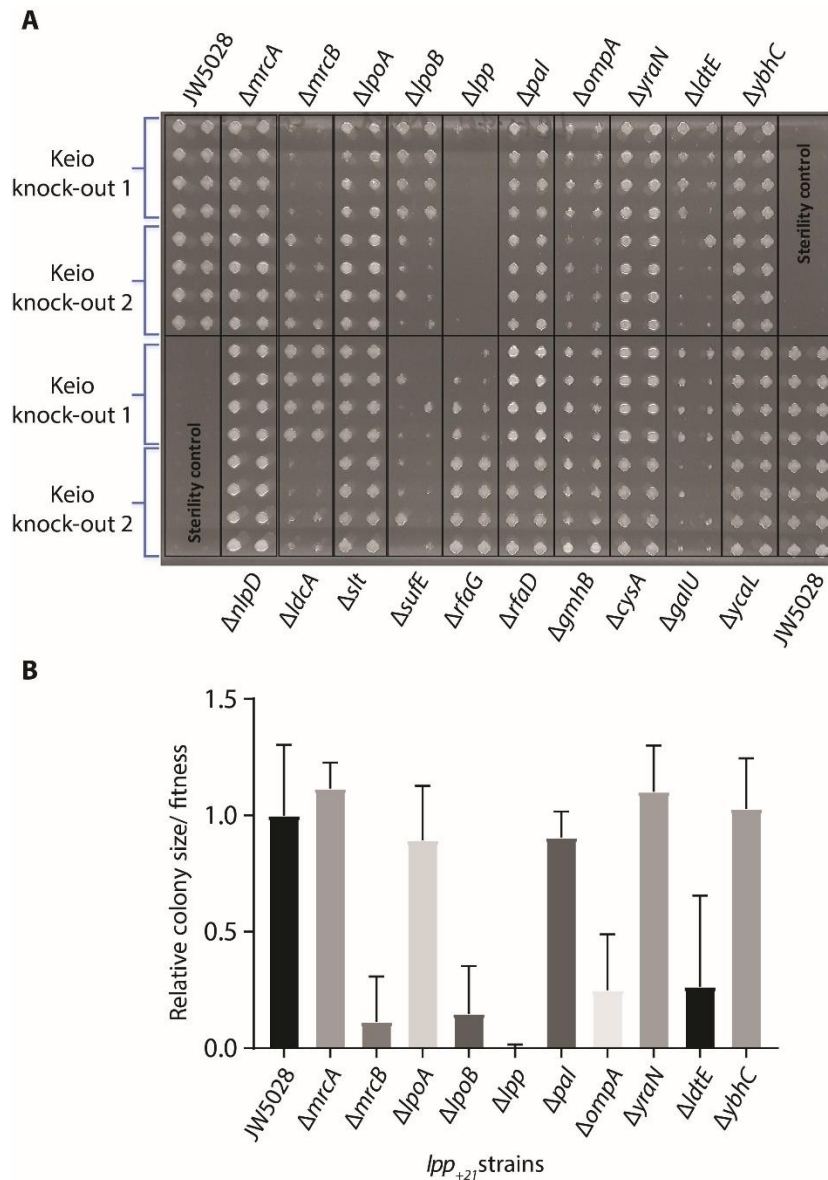


Figure 4.4.2.1: High-throughput verification and colony size quantification of synthetic sick/lethal double recombinants.

(A) A representative mini-screen on osmotically stabilised LB (0.4 M NaCl) of the genes identified from the main screen under the same condition. The self-mating (Δlpp) demonstrates a low false-positive rate, given that the double mutant of the same gene cannot exist in haploid organisms. $\Delta sufE$ and ΔdtE are examples of closely linked genes to lpp that exhibit a false aggravating interaction. Sterility controls were included where the wells were uninoculated in the recipient plate. Both independent Keio gene knock-outs are assessed; $n=8$ for each independent knock-out. (B) Colony size quantification of the double recombinants from the osmotically stabilised LB (0.4 M NaCl) media relative to growth control (JW5028). Values are averaged from independent clones ($n = 8$); error bars depict standard deviation.

Table 4.4.2.1: Non-essential genes enhancing the growth of Lpp₊₂₁ *E. coli*.

Cellular process	Genes identified
Peptidoglycan biosynthesis, maturation, turnover, remodelling	<i>lpoA, lpoB, ldcA, mrcA, mrcB, nlpD, slt, ybhC,</i>
LPS biosynthesis	<i>galU, gmhB/yaeD, rfaD/hldD/wad, ycaL⁵</i>
Drug efflux machinery	<i>acrB, mdtL, mdtK, ydhI, tolC</i>
OM-PG bridging	<i>ompA, pal, yiaD,</i>
Unknown function	<i>yraN</i>

normally non-essential factors regulating PG modification such as LdcA, NlpD, Slt, YbhC become essential. LdcA is an L, D-carboxypeptidase that has an essential function in peptidoglycan recycling/turnover. It is normally essential in stationary phase (Templin *et al.*, 1999), in the Lpp₊₂₁ strain background, it is essential in log-phase too. NlpD is a divisome associated, OM lipoprotein, which activates the PG hydrolase AmiC involved in septal splitting (Uehara *et al.*, 2010). Slt is a 65-kDa lytic transglycosylase (Holtje *et al.*, 1975), which cleaves β -1,4 glycosidic bond between GlcNAc and MurNAc residues of PG. Slt may function as a PG quality control enzyme by the degradation of uncrosslinked glycan strands, preventing their aberrant incorporation into the PG matrix (Cho *et al.*, 2014). YbhC is an OM lipoprotein (Molloy *et al.*, 2000), where phylogenetic analysis places YbhC into a subclade of the carbohydrate esterase 8 (CE8) family, hence its predicted PG modification function (Eklof *et al.*, 2009).

Most of the components of the LPS biosynthetic machinery are essential genes even in the wild-type background and so are not representative in the Keio collection. From the screen, some of the non-essential genes in LPS biosynthesis become essential in Lpp₊₂₁ strain. In addition to the characterised genes *gmhB*, *galU*, and *rfaD* involved in core LPS biogenesis, the gene encoding YcaL was also identified as essential. YcaL belongs to the widely conserved M48 family of metalloproteases (Weski & Ehrmann, 2012), and was shown to modulate LptD (the terminal OM protein involved in LPS shuttling across the periplasm) levels (Soltes *et al.*,

2017). The function of YcaL in quality control of LptD places it in the LPS biosynthesis category.

Genes encoding proteins involved in PG binding (Pal and OmpA) also became essential in Lpp₊₂₁ strain. This was somewhat surprising given that the increased OM-PG distance enforced by Lpp₊₂₁ could have shown Pal and OmpA to be inefficient and dispensible. Instead, this result suggests that Pal and OmpA have become essential factors to keep some local regions of OM-PG at wild-type distance. While not previously demonstrated to function in OM-PG bridging, *yiaD* detection in this screen was of interest. Conserved domain analysis was done that revealed YiaD has a PG-binding domain similar to that of Pal (Figure 4.4.2.2A) and OmpA (Figure 4.4.2.2B) (analysis was performed by Dr Chaille Webb, Monash University). These results categorise YiaD as a PG binding protein, and its essentiality is consistent with the identification of Pal and OmpA.

Another surprising result from the screen was the identification of genes encoding components of drug-efflux pumps. The *tolC*, *acrB*, *mdtL*, *mdtK* and *ydhI* mutants were all diminished in growth in the Lpp₊₂₁ genetic background (Table 4.4.2.1). The increased antibiotic sensitivity phenotype exhibited by Lpp₊₂₁ strain (Table 4.3.2.1), could have been as a result of decreased efflux pump function in Lpp₊₂₁ strain and not due to increase drug permeability. In this case, these data support the hypothesis that the Lpp₊₂₁ strain maintains a semi-functional resistance-modulation-cell division (RND) efflux pump.

Based on the fact that the screen-final selection was on double antibiotics, I speculated that the double drug selection could have had adverse effects on the Lpp₊₂₁ strains. This speculation was ruled out by assessing the growth of the mutants in M9 minimal media without antibiotics. The synthetic growth defects were still observed in the absence of drug selection (Figure 4.4.2.3B). TolC has been proposed to nestle within the PG layer (Meroueh *et al.*, 2006) (Koronakis *et al.*, 2004). The finding that OmpA and Pal are also essential to the Lpp₊₂₁ strain for growth on minimal medium, leads me to suggest that TolC provides a substantial linkage between the OM and PG, and that the stability of this bridge is enhanced by contact with its IM partners to form a functional RND pump. Overall, the essentiality of these genes (OmpA, Pal and TolC) in the Lpp₊₂₁ background, suggest that the Lpp₊₂₁ strain is maintaining at least local areas of closer contact between the OM and PG.

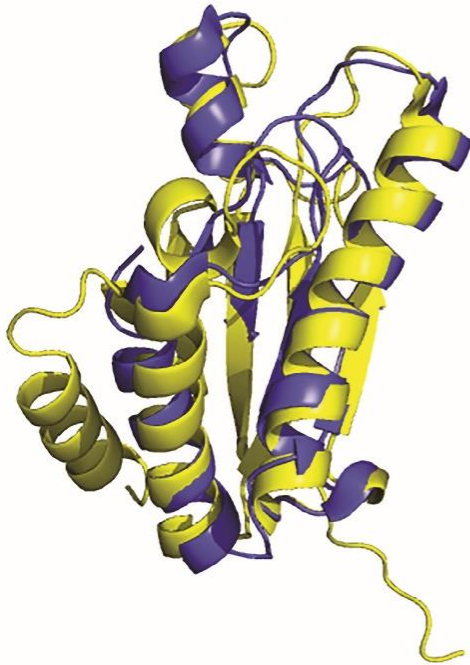
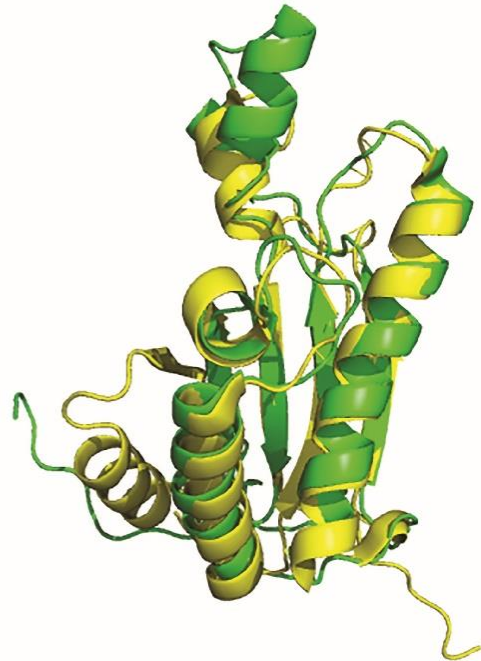
A**B**

Figure 4.4.2.2: A predicted role of YiaD as part of the OM-PG bridge.

(A) A superimposition of YiaD (PDB 2K1S chain A; blue) and the crystal structure of the lipoprotein Pal from *E. coli* (PDB 1OAP; yellow). (B) A superimposition of the structure for YiaD (blue) and the crystal structure of the PG-binding domain of the β -barrel protein OmpA from *E. coli* (PDB 2MQE residues 180 - 325; green). (Analysis was performed by Dr Chaille Webb, Monash University, Australia).

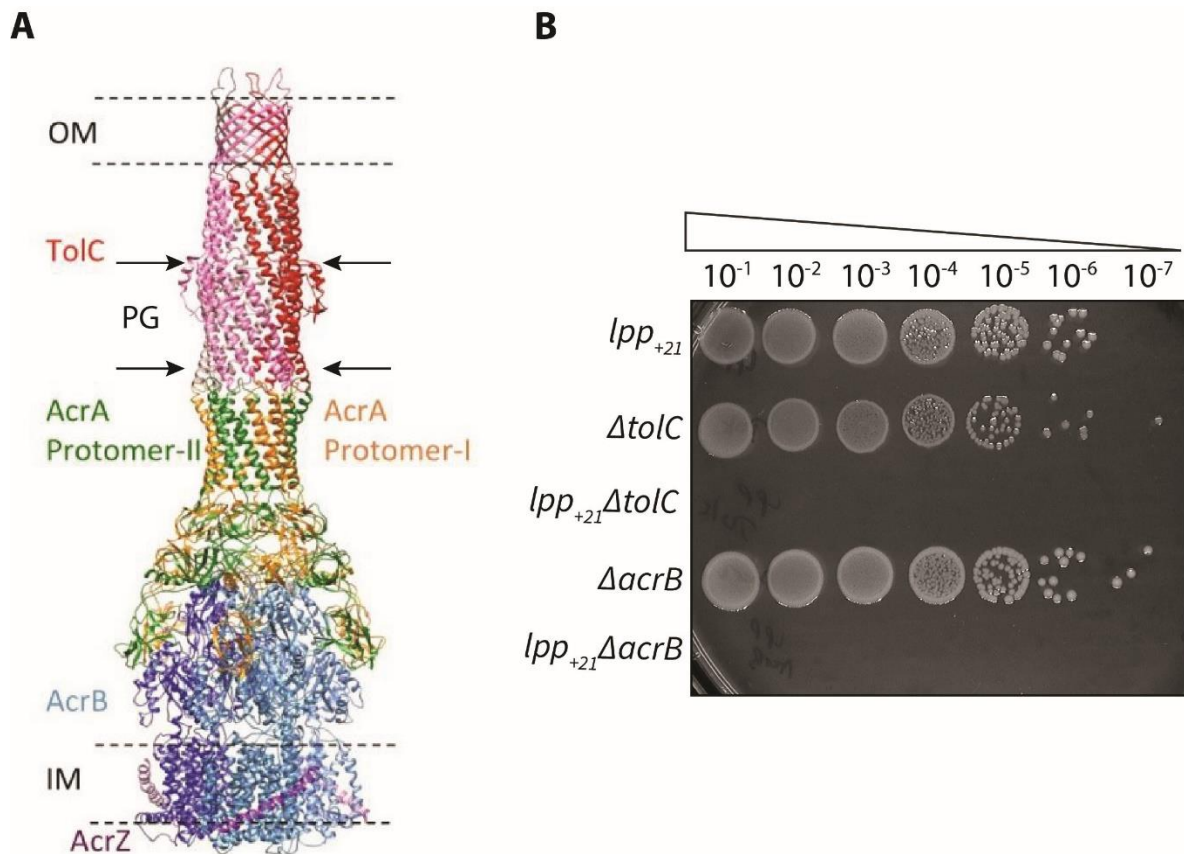


Figure 4.4.2.3: The AcrABZ-TolC complex is predicted to play a structural role as an OM-PG bridge.

(A) The structure of TolC-AcrA-AcrB-AcrZ complex from *E. coli* (PDB 5O66) shows the disposition of the complex across the periplasm. TolC (red and pink) is a trimer, with a β -barrel domain formed from four β -strands from each of the three protomers (Wang *et al.*, 2017). This β -barrel domain is embedded in the outer membrane. An elongated helical domain extends down into the periplasm and has been modelled as engaging with the PG layer (black arrows) (Meroueh *et al.*, 2006). (B) Viability counts of *lpp*₊₂₁ strain, *lpp*₊₂₁ Δ *tolC* strain, Δ *acrB* strain and *lpp*₊₂₁ Δ *acrB* strain in M9 minimal media supplemented with 0.2% (w/v) glucose without antibiotics. The double recombinants exhibit synthetic lethality in this media condition but not in nutrient-rich media. The plates were incubated at 37 °C for 36 hours. Representative data are shown from experiments performed in biological triplicate; n=3.

YraN is the only factor in the screen predicted to reside in the cytoplasm. YraN is predicted to be a Holliday-junction resolvase related protein (Aravind *et al.*, 2000). I speculate that this mutant failed to resolve the merodiploid condition transient in the introduction of the *lpp*₊₂₁ condition to the background strain.

4.4.2.1 Synthetic lethality of PBPs and their cognate lipoprotein activators in the *Lpp*₊₂₁ strain. Table 4.3.2.1 also shows that when one of the redundant pathways is knocked-out by loss of either *lpoA*, *lpoB*, *mrcA* or *mrcB*, the remaining redundant system is not adequate to support PG assembly in *Lpp*₊₂₁ strain to ensure viability of the cell. This phenotype was most severe in nutrient-limited medium (Figure 4.4.2.1.1B) but partially suppressed in nutrient-rich medium (Figure 4.4.2.1.1A).

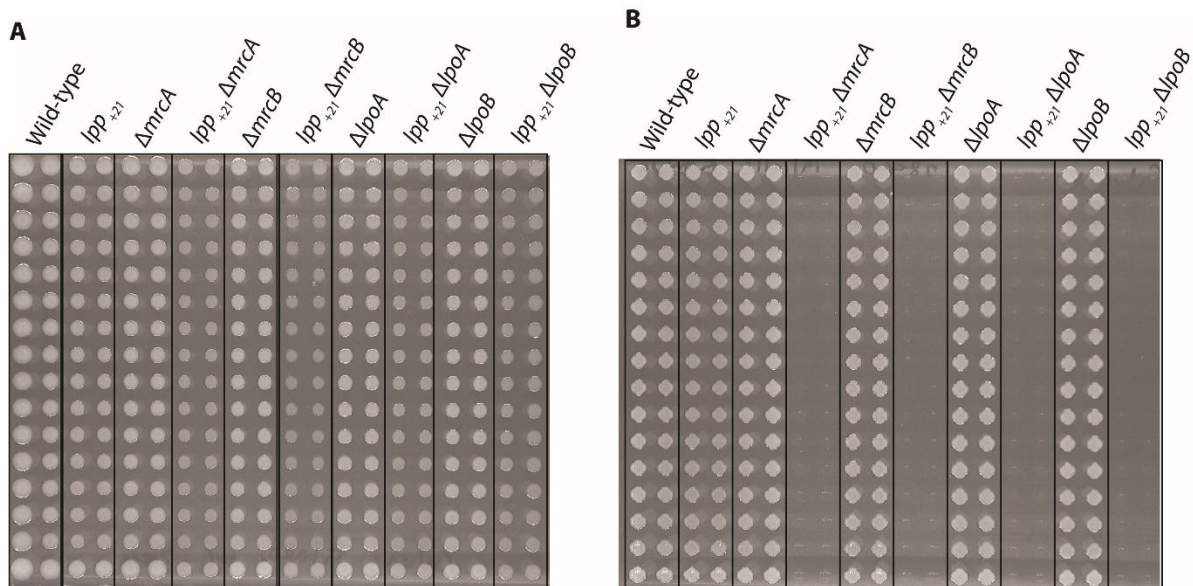


Figure 4.4.2.1.1: Synthetic lethality of major penicillin-binding proteins 1A and 1B, and their cognate lipoprotein activators to *Lpp*₊₂₁.

Panel A shows synthetic lethality results in nutrient-rich (LB) media, while panel B shows results in M9 minimal media supplemented with 0.2% (w/v) glucose. Using the Giant-coli technique, four double recombinants in LB were generated then manually arranged to include more replicas and single deletion controls. The double recombinants and controls were then pinned in both LB, and M9 minimal media and their growth analysed. *lpp*₊₂₁ is synthetically lethal with $\Delta mrcA$, $\Delta mrcB$, $\Delta lpoA$, and $\Delta lpoB$ in minimal media, unlike in nutrient-rich media. Both the single and the double-gene recombinants were pinned at n=32 for each independent knock-out.

4.4.3 Supplementation of amino acids rescues the synthetic lethality displayed by PG assembly proteins in *Lpp₊₂₁*

To explain how rich medium partially suppressed these phenotypes, I supplemented M9 minimal media with other carbon sources, iron (II) sulfate (FeSO₄), or casamino acids. In each of the conditions, the mutants were assessed for viability.

Carbon source supplementation was performed to rule out any defects that may result in impaired glucose import across the periplasm. Glycerol (1% v/v) and maltose (0.2% w/v) were supplemented in M9 minimal media. None of these conditions suppressed the synthetic lethal phenotypes of the double recombinant strains on M9 medium (Table 4.4.3.1).

Table 4.4.3.1: Viability assessment of the *Lpp₊₂₁* double recombinants in M9 minimal media with varying supplements.

Minimal media supplements	Double recombinant strains			
	<i>lpp₊₂₁ ΔmrcA</i>	<i>lpp₊₂₁ ΔmrcB</i>	<i>lpp₊₂₁ ΔlpoA</i>	<i>lpp₊₂₁ ΔlpoB</i>
M9 + Glycerol (1% v/v)	Lethal	Lethal	Lethal	Lethal
M9 + Maltose (0.2% w/v)	Lethal	Lethal	Lethal	Lethal
M9 + Glucose (0.2% w/v)+ FeSO ₄ (1 μM)	Lethal	Lethal	Lethal	Lethal
M9 + Casamino acids (0.2% w/v) + Glucose (0.2% w/v)	Viable	Viable	Viable	Viable

Iron (II) sulfate was supplemented to rule out any iron transport defects in an enlarged periplasm. Iron transport can be regulated by TonB, a component of the energy transducing system that transmits energy from the IM to the OM for the active transport of iron-siderophores complexes. TonB is trans-periplasmic with an N-terminal domain anchored in the IM, and a C-terminal domain that interacts directly with Ton dependent transporters in the OM (Pawelek *et al.*, 2006). For this reason, I hypothesised that increase in OM-IM distance on an *Lpp₊₂₁* background might compromise TonB function in iron import. However, iron (II) sulfate supplementation of M9 minimal media did not suppress the synthetic lethal phenotype (Table 4.4.3.1).

In *E. coli* and other Gram-negative species, the PG peptide contains a myriad of L- and D- amino acids that are linked to MurNAc residues (Egan *et al.*, 2015). The primary peptide

sequence (L-Ala-D-iGlu-*m*-Dap-D-Ala-D-Ala) comprises L- and D-alanine, L-glutamic acid and *meso*-diaminopimelic acid, a derivative of lysine (Egan *et al.*, 2015). As amino acids are known building blocks of PG, I hypothesised that with suboptimal PG assembly systems (in the double recombinants), amino acid limitation could further cripple the system and explain the synthetic lethality. Supplementation of amino acids in the form of casamino acids (0.2% w/v) rescued the synthetic lethal phenotype (Table 4.4.3.1), suggesting that amino acid uptake becomes rate-limiting step in the Lpp₊₂₁ strains.

4.5 Discussion

The artificially long version of Lpp (Lpp₊₂₁) provided an opportunity to study the significance of periplasmic architecture in Gram-negative bacteria, using *E. coli* as a model. Two previous studies in *S. enterica* and *E. coli* used Lpp₊₂₁ to demonstrate that Lpp is the primary determinant of periplasmic size. In the first study, the authors showed that a larger periplasm allowed for correct assembly of artificially longer flagellar rods, restoring motility to a non-motile *flgG* mutant (Cohen *et al.*, 2017), while the second study showed that the OM to IM distance modulates envelope surveillance. In a larger periplasm, the trans-periplasmic protein RcsF could not transfer envelope damage signals to its downstream partner IgaA in the IM (Asmar *et al.*, 2017). In this chapter, a genetic approach was used to investigate other periplasmic processes that might rely on strict control of the OM to IM distance. The genetic screen exploited the Lpp₊₂₁ to screen for non-essential genes that become essential in a larger periplasm.

The screen identified several genes that become essential in the Lpp₊₂₁ strain (Table 4.4.2.1). The genes point to compromised LPS assembly, PG biosynthesis, and OM-PG attachment complexes. Most of the LPS assembly proteins are essential, and the identification of the other, non-essential genes from this cellular process as becoming essential, showed that the LPS transport across the periplasm is impaired in the Lpp₊₂₁ strain. I rationalise that LptA, which acts as a bridge by connecting the IM subcomplex (LptB₂FGC) to the OM subcomplex (LptDE) (Hicks & Jia, 2018) is negatively impacted by increased OM-IM distance. The half-life of unassembled LptA is very short due to turn-over by an unknown protease (Sperandeo *et al.*, 2011) and I speculate that YcaL, previously shown to control turn-over of LptD (Soltes *et al.*, 2017), could be the protease involved in quality control of the unassembled LptA. YcaL was found to be essential in the Lpp₊₂₁ strain. These results are consistent with the SDS sensitivity phenotype exhibited by Lpp₊₂₁ strain (Figure 4.3.2.1). I speculate that the increased OM to IM

distance places an additional load on the LPS periplasmic bridge, limiting LPS translocation and impacting the OM biogenesis process. However, this effect is mild, as shown by the LOS quantification study (Figure 4.3.5.1) and the single-cell imaging assay (Figure 4.3.2.3).

The results also emphasized the importance of strict control of the periplasmic size in PG biosynthesis, maturation, turnover and remodelling. The identification of PG biosynthetic genes was not surprising, as PG is an essential component of the cell envelope with complex redundant systems involved in its biogenesis. This chapter demonstrated that activation of two IM PBP (PBP1A and PBP1B) by their cognate lipoprotein partners (LpoA and LpoB) is impaired in *Lpp₊₂₁* strain. Figure 4.5.1 depicts a model for this hypothesis. In a JW5028 strain, where the OM-IM distance is estimated at 20 nm, the interaction between Lpo and PBP is optimised. As a result, sufficient PG assembly takes place even when one of the redundant pathways are removed. However, in an *Lpp₊₂₁* strain, where the OM-IM distance is estimated at 24 nm, interactivity is sub-optimal. Removal of either of the redundant pathways prevents efficient PG assembly. In nutrient-limited media, the system is crippled, leading to loss of viability. Only by restoring the complete machinery through complementation or by increasing supply of amino acids is it possible to overcome the synthetic lethal phenotype.

The significance of periplasmic architecture in PG biosynthesis is further demonstrated by the identification of other proteins directly or indirectly associated with this cellular process. Among these proteins, LdcA, which is usually essential in stationary phase in a JW5028 background, becomes essential during log phase planktonic growth in *Lpp₊₂₁* background. This is likely due to increased demand to recycle PG for actively diving cells, which is limited due to an impaired PG assembly system in *Lpp₊₂₁* strain. YbhC was also identified as essential in the *Lpp₊₂₁* background. Its structure has revealed that it has the conserved fold of CE8, but the active site is modified to be narrow and hydrophobic and has been suggested to function in PG turnover and cell division (Eklof *et al.*, 2009). Its detection in this screen is consistent with this notion.

The significance of OM-PG binding is demonstrated by the essentiality of other PG binding proteins in an *Lpp₊₂₁* strain. Among them, OmpA and Pal non-covalent PG interaction is well established (Park *et al.*, 2012). Molecular dynamic simulations by Samsudin *et al.*, (2017) show that without Lpp, the OmpA N-terminal and C-terminal linker can extend by 1.5 nm to reach the PG layer. Once the binding occurs, the linker contracts shifting the PG layer upwards by 3 nm. However, in the presence of Lpp, the linker extension is reduced, extending by only 0.5

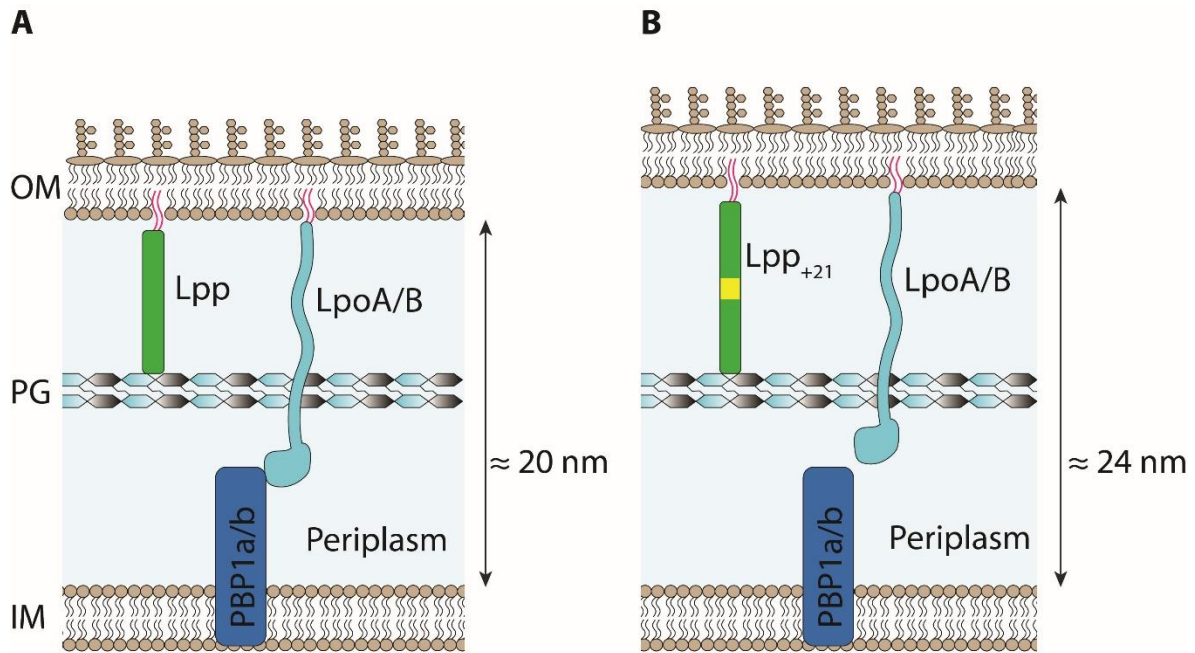


Figure 4.5.1: Schematic representation of the implication of periplasmic architecture on the interaction of PBP1A/B with LpoA/B.

(A) In a JW5028 strain, the periplasmic width is approximately 20 nm (Matias *et al.*, 2003; Asmar *et al.*, 2017; Cohen *et al.*, 2017). At this distance, the interaction of LpoA/B with PBP1A/B is optimal for efficient PG assembly even when one of the redundant pathways is removed. (B) In an Lpp_{+21} strain, the OM-IM distance increases by around 4 nm (Asmar *et al.*, 2017; Cohen *et al.*, 2017), leading to suboptimal interaction of the protein pairs. The inefficient assembly of PG in this background is further compromised when one of the redundant pathways is removed in nutrient-limited media leading to loss of cell viability.

nm. This is in agreement with the ability of Lpp to lift the PG layer upwards and reduce the distance between the PG and the OmpA C-terminal domain. This ability of OmpA linker to extend would allow the OmpA-PG interaction even in the presence of Lpp₊₂₁. However, it remains unknown how much the linker would have to extend in presence on Lpp₊₂₁ and an interesting aspect for further investigation. The realisation that YiaD also has a PG-binding domain similar to that of Pal and OmpA (Figure 4.4.2.2) suggested that YiaD provides additional OM-PG attachment support, a novel function for the protein. This result is consistent with the identification of a YiaD homolog in *A. baumannii* in a screen for OM-PG bridges (Wu *et al.*, 2016).

In addition to YiaD, I also provide evidence that suggests a structural function for TolC. Functional efflux pumps are composed of a TolC trimer embedded in the OM, which is tightly engaged with a hetero-oligomeric pump embedded in the IM. In *E. coli*, the primary pump for many antimicrobial and toxic compounds is composed of the adaptor AcrA, the IM proteins AcrB, and AcrZ (Figure 4.4.2.3A) (Wang *et al.*, 2017). The diameter of the PG-interacting segment of the TolC protein is 7 nm (Koronakis *et al.*, 2004; Lambert *et al.*, 2005), using this information and by modelling, Meroueh *et al.* (2006) showed that TolC is permissively shaped for insertion into the 7 nm pore of the PG. The identification of AcrB in addition to TolC suggests that for TolC to play this structural role, a stable trans-periplasmic structure of the pump is required. The MdtK protein also identified, contributes the IM subunit to a second form of pump docked to TolC, assisting with the efflux of compounds such as ciprofloxacin (Nishino & Yamaguchi, 2001; Pietsch *et al.*, 2017).

To deduce whether the cell wall has enough elasticity to distort and allow these closer interactions to be maintained the information on distribution of Lpp across the *E. coli* cell envelope is crucial. The binding between Lpp and the PG occurs every 10th to 12th mDAP residue of the PG, and only a third of total molecules are covalently bound to PG (Braun & Rehn, 1969). The authors used this information to establish that in about 10⁵ Lpp molecules per cell, the molecules would be distributed 10.3 nm apart along the glycosidic chain of the PG. In addition, using immunoelectron microscopy, it was deduced by visual inspection that the distribution of bound Lpp was homogenous over the entire PG (Hiemstra *et al.*, 1987). This deduction still awaits confirmation using modern approaches. We now have new information that Lpp is found in over one million copies per cell (Li *et al.*, 2014) and we still have gaps in

knowledge that would help us answer this question. For instance, where and when the cross-linking of Lpp to PG occurs, how the cells control the ratio between the free and bound forms of Lpp and if the distribution of bound Lpp is homogenous at the poles and along the cell axis. Unless we have this information, we cannot deduce whether the PG has enough elasticity to distort, as this would be dependent on the proximity of bound Lpp molecules. It is also possible that localisation of Lpp may not be even throughout the cell envelope, resulting in regions of wider and narrower periplasmic width in which Lpp and Pal/OmpA/TolC are located, respectively.

The identification of these proteins indicates that the cell maintains local areas of close contact between the OM and IM, which may explain why the Lpp₊₂₁ strain shows no defect in growth on varying media conditions. The local areas of close contact maintained by these redundant systems would allow systems that require strict control of the periplasmic size to maintain their functionality. These results are consistent with Lpp length analysis observations (Figure 4.2.2), where majority of bacteria display a narrow window of Lpp length, around 78 residues (*E. coli* Lpp length). The highly conserved Lpp length suggest evolution constraints for this specific length. Based on Lpp₊₂₁ synthetic lethal screen results, the constraints are the systems identified, that rely on strict control of the periplasmic size and other PG binding proteins. The longer Lpp variants identified in *Geobacter* and *Vibrio* species suggested the size of the periplasm in these bacteria might be bigger. I hypothesise that other periplasmic systems that required strict control of periplasmic size have coevolved for correct function in this enlarged periplasm. However, this is yet to be investigated.

Chapter 5: Transcriptomic and proteomic profiling to investigate the adaptive response to an enlarged periplasm

5.1 Introduction

The results from the Lpp₊₂₁ synthetic lethal screen in chapter 4 established two major observations. Firstly, the periplasmic size is significant for the normal functioning of a Gram-negative bacterial cell. The number of non-essential genes that become essential in the Lpp₊₂₁ strain (Table 4.4.2.1) support that contention and the active involvement of the genes in various periplasmic processes suggest that the periplasmic machineries and the periplasmic architecture have been evolutionary co-optimised for correct function. Secondly, the data establishes that *E. coli* can effectively adapt to the lengthened Lpp₊₂₁. In Section 4.3, phenotypic characterisation demonstrates that the Lpp₊₂₁ strain exhibited just a few distinctive phenotypes compared to wild-type, and growth profiles were identical in all the conditions tested. The observations lead to a new research question: how does the cell adapt to the changes in periplasmic size for maintenance of periplasmic processes?

Measurements taken in *E. coli* suggest that the periplasm represents 7% of the volume of a bacterial cell (Koch, 1998; Prochnow *et al.*, 2019) yet contains at least 20% of cellular water, and in osmotically stabilised growth medium up to 40%, of the cellular water (Stock *et al.*, 1977). One interpretation of these measurements, given direct measurements show that the water fraction can vary substantially with external osmotic conditions (Cayley *et al.*, 2000; Sochacki *et al.*, 2011), is that the local concentration of a given protein component of the periplasm can vary depending on the volume of the periplasmic compartment. An alternative interpretation would be that mechanisms of proteostasis might be at play in the compartment to correct the protein concentration through transcriptional (i.e. gene activation/repression) or post-transcriptional (i.e. regulated proteolysis) regulation.

In response to its environment, a cell transcribes a specific subset of genes to RNA, which specifies a cell's identity and regulates its sub-cellular processes as an environmentally appropriate phenotype. Collectively termed as the transcriptome, the RNA molecules are essential for interpreting the functional aspects of the genome. Post-transcriptionally, further control can be brought to bear, since (i) some mRNAs will not be translated to protein without additional translational activators, and (ii) protein molecules are subject to turn-over by

proteases, and inhibition/activation of this proteolysis will modulate the steady-state level of each given protein.

In this chapter, I investigated the adaptive response of *E. coli* to an enlarged periplasm. The response was examined at the transcriptional level by RNA sequencing (RNA-seq) analysis and the protein level by mass spectrometry-based proteomic profiling. The following analyses, therefore, consider three scenarios:

- (a) Increased protein steady-state levels in the Lpp₊₂₁ strain, which may indicate transcriptional or post-transcriptional control to restore or enhance the concentration of specific proteins in the enlarged periplasm;
- (b) Unchanged protein steady-state levels in the Lpp₊₂₁ strain, which may indicate that the concentration of those proteins is subject to regulatory mechanisms that do not sense the enlarged periplasm;
- (c) Decreased protein steady-state levels in the Lpp₊₂₁ strain, which may indicate transcriptional or post-transcriptional control to remove specific proteins from the enlarged periplasm.

Quantitative proteomics was used to determine the protein steady-state levels. It relies on mass spectrometry (MS) to identify and quantitate peptides. For this reason, the advancement of MS has contributed immensely to the development of quantitative proteomics by improving the high-throughput identification and quantitation of proteins. However, MS is not inherently quantitative, since proteolytic peptides show considerable variability in physiochemical properties, which in turn results in mass spectrometric variability between runs (Ahrens *et al.*, 2010). Furthermore, mass spectrometers sample a small percentage of the total peptides in a sample (Ahrens *et al.*, 2010). A relative comparison strategy is used to achieve quantification. This strategy compares levels of individual peptides in a sample to those in an identical, but experimentally-modifiable, sample (Wasinger *et al.*, 2013). The approach is termed relative quantitative proteomics. In this study, samples were analysed separately using liquid chromatography-tandem mass spectrometry (LC-MS/MS), then the spectra were compared to determine peptide abundance in one sample relative to another

5.2 Gene regulation in response to an enlarged periplasm

Bacteria have evolved complex regulatory networks in order to adjust their physiology for survival in the face of frequently changing environmental conditions. Maintenance of cellular functions relies on the proper expression of genetic information in which RNA molecules play

a key role. Among the various RNA molecules, mRNA constitutes the molecular link between genes and proteins. Since the concentration of a given transcript depends on the rate of synthesis and degradation, the level of each protein in the cell depends on the control of both transcription and degradation of mRNA (Silva *et al.*, 2011). Unlike DNA, RNA is less stable, particularly mRNAs, where the turnover in prokaryotes is rapid. In *E. coli*, the half-lives of mRNA can vary from less than a minute to over 10 minutes (Mohanty & Kushner, 1999; Selinger *et al.*, 2003). The dynamic nature of mRNA turnover ensures that microorganisms can quickly adapt to sudden environmental changes by modulating the rate of synthesis of either new or existing proteins. This is specifically crucial in prokaryotes, where transcription and translation are coupled, i.e. mRNA is translated while being synthesised.

To date, the method of choice to address the complexity of bacterial transcriptomes, which includes tRNA, rRNA, small regulatory RNA (sRNA) and mRNA, is RNA-seq. In brief, the technique is ultimately an analysis of complementary DNA (cDNA) molecules in high-throughput sequencing (Croucher & Thomson, 2010; Van Vliet, 2010). Following deep sequencing analysis of the cDNA, short sequences (reads) are obtained, which are then mapped onto a reference genome for further data analysis.

5.2.1 Transcriptional response to an enlarged periplasm

RNA-seq was applied to investigate how *E. coli* adjusts to an enlarged periplasm. For consistency with the synthetic lethal screen (Section 4.4), RNA-seq samples were prepared in two conditions: M9 minimal media and M9 minimal media supplemented with 0.5 M sorbitol. Replicates in RNA-seq are crucial for measuring variability and improving data quality, so four biological replicates (quadruplicates) were used for each of the strains, in each test condition. Total RNA was isolated at a mid-log phase of growth, converted to cDNA, and subsequently sequenced on a next-generation sequencing (NGS) platform (Micromon Genomics, Monash University) (Section 2.4.6 and Section 2.4.7 provide detailed information on the procedure).

Following sequencing, read mapping results have to be summarised in terms of read coverage for genomic features of interest before they can be interpreted biologically. One of the most ubiquitous operations that form part of many next-gen analysis pipelines is to count the number of reads overlapping predetermined genomic features of interest, in this case, the gene. Read counts provide an overall summary of the coverage for the genomic feature of interest. In particular, gene-level counts from RNA-seq provide an overall summary of the expression level of the gene (Liao *et al.*, 2013). Figure 5.2.1.1A shows a summary of the percentage of assigned

and unassigned reads for each of the samples processed. The unassigned reads could be due to failure in mapping, poor mapping quality, lack of feature (gene) or due to ambiguity where a read can be assigned to several genes. All the samples showed a high percentage of assigned reads (> 70%) an indicator of a good quality dataset (Liao *et al.*, 2013). Figure 5.2.1.1B shows the number of reads aligned to the reference genome, where all samples have over ten million reads aligned. Ten million reads per sample are considered the benchmark for a reliable dataset (Wang *et al.*, 2011). From the same figure, the samples had good, high mapping rates, where about 94% of the reads mapped to the reference genome.

In addition, the quality of each sequence was analysed. Using the mean score across all bases in the sequence of each sample, the quality score was determined by plotting out the distribution of the means against the base counts. The samples formed a tight distribution with universally high quality, and there were no sequences with low quality. Analysis of the sequence length distribution showed that all samples passed with sequences of a single length (100 bp). The analysis was performed at Monash Bioinformatics platform (Dr Kirill Tsyganov).

Table 5.2.1.1.1 and Table 5.2.1.2.1 document the genes up-regulated and down-regulated, respectively, in the Lpp₊₂₁ strain relative to the isogenic wild-type. At the transcriptome level, very few genes encoding periplasmic processes responded in the Lpp₊₂₁ strain, and no complete regulons were up- or down-regulated.

5.2.1.1 Upregulation of genes in response to an enlarged periplasm

Table 5.2.1.1.1 presents the four genes that passed a minimal criterion for up-regulation (at least Log₂ 1.2-fold change) in the Lpp₊₂₁ strain relative to the wild-type. Of these genes, the Lpp isoform gene (*lpp₊₂₁*) was the most substantial (greater than four-fold) up-regulated in the enlarged periplasm under both osmotic conditions. The up-regulation of *lpp₊₂₁* was a significant result as I had previously observed in western blots that the steady-state level of the Lpp₊₂₁ protein was significantly lower than wild-type levels of Lpp (Figure 4.2.1). The result suggests that diminished Lpp₊₂₁ levels feed-back to upregulate expression of the *lpp₊₂₁* gene in a response directed to achieve levels required for normal functioning of the cells.

yjbT is the only gene of unknown function in this table. The protein is a predicted substrate for the SecYEG IM translocation system and predicted to reside within the periplasm or secreted to the OM (SignalP 5.0). The gene product is of unknown function and, despite being up-regulated in one condition (M9 minimal media), the magnitude of up-regulation is also low at only Log₂ fold change of 1.29.

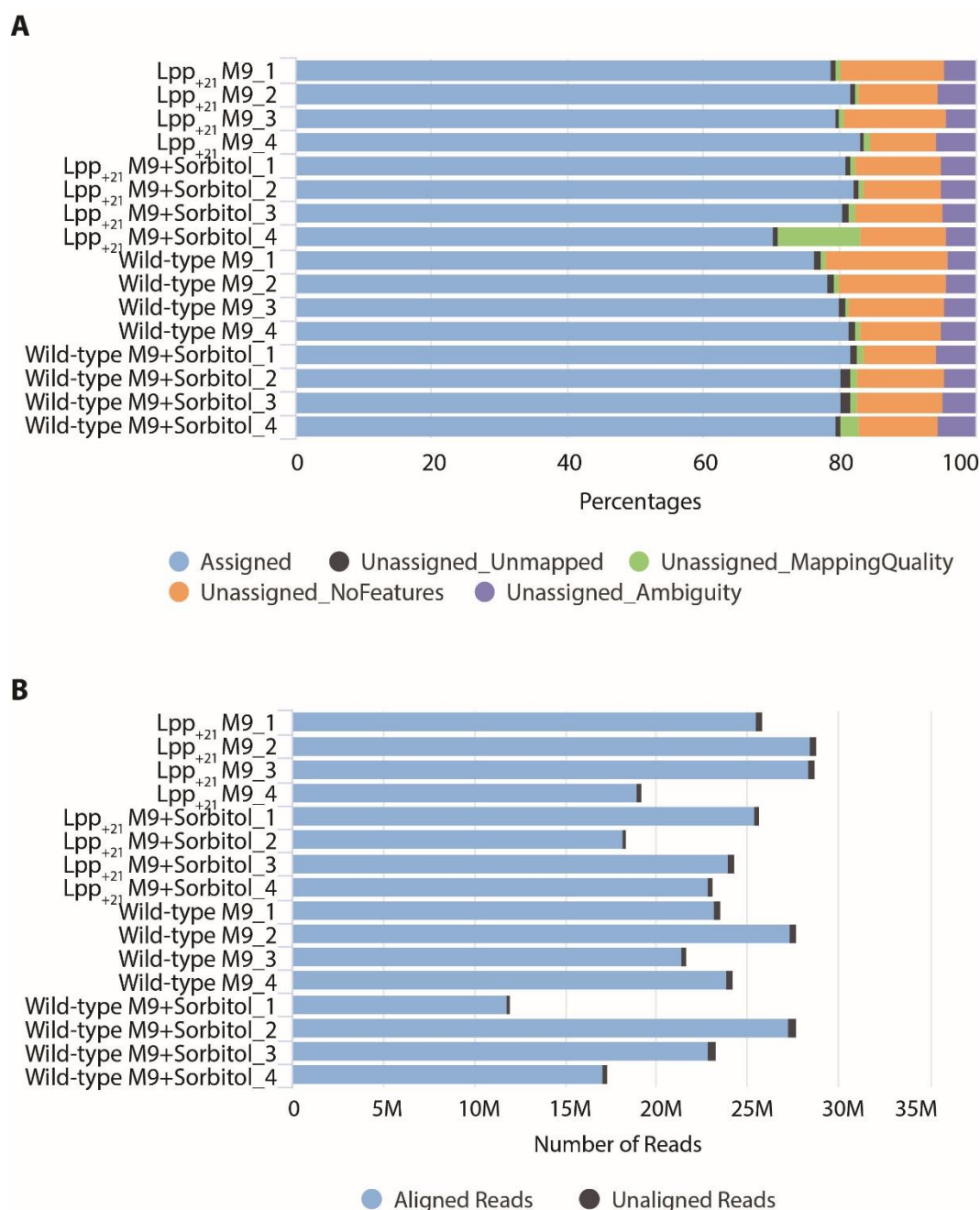


Figure 5.2.1.1. RNA-seq sample quality analysis.

(A) Feature counts (assignments). The graph presents the percentages of assigned or unassigned reads for qualitative assessment of the samples. Over 70% of the reads were assigned to features, indicating a good quality dataset. (B) The bars show the alignment reads per samples, where all reads aligned are over 10 million (M), considered a benchmark for a good dataset. Percentage comparison of the aligned reads versus the unaligned shows that over 94% of the reads mapped to the reference genome (BW25113).

Table 5.2.1.1.1: Genes up-regulated in Lpp₊₂₁ strain.

Gene	Protein	Biological process	Localisation	Growth condition – Log ₂ fold change	
				M9	M9 + sorbitol
<i>lpp₊₂₁</i> [#]	Major outer membrane lipoprotein Lpp	Periplasmic space organisation	Periplasmic space, outer membrane	2.9	3.03
<i>insH-1</i>	CP4-6 prophage; IS5 transposase and trans-activator	transposition of the insertion sequence IS5.	Cytosol	-*	1.42
<i>trpD</i>	anthranilate synthase subunit TrpD	Tryptophan biosynthetic process	Cytosol	-*	1.20
<i>yjbT</i>	PF17089 family protein YjbT	Unknown	Periplasmic space, outer membrane	1.29	-*

- *lpp₊₂₁*. An isoform of *lpp*. The gene contains three extra heptad repeats, hence termed *lpp₊₂₁*

* - No change in gene expression levels

The other two genes in Table 5.2.1.1.1 (*insH1* and *trpD*) are likewise of minimal response in the Lpp₊₂₁ strain. Based on these findings, I theorise that any upregulation in the periplasmic proteome has been achieved through post-transcriptional mechanisms.

5.2.1.2 Genes down-regulated in response to an enlarged periplasm

Table 5.2.1.2.1 presents genes down-regulated at least 4-fold in Lpp₊₂₁ strain relative to the wild-type in either conditions. None of the genes encode proteins predicted to be localised to the periplasm. Notable among these down-regulated genes in Lpp₊₂₁ are the genes involved in galactitol metabolism. The *gatB*, *gatC*, *gatD*, *gatA*, and *gatR* genes are down-regulated at high magnitude in both M9 (-sorbitol) and M9 (+sorbitol) conditions.

ypdK is a small, 23 amino acid long, predicted IM protein. The protein was discovered by Hemm *et al.* (2008) using comparative genomics. The authors noted its expression levels peaked during exponential growth. Although small membrane proteins are difficult to characterise, an increasing body of evidence shows that these polypeptides have important intracellular and intercellular functions (Burkholder *et al.*, 2001; Hemm *et al.*, 2008). Due to

Table 5.2.1.2.1: Genes down-regulated in Lpp₊₂₁ strain.

Gene	Protein	Biological process	Localisation	Log ₂ fold change	
				M9	M9 Sorbitol
<i>gatB</i>	galactitol-specific PTS enzyme IIB component	Galactitol transport	Cytoplasm	6.43	4.35
<i>gatC</i>	galactitol-specific PTS enzyme IIC component	Protein-phosphocysteine-galactitol-phosphotransferase system transporter activity	IM	6.23	4.73
<i>gatD</i>	galactitol-1-phosphate 5-dehydrogenase	Galactitol metabolic process	Cytoplasm	6.06	5.92
<i>gata</i>	galactitol-specific PTS enzyme IIA component	Galactitol transport, galactitol metabolic process	Cytoplasm	5.05	3.08
<i>gatR</i> [#]	DNA binding transcriptional repressor GatR, N-terminal fragment	Repressor of the <i>gat</i> operon for galactitol transport and metabolism	Cytoplasm	2.82	4.38
<i>ypdK</i> ⁺	uncharacterized membrane protein YpdK	Unknown	IM	2.61	-*
<i>srlE</i>	sorbitol-specific PTS enzyme IIBC1 component	Transport of D-sorbitol, D-mannitol PTS permease	IM	-*	2.9
<i>srlB</i>	sorbitol-specific PTS enzyme IIA component	Transport of D-sorbitol, D-mannitol PTS permease	Cytoplasm	-*	2.86
<i>srlD</i>	sorbitol-6-phosphate 2-dehydrogenase	Sorbitol catabolic process	Cytoplasm	-*	2.78
<i>srlM</i>	DNA-binding transcriptional activator SrlM	Positive regulator for sorbitol operon expression.	Cytoplasm	-*	2.74

* - No change in gene expression levels

- Truncated gene. The *gatYZABCD* operon is expressed constitutively due to an insertion of IS3E into the *gatR* gene (Nobelman & Lengeler, 1996)

+ -*YpdK* is highly expressed during exponential growth (Hemm *et al.*, 2008)

time constraints, no further work on YpdK was possible. The reason as to why this *ypdK* is down-regulated in Lpp₊₂₁ strain remains a mystery for now.

Of the ten existing hexitols, only three occur naturally: D-mannitol, D-sorbitol and galactitol (Lengeler, 1975). Bacteria can utilise each of these as a source of carbon and energy. Galactitol is derived from either D - or L- galactose by the reduction of the aldehyde group (Nobelmann & Lengeler, 1996). The three hexitols enter the cell via a specific phosphotransferase system to form the first intracellular species, which is a 6-phospho-derivative (Lengeler, 1975). Of note, due to the symmetry of D-mannitol and galactitol, their phosphorylation at either end produce the same molecule (Lengeler, 1975). D-mannitol and D-sorbitol enter glycolysis via a keto sugar phosphate (β -D-fructofuranose-6-phosphate), unlike galactitol which enters glycolysis via either D-glyceraldehyde-3-phosphate or glycero phosphate, which involves extensive steps to form (Nobelmann & Lengeler, 1996; Nolle *et al.*, 2017). Based on this information, galactitol is uniquely metabolised as a carbon source compared to the other hexitols. The fact that there is no galactitol provided, the reason as to why these genes were significantly down-regulated in Lpp₊₂₁ strain is perplexing. The sorbitol effect is unlikely as the genes are also significantly down-regulated in M9 minimal media without sorbitol.

The other genes down-regulated *srlE*, *srlB*, *srlD*, and *srlM* are involved in sorbitol metabolism and are down-regulated only in the sorbitol condition. Thus, the effect was deemed likely to be sorbitol-specific effects on bacterial physiology. However, since the analysis was comparative to the isogenic wild-type in the same condition, this is reflective of osmotic differences in the wild-type and the Lpp₊₂₁ strains.

5.3 Proteostasis in an enlarged periplasm

In order to survey for post-transcriptional changes, quantitative proteomics was applied. Quantitative proteomics is a powerful technique to understand global proteome dynamics in cells. There are various strategies in which quantitative proteomics can be performed. In this study, a label-free quantitative (LFQ) strategy was used for relative quantification. The technique employs a bottom-up approach, where peptides generated from the proteolytic digestion of protein samples using trypsin are analysed by MS (See Section 2.6.5 for detailed methods). LFQ MS was performed comparing the Lpp₊₂₁ strain to the isogenic wild-type in similar conditions as the RNA-seq experiment. The analysis was performed at Monash Proteomics and Metabolomics Facility.

5.3.1 Proteomics results; data quality analysis

A total of approximately 2000 proteins were identified across all samples falling below a predefined false discovery rate (FDR) threshold of 1%. Figure 5.3.1.1A displays the number

of proteins identified in each sample in both conditions. The number of proteins identified is consistent across the samples, which is an indicator of a high-quality dataset.

For further data quality analysis, a 2-dimensional principal component analysis (PCA) was performed. PCA is a technique used to emphasise variation and bring out strong patterns in a dataset; i.e. the more similar the samples are, the closer they cluster together. Figure 5.3.1.1B displays the PCA results of the samples from this study. Clustering was both condition and replicate dependent, and as expected, biological replicates exhibited a tight clustering.

Data quality was also assessed by determining the distribution of the coefficient of variation (CV); the smaller the median CV, the better the quantitative similarity between the replicates. The density histogram and box plot generated (not shown) showed good separation and a small CV (<10%), indicating high-quality dataset.

5.3.2 Changes to the proteome in an enlarged periplasm

Relative quantitative analysis was performed between the two strains and also across the two conditions. To be consistent with the transcriptome analyses, all samples were prepared in M9 growth medium, after growth to mid-log phase. Two analyses are presented here: a comparison of the Lpp₊₂₁ strain to the wild-type strain in hypo-osmotic (M9) medium, and a comparison of the Lpp₊₂₁ strain to the wild-type strain in a more iso-osmotic (M9+sorbitol) medium. In order to control for strictly osmotic effects, an equivalent analysis was made of the wild-type strain grown in M9 and in M9+sorbitol. The data for the control analysis is presented in Appendix 5, and those few proteins identified in that analysis that were also identified in the analysis of the Lpp₊₂₁ strain are indicated and discussed.

A two-sample T-test was performed to identify the protein change in the datasets. Significant changes are indicated as black dots on the volcano plots (Figure 5.3.2.1.1 and 5.3.2.2.1). Volcano plots used for data visualisation were generated by plotting the student's T-test difference (the log₂ fold change) on the X-axis, against the -log₁₀ of the student's T-test p-value on the Y-axis. The log₂ fold changes were derived by calculating the difference between the LFQ intensities of two conditions, i.e. wild-type vs Lpp₊₂₁. The values were arranged in descending order where the negative values represent decreased steady-state proteins and the

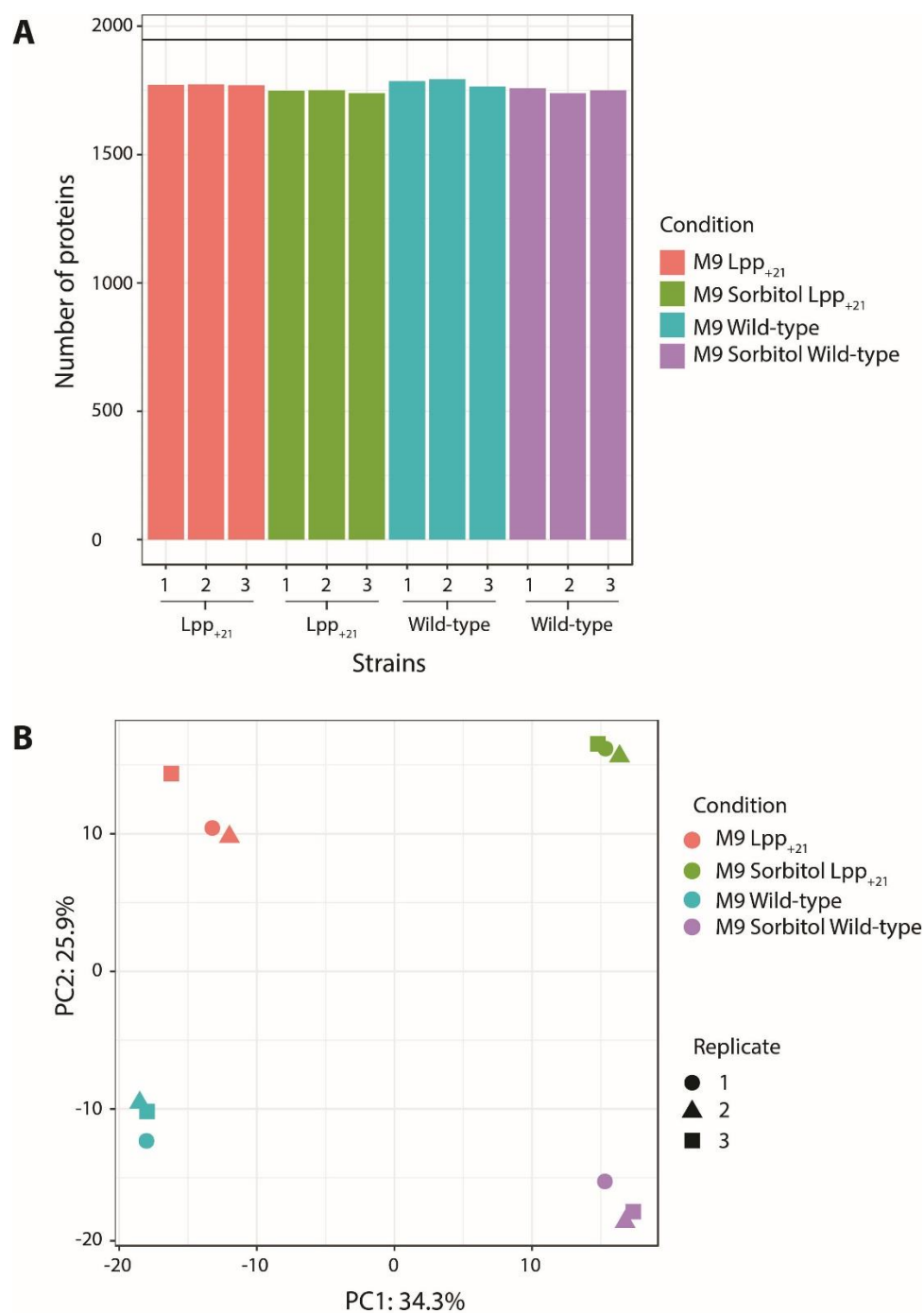


Figure 5.3.1.1: Data quality plots.

(A) Protein numbers plot. The figure shows the number of proteins identified in each sample. The strains, wild-type and Lpp₊₂₁ are in three biological replicates (n=3). (B) Principle component analysis where the three shapes represent biological replicates of the same sample. Different colours represent different conditions.

positive values represent increased steady-state proteins (Table 5.3.2.1.1 and 5.3.2.2.1). The student T-test P-value was used to check whether the observed change was significant, where the smaller the value, the more significant the change in protein steady-state level.

Volcano plots in Figures 5.3.2.1.1 and 5.3.2.2.1 document changes to the proteome in Lpp₊₂₁ vs wild-type and Lpp₊₂₁ (+sorbitol) vs wild-type (+sorbitol) datasets, respectively. In these plots, the left side displays the proteins decreased in steady-state level in Lpp₊₂₁ strain while the right side displays the proteins increased in steady-state level in Lpp₊₂₁ strain. Any proteins that were also subject to changes in the wild-type samples +/- sorbitol are labelled green in the Volcano plots.

5.3.2.1 Changes to the proteome in an enlarged periplasm in hypo-osmotic medium

The largest decrease in protein levels in the Lpp₊₂₁ strain was the level of Lpp₊₂₁ compared to Lpp, as documented in Figure 4.2.1C, and the diminished steady-state levels for the various proteins of the galactitol pathway. Transcriptome analysis (Table 5.2.1.2.1) documented the transcriptional down-regulation of the *gat* genes, and the proteomics analysis verifies that this is transformed into a down-regulation in protein steady-state for GatA, GatB, GatC, GatD and GatZ. As supported earlier (Section 5.2.1), the diminished amount of the Lpp isoform in the Lpp₊₂₁ strain appears to stimulate an up-regulation of the *lpp* gene, which is not fully effective in restoring Lpp₊₂₁ levels to anything similar to the Lpp level maintained in the wild-type strain (Figure 5.3.2.1.1).

A significant increase in protein levels in the Lpp₊₂₁ strain were the levels of well characterised Rsd and AmiA. Rsd is a regulator of sigma factor D (RpoD), the ‘housekeeping’ sigma factor in *E. coli* that keeps essential genes and pathway operating (Jishage & Ishihama, 1998; Gruber & Gross, 2003). Rsd binds to RpoD with a stoichiometry of 1:1 (Westblade *et al.*, 2004), and has been postulated to thereby displace the RNAP core enzyme, which becomes accessible for starvation/stationary phase (RpoS) sigma factor (Mitchell *et al.*, 2007). The role of Rsd in modulating the function of these sigma factors, directly or indirectly, is a statement of the need to adapt to the stress induced by the enlarged periplasm, which would likely mask any phenotype as the cell adapts. AmiA is well-characterised PG hydrolase localised in the periplasm. Its increased steady-state levels would have an effect of increased remodelling activity of the PG, liberating more tripeptides into the periplasm (Heidrich *et al.*, 2001; Priyadarshini *et al.*, 2007). Its identification correlates with results from Lpp₊₂₁ synthetic lethal

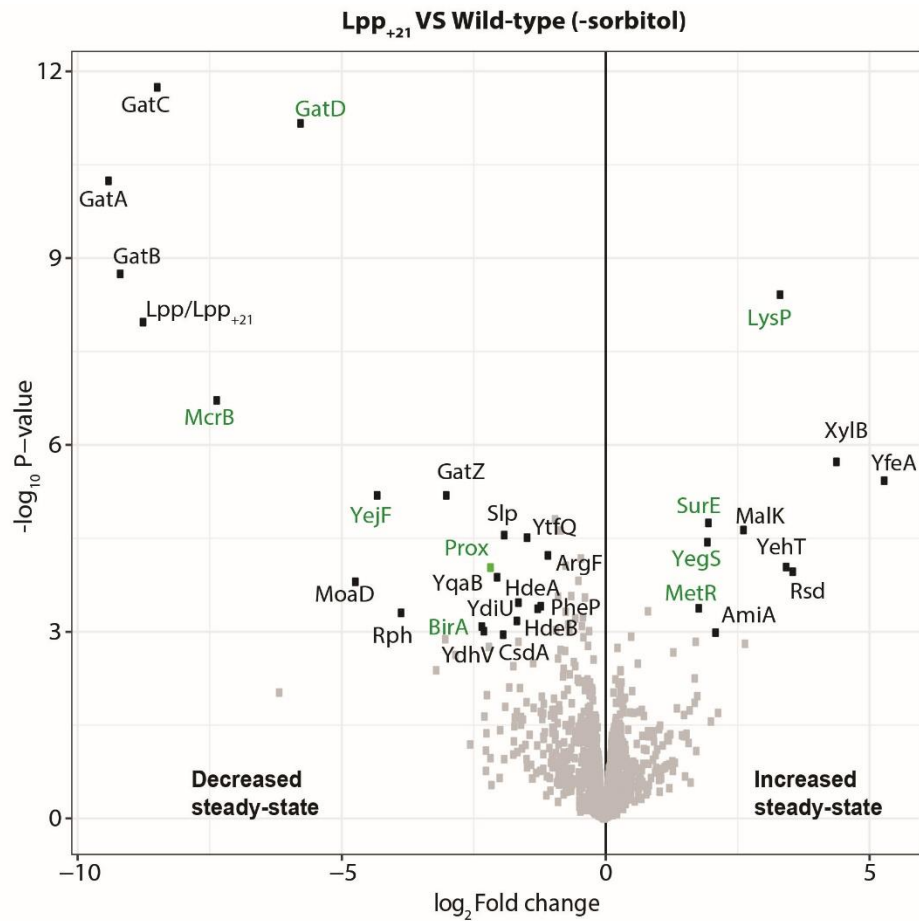


Figure 5.3.2.1.1: Changes to proteome in an enlarged periplasm; Lpp₊₂₁ vs wild-type in M9 minimal media.

The volcano plot was generated by plotting the student's T-test difference (the log₂ fold change), on the X-axis against the $-\log_{10}$ of the student's T-test p-value on the Y-axis. The log₂ fold changes were derived by calculating the difference between the LFQ intensities of the two conditions. Proteins that exhibited significant changes in steady-state levels are located in the left and right upper quadrants (labelled). The left side displays the decreased steady-state proteins in Lpp₊₂₁ strain while the right side displays the increased steady-state proteins in Lpp₊₂₁ strain.

screen, where non-essential genes involved in PG biogenesis became essential in *Lpp*₊₂₁ background.

Despite displaying significant changes in steady-state levels, LysP, SurE, YegS, MetR, McrB, YegF, ProX and BirA (labelled green, Figure 5.3.2.1.1) were also identified in the control analysis (wild-type samples +/- sorbitol) (Appendix 5) and are therefore likely to be a result of osmotic adaptation.

Table 5.3.2.1.1: Over-expressed and diminished proteins in M9 minimal media; *Lpp*₊₂₁ vs wild-type.

Proteins were sorted based on the log₂ fold change (FC) in descending order. Negative values represent diminished/ down-expressed proteins, while positive values depict over-expressed proteins. The volcano plots in the figure above were generated from the log₂ fold change, and the P-values included in the tables.

Gene Name	Protein IDs	log ₂ FC	P-value	P-value adjusted	Protein names
					Decreased steady-state
<i>gatA</i>	P69828	-9.2	1.86E-08	9.05E-06	Galactitol-specific phosphotransferase enzyme IIA component
<i>Lpp</i> ₊₂₁	P69776	-8.86	3.13E-09	2.40E-06	Major outer membrane lipoprotein Lpp
<i>gatB</i>	P37188	-8.33	1.28E-09	2.40E-06	Galactitol-specific phosphotransferase enzyme IIB component
<i>gatC</i>	P69831	-8.14	3.69E-09	2.40E-06	Galactitol permease IIC component
<i>mcrB</i>	P15005	-6.52	4.20E-07	0.000102	5-methylcytosine-specific restriction enzyme B
<i>gatD</i>	P0A9S3	-5.9	8.25E-08	2.68E-05	Galactitol-1-phosphate 5-dehydrogenase
<i>yefF</i>	P33916	-4.59	1.03E-05	0.00144	Putative oligopeptide ABC transporter ATP binding subunit YefF
<i>moaD</i>	P30748	-4.12	0.000788	0.0365	Molybdopterin synthase sulfur carrier subunit
<i>lsrB</i>	P76142	-3.26	1.06E-06	0.000206	Autoinducer 2-binding protein LsrB
<i>gatZ</i>	P0C8J8	-2.99	6.67E-06	0.000999	D-tagatose-1,6-bisphosphate aldolase subunit GatZ
<i>birA</i>	P06709	-2.35	0.00113	0.048	Bifunctional ligase/repressor BirA
<i>proX</i>	P0AFM2	-2.2	0.000171	0.0128	Glycine betaine-binding periplasmic protein
<i>slp</i>	P37194	-1.9	2.90E-05	0.00314	Starvation lipoprotein
<i>ftnB</i>	P0A9A2	-1.83	0.000774	0.0365	Bacterial non-heme ferritin-like protein
<i>ydiU</i>	P77649	-1.72	9.91E-05	0.00805	UPF0061 protein YdiU
<i>hdeA</i>	P0AES9	-1.63	0.000351	0.0214	Acid stress chaperone HdeA

<i>ychJ</i>	P0AB03	-1.46	0.000609	0.0311	Uncharacterized protein YcbJ
<i>ytfQ</i>	P39325	-1.46	3.16E-05	0.00324	Galactofuranose ABC transporter periplasmic binding protein
					Increased steady-state
<i>yfeA</i>	P23842	6.35	3.92E-07	0.000102	Uncharacterized protein YfeA
<i>yehT</i>	P0AFT5	3.58	2.53E-05	0.0029	Transcriptional regulatory protein YehT
<i>xylB</i>	P09099	3.24	0.000516	0.0287	Xylulose kinase
<i>rsd</i>	P0AFX4	3.15	0.000567	0.0307	Regulator of sigma D
<i>lysP</i>	P25737	2.5	7.82E-07	0.000169	Lysine-specific permease
<i>surE</i>	P0A840	2.08	5.46E-06	0.000966	5/3-nucleotidase SurE
<i>amiA</i>	P36548	1.95	4.79E-08	1.87E-05	N-acetylmuramoyl-L-alanine amidase AmiA
<i>yegS</i>	P76407	1.69	0.000336	0.0211	Lipid kinase YegS
<i>metR</i>	P0A9F9	1.6	4.67E-05	0.00455	HTH-type transcriptional regulator MetR

5.3.2.2 Changes to the proteome in an enlarged periplasm in iso-osmotic medium

In the presence of sorbitol, providing osmotic support, the largest decrease in protein levels in the Lpp₊₂₁ strain was again the level of Lpp₊₂₁ compared to Lpp, and the diminished steady-state levels for the various proteins of the galactitol pathway: GatA, GatB, GatC and GatD. These factors were also seen in the absence of sorbitol (Figure 5.3.2.1.1). Two components of the sorbitol transport pathway (SrlE and SrlD) was also decreased, reflecting the decreased transcriptional activation of *srlE*, *srlD* and other components of this pathway seen in the RNA-seq analysis (Table 5.2.1.2.1). The two proteins are reduced in steady-state only in the sorbitol condition, as seen in the RNA-seq analysis, a further testament of osmotic differences between the two strains.

In addition to AmiA identified in M9 (-sorbitol) condition, AmiC, another PG hydrolyse was increased in steady state-levels. AmiC interacts directly with NlpD cleaves the peptide moiety from the MurNac residue, removing murein crosslinks which would have an effect of increased remodelling activity of the PG, in turn, liberating more tripeptides into the periplasm (Uehara *et al.*, 2010; Rocaboy *et al.*, 2013). Consistent with these results was the observation that subunits of the oligopeptide transporter system (OppA, OppB, OppC, OppD and OppF) were all present at relatively high levels in Lpp₊₂₁ strain (Figure 5.3.2.2.1 and Table 5.3.2.2.1). This transporter is highly active in sequestering peptides liberated during PG remodelling into the periplasm (Park & Uehara, 2008). Nascent AmiA and AmiC are known to be translocated into the periplasm via the TAT complex (Bernhardt & de Boer, 2003), and TatA was also found at

an increased steady-state level in the Lpp₊₂₁ strain, possibly to accommodate the increase in its number of substrates. However, the steady-state levels of TatB and TatC, the other components of the TAT complex remained unchanged. AmiC is even more interesting because it interacts directly with NlpD (an OM lipoprotein) for its activation (Uehara *et al.*, 2010). NlpD displayed synthetic lethal phenotype in an Lpp₊₂₁ background (Table 4.4.2.1).

The analysis also identified three other related proteins (FeoA, FeoB and FhuE) all increased in steady-state levels in the Lpp₊₂₁ strain. FeoA is part of the conserved ferrous iron transport system, shown to form a complex with the FeoB transmembrane transporter component which catalyses slow GTP hydrolysis and a fast GDP release required for Fe²⁺ uptake (Marlovits *et al.*, 2002; Lau *et al.*, 2013). FhuE, localised in the OM, serves as a receptor for ferric-coprogen and ferric rhodotorulic acid. The roles of these proteins suggest Lpp₊₂₁ strain upregulates iron import. Common among these three proteins is TonB, required for energy transduction from the IM to the OM for uptake of substrates (Bitter *et al.*, 1994). It is likely that enlargement of the periplasm is a functional constraint to TonB (discussed further in Section 5.5).

DsbA, a periplasmic thiol: disulphide oxidoreductase, that promotes the formation of disulphide bonds was diminished in steady-state levels in the Lpp₊₂₁ strain. Phenotypic assessment of the Lpp₊₂₁ strain (Section 4.3.3) in the presence of DTT, revealed increased sensitivity of the Lpp₊₂₁ strain towards the reductant, relative to the wild-type (Figure 4.3.3.1). Since DsbA is a primary folding catalyst in the periplasm, the sensitivity exhibited by Lpp₊₂₁ could likely be as a result of impaired disulphide bond-forming system. The basis for DsbA reduced steady-state levels as an adaptation to an enlarged periplasm is unknown but intriguing and worth further investigation.

Two crucial proteins in LPS biogenesis were identified, LptA and LpxB, both diminished in steady-state levels in Lpp₊₂₁ strain. Both of the proteins are required for cell viability. LptA forms the periplasmic component of the LPS transport system, which functions to bridge the IM components (LptB₂FGC) and OM components (LptDE) (Figure 1.5.3.1). LpxB (Lipid A disaccharide synthase) catalyses the formation of lipid A disaccharide, an intermediate in the lipid IV_A biosynthesis pathway (Radika & Raetz, 1988). The β,1'-6 linked disaccharide generated by LpxB is characteristic of all lipid A molecules owing to its essentiality in *E. coli* (Ray *et al.*, 1984). The two proteins are likely linked to SDS sensitivity phenotype observed in Lpp₊₂₁ strain (Figure 4.3.2.1) based on their crucial role in maintaining envelope integrity.

On further analysis of the results from the two conditions, it was realised that all the identified lipoproteins displayed diminished steady-state levels in the *Lpp₊₂₁* strain. These included: Slp, BamE, YgeR, and YgdI. Slp (starvation lipoprotein) is tethered to the OM (Stenberg *et al.*, 2005) and was shown to take part in acid resistance as its expression increased when cells were grown at pH 5.5 and 4.5 compared to pH 7.4 (Tucker *et al.*, 2002). BamE is an accessory lipoprotein of the BAM complex, shown to function with other components of the BAM in the assembly of OMPs (Sklar *et al.*, 2007a; Kim *et al.*, 2011b; Knowles *et al.*, 2011). YgeR, LysM domain-containing putative peptidase, is a verified lipoprotein (Juncker *et al.*, 2003), that upon deletion exacerbates the cell division phenotype of an *envC nlpD* double deletion mutant (Uehara *et al.*, 2009), hence its implication in cell division. YgdI is a verified lipoprotein of unknown function (Juncker *et al.*, 2003). The diminished steady-state levels of these lipoproteins are likely due to a pathway effect rather than a functional effect as they share the same biogenesis pathway with *Lpp₊₂₁*, which I hypothesise may have an indirect negative effect on their biogenesis (discussed further in Section 5.5).

Table 5.3.2.2.1: Over-expressed and diminished proteins in M9 minimal media supplemented with 0.5 M sorbitol; *Lpp₊₂₁* vs wild-type.

Proteins were sorted based on the log₂ fold change (FC) in descending order. Negative values represent diminished/down-expressed proteins, while positive values depict over-expressed proteins. The volcano plots in the figure above were generated from the log₂ fold change, and the P-values included in the tables.

Gene Name	Protein IDs	log ₂ FC	P-value	P-value adjusted	Protein names
					Decreased steady-state
<i>Lpp₊₂₁</i>	P69776	-9.38	1.83E-09	1.78E-06	Major outer membrane lipoprotein Lpp
<i>glpQ</i>	P09394	-7.43	0.00299	0.0247	Glycerophosphoryl diester phosphodiesterase
<i>gatA</i>	P69828	-7.33	1.50E-07	3.24E-05	Galactitol-specific phosphotransferase enzyme IIA component
<i>gatB</i>	P37188	-7.27	4.51E-09	2.93E-06	Galactitol-specific phosphotransferase enzyme IIB component
<i>gatC</i>	P69831	-7.02	1.46E-08	7.12E-06	Galactitol permease IIC component
<i>ygdI</i>	P65292	-5.43	3.30E-06	0.000215	Uncharacterized lipoprotein YgdI
<i>bamE</i>	P0A937	-4.51	1.34E-07	3.24E-05	OM protein assembly factor BamE
<i>gatD</i>	P0A9S3	-4.07	2.29E-06	0.000178	Galactitol-1-phosphate 5-dehydrogenase

<i>srlE</i>	P56580	-3.58	0.000139	0.00361	Glucitol/sorbitol-specific phosphotransferase enzyme IIB component
<i>lpxB</i>	P10441	-3.47	0.000797	0.0115	Lipid-A-disaccharide synthase
<i>dcuA</i>	P0ABN5	-3.22	2.04E-05	0.000811	Anaerobic C4-dicarboxylate transporter DcuA
<i>ygeR</i>	Q46798	-3.03	0.000319	0.00634	Uncharacterized lipoprotein YgeR
<i>sbcD</i>	P0AG76	-2.82	3.92E-06	0.000231	Nuclease SbcCD subunit D
<i>aroP</i>	P15993	-2.57	0.00558	0.038	Aromatic amino acid transport protein AroP
<i>dnaG</i>	P0ABS5	-2.55	3.53E-05	0.00124	DNA primase
<i>pabB</i>	P05041	-2.39	0.000738	0.0112	Aminodeoxychorismate synthase component 1
<i>yceA</i>	P24188	-2.28	0.000493	0.00842	UPF0176 protein YceA
<i>ubiJ</i>	P0ADP7	-2.23	0.00336	0.0267	Ubiquinone biosynthesis protein UbiJ
<i>yfaZ</i>	P76471	-2.03	0.00344	0.027	Uncharacterized protein YfaZ
<i>rluA</i>	P0AA37	-2	0.00767	0.0475	Ribosomal large subunit pseudouridine synthase A
<i>rseP</i>	P0AEH1	-1.98	0.00711	0.0457	Regulator of sigma-E protease RseP
<i>yhdP</i>	P46474	-1.92	0.00265	0.0234	Uncharacterized protein YhdP
<i>ydeN</i>	P77318	-1.89	0.000486	0.00838	Uncharacterized sulfatase YdeN
<i>yjiN</i>	P39385	-1.88	0.00385	0.0287	Uncharacterized protein YjiN
<i>mutM</i>	P05523	-1.82	0.00106	0.0133	Formamidopyrimidine-DNA glycosylase
<i>ybdR</i>	P77316	-1.82	0.000256	0.00546	Uncharacterized zinc-type alcohol dehydrogenase-like protein YbdR
<i>mglB</i>	P0AEE5	-1.77	3.20E-07	4.15E-05	D-galactose-binding periplasmic protein
<i>osmY</i>	P0AFH8	-1.75	2.93E-05	0.0011	Osmotically-inducible protein Y
<i>ydcS</i>	P76108	-1.73	9.75E-07	1.00E-04	Putative ABC transporter periplasmic-binding protein YdcS
<i>yjbB</i>	P0AF43	-1.7	0.000511	0.00866	Uncharacterized protein YjbB
<i>yahO</i>	P75694	-1.69	4.90E-06	0.000281	Uncharacterized protein YahO
<i>rluE</i>	P75966	-1.64	0.00455	0.0328	Ribosomal large subunit pseudouridine synthase E
<i>xerC</i>	P0A8P6	-1.63	0.00154	0.0161	Tyrosine recombinase XerC
<i>iscR</i>	P0AGK8	-1.58	0.00269	0.0235	HTH-type transcriptional regulator IscR
<i>srlD</i>	P05707	-1.56	0.00103	0.0132	Sorbitol-6-phosphate 2-dehydrogenase
<i>ytfQ</i>	P39325	-1.54	2.04E-05	0.000811	ABC transporter periplasmic-binding protein YtfQ
<i>Ivy</i>	P0AD59	-1.49	1.38E-06	0.000128	Inhibitor of vertebrate lysozyme
<i>ubiA</i>	P0AGK1	-1.47	0.000241	0.00527	4-hydroxybenzoate octaprenyltransferase
<i>hdeB</i>	P0AET2	-1.46	0.000148	0.00379	Acid stress chaperone HdeB
<i>rbsB</i>	P02925	-1.43	2.81E-07	3.93E-05	D-ribose-binding periplasmic protein
<i>ydeI</i>	P31130	-1.43	1.41E-05	0.000669	Uncharacterized protein YdeI
<i>lptA</i>	P0ADV1	-1.41	0.00136	0.0154	Lipopolysaccharide export system protein LptA
<i>lsrB</i>	P76142	-1.37	0.00113	0.0137	Autoinducer 2-binding protein LsrB

<i>srlB</i>	P05706	-1.36	0.00204	0.0192	Glucitol/sorbitol-specific phosphotransferase enzyme IIA component
<i>metR</i>	P0A9F9	-1.34	0.000186	0.00432	HTH-type transcriptional regulator MetR
<i>dsbA</i>	P0AEG4	-1.22	1.38E-07	3.24E-05	Thiol:disulphide interchange protein DsbA
					Increased steady-state
<i>cspG</i>	P0A978	6.12	1.47E-05	0.000683	Cold shock-like protein CspG
<i>garL</i>	P23522	5.69	9.26E-10	1.78E-06	5-keto-4-deoxy-D-glucarate aldolase
<i>nsrR</i>	P0AF63	4.56	0.000118	0.00319	HTH-type transcriptional repressor NsrR
<i>dmsB</i>	P18776	4.37	0.000586	0.00982	Anaerobic dimethyl sulfoxide reductase chain B
<i>amiC</i>	P63883	4.1	9.83E-05	0.00282	N-acetylmuramoyl-L-alanine amidase AmiC
<i>ynfF</i>	P77783	3.74	0.000454	0.00808	Probable dimethyl sulfoxide reductase chain YnfF
<i>bioD2</i>	P0A6E9	3.63	0.00295	0.0247	ATP-dependent dethiobiotin synthetase BioD 2
<i>zraP</i>	P0AAA9	3.54	5.94E-05	0.0019	Zinc resistance-associated protein
<i>hypB</i>	P0AAN3	3.42	0.00559	0.038	Hydrogenase isoenzymes nickel incorporation protein HypB
<i>hiuH</i>	P76341	3.39	0.00201	0.0191	5-hydroxyisourate hydrolase
<i>yciI</i>	P0AB55	3.35	1.82E-07	3.55E-05	Protein YciI
<i>yhbT</i>	P64599	3.14	0.00126	0.0148	Uncharacterized protein YhbT
<i>feoA</i>	P0AEL3	3.09	0.000744	0.0112	Ferrous iron transport protein A
<i>elaA</i>	P0AEH3	3.03	0.00266	0.0234	Protein ElaA
<i>yjiI</i>	P37342	2.93	0.00326	0.0262	Uncharacterized protein YjiI
<i>birA</i>	P06709	2.78	0.000344	0.00658	Bifunctional ligase/repressor BirA
<i>ydfK</i>	P76154	2.7	4.75E-05	0.00154	Uncharacterized protein YdfK;Uncharacterized protein YnaE
<i>yehT</i>	P0AFT5	2.33	0.00069	0.0107	Transcriptional regulatory protein YehT
<i>gldA</i>	P0A9S5	2.28	0.000869	0.0121	Glycerol dehydrogenase
<i>feoB</i>	P33650	2.26	0.00159	0.0164	Ferrous iron transport protein B
<i>hypE</i>	P24193	2.09	1.69E-05	0.000766	Hydrogenase isoenzymes formation protein HypE
<i>oppD</i>	P76027	2.08	0.000636	0.0103	Oligopeptide transport ATP-binding protein OppD
<i>yfcG</i>	P77526	2.04	0.000691	0.0107	Disulfide-bond oxidoreductase YfcG
<i>glnK</i>	P0AC55	1.99	0.000793	0.0115	Nitrogen regulatory protein P-II 2
<i>phoB</i>	P0AFJ5	1.98	4.56E-08	1.78E-05	Phosphate regulon transcriptional regulatory protein PhoB
<i>oppF</i>	P77737	1.92	0.00134	0.0153	Oligopeptide transport ATP-binding protein OppF
<i>oppC</i>	P0AFH6	1.91	0.000643	0.0103	Oligopeptide transport system permease protein OppC
<i>sdaA</i>	P16095	1.84	0.000105	0.00289	L-serine dehydratase 1
<i>fhuE</i>	P16869	1.82	0.000243	0.00527	FhuE receptor
<i>frdB</i>	P0AC47	1.76	0.00152	0.0161	Fumarate reductase iron-sulfur subunit

<i>asnA</i>	P00963	1.7	7.90E-05	0.00237	Aspartate--ammonia ligase
<i>cspA</i>	P0A9X9	1.7	0.00383	0.0287	Cold shock protein CspA
<i>dgoR</i>	P31460	1.7	3.55E-05	0.00124	Galactonate operon transcriptional repressor
<i>frdA</i>	P00363	1.69	0.000779	0.0114	Fumarate reductase flavoprotein subunit
<i>tatA</i>	P69428	1.65	0.000323	0.00635	Sec-independent protein translocase protein TatA
<i>frdC</i>	P0A8Q0	1.62	0.000747	0.0112	Fumarate reductase subunit C
<i>oppB</i>	P0AFH2	1.62	0.00158	0.0164	Oligopeptide transport system permease protein OppB
<i>ydaM</i>	P77302	1.61	0.00081	0.0116	Probable diguanylate cyclase YdaM
<i>pepE</i>	P0A7C6	1.57	0.00305	0.025	Peptidase E
<i>pflB</i>	P09373	1.55	0.00443	0.0321	Formate acetyltransferase 1
<i>pspE</i>	P23857	1.54	0.0042	0.0308	Thiosulfate sulfurtransferase PspE
<i>cspI</i>	P0A986	1.52	0.000179	0.00429	Cold shock-like protein CspI
<i>rihC</i>	P22564	1.5	0.00251	0.0227	Non-specific ribonucleoside hydrolase RihC
<i>oppA</i>	P23843	1.49	0.00775	0.0476	Periplasmic oligopeptide-binding protein
<i>nrdD</i>	P28903	1.47	0.000203	0.0046	Anaerobic ribonucleoside-triphosphate reductase
<i>fnbB</i>	P0A9A2	1.46	0.00343	0.027	Bacterial non-heme ferritin-like protein
<i>ycbJ</i>	P0AB03	1.44	0.000647	0.0103	Uncharacterized protein YcbJ
<i>edd</i>	P0ADF6	1.4	0.00115	0.0138	Phosphogluconate dehydratase
<i>rmf</i>	P0AFW2	1.35	1.84E-05	0.000766	Ribosome modulation factor
<i>rraB</i>	P0AF90	1.32	0.000159	0.00403	Regulator of ribonuclease activity B

Analysis of the control for sorbitol (wild-type +/- sorbitol) results identified the three major proteins involved in response to osmotic stress, ProX, ProW and ProV (Gowrishankar, 1989; Gul & Poolman, 2013) (Appendix 5). The proteins are subunits of an osmoresponsive ProU ABC transport system involved in the import of glycine betaine and proline as a mechanism to cope with an osmotic upshift (Gul & Poolman, 2013). The three proteins displayed significant increase in protein steady-state levels in wild-type (+sorbitol) strain. In addition, ProP, OsmY, and OsmX, also involved in osmoregulation, were also present in increased steady steady-levels in wild-type (+sorbitol) condition. The results were significant, acting as internal controls, i.e. the wild-type (+sorbitol) condition was expected to exhibit changes that indicate osmoregulation, thus, a validation for the rest of the proteomics data extracted from the other conditions.

In addition to the proteins listed on the tables, extra proteins were identified by manual comparison of LFQ intensities of each protein, also termed as the unimputed matrix. These proteins were not identified as hits from the student's t-test statistical analysis as they were missing in one of the samples. For this reason, a relative quantification could not be made. Among these proteins, two were of interest, LdtC and MraY. Both were increased in steady-state levels in Lpp₊₂₁ strain. LdtC is an L, D-transpeptidase that covalently anchors Lpp to the PG (Magnet *et al.*, 2007). The enzyme is redundant to two other enzymes, LdtA and LdtB. Its identification is likely related to diminished Lpp₊₂₁ levels as a compensatory mechanism. MraY increase implies further adaptation to impaired PG biogenesis process in the Lpp₊₂₁ strain. It catalyses the first step of the lipid cycle reaction in the biosynthetic pathway of PG (see section 1.3.1.2) and hence is essential for cell viability.

5.4 Rational design to investigate Lpp₊₂₁ synthetic lethality of over-expressed proteins

Proteomics results presented in the previous section show how Lpp₊₂₁ strain adapts to the increased periplasmic size through dynamic readjustment of a myriad of cellular processes, exhibited by the high number of proteins displaying significant changes in steady-state levels. Among them, the over-expressed proteins indicate a higher demand for the proteins in the Lpp₊₂₁ background. Based on this insight, I hypothesised that removal of some of these proteins in Lpp₊₂₁ background would likely result in a lethal phenotype.

A rational design experiment was set-up. Unlike the screen where all genes were analysed, in this case, few selected genes were chosen for Lpp₊₂₁ synthetic lethal analyses. The genes of interest coded for proteins increased in steady-state levels in the Lpp₊₂₁ strain. *lpp₊₂₁* was introduced into the single gene deletion (replacing *lpp*) strains using the same technique (Giant-colli) used in the Lpp₊₂₁ synthetic lethal screen. The two gene modifications of the strains shown in Table 5.4.1 were verified using colony PCR and growth was assessed in M9 minimal media.

As anticipated, most of the double recombinants exhibited synthetic lethality. Of immense interest, AmiA and AmiC were some of the most significant over-expressed proteins in Lpp₊₂₁. The two proteins are vital for PG turnover and recycling; a process likely up-regulated to compensate for suboptimal assembly of new PG. However, knock-out of *tatA* in Lpp₊₂₁ background did not result in loss of viability in M9 minimal media. TatA is a nonessential component of the TAT complex shown to translocate AmiA and AmiC across the IM.

Table 5.4.1: Synthetic lethality evaluation of over-expressed non-essential proteins in nutrient-limited media conditions (M9 Minimal media 0.2% (w/v) glucose).

Double recombinant strain	Protein function	Viability in M9 minimal media
<i>lpp₊₂₁ ΔamiA</i>	PG hydrolase – PG turnover and recycling	Synthetic lethal
<i>lpp₊₂₁ ΔamiC</i>	PG hydrolase – PG turnover and recycling	Synthetic lethal
<i>lpp₊₂₁ ΔldtC</i>	Linkage of Lpp to peptidoglycan	Viable
<i>lpp₊₂₁ ΔoppD</i>	Oligopeptide transporter subunit – PG remodelling	Synthetic lethal
<i>lpp₊₂₁ ΔoppF</i>	Oligopeptide transporter subunit – PG remodelling	Synthetic lethal
<i>lpp₊₂₁ ΔtatA</i>	Translocation of proteins across the IM (Component of the TAT complex)	Viable
<i>lpp₊₂₁ ΔfeoB</i>	Component of the ferrous iron uptake system	Synthetic lethal
<i>lpp₊₂₁ ΔfhuE</i>	Receptor for ferric-coprogen - Iron uptake system	Synthetic lethal

Other proteins investigated were the oligopeptide transporter subunits OppD and OppF, chosen due to significant fold changes in increased expression in *Lpp₊₂₁* strain compared to the other subunits. As anticipated, the two strains (*lpp₊₂₁ ΔoppD* and *lpp₊₂₁ ΔoppF*) lost viability in M9 (-sorbitol) despite being identified in M9 (+sorbitol). The result depicts how PG recycling is vital for cells viability in an enlarged periplasm and also shows the major contributor to the phenotype is the limited supply of nutrients where sorbitol exacerbates the condition.

Synthetic lethality was also observed when growth of *lpp₊₂₁ ΔfeoB* and *lpp₊₂₁ ΔfhuE* strains was assessed. FeoB and FhuE are increased in steady-state in the *Lpp₊₂₁* strain in the M9 (+sorbitol) condition. The two proteins play vital roles in iron import, indirectly involving TonB, whose periplasmic domain functionality is under periplasmic size constraint.

LdtC is one among three redundant L, D-transpeptidases involved in the attachment of Lpp to PG. The fact that mutants with all three L, D-transpeptidases deleted (*ΔldtAΔldtBΔldtC*) or *lpp_{ΔK58}* that does not attach to PG, remain viable under standard laboratory conditions (Magnet

et al., 2008; Asmar *et al.*, 2017), the viability of *lpp₊₂₁ ΔldtC* in M9 minimal media was not surprising despite its over-expression. Overall, the rational design study was significant in justification of the proteome changes taking place in *Lpp₊₂₁* strain. However, the genes were not identified from the *Lpp₊₂₁* synthetic lethal screen; a possible reason could have been that they were among the false-negatives, which is not unusual when screening a huge library of mutants.

5.5 Discussion

The periplasm of Gram-negatives harbours a myriad of processes as an adaptation to the challenges introduced by the evolution of the bilayer. In this chapter, I investigated the adaptive response of *E. coli* to an enlarged periplasm. The response was examined at the transcriptional level by RNA-seq analysis and the protein level by mass spectrometry-based proteomic profiling. The study was based on *Lpp₊₂₁* phenotypic characterisations, where the strain exhibited just a few distinctive phenotypes compared to wild-type, and growth profiles were identical in all conditions tested (section 4.3). Using the genes and proteins identified in both hypo-osmotic and iso-osmotic (osmotic stabilised) conditions, I address the study question on how the cell adapts to the changes in periplasmic size. It is also worth mentioning that the use of RNA-seq and mass spectrometry-based proteomic profiling was advantageous in that essential genes/proteins were investigated unlike in *Lpp₊₂₁* synthetic lethal screen.

A general observation noted was that the number of proteins identified as changing in steady-state level in the *Lpp₊₂₁* cells was more than the number of genes displaying changes in transcript levels. The variation is more pronounced in osmotically stabilised medium (+sorbitol) than hypo-osmotic medium (-sorbitol), an indicator of increased osmotic stress in the periplasm. In other words, this data implies a mild response to an enlarged periplasm at the transcriptional level, but a strong response at the proteome level. The variation observed between the transcriptional and the post-transcriptional level is reasonable in the sense that proteins steady-state levels depends on factors controlling the rate of protein synthesis and rate of protein degradation. Despite the variance at the transcriptional and the proteome level, some consistency was observed where up-regulated genes, led to more proteins levels and vice versa.

Most notable was the change at both gene and protein levels for *Lpp₊₂₁*. This protein was significantly decreased in steady-state levels which was also documented in western blot analysis (Figure 4.2.1C). The justification for these results is probably due to a compensatory mechanism. The *Lpp₊₂₁* isoform, is likely more prone to protease degradation, due to its

elongated artificial nature. The diminished amount of the isoform in Lpp₊₂₁ strain appears to stimulate an up-regulation of the *lpp* gene (Table 5.2.1.1.1), which is not fully effective in restoring Lpp₊₂₁ levels to anything similar to the Lpp level maintained in the wild-type strain. The likely adaptation by the Lpp₊₂₁ strain to bridge this gap was to increase the steady-state levels of LdtC. LdtC covalently binds Lpp to PG (Magnet *et al.*, 2007) and by having more of LdtC, the cell ensures the less available Lpp₊₂₁ is bound to PG. LdtC over-expression further signifies the importance of OM-PG tethering by Lpp.

Galactitol metabolism genes (*gatA*, *gatB*, *gatC*, *gatD*), were not only significantly down-regulated at the transcriptional level, but also the proteins they code for (GatA, GatB, GatC and GatD) exhibited diminished steady-state levels in Lpp₊₂₁ strain, implying that Lpp₊₂₁ strain is likely using alternative carbon sources and does not need galactitol. It is not known whether sorbitol turns off galactitol use (competitively), but in case it does, it does not justify why the genes were also down-regulated in M9 (-sorbitol) condition. For these reasons, I could not directly or indirectly link galactitol use as a response to an enlarged periplasm, complicated further by the fact that galactitol was not supplemented in either condition. This question remains open for further investigation.

In Chapter 4, results on Lpp₊₂₁ synthetic lethal screen identified genes involved in PG biosynthesis, maturation, turnover, and remodelling an indication that these processes rely on strict control of the periplasmic size. Consistent with those results, more proteins were identified from the proteomics study pointing to this essential cellular process. Among the identified were AmiA and AmiC, whose steady-state levels were increased in Lpp₊₂₁ strain. Both proteins are PG hydrolases active in PG remodelling, releasing liberating more tripeptides in the periplasm (Heidrich *et al.*, 2001; Priyadarshini *et al.*, 2007). Consistent with increased levels of oligopeptides in the periplasm, the subunits of the oligopeptide transporter system (OppA, OppB, OppC, OppD and OppF) were all increased in steady-state levels in Lpp₊₂₁. This system is highly active in sequestering peptides liberated from PG hydrolysis (Park & Uehara, 2008). As further proof of this rationale, TatA was also increased in steady-state levels. AmiA and AmiC are both translocated to the periplasm via the TAT system (Bernhardt & de Boer, 2003) and increased TatA levels likely improve the translocation efficiency of the TAT complex. TatA forms the translocation pore of the TAT complex (Rodriguez *et al.*, 2013) but it can interact with substrate molecules independent of the TatBC components, which enhances Tat transport (Taubert *et al.*, 2015). This is also a possible reason as to why the other components did not display changes in steady-state levels. A further adaptation in PG

biosynthesis process was the increase in steady-state levels of MraY. This protein catalyses the first step in lipid cycle reactions in the biosynthetic pathway of PG, which would increase the overall rate of the process. These adaptations are summarised as a cartoon model shown in Figure 5.5.1.

The decrease in steady-state levels of two essential proteins, LptA and LpxB, both involved in LPS biogenesis was consistent with previous results: The increased sensitivity of Lpp₊₂₁ to SDS and the identification of genes, involved in LPS biogenesis, that were synthetic lethal to Lpp₊₂₁. In case of LptA, the increased OM-IM distance may impact negatively on the assembly of LptA bridge structure that it forms, with the unassembled LptA being subject to proteolysis. In the case of LpxB, as a crucial enzyme in Lipid A biosynthesis, its inhibition would decrease the overall rate of LPS biosynthesis consequently avoiding the accumulation of LPS along its transport route, which in this case is likely defective due to periplasmic size constraints imposed on the trans-periplasmic Lpt bridge. By doing so, the cell ensures a balance on the number of LPS molecules that are made available for export via the suboptimal Lpt export system (Figure 5.5.1).

Another system that adapted to the enlarged periplasm was the iron import system. This was based on the identification of FeoA, FeoB and FhuE all increased in steady-levels in Lpp₊₂₁ strain. Uptake of iron is essential and depends on efficient interactions between the TonB complex in the IM with the TonB-dependent transporters in the OM (Pawelek *et al.*, 2006). This interaction is mediated by an elongated periplasmic domain of TonB that extends from the IM to the OM. This structural feature of TonB subjects its function to periplasmic size constraints, and it is reasonable that these proteins are increased in steady-state levels as a compensatory mechanism for impaired TonB function.

Another group of proteins worth mentioning are the proteins of unknown function, known or predicted to localise in the OM or the periplasm (YfaZ, YhdP, YdeN, YahO, YdeI, YgdI, YgeR, and others listed in Table 5.3.2.1.1 and Table 5.3.2.2.1. These proteins were identified in either increased or diminished levels exhibiting significant fold changes. The proteins, though of unknown functions, could play significant roles in the adaption of Lpp₊₂₁ to an enlarged periplasm and are subject to further investigations.

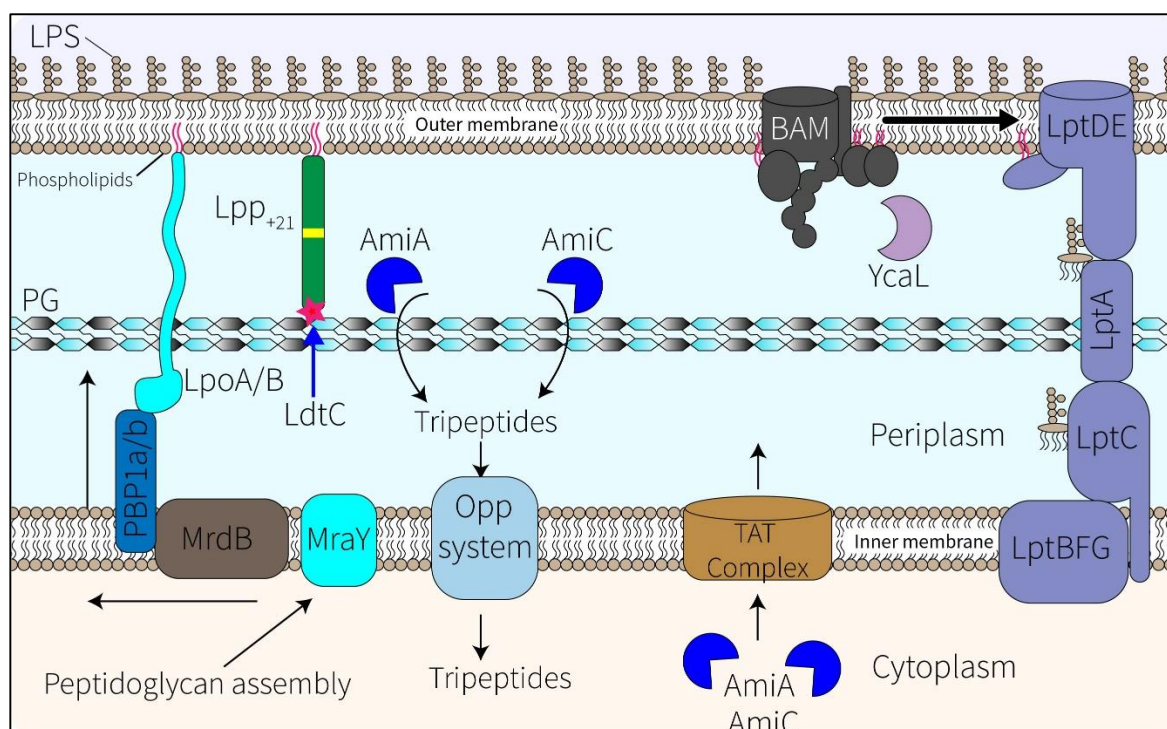


Figure 5.5.1: Proteostasis in an enlarged periplasm; quantitative proteomics summary.

The figure depicts some of the significant periplasmic modifications identified from the relative quantitative proteomics study as an adaptation to an enlarged periplasm. The adaptation is seen in the steady-state levels of Lpp₊₂₁, which were relatively less compared to the wild-type Lpp levels. An increase in the steady-state level of LdtC is consistent with the need to promote OM-PG linkage in response to diminished availability of Lpp₊₂₁. The cell adapts to impaired PG assembly (established by Lpp₊₂₁ synthetic lethal screen) by over-expression of MraY, involved in the first in lipid cycle reactions in the in PG biosynthesis and MrdB another key enzyme in PG biosynthesis pathway. PG hydrolyses, AmiA and AmiC are over-expressed relative to the wild-type, which would increase PG remodelling and recycling, liberating more tripeptides in the periplasm. Consistent with increased levels of oligopeptides in the periplasm, the subunits of the oligopeptide transporter system (OppA, OppB, OppC, OppD and OppF) were all in increased steady-state levels in Lpp₊₂₁ strain. This system is highly active in transporting peptides liberated from PG hydrolysis. AmiA and AmiC are known to be translocated into the periplasm via TAT complex. TatA, a component of TAT complex, was over-expressed in Lpp₊₂₁ strain. LptA, a crucial protein in the formation of the Lpt trans-periplasmic bridge for export of LPS and LpxB crucial enzyme in Lipid A biosynthesis were diminished in steady-state levels implying a suboptimal LPS biosynthesis process. Diminished LptA is likely due to proteolysis when it fails to assemble into the bridge structure it forms.

The adaptation to an enlarged periplasm may not have only been via protein functional aspect but rather a pathway effect. This notion is based on the steady-state levels of lipoproteins identified, e.g. Slp, BamE, YgdI, YgeR. The lipoproteins were all diminished in steady-state levels in Lpp₊₂₁ strain. Further analysis of the proteomics results revealed that no lipoprotein exhibited increased steady-state levels. The results point to a deficient lipoprotein biosynthesis pathway. Knowing that Lpp is the most abundant protein in the cell (Li *et al.*, 2014) and that there is an increased demand for the isoform, a reasonable explanation for the diminished levels of these lipoproteins is an overwhelmed lipoprotein biosynthesis pathway. To meet the demand for Lpp₊₂₁, the cell up-regulates *lpp₊₂₁* and dedicates the lipoprotein biosynthetic machinery to biosynthesis of Lpp₊₂₁, negatively impacting on the biosynthesis of other lipoproteins.

In conclusion, transcriptomic and proteomic profiling studies on Lpp₊₂₁ have shed some light on how a Gram-negative cell adapts to an enlarged periplasm. Based on the number of proteins identified compared to the genes, the response is mainly at the protein level and is dynamic involving the whole cell and not confined in the periplasm. In this study, only a few proteins were chosen for further analysis. However, there are many other mechanisms at play that can further be studied, i.e. the number of unknown periplasmic proteins identified is compelling.

In summary, the results show that the size of the periplasm is significant for normal functioning of the cell and that periplasmic processes have been evolutionary co-optimised for optimal function on a precise periplasmic size. On the other hand, the cell has evolved mechanisms to cope with periplasmic size changes to maintain optimal functionality of the periplasmic processes, some required for cell viability. The adaptation is crucial since Gram-negatives occupy different environmental niches, e.g. *E. coli*, which thrives in varying osmotic conditions, i.e. the gut and the laboratory media.

References

- Ahrens, C. H., Brunner, E., Qeli, E., Basler, K., & Aebersold, R. (2010). Generating and navigating proteome maps using mass spectrometry. *Nat Rev Mol Cell Biol*, 11(11), 789-801.
- Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., & Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia coli*. *BMC Syst Biol*, 8, 133.
- Albrecht, R., & Zeth, K. (2011). Structural basis of outer membrane protein biogenesis in bacteria. *J Biol Chem*, 286(31), 27792-27803.
- Anwari, K., Poggio, S., Perry, A., Gatsos, X., Ramarathinam, S. H., Williamson, N. A., Noinaj, N., Buchanan, S., Gabriel, K., Purcell, A. W., Jacobs-Wagner, C., & Lithgow, T. (2010). A modular BAM complex in the outer membrane of the alpha-proteobacterium *Caulobacter crescentus*. *PLoS One*, 5(1), e8619.
- Aravind, L., Makarova, K. S., & Koonin, E. V. (2000). SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res*, 28(18), 3417-3432.
- Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology*, 37(4), 420.
- Asmar, A. T., & Collet, J. F. (2018). Lpp, the Braun lipoprotein, turns 50-major achievements and remaining issues. *FEMS Microbiol Lett*, 365(18).
- Asmar, A. T., Ferreira, J. L., Cohen, E. J., Cho, S. H., Beeby, M., Hughes, K. T., & Collet, J. F. (2017). Communication across the bacterial cell envelope depends on the size of the periplasm. *PLoS Biol*, 15(12), e2004303.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*, 2, 2006, 0008.
- Babu, M., Arnold, R., Bundalovic-Torma, C., Gagarinova, A., Wong, K. S., Kumar, A., Stewart, G., Samanfar, B., Aoki, H., Wagih, O., Vlasblom, J., Phanse, S., Lad, K., Yeou Hsiung Yu, A., Graham, C., Jin, K., Brown, E., Golshani, A., Kim, P., Moreno-Hagelsieb, G., Greenblatt, J., Houry, W. A., Parkinson, J., & Emili, A. (2014). Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in *Escherichia coli*. *PLoS Genet*, 10(2), e1004120.
- Babu, M., Diaz-Mejia, J. J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., Yousif, F., Ding, H., Xiong, X., Nazarians-Armavil, A., Alamgir, M., Ali, M., Pogoutse, O., Pe'er, A., Arnold, R., Michaut, M., Parkinson, J., Golshani, A., Whitfield, C., Wodak, S. J., Moreno-Hagelsieb, G., Greenblatt, J. F., & Emili, A. (2011). Genetic interaction maps in *Escherichia coli* reveal functional crosstalk among cell envelope biogenesis pathways. *PLoS Genet*, 7(11), e1002377.
- Babu, M., Musso, G., Diaz-Mejia, J. J., Butland, G., Greenblatt, J. F., & Emili, A. (2009). Systems-level approaches for identifying and analyzing genetic interaction networks in *Escherichia coli* and extensions to other prokaryotes. *Mol Biosyst*, 5(12), 1439-1455.
- Bachmann, B. J. (1996). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. *Escherichia coli and Salmonella: cellular and molecular biology*, 2nd ed. ASM Press, Washington, DC, 2460-2488.
- Baldi, D. L., Higginson, E. E., Hocking, D. M., Praszkie, J., Cavaliere, R., James, C. E., Bennett-Wood, V., Azzopardi, K. I., Turnbull, L., & Lithgow, T. (2012a). The type II secretion system and its ubiquitous lipoprotein substrate, SslE, are required for biofilm formation and virulence of enteropathogenic *Escherichia coli*. *Infection and immunity*, 80(6), 2042-2052.

- Baldi, D. L., Higginson, E. E., Hocking, D. M., Praszkie, J., Cavaliere, R., James, C. E., Bennett-Wood, V., Azzopardi, K. I., Turnbull, L., Lithgow, T., Robins-Browne, R. M., Whitchurch, C. B., & Tauschek, M. (2012b). The type II secretion system and its ubiquitous lipoprotein substrate, SslE, are required for biofilm formation and virulence of enteropathogenic *Escherichia coli*. *Infect Immun*, 80(6), 2042-2052.
- Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., & Blanot, D. (2008). Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol Rev*, 32(2), 168-207.
- Bateman, A., & Bycroft, M. (2000). The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). *J Mol Biol*, 299(4), 1113-1119.
- Benedet, M., Falchi, F. A., Puccio, S., Di Benedetto, C., Peano, C., Polissi, A., & Deho, G. (2016). The Lack of the Essential LptC Protein in the Trans-Envelope Lipopolysaccharide Transport Machine Is Circumvented by Suppressor Mutations in LptF, an Inner Membrane Component of the *Escherichia coli* Transporter. *PLoS One*, 11(8), e0161354.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.
- Berney, M., Hammes, F., Bosshard, F., Weilenmann, H. U., & Egli, T. (2007). Assessment and interpretation of bacterial viability by using the LIVE/DEAD BacLight Kit in combination with flow cytometry. *Appl Environ Microbiol*, 73(10), 3283-3290.
- Bernhardt, T. G., & de Boer, P. A. (2003). The *Escherichia coli* amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol Microbiol*, 48(5), 1171-1182.
- Bitter, W., van Leeuwen, I. S., de Boer, J., Zomer, H. W., Koster, M. C., Weisbeek, P. J., & Tommassen, J. (1994). Localization of functional domains in the *Escherichia coli* coprogen receptor FhuE and the *Pseudomonas putida* ferric-pseudobactin 358 receptor PupA. *Mol Gen Genet*, 245(6), 694-703.
- Bladen, H. A., & Mergenhagen, S. E. (1964). Ultrastructure of Veillonella and Morphological Correlation of an Outer Membrane with Particles Associated with Endotoxic Activity. *J Bacteriol*, 88, 1482-1492.
- Bouhss, A., Trunkfield, A. E., Bugg, T. D., & Mengin-Lecreulx, D. (2008). The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev*, 32(2), 208-233.
- Boyle, D. S., & Donachie, W. D. (1998). *mraY* is an essential gene for cell growth in *Escherichia coli*. *J Bacteriol*, 180(23), 6429-6432.
- Braun, V. (1975). Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim Biophys Acta*, 415(3), 335-377.
- Braun, V., & Hantke, K. (2019). Lipoproteins: Structure, Function, Biosynthesis. *Subcell Biochem*, 92, 39-77.
- Braun, V., & Rehn, K. (1969). Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the *E. coli* cell wall. The specific effect of trypsin on the membrane structure. *Eur J Biochem*, 10(3), 426-438. doi:10.1111/j.1432-1033.1969.tb00707.
- Braun, V., & Wolff, H. (1970). The murein-lipoprotein linkage in the cell wall of *Escherichia coli*. *Eur J Biochem*, 14(2), 387-391.
- Budd, A., Blandin, S., Levashina, E. A., & Gibson, T. J. (2004). Bacterial alpha2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol*, 5(6), R38.
- Buddelmeijer, N. (2015). The molecular mechanism of bacterial lipoprotein modification--how, when and why? *FEMS Microbiol Rev*, 39(2), 246-261.

- Burkholder, W. F., Kurtser, I., & Grossman, A. D. (2001). Replication initiation proteins regulate a developmental checkpoint in *Bacillus subtilis*. *Cell*, 104(2), 269-279.
- Butland, G., Babu, M., Diaz-Mejia, J. J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A. G., Pogoutse, O., Mori, H., Wanner, B. L., Lo, H., Wasniewski, J., Christopolous, C., Ali, M., Venn, P., Safavi-Naini, A., Sourour, N., Caron, S., Choi, J. Y., Laigle, L., Nazarians-Armavil, A., Deshpande, A., Joe, S., Datsenko, K. A., Yamamoto, N., Andrews, B. J., Boone, C., Ding, H., Sheikh, B., Moreno-Hagelseib, G., Greenblatt, J. F., & Emili, A. (2008). eSGA: *E. coli* synthetic genetic array analysis. *Nat Methods*, 5(9), 789-795.
- Butler, E. K., Tan, W. B., Joseph, H., & Ruiz, N. (2014). Charge requirements of lipid II flippase activity in *Escherichia coli*. *J Bacteriol*, 196(23), 4111-4119.
- Cascales, E., Bernadac, A., Gavioli, M., Lazzaroni, J. C., & Lloubes, R. (2002). Pal lipoprotein of *Escherichia coli* plays a major role in outer membrane integrity. *J Bacteriol*, 184(3), 754-759.
- Cayley, D. S., Guttman, H. J., & Record, M. T., Jr. (2000). Biophysical characterization of changes in amounts and activity of *Escherichia coli* cell and compartment water and turgor pressure in response to osmotic stress. *Biophys J*, 78(4), 1748-1764.
- Charlson, E. S., Werner, J. N., & Misra, R. (2006). Differential effects of YfgL mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *J Bacteriol*, 188(20), 7186-7194.
- Chart, H., Smith, H. R., La Ragione, R. M., & Woodward, M. J. (2000). An investigation into the pathogenic properties of *Escherichia coli* strains BLR, BL21, DH5alpha and EQ1. *J Appl Microbiol*, 89(6), 1048-1058.
- Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene*, 158(1), 9-14.
- Cherier, D., Giacomucci, S., Patin, D., Bouhss, A., Touze, T., Blanot, D., Mengin-Lecreulx, D., & Barreteau, H. (2016). Pectocin M1 (PcaM1) Inhibits *Escherichia coli* Cell Growth and Peptidoglycan Biosynthesis through Periplasmic Expression. *Antibiotics (Basel)*, 5(4).
- Cho, H., Uehara, T., & Bernhardt, T. G. (2014). Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell*, 159(6), 1300-1311.
- Chumley, F. G., Menzel, R., & Roth, J. R. (1979). Hfr formation directed by tn10. *Genetics*, 91(4), 639-655.
- Chung, C. T., Niemela, S. L., & Miller, R. H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A*, 86(7), 2172-2175.
- Clarke, B. R., Bronner, D., Keenleyside, W. J., Severn, W. B., Richards, J. C., & Whitfield, C. (1995). Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from *Klebsiella pneumoniae* serotype O1. *J Bacteriol*, 177(19), 5411-5418.
- Cohen, E. J., Ferreira, J. L., Ladinsky, M. S., Beeby, M., & Hughes, K. T. (2017). Nanoscale-length control of the flagellar driveshaft requires hitting the tethered outer membrane. *Science*, 356(6334), 197-200.
- Collins, S. R., Schuldiner, M., Krogan, N. J., & Weissman, J. S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol*, 7(7), R63.
- Cowles, C. E., Li, Y., Semmelhack, M. F., Cristea, I. M., & Silhavy, T. J. (2011). The free and bound forms of Lpp occupy distinct subcellular locations in *Escherichia coli*. *Mol Microbiol*, 79(5), 1168-1181.

- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*, 10(4), 1794-1805.
- Croucher, N. J., & Thomson, N. R. (2010). Studying bacterial transcriptomes using RNA-seq. *Current opinion in microbiology*, 13(5), 619-624.
- Database resources of the National Center for Biotechnology Information. (2018). *Nucleic Acids Res*, 46(D1), D8-d13.
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, 97(12), 6640-6645.
- Datta, A. K., & Basu, S. (1999). Chemical characterization of the lipopolysaccharides from enteropathogenic *Escherichia coli* O142 and O158. *Indian J Biochem Biophys*, 36(1), 55-58.
- De Geyter, J., Tsirigotaki, A., Orfanoudaki, G., Zorzini, V., Economou, A., & Karamanou, S. (2016). Protein folding in the cell envelope of *Escherichia coli*. *Nat Microbiol*, 1(8), 16107.
- Demchick, P., & Koch, A. L. (1996). The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol*, 178(3), 768-773.
- d'Enfert, C., Ryter, A., & Pugsley, A. P. (1987). Cloning and expression in *Escherichia coli* of the *Klebsiella pneumoniae* genes for production, surface localization and secretion of the lipoprotein pullulanase. *EMBO J*, 6(11), 3531-3538.
- Deonier, R. C., & Mirels, L. (1977). Excision of F plasmid sequences by recombination at directly repeated insertion sequence 2 elements: involvement of recA. *Proc Natl Acad Sci U S A*, 74(9), 3965-3969.
- Desvaux, M., Parham, N. J., & Henderson, I. R. (2004). Type V protein secretion: simplicity gone awry? *Curr Issues Mol Biol*, 6(2), 111-124.
- Dijkstra, A. J., & Keck, W. (1996). Peptidoglycan as a barrier to transenvelope transport. *J Bacteriol*, 178(19), 5555-5562.
- Dmitriev, B. A., Ehlers, S., & Rietschel, E. T. (1999). Layered murein revisited: a fundamentally new concept of bacterial cell wall structure, biogenesis and function. *Med Microbiol Immunol*, 187(3), 173-181.
- Dmitriev, B. A., Toukach, F. V., Schaper, K. J., Holst, O., Rietschel, E. T., & Ehlers, S. (2003). Tertiary structure of bacterial murein: the scaffold model. *J Bacteriol*, 185(11), 3458-3468.
- Domingo Meza-Aguilar, J., Fromme, P., Torres-Larios, A., Mendoza-Hernandez, G., Hernandez-Chinas, U., Arreguin-Espinosa de Los Monteros, R. A., Eslava Campos, C. A., & Fromme, R. (2014). X-ray crystal structure of the passenger domain of plasmid-encoded toxin(Pet), an autotransporter enterotoxin from enteroaggregative *Escherichia coli* (EAEC). *Biochem Biophys Res Commun*, 445(2), 439-444.
- Dong, C., Hou, H. F., Yang, X., Shen, Y. Q. & Dong, Y. H. (2012a). Structure of *Escherichia coli* BamD and its functional implications in outer membrane protein assembly. *Acta Crystallogr D Biol Crystallogr*, 68(Pt 2), 95-101.
- Dong, C., Yang, X., Hou, H. F., Shen, Y. Q. & Dong, Y. H. (2012b). Structure of *Escherichia coli* BamB and its interaction with POTRA domains of BamA. *Acta Crystallogr D Biol Crystallogr*, 68(Pt 9), 1134-1139.
- Dramsi, S., Davison, S., Magnet, S., & Arthur, M. (2008). Surface proteins covalently attached to peptidoglycan: examples from both gram-positive and gram-negative bacteria. *FEMS Microbiol Rev*.
- Driessen, A. J., & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem*, 77, 643-667.

- Dutton, R. J., Boyd, D., Berkmen, M., & Beckwith, J. (2008). Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *Proc Natl Acad Sci U S A*, 105(33), 11933-11938.
- Economou, A., & Wickner, W. (1994). SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell*, 78(5), 835-843.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 32(5), 1792-1797.
- Egan, A. J. F. (2018). Bacterial outer membrane constriction. *Mol Microbiol*, 107(6), 676-687.
- Egan, A. J., Biboy, J., van't Veer, I., Breukink, E., & Vollmer, W. (2015). Activities and regulation of peptidoglycan synthases. *Philos Trans R Soc Lond B Biol Sci*, 370(1679).
- Egan, A. J., Jean, N. L., Koumoutsis, A., Bougault, C. M., Biboy, J., Sassine, J., Solovyova, A. S., Breukink, E., Typas, A., Vollmer, W., & Simorre, J. P. (2014). Outer-membrane lipoprotein LpoB spans the periplasm to stimulate the peptidoglycan synthase PBP1B. *Proc Natl Acad Sci U S A*, 111(22), 8197-8202.
- Eklof, J. M., Tan, T. C., Divne, C., & Brumer, H. (2009). The crystal structure of the outer membrane lipoprotein YbhC from *Escherichia coli* sheds new light on the phylogeny of carbohydrate esterase family 8. *Proteins*, 76(4), 1029-1036.
- Endo, T., Kawano, S., & Yamano, K. (2011). BamE structure: the assembly of beta-barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO Rep*, 12(2), 94-95.
- Erickson, H. P. (2017). How bacterial cell division might cheat turgor pressure - a unified mechanism of septal division in Gram-positive and Gram-negative bacteria. *BioEssays*, 39(8).
- Eslava, C., Navarro-Garcia, F., Czczulin, J. R., Henderson, I. R., Cravioto, A., & Nataro, J. P. (1998). Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun*, 66(7), 3155-3163.
- Estrada Mallarino, L., Fan, E., Odermatt, M., Muller, M., Lin, M., Liang, J., Heinzelmann, M., Fritsche, F., Apell, H. J., & Welte, W. (2015). TtOmp85, a beta-barrel assembly protein, functions by barrel augmentation. *Biochemistry*, 54(3), 844-852.
- Ewels, P., Magnusson, M., Lundin, S., & Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047-3048.
- Fairman, J. W., Noinaj, N., & Buchanan, S. K. (2011). The structural biology of beta-barrel membrane proteins: a summary of recent reports. *Curr Opin Struct Biol*, 21(4), 523-531.
- Falchi, F. A., Maccagni, E. A., Puccio, S., Peano, C., De Castro, C., Palmigiano, A., Garozzo, D., Martorana, A. M., Polissi, A., Deho, G., & Sperandio, P. (2018). Mutation and Suppressor Analysis of the Essential Lipopolysaccharide Transport Protein LptA Reveals Strategies To Overcome Severe Outer Membrane Permeability Defects in *Escherichia coli*. *J Bacteriol*, 200(2).
- Fan, E., Chauhan, N., Udatha, D., Leo, J. C., & Linke, D. (2016). Type V Secretion Systems in Bacteria. *Microbiol Spectr*, 4(1).
- Figueroa-Cuilan, W. M., & Brown, P. J. B. (2018). Cell Wall Biogenesis During Elongation and Division in the Plant Pathogen *Agrobacterium tumefaciens*. *Curr Top Microbiol Immunol*, 418, 87-110.
- Freundlieb, S., & Boos, W. (1986). Alpha-amylase of *Escherichia coli*, mapping and cloning of the structural gene, malS, and identification of its product as a periplasmic protein. *J Biol Chem*, 261(6), 2946-2953.

- Firdich, E., & Whitfield, C. (2005). Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the Enterobacteriaceae. *J Endotoxin Res*, 11(3), 133-144.
- Gagarinova, A., & Emili, A. (2012). Genome-scale genetic manipulation methods for exploring bacterial molecular biology. *Mol Biosyst*, 8(6), 1626-1638.
- Gagarinova, A., Babu, M., Greenblatt, J., & Emili, A. (2012). Mapping bacterial functional networks and pathways in *Escherichia coli* using synthetic genetic arrays. *JoVE (Journal of Visualized Experiments)*(69), e4056.
- Gan, L., Chen, S., & Jensen, G. J. (2008). Molecular organization of Gram-negative peptidoglycan. *Proc Natl Acad Sci U S A*, 105(48), 18953-18957.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, 31(13), 3784-3788.
- Ge, X., Wang, R., Ma, J., Liu, Y., Ezemaduka, A. N., Chen, P. R., Fu, X., & Chang, Z. (2014). DegP primarily functions as a protease for the biogenesis of beta-barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *FEBS J*, 281(4), 1226-1240.
- Gerding, M. A., Liu, B., Bendezu, F. O., Hale, C. A., Bernhardt, T. G., & de Boer, P. A. (2009). Self-enhanced accumulation of FtsN at Division Sites and Roles for Other Proteins with a SPOR domain (DamX, DedD, and RlpA) in *Escherichia coli* cell constriction. *J Bacteriol*, 191(24), 7383-7401.
- Ghatak, S., King, Z. A., Sastry, A., & Palsson, B. O. (2019). The y-ome defines the 35% of *Escherichia coli* genes that lack experimental evidence of function. *Nucleic Acids Res*, 47(5), 2446-2454.
- Gibbs, R. J., Stewart, J., & Poxton, I. R. (2004). The distribution of, and antibody response to, the core lipopolysaccharide region of *Escherichia coli* isolated from the faeces of healthy humans and cattle. *J Med Microbiol*, 53(Pt 10), 959-964.
- Glauner, B. (1988). Separation and quantification of mucopeptides with high-performance liquid chromatography. *Anal Biochem*, 172(2), 451-464.
- Glauner, B., Holtje, J. V., & Schwarz, U. (1988b). The composition of the murein of *Escherichia coli*. *J Biol Chem*, 263(21), 10088-10095.
- Glauner, B., Hölte, J., & Schwarz, U. (1988a). The composition of the murein of *Escherichia coli*. *Journal of Biological Chemistry*, 263(21), 10088-10095.
- Gowrishankar, J. (1989). Nucleotide sequence of the osmoregulatory *proU* operon of *Escherichia coli*. *J Bacteriol*, 171(4), 1923-1931.
- Greenfield, L. K., & Whitfield, C. (2012). Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways. *Carbohydr Res*, 356, 12-24.
- Grover, R. K., Cheng, J., Peng, Y., Jones, T. M., Ruiz, D. I., Ulevitch, R. J., Glass, J. I., Dennis, E. A., Salomon, D. R., & Lerner, R. A. (2012). The costimulatory immunogen LPS induces the B-Cell clones that infiltrate transplanted human kidneys. *Proc Natl Acad Sci U S A*, 109(16), 6036-6041.
- Gruber, T. M., & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol*, 57, 441-466.
- Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W., & Dong, C. (2016). Structural basis of outer membrane protein insertion by the BAM complex. *Nature*, 531(7592), 64-69.
- Gul, N., & Poolman, B. (2013). Functional reconstitution and osmoregulatory properties of the ProU ABC transporter from *Escherichia coli*. *Mol Membr Biol*, 30(2), 138-148.
- Gunasinghe, S. D., Shiota, T., Stubenrauch, C. J., Schulze, K. E., Webb, C. T., Fulcher, A. J., Dunstan, R. A., Hay, I. D., Naderer, T., Whelan, D. R., Bell, T. D. M., Elgass, K. D., Strugnell, R. A., & Lithgow, T. (2018). The WD40 Protein BamB Mediates Coupling

- of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. *Cell Rep*, 23(9), 2782-2794.
- Gupta, R. S. (2011). Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes. *Antonie Van Leeuwenhoek*, 100(2), 171-182.
- Hagan, C. L., Kim, S., & Kahne, D. (2010). Reconstitution of outer membrane protein assembly from purified components. *Science*, 328(5980), 890-892.
- Hagan, C. L., Silhavy, T. J., & Kahne, D. (2011). beta-Barrel membrane protein assembly by the Bam complex. *Annu Rev Biochem*, 80, 189-210.
- Halter, R., Pohlner, J., & Meyer, T. F. (1984). IgA protease of *Neisseria gonorrhoeae*: isolation and characterization of the gene and its extracellular product. *EMBO J*, 3(7), 1595-1601.
- Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y., & Huang, Y. (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat Struct Mol Biol*, 23(3), 192-196.
- Han, W., Wu, B., Li, L., Zhao, G., Woodward, R., Pettit, N., Cai, L., Thon, V., & Wang, P. G. (2012). Defining function of lipopolysaccharide O-antigen ligase WaaL using chemoenzymatically synthesized substrates. *J Biol Chem*, 287(8), 5357-5365.
- Hegde, R. S., & Bernstein, H. D. (2006). The surprising complexity of signal sequences. *Trends Biochem Sci*, 31(10), 563-571.
- Heidrich, C., Templin, M. F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M. A., & Holtje, J. V. (2001). Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of *Escherichia coli*. *Mol Microbiol*, 41(1), 167-178.
- Hemm, M. R., Paul, B. J., Schneider, T. D., Storz, G., & Rudd, K. E. (2008). Small membrane proteins found by comparative genomics and ribosome binding site models. *Mol Microbiol*, 70(6), 1487-1501.
- Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C., & Ala'Aldeen, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev*, 68(4), 692-744.
- Heuck, A., Schleiffer, A., & Clausen, T. (2011). Augmenting beta-augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J Mol Biol*, 406(5), 659-666.
- Hicks, G., & Jia, Z. (2018). Structural Basis for the Lipopolysaccharide Export Activity of the Bacterial Lipopolysaccharide Transport System. *Int J Mol Sci*, 19(9).
- Hiemstra, H., Nanninga, N., Woldringh, C. L., Inouye, M., & Witholt, B. (1987). Distribution of newly synthesized lipoprotein over the outer membrane and the peptidoglycan sacculus of an *Escherichia coli* lac-lpp strain. *J Bacteriol*, 169(12), 5434-5444. doi:10.1128/jb.169.12.5434-5444.1987
- Hirashima, A., & Inouye, M. (1973). Specific biosynthesis of an envelope protein of *Escherichia coli*. *Nature*, 242(5397), 405-407.
- Hirota, Y., Suzuki, H., Nishimura, Y., & Yasuda, S. (1977). On the process of cellular division in *Escherichia coli*: a mutant of *E. coli* lacking a murein-lipoprotein. *Proc Natl Acad Sci U S A*, 74(4), 1417-1420.
- Holm, L. (2019). Benchmarking Fold Detection by DaliLite v.5. *Bioinformatics*.
- Holm, L., & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J Mol Biol*, 233(1), 123-138. doi:10.1006/jmbi.1993.1489
- Holtje, J. V., Mirelman, D., Sharon, N., & Schwarz, U. (1975). Novel type of murein transglycosylase in *Escherichia coli*. *J Bacteriol*, 124(3), 1067-1076.

- Hwang, H., Paracini, N., Parks, J. M., Lakey, J. H., & Gumbart, J. C. (2018). Distribution of mechanical stress in the *Escherichia coli* cell envelope. *Biochim Biophys Acta Biomembr*, 1860(12), 2566-2575. doi:10.1016/j.bbamem.2018.09.020
- Ingledeu, W. J., & Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiol Rev*, 48(3), 222-271.
- Inouye, M., Shaw, J., & Shen, C. (1972). The assembly of a structural lipoprotein in the envelope of *Escherichia coli*. *J Biol Chem*, 247(24), 8154-8159.
- Jain, S., & Goldberg, M. B. (2007). Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol*, 189(14), 5393-5398.
- Jean, N. L., Bougault, C. M., Egan, A. J., Vollmer, W., & Simorre, J. P. (2015b). Solution NMR assignment of LpoB, an outer-membrane anchored Penicillin-Binding Protein activator from *Escherichia coli*. *Biomol NMR Assign*, 9(1), 123-127.
- Jean, N. L., Bougault, C. M., Lodge, A., Derouaux, A., Callens, G., Egan, A. J., Ayala, I., Lewis, R. J., Vollmer, W., & Simorre, J. P. (2014). Elongated structure of the outer-membrane activator of peptidoglycan synthesis LpoA: implications for PBP1A stimulation. *Structure*, 22(7), 1047-1054.
- Jean, N. L., Bougault, C., Derouaux, A., Callens, G., Vollmer, W., & Simorre, J. P. (2015a). Backbone and side-chain (1)H, (13)C, and (15)N NMR assignments of the N-terminal domain of *Escherichia coli* LpoA. *Biomol NMR Assign*, 9(1), 65-69.
- Jefferies, D., Shearer, J., & Khalid, S. (2019). Role of O-Antigen in Response to Mechanical Stress of the *E. coli* Outer Membrane: Insights from Coarse-Grained MD Simulations. *J Phys Chem B*, 123(17), 3567-3575. doi:10.1021/acs.jpcb.8b12168
- Jishage, M., & Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc Natl Acad Sci U S A*, 95(9), 4953-4958.
- Johnston, C., Caymaris, S., Zomer, A., Bootsma, H. J., Prudhomme, M., Granadel, C., Hermans, P. W., Polard, P., Martin, B., & Claverys, J. P. (2013). Natural genetic transformation generates a population of merodiploids in *Streptococcus pneumoniae*. *PLoS Genet*, 9(9), e1003819.
- Joiner, K. A. (1988). Complement evasion by bacteria and parasites. *Annu Rev Microbiol*, 42, 201-230.
- Jones, C. J., Macnab, R. M., Okino, H., & Aizawa, S. (1990). Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*. *J Mol Biol*, 212(2), 377-387.
- Jorgensen, J., & Turnidge, J. (2015). CLSI: Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. *Manual of clinical microbiol*, 1253-1273.
- Jose, J., & Meyer, T. F. (2007). The autodisplay story, from discovery to biotechnical and biomedical applications. *Microbiol Mol Biol Rev*, 71(4), 600-619.
- Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci*, 12(8), 1652-1662.
- Karp, P. D., Keseler, I. M., Shearer, A., Latendresse, M., Krummenacker, M., Paley, S. M., Paulsen, I., Collado-Vides, J., Gama-Castro, S., Peralta-Gil, M., Santos-Zavaleta, A., Penaloza-Spinola, M. I., Bonavides-Martinez, C., & Ingraham, J. (2007). Multidimensional annotation of the *Escherichia coli* K-12 genome. *Nucleic Acids Res*, 35(22), 7577-7590.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modelling, prediction and analysis. *Nat Protoc*, 10(6), 845-858.
- Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., & Krummenacker, M. (2016). The

- EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic acids research*, 45(D1), D543-D550.
- Kim, K. H., Aulakh, S., & Paetzel, M. (2011a). Crystal structure of beta-barrel assembly machinery BamCD protein complex. *J Biol Chem*, 286(45), 39116-39121.
- Kim, K. H., Kang, H. S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P., & Paetzel, M. (2011b). Structural characterization of *Escherichia coli* BamE, a lipoprotein component of the beta-barrel assembly machinery complex. *Biochemistry*, 50(6), 1081-1090.
- Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., & Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science*, 317(5840), 961-964.
- Kjaergaard, K., Hasman, H., Schembri, M. A., & Klemm, P. (2002). Antigen 43-mediated autotransporter display, a versatile bacterial cell surface presentation system. *J Bacteriol*, 184(15), 4197-4204.
- Knowles, T. J., Browning, D. F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D. L., Squire, D., Chaudhuri, R. R., Viant, M. R., Cunningham, A. F., Henderson, I. R., & Overduin, M. (2011). Structure and function of BamE within the outer membrane and the beta-barrel assembly machine. *EMBO Rep*, 12(2), 123-128.
- Knowles, T. J., Scott-Tucker, A., Overduin, M., & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol*, 7(3), 206-214.
- Koch, A. L. (1998). The biophysics of the gram-negative periplasmic space. *Crit Rev Microbiol*, 24(1), 23-59.
- Koch, A. L., & Woeste, S. (1992). Elasticity of the sacculus of *Escherichia coli*. *J Bacteriol*, 174(14), 4811-4819.
- Koch, A. L., Lane, S. L., Miller, J. A., & Nickens, D. G. (1987). Contraction of filaments of *Escherichia coli* after disruption of cell membrane by detergent. *J Bacteriol*, 169(5), 1979-1984.
- Koebnik, R., Locher, K. P., & Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol*, 37(2), 239-253.
- Konovalova, A., & Silhavy, T. J. (2015). Outer membrane lipoprotein biogenesis: Lol is not the end. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1679), 20150030.
- Koronakis, V., Eswaran, J., & Hughes, C. (2004). Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem*, 73, 467-489.
- Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 305(3), 567-580.
- Kudva, R., Denks, K., Kuhn, P., Vogt, A., Muller, M., & Koch, H. G. (2013). Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. *Res Microbiol*, 164(6), 505-534.
- Labischinski, H., Goodell, E. W., Goodell, A., & Hochberg, M. L. (1991). Direct proof of a "more-than-single-layered" peptidoglycan architecture of *Escherichia coli* W7: a neutron small-angle scattering study. *J Bacteriol*, 173(2), 751-756.
- Lambert, O., Benabdelhak, H., Chami, M., Jouan, L., Nouaille, E., Ducruix, A., & Brisson, A. (2005). Trimeric structure of OprN and OprM efflux proteins from *Pseudomonas aeruginosa*, by 2D electron crystallography. *J Struct Biol*, 150(1), 50-57.
- Lau, C. K., Ishida, H., Liu, Z., & Vogel, H. J. (2013). Solution structure of *Escherichia coli* FeoA and its potential role in bacterial ferrous iron transport. *J Bacteriol*, 195(1), 46-55.

- Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*, 15(2), R29.
- Lawley, T. D., Klimke, W. A., Gubbins, M. J., & Frost, L. S. (2003). F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett*, 224(1), 1-15.
- Lengeler, J. (1975). Nature and properties of hexitol transport systems in *Escherichia coli*. *J Bacteriol*, 124(1), 39-47.
- Lewenza, S., Vidal-Ingigliardi, D., & Pugsley, A. P. (2006). Direct visualization of red fluorescent lipoproteins indicates conservation of the membrane sorting rules in the family Enterobacteriaceae. *J Bacteriol*, 188(10), 3516-3524.
- Leyton, D. L., Belousoff, M. J., & Lithgow, T. (2015). The beta-Barrel Assembly Machinery Complex. *Methods Mol Biol*, 1329, 1-16.
- Leyton, D. L., De Luna, M. d. G., Sevastsyanovich, Y. R., Tveen Jensen, K., Browning, D. F., Scott-Tucker, A., & Henderson, I. R. (2010). The unusual extended signal peptide region is not required for secretion and function of an *Escherichia coli* autotransporter. *FEMS microbiology letters*, 311(2), 133-139.
- Leyton, D. L., Johnson, M. D., Thapa, R., Huysmans, G. H., Dunstan, R. A., Celik, N., Shen, H. H., Loo, D., Belousoff, M. J., Purcell, A. W., Henderson, I. R., Beddoe, T., Rossjohn, J., Martin, L. L., Strugnell, R. A., & Lithgow, T. (2014). A mortise-tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nat Commun*, 5, 4239.
- Leyton, D. L., Sevastsyanovich, Y. R., Browning, D. F., Rossiter, A. E., Wells, T. J., Fitzpatrick, R. E., Overduin, M., Cunningham, A. F., & Henderson, I. R. (2011). Size and conformation limits to secretion of disulfide-bonded loops in autotransporter proteins. *J Biol Chem*, 286(49), 42283-42291.
- Li, G. W., Burkhardt, D., Gross, C., & Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3), 624-635.
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.
- Liao, Y., Smyth, G. K., & Shi, W. (2013). featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-930.
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-930.
- Loos, M. S., Ramakrishnan, R., Vranken, W., Tsirigotaki, A., Tsare, E.-P., Zorzini, V., Geyter, J. D., Yuan, B., Tsamardinos, I., Klappa, M., Schymkowitz, J., Rousseau, F., Karamanou, S., & Economou, A. (2019). Structural Basis of the Subcellular Topology Landscape of *Escherichia coli*. *Frontiers in Microbiology*, 10(1670).
- Lucht, J. M., & Bremer, E. (1994). Adaptation of *Escherichia coli* to high osmolarity environments: osmoregulation of the high-affinity glycine betaine transport system proU. *FEMS Microbiol Rev*, 14(1), 3-20.
- Lupoli, T. J., Lebar, M. D., Markovski, M., Bernhardt, T., Kahne, D., & Walker, S. (2014). Lipoprotein activators stimulate *Escherichia coli* penicillin-binding proteins by different mechanisms. *J Am Chem Soc*, 136(1), 52-55.
- Magnet, S., Bellais, S., Dubost, L., Fourgeaud, M., Mainardi, J. L., Petit-Frere, S., Marie, A., Mengin-Lecreulx, D., Arthur, M., & Gutmann, L. (2007). Identification of the L, D-transpeptidases responsible for attachment of the Braun lipoprotein to *Escherichia coli* peptidoglycan. *J Bacteriol*, 189(10), 3927-3931.
- Magnet, S., Dubost, L., Marie, A., Arthur, M., & Gutmann, L. (2008). Identification of the L, D-transpeptidases for peptidoglycan cross-linking in *Escherichia coli*. *J Bacteriol*, 190(13), 4782-4785.

- Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R., & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol*, 61(1), 151-164.
- Marlovits, T. C., Haase, W., Herrmann, C., Aller, S. G., & Unger, V. M. (2002). The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proc Natl Acad Sci U S A*, 99(25), 16243-16248.
- Marolda, C. L., Li, B., Lung, M., Yang, M., Hanuszkiewicz, A., Rosales, A. R., & Valvano, M. A. (2010). Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. *J Bacteriol*, 192(23), 6160-6171.
- Martorana, A. M., Sperandio, P., Polissi, A., & Deho, G. (2011). Complex transcriptional organization regulates an *Escherichia coli* locus implicated in lipopolysaccharide biogenesis. *Res Microbiol*, 162(5), 470-482.
- Martynowski, D., Grochulski, P., & Howard, P. S. (2013). Structure of a periplasmic domain of the EpsAB fusion protein of the *Vibrio vulnificus* type II secretion system. *Acta Crystallogr D Biol Crystallogr*, 69(Pt 2), 142-149.
- Matias, V. R., Al-Amoudi, A., Dubochet, J., & Beveridge, T. J. (2003). Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. *J Bacteriol*, 185(20), 6112-6118.
- Matlack, K. E., Mothes, W., & Rapoport, T. A. (1998). Protein translocation: tunnel vision. *Cell*, 92(3), 381-390.
- Meroueh, S. O., Bencze, K. Z., Heseck, D., Lee, M., Fisher, J. F., Stemmler, T. L., & Mobashery, S. (2006). Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A*, 103(12), 4404-4409.
- Miller, S. I., & Salama, N. R. (2018). The gram-negative bacterial periplasm: Size matters. *PLoS Biol*, 16(1), e2004935.
- Misra, R., Stikeleather, R., & Gabriele, R. (2015). In vivo roles of BamA, BamB and BamD in the biogenesis of BamA, a core protein of the beta-barrel assembly machine of *Escherichia coli*. *J Mol Biol*, 427(5), 1061-1074.
- Missiakas, D., & Raina, S. (1997). Protein folding in the bacterial periplasm. *J Bacteriol*, 179(8), 2465-2471.
- Missiakas, D., Georgopoulos, C., & Raina, S. (1993). Identification and characterization of the *Escherichia coli* gene dsbB, whose product is involved in the formation of disulfide bonds in vivo. *Proc Natl Acad Sci U S A*, 90(15), 7084-7088.
- Mitchell, J. E., Oshima, T., Piper, S. E., Webster, C. L., Westblade, L. F., Karimova, G., Ladant, D., Kolb, A., Hobman, J. L., Busby, S. J., & Lee, D. J. (2007). The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70. *J Bacteriol*, 189(9), 3489-3495.
- Moffatt, J. H., Harper, M., Harrison, P., Hale, J. D., Vinogradov, E., Seemann, T., Henry, R., Crane, B., St Michael, F., Cox, A. D., Adler, B., Nation, R. L., Li, J., & Boyce, J. D. (2010). Colistin resistance in *Acinetobacter baumannii* is mediated by complete loss of lipopolysaccharide production. *Antimicrob Agents Chemother*, 54(12), 4971-4977.
- Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Disteche, M., de Kruijff, B., & Breukink, E. (2011). Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J*, 30(8), 1425-1432.
- Mohanty, B. K., & Kushner, S. R. (1999). Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism. *Mol Microbiol*, 34(5), 1094-1108.

- Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T., Nouwens, A. S., Williams, K. L., & Gooley, A. A. (2000). Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem*, 267(10), 2871-2881.
- Nakamura, K., & Inouye, M. (1979). DNA sequence of the gene for the outer membrane lipoprotein of *E. coli*: an extremely AT-rich promoter. *Cell*, 18(4), 1109-1117.
- Neu, H. C., & Heppel, L. A. (1964). Some Observations on the "Latent" Ribonuclease of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 51, 1267-1274.
- Nichols, R. J., Sen, S., Choo, Y. J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K. M., Lee, K. J., Wong, A., Shales, M., Lovett, S., Winkler, M. E., Krogan, N. J., Typas, A., & Gross, C. A. (2011). Phenotypic landscape of a bacterial cell. *Cell*, 144(1), 143-156.
- Nikaido, H. (1976). Outer membrane of *Salmonella typhimurium*: transmembrane diffusion of some hydrophobic substances. *Biochimica Et Biophysica Acta (BBA)-Biomembranes*, 433(1), 118-132.
- Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev*, 67(4), 593-656.
- Nishino, K., & Yamaguchi, A. (2001). Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J Bacteriol*, 183(20), 5803-5812.
- Nobelmann, B., & Lengeler, J. W. (1996). Molecular analysis of the gat genes from *Escherichia coli* and of their roles in galactitol transport and metabolism. *J Bacteriol*, 178(23), 6790-6795.
- Noinaj, N., Gumbart, J. C., & Buchanan, S. K. (2017). The β -barrel assembly machinery in motion. *Nature Reviews Microbiology*, 15(4), 197-204.
- Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C., & Buchanan, S. K. (2014). Lateral opening and exit pore formation are required for BamA function. *Structure*, 22(7), 1055-1062.
- Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T., & Buchanan, S. K. (2013). Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature*, 501(7467), 385-390.
- Noinaj, N., Rollauer, S. E., & Buchanan, S. K. (2015). The beta-barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr Opin Struct Biol*, 31, 35-42.
- Nolle, N., Felsl, A., Heermann, R., & Fuchs, T. M. (2017). Genetic Characterization of the Galactitol Utilization Pathway of *Salmonella enterica* Serovar Typhimurium. *J Bacteriol*, 199(4).
- Okuda, S., Sherman, D. J., Silhavy, T. J., Ruiz, N., & Kahne, D. (2016). Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nat Rev Microbiol*, 14(6), 337-345.
- Onufryk, C., Crouch, M. L., Fang, F. C., & Gross, C. A. (2005). Characterization of six lipoproteins in the sigmaE regulon. *J Bacteriol*, 187(13), 4552-4561.
- Paetzel, M., Karla, A., Strynadka, N. C., & Dalbey, R. E. (2002). Signal peptidases. *Chem Rev*, 102(12), 4549-4580.
- Paradis-Bleau, C., Markovski, M., Uehara, T., Lupoli, T. J., Walker, S., Kahne, D. E., & Bernhardt, T. G. (2010). Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. *Cell*, 143(7), 1110-1120.
- Park, B. S., & Lee, J. O. (2013). Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp Mol Med*, 45, e66.
- Park, J. S., Lee, W. C., Yeo, K. J., Ryu, K. S., Kumarasiri, M., Hesek, D., Lee, M., Mobashery, S., Song, J. H., Kim, S. I., Lee, J. C., Cheong, C., Jeon, Y. H., & Kim, H. Y. (2012). Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. *FASEB J*, 26(1), 219-228.

- Park, J. T., & Uehara, T. (2008). How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). *Microbiol Mol Biol Rev*, 72(2), 211-227, table of contents.
- Pawelek, P. D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C. M., Moiseeva, N., Allaire, M., & Coulton, J. W. (2006). Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science*, 312(5778), 1399-1402.
- Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., & Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*, 24(1), 79-88.
- Peng, D., Hong, W., Choudhury, B. P., Carlson, R. W., & Gu, X. X. (2005). *Moraxella catarrhalis* bacterium without endotoxin, a potential vaccine candidate. *Infect Immun*, 73(11), 7569-7577.
- Pietsch, F., Bergman, J. M., Brandis, G., Marcusson, L. L., Zorzet, A., Huseby, D. L., & Hughes, D. (2017). Ciprofloxacin selects for RNA polymerase mutations with pleiotropic antibiotic resistance effects. *J Antimicrob Chemother*, 72(1), 75-84.
- Pilizota, T., & Shaevitz, J. W. (2013). Plasmolysis and cell shape depend on solute outer-membrane permeability during hyperosmotic shock in *E. coli*. *Biophys J*, 104(12), 2733-2742.
- Plummer, A. M., & Fleming, K. G. (2016). From Chaperones to the Membrane with a BAM! *Trends Biochem Sci*, 41(10), 872-882.
- Pohlner, J., Halter, R., Beyreuther, K., & Meyer, T. F. (1987). Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature*, 325(6103), 458-462.
- Polissi, A., & Georgopoulos, C. (1996). Mutational analysis and properties of the *msbA* gene of *Escherichia coli*, coding for an essential ABC family transporter. *Mol Microbiol*, 20(6), 1221-1233.
- Polissi, A., & Sperandio, P. (2014). The lipopolysaccharide export pathway in *Escherichia coli*: structure, organization and regulated assembly of the Lpt machinery. *Mar Drugs*, 12(2), 1023-1042.
- Prats, R., & de Pedro, M. A. (1989). Normal growth and division of *Escherichia coli* with a reduced amount of murein. *J Bacteriol*, 171(7), 3740-3745.
- Preiss, J. (1984). Bacterial glycogen synthesis and its regulation. *Annu Rev Microbiol*, 38, 419-458.
- Priyadarshini, R., de Pedro, M. A., & Young, K. D. (2007). Role of peptidoglycan amidases in the development and morphology of the division septum in *Escherichia coli*. *J Bacteriol*, 189(14), 5334-5347.
- Prochnow, H., Fetz, V., Hotop, S. K., Garcia-Rivera, M. A., Heumann, A., & Bronstrup, M. (2019). Subcellular Quantification of Uptake in Gram-Negative Bacteria. *Anal Chem*, 91(3), 1863-1872.
- Prozorov, A. A. (2003). [Conjugation in bacilli]. *Mikrobiologiya*, 72(5), 581-593.
- Pucciarelli, M. G., & Garcia-del Portillo, F. (2003). Protein-peptidoglycan interactions modulate the assembly of the needle complex in the *Salmonella* invasion-associated type III secretion system. *Mol Microbiol*, 48(2), 573-585.
- Radika, K., & Raetz, C. R. (1988). Purification and properties of lipid A disaccharide synthase of *Escherichia coli*. *J Biol Chem*, 263(29), 14859-14867.
- Raha, M., Kawagishi, I., Muller, V., Kihara, M., & Macnab, R. M. (1992). *Escherichia coli* produces a cytoplasmic alpha-amylase, AmyA. *J Bacteriol*, 174(20), 6644-6652.
- Raina, S., & Missiakas, D. (1997). Making and breaking disulfide bonds. *Annu Rev Microbiol*, 51, 179-202. 9

- Ray, B. L., Painter, G., & Raetz, C. R. (1984). The biosynthesis of gram-negative endotoxin. Formation of lipid A disaccharides from monosaccharide precursors in extracts of *Escherichia coli*. *J Biol Chem*, 259(8), 4852-4859.
- Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zahringer, U., Seydel, U., Di Padova, F., & et al. (1994). Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J*, 8(2), 217-225
- Rigel, N. W., Schwalm, J., Ricci, D. P., & Silhavy, T. J. (2012). BamE modulates the *Escherichia coli* beta-barrel assembly machine component BamA. *J Bacteriol*, 194(5), 1002-1008.
- Rizzitello, A. E., Harper, J. R., & Silhavy, T. J. (2001). Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J Bacteriol*, 183(23), 6794-6800.
- Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*, 11(3), R25.
- Rocaboy, M., Herman, R., Sauvage, E., Remaut, H., Moonens, K., Terrak, M., Charlier, P., & Kerff, F. (2013). The crystal structure of the cell division amidase AmiC reveals the fold of the AMIN domain, a new peptidoglycan binding domain. *Mol Microbiol*, 90(2), 267-277.
- Rodriguez, F., Rouse, S. L., Tait, C. E., Harmer, J., De Riso, A., Timmel, C. R., Sansom, M. S., Berks, B. C., & Schnell, J. R. (2013). Structural model for the protein-translocating element of the twin-arginine transport system. *Proc Natl Acad Sci U S A*, 110(12), E1092-1101.
- Rogers, H., Perkins, H., & Ward, J. (1980). Biosynthesis of peptidoglycan. In *Microbial cell walls and membranes* (pp. 239-297): Springer.
- Rojas, E. R., Billings, G., Odermatt, P. D., Auer, G. K., Zhu, L., Miguel, A., Chang, F., Weibel, D. B., Theriot, J. A., & Huang, K. C. (2018). The outer membrane is an essential load-bearing element in Gram-negative bacteria. *Nature*, 559(7715), 617-621.
- Rollauer, S. E., Soorashjani, M. A., Noinaj, N., & Buchanan, S. K. (2015). Outer membrane protein biogenesis in Gram-negative bacteria. *Philos Trans R Soc Lond B Biol Sci*, 370(1679).
- Rondelet, A., & Condemine, G. (2013). Type II secretion: the substrates that won't go away. *Research in microbiology*, 164(6), 556-561.
- Rossiter, A. E., Leyton, D. L., Tveen-Jensen, K., Browning, D. F., Sevastyanovich, Y., Knowles, T. J., Nichols, K. B., Cunningham, A. F., Overduin, M., Schembri, M. A., & Henderson, I. R. (2011). The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J Bacteriol*, 193(16), 4250-4253.
- Ruiz, N. (2008). Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 105(40), 15553-15557.
- Ruiz, N. (2016). Filling holes in peptidoglycan biogenesis of *Escherichia coli*. *Curr Opin Microbiol*, 34, 1-6.
- Ruiz-Perez, F., Henderson, I. R., & Nataro, J. P. (2010). Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes*, 1(5), 339-344.
- Ruiz-Perez, F., Henderson, I. R., Leyton, D. L., Rossiter, A. E., Zhang, Y., & Nataro, J. P. (2009). Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of Enterobacteriaceae. *J Bacteriol*, 191(21), 6571-6583.
- Samsudin, F., Boags, A., Piggot, T. J., & Khalid, S. (2017). Braun's Lipoprotein Facilitates OmpA Interaction with the *Escherichia coli* Cell Wall. *Biophys J*, 113(7), 1496-1504. doi:10.1016/j.bpj.2017.08.011

- Sankaran, K., & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J Biol Chem*, 269(31), 19701-19706.
- Santambrogio, C., Sperandio, P., Villa, R., Sobott, F., Polissi, A., & Grandori, R. (2013). LptA assembles into rod-like oligomers involving disorder-to-order transitions. *J Am Soc Mass Spectrom*, 24(10), 1593-1602.
- Sauri, A., Soprova, Z., Wickstrom, D., de Gier, J. W., Van der Schors, R. C., Smit, A. B., Jong, W. S., & Luirink, J. (2009). The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology*, 155(Pt 12), 3982-3991.
- Scheurwater, E. M., & Burrows, L. L. (2011). Maintaining network security: how macromolecular structures cross the peptidoglycan layer. *FEMS Microbiol Lett*, 318(1), 1-9.
- Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M., & Rosenow, C. (2003). Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res*, 13(2), 216-223.
- Selkrig, J., Leyton, D. L., Webb, C. T., & Lithgow, T. (2014). Assembly of beta-barrel proteins into bacterial outer membranes. *Biochim Biophys Acta*, 1843(8), 1542-1550.
- Selkrig, J., Mosbahi, K., Webb, C. T., Belousoff, M. J., Perry, A. J., Wells, T. J., Morris, F., Leyton, D. L., Totsika, M., Phan, M. D., Celik, N., Kelly, M., Oates, C., Hartland, E. L., Robins-Browne, R. M., Ramarathinam, S. H., Purcell, A. W., Schembri, M. A., Strugnelli, R. A., Henderson, I. R., Walker, D., & Lithgow, T. (2012). Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol*, 19(5), 506-510, S501.
- Sevast'yanovich, Y. R., Leyton, D. L., Wells, T. J., Wardius, C. A., Tveen-Jensen, K., Morris, F. C., Knowles, T. J., Cunningham, A. F., Cole, J. A., & Henderson, I. R. (2012). A generalised module for the selective extracellular accumulation of recombinant proteins. *Microb Cell Fact*, 11, 69.
- Sham, L. T., Butler, E. K., Lebar, M. D., Kahne, D., Bernhardt, T. G., & Ruiz, N. (2014). Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science*, 345(6193), 220-222.
- Shu, W., Liu, J., Ji, H., & Lu, M. (2000). Core structure of the outer membrane lipoprotein from *Escherichia coli* at 1.9 Å resolution. *J Mol Biol*, 299(4), 1101-1112.
- Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., & Remington, S. J. (2006). Novel chromophores and buried charges control color in mFruits. *Biochemistry*, 45(32), 9639-9647.
- Shuman, H. A., & Silhavy, T. J. (2003). The art and design of genetic screens: *Escherichia coli*. *Nat Rev Genet*, 4(6), 419-431.
- Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harb Perspect Biol*, 2(5), a000414.
- Silva, I. J., Saramago, M., Dressaire, C., Domingues, S., Viegas, S. C., & Arraiano, C. M. (2011). Importance and key events of prokaryotic RNA decay: the ultimate fate of an RNA molecule. *Wiley Interdiscip Rev RNA*, 2(6), 818-836.
- Simpson, R. J. (2006). SDS-PAGE of proteins. *Cold Spring Harbor Protocols*, 2006(1), pdb.prot4313-pdb.prot4313.
- Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D., & Silhavy, T. J. (2007a). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 104(15), 6400-6405.
- Sklar, J. G., Wu, T., Kahne, D., & Silhavy, T. J. (2007b). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev*, 21(19), 2473-2484.
- Sochacki, K. A., Shkel, I. A., Record, M. T., & Weisshaar, J. C. (2011). Protein diffusion in the periplasm of *E. coli* under osmotic stress. *Biophys J*, 100(1), 22-31.

- Soltes, G. R., Martin, N. R., Park, E., Sutterlin, H. A., & Silhavy, T. J. (2017). Distinctive Roles for Periplasmic Proteases in the Maintenance of Essential Outer Membrane Protein Assembly. *J Bacteriol*, 199(20).
- Sperandeo, P., Martorana, A. M., & Polissi, A. (2017). Lipopolysaccharide biogenesis and transport at the outer membrane of Gram-negative bacteria. *Biochim Biophys Acta Mol Cell Biol Lipids*, 1862(11), 1451-1460.
- Sperandeo, P., Villa, R., Martorana, A. M., Samalikova, M., Grandori, R., Deho, G., & Polissi, A. (2011). New insights into the Lpt machinery for lipopolysaccharide transport to the cell surface: LptA-LptC interaction and LptA stability as sensors of a properly assembled transenvelope complex. *J Bacteriol*, 193(5), 1042-1053.
- Spieß, C., Beil, A., & Ehrmann, M. (1999). A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell*, 97(3), 339-347.
- Steehgs, L., den Hartog, R., den Boer, A., Zomer, B., Roholl, P., & van der Ley, P. (1998). Meningitis bacterium is viable without endotoxin. *Nature*, 392(6675), 449.
- Stenberg, F., Chovanec, P., Maslen, S. L., Robinson, C. V., Ilag, L. L., von Heijne, G., & Daley, D. O. (2005). Protein complexes of the *Escherichia coli* cell envelope. *J Biol Chem*, 280(41), 34409-34419.
- Stenutz, R., Weintraub, A., & Widmalm, G. (2006). The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol Rev*, 30(3), 382-403.
- Stock, J. B., Rauch, B., & Roseman, S. (1977). Periplasmic space in *Salmonella typhimurium* and *Escherichia coli*. *J Biol Chem*, 252(21), 7850-7861.
- Suginaka, H., Blumberg, P. M., & Strominger, J. L. (1972). Multiple penicillin-binding components in *Bacillus subtilis*, *Bacillus cereus*, *Staphylococcus aureus*, and *Escherichia coli*. *J Biol Chem*, 247(17), 5279-5288.
- Suits, M. D., Sperandeo, P., Deho, G., Polissi, A., & Jia, Z. (2008). Novel structure of the conserved gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *J Mol Biol*, 380(3), 476-488.
- Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol*, 18(10), 464-470.
- Sweeney, R. P., & Lowary, T. L. (2019). New insights into lipopolysaccharide assembly and export. *Curr Opin Chem Biol*, 53, 37-43.
- Tapia-Pastrana, G., Chavez-Dueñas, L., Lanz-Mendoza, H., Teter, K., & Navarro-Garcia, F. (2012). VirK is a periplasmic protein required for efficient secretion of plasmid-encoded toxin from enteroaggregative *Escherichia coli*. *Infection and immunity*, 80(7), 2276-2285.
- Taubert, J., Hou, B., Risselada, H. J., Mehner, D., Lunsdorf, H., Grubmüller, H., & Bruser, T. (2015). TatBC-independent TatA/Tat substrate interactions contribute to transport efficiency. *PLoS One*, 10(3), e0119761.
- Tellez, R., Jr., & Misra, R. (2012). Substitutions in the BamA β -barrel domain overcome the conditional lethal phenotype of a DeltabamB DeltabamE strain of *Escherichia coli*. *J Bacteriol*, 194(2), 317-324.
- Templin, M. F., Ursinus, A., & Holtje, J. V. (1999). A defect in cell wall recycling triggers autolysis during the stationary growth phase of *Escherichia coli*. *EMBO J*, 18(15), 4108-4117.
- Terada, M., Kuroda, T., Matsuyama, S. I., & Tokuda, H. (2001). Lipoprotein sorting signals evaluated as the LolA-dependent release of lipoproteins from the cytoplasmic membrane of *Escherichia coli*. *J Biol Chem*, 276(50), 47690-47694.
- Tian, W., Zhang, L. V., Tasan, M., Gibbons, F. D., King, O. D., Park, J., Wunderlich, Z., Cherry, J. M., & Roth, F. P. (2008). Combining guilt-by-association and guilt-by-

- profiling to predict *Saccharomyces cerevisiae* gene function. *Genome Biol*, 9 Suppl 1, S7.
- Tocheva, E. I., Ortega, D. R., & Jensen, G. J. (2016). Sporulation, bacterial cell envelopes and the origin of life. *Nat Rev Microbiol*, 14(8), 535-542.
- Tokuda, H., & Matsuyama, S. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta*, 1694(1-3), IN1-9.
- Tokunaga, M., Tokunaga, H., & Wu, H. C. (1982). Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. *Proc Natl Acad Sci U S A*, 79(7), 2255-2259.
- Tsang, M. J., Yakhnina, A. A., & Bernhardt, T. G. (2017). NlpD links cell wall remodelling and outer membrane invagination during cytokinesis in *Escherichia coli*. *PLoS Genet*, 13(7), e1006888.
- Tsyganov, K., Perry, A. J., Archer, S. K., & Powell, D. (2018). RNAsik: A Pipeline for complete and reproducible RNA-seq analysis that runs anywhere with speed and ease. *J. Open Source Software*, 3(28), 583.
- Tucker, D. L., Tucker, N., & Conway, T. (2002). Gene expression profiling of the pH response in *Escherichia coli*. *J Bacteriol*, 184(23), 6551-6558.
- Turner, R. D., Mesnage, S., Hobbs, J. K., & Foster, S. J. (2018). Molecular imaging of glycan chains couples cell-wall polysaccharide architecture to bacterial cell morphology. *Nat Commun*, 9(1), 1263. doi:10.1038/s41467-018-03551-y
- Tyanova, S., Temu, T., & Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*, 11(12), 2301-2319.
- Typas, A., Banzhaf, M., Gross, C. A., & Vollmer, W. (2011). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol*, 10(2), 123-136.
- Typas, A., Banzhaf, M., van den Berg van Saparoea, B., Verheul, J., Biboy, J., Nichols, R. J., Zietek, M., Beilharz, K., Kannenberg, K., von Rechenberg, M., Breukink, E., den Blaauwen, T., Gross, C. A., & Vollmer, W. (2010). Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell*, 143(7), 1097-1109.
- Typas, A., Nichols, R. J., Siegele, D. A., Shales, M., Collins, S. R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B. L., Mori, H., Weissman, J. S., Krogan, N. J., & Gross, C. A. (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nat Methods*, 5(9), 781-787.
- Uehara, T., Dinh, T., & Bernhardt, T. G. (2009). LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in *Escherichia coli*. *J Bacteriol*, 191(16), 5094-5107.
- Uehara, T., Parzych, K. R., Dinh, T., & Bernhardt, T. G. (2010). Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. *EMBO J*, 29(8), 1412-1422.
- Ureta, A. R., Endres, R. G., Wingreen, N. S., & Silhavy, T. J. (2007). Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. *J Bacteriol*, 189(2), 446-454.
- van den Bogaart, G., Hermans, N., Krasnikov, V., & Poolman, B. (2007). Protein mobility and diffusive barriers in *Escherichia coli*: consequences of osmotic stress. *Mol Microbiol*, 64(3), 858-871.
- van Heijenoort, J. (2001). Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod Rep*, 18(5), 503-519.
- van Heijenoort, J. (2007). Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev*, 71(4), 620-635.

- Van Vliet, A. H. (2010). Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS microbiology letters*, 302(1), 1-7.
- Vazquez-Laslop, N., Lee, H., Hu, R., & Neyfakh, A. A. (2001). Molecular sieve mechanism of selective release of cytoplasmic proteins by osmotically shocked *Escherichia coli*. *J Bacteriol*, 183(8), 2399-2404.
- Vogt, J., & Schulz, G. E. (1999). The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. *Structure*, 7(10), 1301-1309.
- Volkmer, B., & Heinemann, M. (2011). Condition-dependent cell volume and concentration of *Escherichia coli* to facilitate data conversion for systems biology modeling. *PLoS One*, 6(7), e23126.
- Vollmer, W. (2008). Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiol Rev*, 32(2), 287-306.
- Vollmer, W., & Bertsche, U. (2008). Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochim Biophys Acta*, 1778(9), 1714-1734.
- Vollmer, W., & Holtje, J. V. (2004). The architecture of the murein (peptidoglycan) in gram-negative bacteria: vertical scaffold or horizontal layer(s)? *J Bacteriol*, 186(18), 5978-5987.
- Vollmer, W., & Seligman, S. J. (2010). Architecture of peptidoglycan: more data and more models. *Trends in microbiology*, 18(2), 59-66.
- Vollmer, W., Blanot, D., & de Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiol Rev*, 32(2), 149-167.
- Vuong, P., Bennion, D., Mantei, J., Frost, D., & Misra, R. (2008). Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *J Bacteriol*, 190(5), 1507-1517.
- Wagner, J. K., Heindl, J. E., Gray, A. N., Jain, S., & Goldberg, M. B. (2009). Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of *Shigella flexneri*. *J Bacteriol*, 191(3), 815-821.
- Wang, Y., Ghaffari, N., Johnson, C. D., Braga-Neto, U. M., Wang, H., Chen, R., & Zhou, H. (2011). Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics*, 12 Suppl 10, S5.
- Wang, Z., Fan, G., Hryc, C. F., Blaza, J. N., Serysheva, II, Schmid, M. F., Chiu, W., Luisi, B. F., & Du, D. (2017). An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *Elife*, 6.
- Wasinger, V. C., Zeng, M., & Yau, Y. (2013). Current status and advances in quantitative proteomic mass spectrometry. *Int J Proteomics*, 2013, 180605.
- Webb, C. T., Selkirk, J., Perry, A. J., Noinaj, N., Buchanan, S. K., & Lithgow, T. (2012). Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J Mol Biol*, 422(4), 545-555.
- Weber, A., Kogl, S. A., & Jung, K. (2006). Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*. *J Bacteriol*, 188(20), 7165-7175.
- Weidel, W., & Pelzer, H. (1964). Bag-shaped Macromolecules-a New Outlook on Bacterial Cell Walls. *Adv Enzymol Relat Subj Biochem*, 26, 193-232.
- Weiner, J. H., & Li, L. (2008). Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim Biophys Acta*, 1778(9), 1698-1713.
- Weski, J., & Ehrmann, M. (2012). Genetic analysis of 15 protein folding factors and proteases of the *Escherichia coli* cell envelope. *J Bacteriol*, 194(12), 3225-3233.

- Westblade, L. F., Ilag, L. L., Powell, A. K., Kolb, A., Robinson, C. V., & Busby, S. J. (2004). Studies of the *Escherichia coli* Rsd-sigma70 complex. *J Mol Biol*, 335(3), 685-692.
- Whitfield, C., & Trent, M. S. (2014a). Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem*, 83, 99-128.
- Whitfield, C., & Trent, M. S. (2014b). Biosynthesis and export of bacterial lipopolysaccharides. *Annual review of biochemistry*, 83, 99-128.
- Wimley, W. C. (2003). The versatile beta-barrel membrane protein. *Curr Opin Struct Biol*, 13(4), 404-411.
- Wu, E. L., Engstrom, O., Jo, S., Stuhlsatz, D., Yeom, M. S., Klauda, J. B., Widmalm, G., & Im, W. (2013). Molecular dynamics and NMR spectroscopy studies of *E. coli* lipopolysaccharide structure and dynamics. *Biophys J*, 105(6), 1444-1455.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell*, 121(2), 235-245.
- Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S. S., Silhavy, T. J., & Kahne, D. (2006). Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 103(31), 11754-11759.
- Wu, X., Chavez, J. D., Schweppe, D. K., Zheng, C., Weisbrod, C. R., Eng, J. K., Murali, A., Lee, S. A., Ramage, E., Gallagher, L. A., Kulasekara, H. D., Edrozo, M. E., Kamischke, C. N., Brittnacher, M. J., Miller, S. I., Singh, P. K., Manoil, C., & Bruce, J. E. (2016). In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in *Acinetobacter baumannii* strain AB5075. *Nat Commun*, 7, 13414.
- Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., & Tokuda, H. (2000). A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat Cell Biol*, 2(4), 212-218.
- Yamaguchi, K., Yu, F., & Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell*, 53(3), 423-432.
- Yousif, S. Y., Broome-Smith, J. K., & Spratt, B. G. (1985). Lysis of *Escherichia coli* by beta-lactam antibiotics: deletion analysis of the role of penicillin-binding proteins 1A and 1B. *J Gen Microbiol*, 131(10), 2839-2845.
- Zhang, W. Y., & Wu, H. C. (1992). Alterations of the carboxyl-terminal amino acid residues of *Escherichia coli* lipoprotein affect the formation of murein-bound lipoprotein. *J Biol Chem*, 267(27), 19560-19564.
- Zhou, Z., White, K. A., Polissi, A., Georgopoulos, C., & Raetz, C. R. (1998). Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. *J Biol Chem*, 273(20), 12466-12475.
- Zückert, W. R. (2014). Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1843(8), 1509-1516.

Appendices

Appendix 1.0

lpp₊₂₁ gene block

GTGTAATACTTGTAACGCTACATGGAGATTAAGTCAATCTAGAGGGTATTAATAA
TGAAAGCTACTAAACTGGTACTGGGCGCGGTAATCCTGGGTTCTACTCTGCTGGC
AGGTTGCTCCAGCAACGCTAAAATCGATCAGCTGTCTTCTGACGTTTCAGACTCTG
AACGCTAAAGTTGACACCCTGAGCGCGAAAGTGGAACAGCTGAGCAACGATGTG
AACGCGATGCGCAGCGATGTGGATCAGCTGAGCAACGACGTGAACGCAATGCGT
TCCGACGTTTCAGGCTGCTAAAGATGACGCAGCTCGTGCTAACCAGCGTCTGGAC
AACATGGCTACTAAATACCGCAAGTAAATATGAATATCCTCCTTAGTTCCTATT
CCGAAGTTCCTATTCTCTAGAAAG

Black letters - Original *Lpp* sequence

Red letters - Insertion sequence for the extra 21 base pairs

Letters on green background – Complementary sequence upstream of *Lpp* (-54 to -1 bp).

Letters on yellow background – Complementary sequence to the start of *cat* gene

Appendix 2.0

A comparative analysis of Lpp length variants in other bacterial species.

Non-redundant Lpp sequences were identified in 757 publicly available bacterial genomes, and the sequence length was extracted. The number of Lpp representative sequences (number of entries) available is also shown. The longer Lpp variants compared to *E. coli* Lpp using multiple sequence alignment (Figure 4.2.2B) are shown on a yellow background.

Shared name	Number of IDs rep node	Sequence length	Organism
Q5PH62	58	79	<i>Salmonella paratyphi A</i> (strain ATCC 9150 / SARB42)
A3ULK1	41	88	<i>Vibrio splendidus</i> (strain I2B01)
zzz1037	40	80	<i>Aeromonas salmonicida</i> (strain A449)
zzz1022	31	88	<i>Shewanella</i> sp. (strain ANA-3)
zzz1011	23	84	<i>Vibrio parahaemolyticus</i> serotype O3:K6 (strain AQ3810)
zzz1058	22	80	<i>Salmonella enteritidis</i>
D3VD61	19	78	<i>Xenorhabdus nematophila</i> (strain ATCC 19061 / DSM 3370 / LMG 1036 / NCIB 9965 / AN6)
A0A086WU78	16	82	<i>Vibrio</i> sp. B183
A7N6D3	15	82	<i>Vibrio campbellii</i> (strain ATCC BAA-1116 / BB120)
H2IT70	15	78	<i>Rahnella aquatilis</i> (strain ATCC 33071 / DSM 4594 / JCM 1683 / NBRC 105701 / NCIMB 13365 / CIP 78.65)
zzz1041	15	78	<i>Providencia stuartii</i> ATCC 25827
E5YFX6	14	78	<i>Enterobacteriaceae</i> bacterium 9_2_54FAA
C7BHQ9	13	78	<i>Photobacterium asymbiotica</i> subsp. <i>asymbiotica</i> (strain ATCC 43949 / 3105-77) (<i>Xenorhabdus luminescens</i> (strain 2))
A0A0B7JG20	12	87	<i>Photobacterium phosphoreum</i> ANT-2200
A0A090QK38	11	87	<i>Photobacterium aphoticum</i>
A0A085PVQ9	11	82	<i>Vibrio cholerae</i>
B4EWN9	10	79	<i>Proteus mirabilis</i> (strain HI4320)
Q1ZVT3	9	86	<i>Photobacterium angustum</i> (strain S14 / CCUG 15956) (<i>Vibrio</i> sp. (strain S14 / CCUG 15956))
A6CVF6	9	82	<i>Vibrio shilonii</i> AK1
D3V1T4	9	78	<i>Xenorhabdus bovienii</i> (strain SS-2004)
C6C486	9	78	<i>Dickeya paradisiaca</i> (strain Ech703) (<i>Dickeya dadantii</i> (strain Ech703))
C5BGU3	8	78	<i>Edwardsiella ictaluri</i> (strain 93-146)

Q07YU0	7	88	<i>Shewanella frigidimarina</i> (strain NCIMB 400)
A0A2N7CEX0	6	89	<i>Vibrio splendidus</i>
B1EGS5	6	85	<i>Escherichia albertii</i> (strain TW07627)
H2FWL8	6	80	<i>Oceanimonas</i> sp. (strain GK1)
M1SWN3	6	78	<i>Morganella morganii</i> subsp. <i>morganii</i> KT
zzz1057	6	77	<i>Salmonella agona</i> (strain SL483)
A0A191W9R5	5	89	<i>Vibrio anguillarum</i> (<i>Listonella anguillarum</i>)
A0A0W8JBW2	5	89	<i>Vibrio</i> sp. MEBiC08052
D0Z2W3	5	84	<i>Photobacterium damsela</i> subsp. <i>damsela</i> CIP 102761
F9RE98	5	82	<i>Vibrio</i> sp. (strain N418)
B1EGS4	5	82	<i>Escherichia albertii</i> (strain TW07627)
A0A0F5AMD0	5	81	<i>Salinivibrio</i> sp. KP-1
zzz1007	5	76	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Give str. S5-487
A0A087IRR1	4	92	<i>Vibrio vulnificus</i>
F0LZB4	4	89	<i>Vibrio furnissii</i> (strain DSM 14383 / NCTC 11218)
A0A0B9H4W1	4	87	<i>Photobacterium gaetbulicola</i>
Q1YYZ8	4	87	<i>Photobacterium profundum</i> 3TCK
zzz1008	4	82	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Give str. S5-487
R1J0X1	4	81	<i>Grimontia indica</i>
D0ISE6	4	81	<i>Grimontia hollisae</i> CIP 101886
A0A1N6MYG9	4	79	<i>Xenorhabdus innexi</i>
N9TW73	4	79	<i>Aeromonas diversa</i> CDC 2478-85
E5YFX5	4	78	<i>Enterobacteriaceae</i> bacterium 9_2_54FAA
E9XGQ8	4	76	<i>Escherichia coli</i> TW10509
zzz1064	4	70	<i>Pragia fontium</i>
Q74GR3	3	100	<i>Geobacter sulfurreducens</i> (strain ATCC 51573 / DSM 12127 / PCA)
A0A099LXQ1	3	99	<i>Vibrio navarrensis</i>
U4EAY7	3	96	<i>Vibrio nigripulchritudo</i> FTn2
B1KHF0	3	88	<i>Shewanella woodyi</i> (strain ATCC 51908 / MS32)
A8H160	3	88	<i>Shewanella pealeana</i> (strain ATCC 700345 / ANG-SQ1)
A0A0J8XY86	3	87	<i>Photobacterium swingsii</i>
B7LPY2	3	86	<i>Escherichia fergusonii</i> (strain ATCC 35469 / DSM 13698 / CDC 0568-73)
S0TZC0	3	85	<i>Escherichia</i> sp. KTE114
A0A099KXT4	3	85	<i>Colwellia psychrerythraea</i> (<i>Vibrio psychroerythus</i>)
S0UG29	3	82	<i>Escherichia</i> sp. KTE114
A0A1E5CI50	3	81	<i>Enterovibrio calviensis</i> FF-85
Q2NT13	3	78	<i>Sodalis glossinidius</i> (strain morsitans)

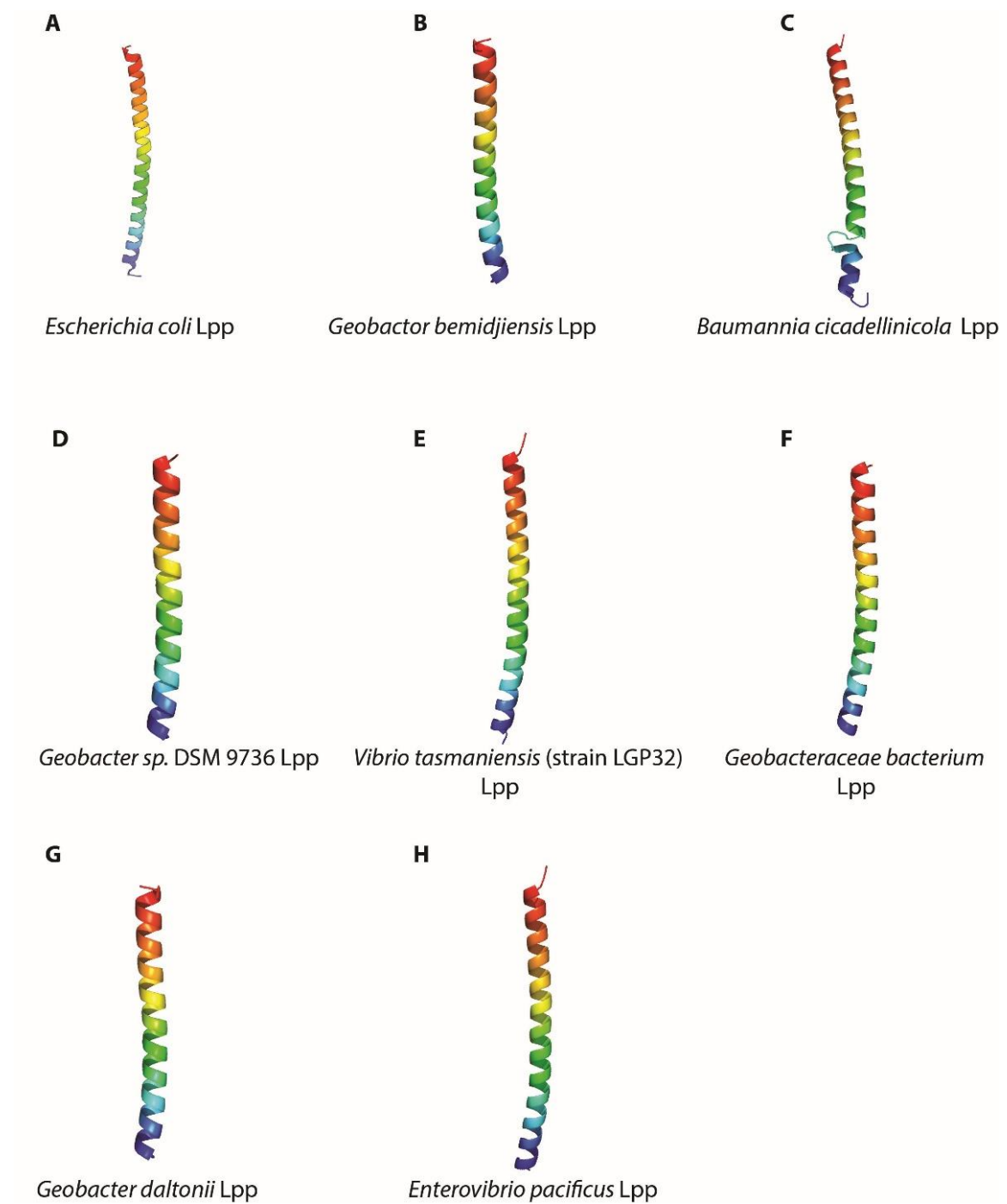
A0A0A1E782	2	96	<i>Vibrio tapetis</i>
A0A099KHZ0	2	95	<i>Thalassotalea</i> sp. ND16A
A0A193KHR2	2	94	<i>Vibrio breoganii</i>
A0A0B8P8Y0	2	94	<i>Vibrio ishigakensis</i>
U3CEK8	2	94	<i>Vibrio ezurae</i> NBRC 102218
C0B489	2	94	<i>Proteus penneri</i> ATCC 35198
A0A2A4VEE4	2	91	<i>Gammaproteobacteria</i> bacterium
A0A2N7H2J0	2	89	<i>Vibrio</i> sp. 10N.261.52.E5
A0A2J8GXY6	2	89	<i>Vibrio diazotrophicus</i>
A0A1M6BGH4	2	89	<i>Vibrio aerogenes</i> CECT 7868
A6FHJ6	2	89	<i>Moritella</i> sp. PE36
A0A1B9R2V5	2	88	<i>Vibrio genomosp.</i> F10
A0A0A1E9S6	2	88	<i>Vibrio aestuarianus</i>
A0A090IDP0	2	88	<i>Moritella viscosa</i>
A0A0S2JIH5	2	85	<i>Colwellia</i> sp. MT41
G4C2F3	2	83	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Infantis</i> str. SARB27
C9QFZ8	2	82	<i>Vibrio orientalis</i> CIP 102891 = ATCC 33934
A0A090P116	2	82	<i>Vibrio ponticus</i>
U3AV12	2	81	<i>Vibrio azureus</i> NBRC 104587
H2IMF9	2	81	<i>Vibrio</i> sp. (strain EJY3)
A0A1J4QFS7	2	80	<i>Oceanisphaera psychrotolerans</i>
zzz1032	2	79	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Cubana</i> str. 76814
W1IMQ8	2	78	<i>Xenorhabdus cabanillasii</i> JM26
L0MW00	2	78	<i>Candidatus Blochmannia chromaiodes</i> str. 640
D2TW17	2	78	<i>Arsenophonus nasoniae</i> (son-killer infecting <i>Nasonia vitripennis</i>)
zzz1006	2	78	<i>Serratia</i> sp. DD3
A0A0C1VJD1	2	76	<i>Candidatus Riesia pediculischaeffi</i> PTSU
R5EUY4	2	75	<i>Succinatimonas</i> sp. CAG:777
A0A085H8H4	2	75	<i>Khuyvera ascorbata</i> ATCC 33433
A0A1V3HT63	2	74	<i>Salinivibrio</i> sp. IB872
R8AVC8	2	74	<i>Plesiomonas shigelloides</i> 302-73
A0A0C5W1C8	2	62	<i>Enterobacteriaceae</i> bacterium bta3-1
M7RFY7	2	60	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Dublin</i> str. UC16
B5EG68	1	121	<i>Geobacter bemidjiensis</i> (strain <i>Bem</i> / ATCC BAA-1014 / DSM 16622)
Q1LTZ8	1	114	<i>Baumannia cicadellinica</i> subsp. <i>Homalodisca coagulata</i>

A0A212PF75	1	113	<i>Geobacter sp. DSM 9736</i>
B7VQG8	1	110	<i>Vibrio tasmaniensis</i> (strain LGP32) (<i>Vibrio splendidus</i> (strain Mel32))
A0A1G0M4A6	1	106	<i>Geobacteraceae bacterium GWC2_55_20</i>
B9M8H1	1	106	<i>Geobacter daltonii</i> (strain DSM 22248 / JCM 15807 / FRC-32)
A0A1C3EKM1	1	101	<i>Enterovibrio pacificus</i>
A0A1G2W8J8	1	100	<i>Desulfuromonadaceae bacterium GWB2_53_15</i>
A0A0K2BKH0	1	100	<i>Candidatus Baumannia cicadellinicola</i>
Q39PY3	1	100	<i>Geobacter metallireducens</i> (strain GS-15 / ATCC 53774 / DSM 7210)
F7YQS6	1	100	<i>Vibrio anguillarum</i> (strain ATCC 68554 / 775) (<i>Listonella anguillarum</i>)
A1ALI3	1	100	<i>Pelobacter propionicus</i> (strain DSM 2379 / NBRC 103807 / OttBd1)
A0A0A8WSX7	1	99	<i>Geobacter sp. OR-1</i>
Q57PQ8	1	97	<i>Salmonella choleraesuis</i> (strain SC-B67)
A5KU37	1	96	<i>Vibrionales bacterium</i> (strain SWAT-3)
A0A2M8H032	1	95	<i>Vibrio sp. HA2012</i>
A0A244CNC1	1	95	<i>Pseudoalteromonas ulvae</i>
A0A2I0F801	1	94	<i>Moritella sp. Urea-trap-13</i>
A0A1A6KWS4	1	94	<i>Vibrio sp. UCD-FRSSP16_10</i>
A0A0B5BDI4	1	92	<i>Geobacter pickeringii</i>
G4C2F4	1	92	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Infantis</i> str. SARB27
A0A1X1MMW4	1	91	<i>Vibrio sp. qd031</i>
A0A1X7MA93	1	90	<i>methanotrophic bacterial endosymbiont of Bathymodiolus sp.</i>
A0A2N7GEM2	1	89	<i>Vibrio sp. 10N.261.55.A7</i>
A0A1Y6IVV0	1	89	<i>Vibrio mangrovi</i>
A0A1T4Q593	1	89	<i>Vibrio cincinnatiensis</i> DSM 19608
A0A1R4B665	1	89	<i>Vibrio palustris</i>
A0A1E5CNJ0	1	89	<i>Vibrio genomosp. F6 str. FF-238</i>
A0A2N7DE01	1	88	<i>Vibrio sp. 10N.286.49.B3</i>
A0A2N7CMB2	1	88	<i>Shewanella sp. 10N.286.51.B7</i>
A0A066UT84	1	88	<i>Vibrio fortis</i>
R9PNI8	1	88	<i>Agarivorans albus</i> MKT 106
K2IUG9	1	88	<i>Gallaecimonas xiamenensis</i> 3-C-1
B0TIP5	1	88	<i>Shewanella halifaxensis</i> (strain HAW-EB4)
A0A1M5NVC6	1	87	<i>Ferrimonas marina</i>

A0A0J1H8G5	1	87	<i>Photobacterium ganghwense</i>
A0A0J1H0Y9	1	87	<i>Photobacterium aquae</i>
L8JDT7	1	87	<i>Photobacterium marinum</i>
A0A090QJP1	1	87	<i>Photobacterium aphoticum</i>
A1S3T4	1	87	<i>Shewanella amazonensis</i> (strain ATCC BAA-1098 / SB2B)
A0A1G8AFV7	1	86	<i>Vibrio xiamenensis</i>
E1SS57	1	86	<i>Ferrimonas balearica</i> (strain DSM 9799 / CCM 4581 / PAT)
A0A090R6A4	1	86	<i>Vibrio</i> sp. C7
A0A2N1I4B5	1	85	<i>Colwellia</i> sp. 12G3
A0A1Y5EDR6	1	85	<i>Colwellia psychrerythraea</i> (<i>Vibrio psychroerythus</i>)
A0A1G8UCA1	1	85	<i>Ferrimonas sediminum</i>
A0A099KF85	1	85	<i>Colwellia psychrerythraea</i> (<i>Vibrio psychroerythus</i>)
A0A0Q0NB62	1	83	<i>Vibrio metoecus</i>
A0A0L6ZWR8	1	83	<i>Escherichia coli</i>
Q8L1C4	1	83	<i>Photobacterium damsela</i> subsp. <i>piscicida</i> (<i>Pasteurella piscicida</i>)
E8LLB1	1	83	<i>Succinatimonas hippei</i> (strain DSM 22608 / JCM 16073 / KCTC 15190 / YIT 12066)
A0A2N7DD41	1	82	<i>Vibrio</i> sp. 10N.286.49.C2
U3BEX2	1	82	<i>Vibrio proteolyticus</i> NBRC 13287
C9P848	1	82	<i>Vibrio metschnikovii</i> CIP 69.14
C4L7B0	1	82	<i>Tolumonas auensis</i> (strain DSM 9187 / TA4)
A0A2N1DRC4	1	81	<i>Psychromonas</i> sp. MB-3u-54
A0A1V3H981	1	81	<i>Salinivibrio</i> sp. MA351
A0A0H2YJB4	1	81	<i>Yersinia pestis</i> bv. <i>Antiqua</i> (strain Nepal516)
A1SXX7	1	81	<i>Psychromonas ingrahamii</i> (strain 37)
K8WIU4	1	80	<i>Providencia sneebia</i> DSM 19967
A4CAR0	1	80	<i>Pseudoalteromonas tunicata</i> D2
A0A1Y0D1W1	1	79	<i>Oceanisphaera profunda</i>
A0A1Y0CUN6	1	79	<i>Oceanisphaera avium</i>
A0A1S1HL96	1	79	<i>Providencia stuartii</i>
A0A133LC04	1	79	<i>Citrobacter freundii</i>
E2X7C4	1	79	<i>Shigella dysenteriae</i> 1617
A0A2I5T7D2	1	78	<i>Serratia</i> sp. (strain ATCC 39006)
A0A2G8BUL7	1	78	<i>Tatumella</i> sp. OPLPL6
A0A1Y2SDF5	1	78	<i>Xenorhabdus vietnamensis</i>
A0A068RAP6	1	78	<i>Serratia symbiotica</i> SCt-VLC
A0A1B7JH63	1	78	<i>Proteus hauseri</i> ATCC 700826
A0A151ULG4	1	78	<i>Sodalis</i> -like endosymbiont of <i>Proechinophthirus fluctus</i>

A0A143WSP4	1	78	<i>Candidatus Doolittlea endobia</i>
A0A0X9VR02	1	78	<i>Candidatus Arsenophonus lipoptenae</i>
P09461	1	78	<i>Proteus mirabilis</i>
A0A0E3Y6I7	1	78	<i>Blochmannia endosymbiont of Polyrhachis (Hedomyrma) turneri</i>
R4HZU6	1	78	<i>Serratia symbiotica</i> str. 'Cinara cedri'
Q7VR58	1	78	<i>Blochmannia floridanus</i>
K8WYB4	1	78	<i>Providencia burhodogranariea</i> DSM 19968
H6Q4W1	1	78	<i>Wigglesworthia glossinidia endosymbiont of Glossina morsitans morsitans</i> (Yale colony)
G2H248	1	78	<i>Candidatus Regiella insecticola</i> 5.15
E8Q667	1	78	<i>Blochmannia vafer</i> (strain BVAf)
Q8D2J9	1	77	<i>Wigglesworthia glossinidia brevipalpis</i>
D4G7U1	1	77	<i>Riesia pediculicola</i> (strain USDA)
P02938	1	77	<i>Serratia marcescens</i>
A0A1V0HPH6	1	76	<i>Candidatus Riesia</i> sp. GBBU
D7XFN1	1	73	<i>Escherichia coli</i> MS 198-1
E3BJG3	1	71	<i>Vibrio caribbeanicus</i> ATCC BAA-2122
A0A227J9P2	1	70	<i>Vibrio parahaemolyticus</i>
A0A2C6DPH4	1	69	<i>Budvicia aquatica</i>
U3U387	1	69	<i>Candidatus Pantoea carbekii</i>
A0A1E5FEB6	1	67	<i>Vibrio splendidus</i> 12E03
A0A085GAH7	1	62	<i>Ewingella americana</i> ATCC 33852
A0A1E8DXA0	1	60	<i>Salmonella enterica</i> I
G5LPG3	1	58	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Alachua str. R6-377
G5L9T8	1	58	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Adelaide str. A4-669
A0A2J4Y9F5	1	48	<i>Klebsiella michiganensis</i>
I6H3K6	1	38	<i>Shigella flexneri</i> 1235-66

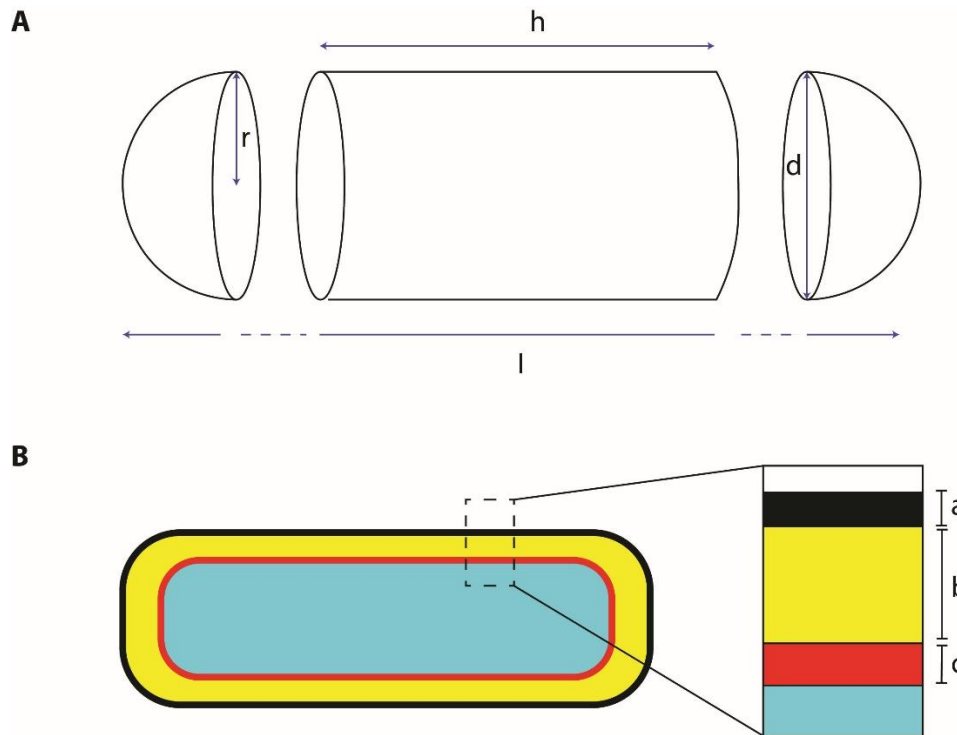
Appendix 3.0



Phyre2 prediction structures of longer Lpp homologues from other species.

(A) *E. coli* Lpp monomer (PDB 1EQ7) in comparison with phyre2 predicted structures of longer Lpp homologues from the indicated species in panel B-H. The homologues are predicted to adopt an α -helical structure similar to *E. coli* Lpp with high confidence (>90%).

Appendix 4.0



C

Length of cylinder (h): $l - (2r)$

Volume of cylinder: $\pi r^2 h$

Volume of the hemisphere $\frac{2}{3} \pi r^3$

Total volume (OM+PP+IM+CP) = $(\pi r^2 h) + (\frac{4}{3} \pi r^3)$

Volume (PP+IM+CP) = $(\pi [r-a]^2 h) + (\frac{4}{3} \pi [r-a]^3)$

Volume (IM+CP) = $(\pi [r-a-b]^2 h) + (\frac{4}{3} \pi [r-a-b]^3)$

Volume (CP) = $(\pi [r-a-b-c]^2 h) + (\frac{4}{3} \pi [r-a-b-c]^3)$

D

E. coli (BW25113) single cell dimensions

Diameter (d) * = $1.3 \mu\text{m}$

Radius (r) = $d/2 = 0.65 \mu\text{m}$

$$\text{Total length (I)}^* = 3.9 \mu\text{m}$$

$$\text{Length cylinder (h)} = 2.6 \mu\text{m}$$

$$\text{Thickness OM (a)}^{**} = 0.007$$

$$\text{Thickness PP (b)}^{**} = 0.021 \mu\text{m} \text{ (wild-type); } 0.025 \mu\text{m}^{***} \text{ (Lpp}_{+21}\text{)}$$

$$\text{Thickness IM (c)}^{**} = 0.006 \mu\text{m}$$

E

Wild-type

$$\text{Volume (PP+IM+CP): } (\pi [r-a]^2 h) + (4/3 \pi [r-a]^3)$$

$$= (3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m}]^2 \times 2.6 \mu\text{m}) + (4/3 \times 3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m}]^3)$$

$$= 4.49 \mu\text{m}^3$$

$$\text{Volume (IM+CP): } (\pi [r-a-b]^2 h) + (4/3 \pi [r-a-b]^3)$$

$$= (3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m} - 0.021 \mu\text{m}]^2 \times 2.6 \mu\text{m}) + (4/3 \times 3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m} - 0.021 \mu\text{m}]^3)$$

$$= 4.17 \mu\text{m}^3$$

$$\text{PP} = (\text{PP+IM+CP}) - (\text{IM+CP})$$

$$= 4.49 \mu\text{m}^3 - 4.17 \mu\text{m}^3$$

$$\# = \underline{\underline{0.32 \mu\text{m}^3}}$$

Lpp₊₂₁

$$\text{Volume (PP+IM+CP): } (\pi [r-a]^2 h) + (4/3 \pi [r-a]^3)$$

$$= (3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m}]^2 \times 2.6 \mu\text{m}) + (4/3 \times 3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m}]^3)$$

$$= 4.49 \mu\text{m}^3$$

$$\text{Volume (IM+CP): } (\pi [r-a-b]^2 h) + (4/3 \pi [r-a-b]^3)$$

$$\begin{aligned}
&= (3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m} - 0.024 \mu\text{m}]^2 \times 2.6 \mu\text{m}) + (4/3 \times 3.14 \times [0.65 \mu\text{m} - 0.007 \mu\text{m} - 0.024 \mu\text{m}]^3) \\
&= 4.10 \mu\text{m}^3
\end{aligned}$$

$$\begin{aligned}
\text{PP} &= (\text{PP} + \text{IM} + \text{CP}) - (\text{IM} + \text{CP}) \\
&= 4.49 \mu\text{m}^3 - 4.10 \mu\text{m}^3 \\
&= \underline{\underline{0.39 \mu\text{m}^3}}
\end{aligned}$$

Periplasm volume calculations for the wild-type and Lpp₊₂₁ *E. coli* strains

(A) The rod shape of *E. coli* is considered a cylinder flanked by two hemispheres. The dimensions for total volume calculations are shown (Prochnow *et al.* 2019). (B) Model depicting the distributions of the bacterial compartments (Prochnow *et al.* 2019). (C) How the equations for calculating the various compartmental volumes were derived based on panel A cylinder depiction. The equations for each compartment volume are shown (Prochnow *et al.* 2019). (D) Cell dimensions of *E. coli* BW25113 (same strain used in this study). * values according to Volkmer and Heinemann *et al.* (2011); ** Values according to Matias *et al.* (2003), *** Values according to Asmar *et al.* (2017) and Cohen *et al.* (2017). (E) Periplasm volumes calculations using the parameters depicted in (C) and (D) for the wild-type and Lpp₊₂₁ strains. [#] Reported wild-type periplasmic volume for *E. coli* (BW25113) (Prochnow *et al.* 2019). The increase in size of the periplasm in Lpp₊₂₁ is incorporated in the calculations, shown as text in blue background. Abbreviations: OM, outer membrane; PP, periplasm; IM, inner membrane; CP, cytoplasm.

Appendix 5.0

Over-expressed and diminished proteins in *E. coli* wild-type; M9 (+sorbitol) vs M9 (-sorbitol)

Proteins were sorted based on the log₂ fold change (FC) in descending order. Negative values represent diminished/ down-expressed proteins, while positive values depict over-expressed proteins. The volcano plots in the figure above were generated from the log₂ fold change, and the P-values included in the tables.

Gene Name	Protein IDs	log ₂ FC	P-value	P-value adjusted	Protein names
					Decreased steady-state
<i>proV</i>	P14175	-8.65	3.09E-08	5.01E-06	Glycine betaine/L-proline transport ATP-binding protein ProV
<i>proW</i>	P14176	-7.21	2.27E-09	8.84E-07	Glycine betaine/L-proline transport system permease protein ProW
<i>proX</i>	P0AFM2	-6.2	2.16E-08	4.20E-06	Glycine betaine-binding periplasmic protein
<i>ygdI</i>	P65292	-5.1	5.71E-06	0.000179	Uncharacterized lipoprotein YgdI
<i>srlD</i>	P05707	-4.84	8.78E-08	9.00E-06	Sorbitol-6-phosphate 2-dehydrogenase
<i>srlB</i>	P05706	-4.75	7.57E-08	8.67E-06	Glucitol/sorbitol-specific phosphotransferase enzyme IIA component
<i>metR</i>	P0A9F9	-4.17	1.00E-08	2.44E-06	HTH-type transcriptional regulator MetR
<i>nrdE</i>	P39452	-3.57	2.07E-09	8.84E-07	Ribonucleoside-diphosphate reductase 2 subunit alpha
<i>lysP</i>	P25737	-3.26	6.94E-08	8.45E-06	Lysine-specific permease
<i>srlE</i>	P56580	-3.26	0.000284	0.00237	Glucitol/sorbitol-specific phosphotransferase enzyme IIB component
<i>tam</i>	P76145	-2.97	0.00282	0.0122	Trans-aconitate 2-methyltransferase
<i>yhdP</i>	P46474	-2.94	0.000133	0.00151	Uncharacterized protein YhdP
<i>betA</i>	P17444	-2.85	2.62E-05	0.00049	Oxygen-dependent choline dehydrogenase
<i>mgo</i>	P33940	-2.79	0.000444	0.00319	Malate:quinone oxidoreductase
<i>nuoK</i>	P0AFE4	-2.74	0.00335	0.0136	NADH-quinone oxidoreductase subunit K
<i>dsbD</i>	P36655	-2.63	0.00188	0.00905	Thiol:disulfide interchange protein DsbD
<i>ygeR</i>	Q46798	-2.55	0.00109	0.00596	Uncharacterized lipoprotein YgeR
<i>ybhB</i>	P12994	-2.54	0.00744	0.0248	UPF0098 protein YbhB
<i>otsB</i>	P31678	-2.33	0.00192	0.00914	Trehalose-6-phosphate phosphatase
<i>yceA</i>	P24188	-2.33	0.000419	0.00307	UPF0176 protein YceA

<i>ispU</i>	P60472	-2.28	0.00619	0.0215	Ditrans, polycis-undecaprenyl-diphosphate synthase ((2E,6E)-farnesyl-diphosphate specific)
<i>holE</i>	P0ABS8	-2.21	0.000559	0.00369	DNA polymerase III subunit theta
<i>nrdF</i>	P37146	-2.21	0.00237	0.0107	Ribonucleoside-diphosphate reductase 2 subunit beta
<i>yfaZ</i>	P76471	-2.17	0.00225	0.0103	Uncharacterized protein YfaZ
<i>adhP</i>	P39451	-2.15	0.000608	0.0039	Alcohol dehydrogenase, propanol-preferring
<i>artQ</i>	P0AE34	-2.06	0.00777	0.0253	Arginine ABC transporter permease protein ArtQ
<i>surE</i>	P0A840	-2.00	7.85E-06	0.000225	5/3-nucleotidase SurE
					Increased steady-state
<i>mcrB</i>	P15005	6.58	3.85E-07	2.88E-05	5-methylcytosine-specific restriction enzyme B
<i>cusC</i>	P77211	6.15	2.98E-06	0.000114	Cation efflux system protein CusC
<i>fucI</i>	P69922	5.33	0.00186	0.009	L-fucose isomerase
<i>garL</i>	P23522	5.02	3.03E-09	9.85E-07	5-keto-4-deoxy-D-glucarate aldolase
<i>cusF</i>	P77214	4.92	0.00016	0.00172	Cation efflux system protein CusF
<i>fadM</i>	P77712	4.74	2.36E-06	9.98E-05	Long-chain acyl-CoA thioesterase FadM
<i>dmsB</i>	P18776	4.5	0.00047	0.00332	Anaerobic dimethyl sulfoxide reductase chain B
<i>dcuC</i>	P0ABP3	4.34	0.0011	0.00603	Anaerobic C4-dicarboxylate transporter DcuC
<i>cobS</i>	P36561	4.18	7.89E-05	0.00105	Adenosylcobinamide-GDP ribazoletransferase
<i>yejF</i>	P33916	4.07	2.82E-05	0.000518	Uncharacterized ABC transporter ATP-binding protein YejF
<i>nirB</i>	P08201	3.98	0.00953	0.0297	Nitrite reductase (NADH) large subunit
<i>amiB</i>	P26365	3.9	5.19E-06	0.000171	N-acetylmuramoyl-L-alanine amidase AmiB
<i>cusB</i>	P77239	3.8	0.0128	0.0373	Cation efflux system protein CusB
<i>yciI</i>	P0AB55	3.79	5.91E-08	7.77E-06	Protein YciI
<i>ynfF</i>	P77783	3.73	0.000461	0.00328	Probable dimethyl sulfoxide reductase chain YnfF
<i>bioD2</i>	P0A6E9	3.66	0.00283	0.0122	ATP-dependent dethiobiotin synthetase BioD 2
<i>puuA</i>	P78061	3.38	6.35E-06	0.000192	Gamma-glutamylputrescine synthetase PuuA
<i>hypB</i>	P0AAN3	3.31	0.00678	0.0232	Hydrogenase isoenzymes nickel incorporation protein HypB
<i>pspE</i>	P23857	3.31	1.52E-05	0.000336	Thiosulfate sulfurtransferase PspE
<i>focA</i>	P0AC23	3.29	0.000116	0.00141	Probable formate transporter 1
<i>napA</i>	P33937	3.28	2.45E-05	0.000478	Periplasmic nitrate reductase
<i>nsrR</i>	P0AF63	3.21	0.00152	0.00773	HTH-type transcriptional repressor NsrR
<i>puuC</i>	P23883	3.2	1.32E-05	0.000307	Aldehyde dehydrogenase PuuC
<i>yhbT</i>	P64599	3.16	0.0012	0.0064	Uncharacterized protein YhbT
<i>fumB</i>	P14407	3.14	0.00322	0.0132	Fumarate hydratase class I, anaerobic

<i>puuD</i>	P76038	3.1	8.45E-08	9.00E-06	Gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD
<i>macA</i>	P75830	2.98	5.66E-06	0.000179	Macrolide export protein MacA
<i>glpD</i>	P13035	2.89	0.00138	0.00716	Aerobic glycerol-3-phosphate dehydrogenase
<i>gcvH</i>	P0A6T9	2.71	0.000195	0.00196	Glycine cleavage system H protein
<i>ydhV</i>	P76192	2.63	0.00192	0.00914	Uncharacterized oxidoreductase YdhV
<i>moaD</i>	P30748	2.53	0.0148	0.0417	Molybdopterin synthase sulfur carrier subunit
<i>yjiI</i>	P37342	2.49	0.0084	0.0269	Uncharacterized protein YjiI
<i>hypE</i>	P24193	2.48	3.97E-06	0.000138	Hydrogenase isoenzymes formation protein HypE
<i>rbsA</i>	P04983	2.47	1.61E-06	7.83E-05	Ribose import ATP-binding protein RbsA
<i>plaP</i>	P0AA47	2.44	0.000175	0.00184	Low-affinity putrescine importer PlaP
<i>rfaJ</i>	P27129	2.42	9.07E-06	0.000245	Lipopolysaccharide 1,2-glucosyltransferase
<i>cytR</i>	P0ACN7	2.35	0.000107	0.00133	HTH-type transcriptional repressor CytR
<i>feoA</i>	P0AEL3	2.3	0.00506	0.0184	Ferrous iron transport protein A
<i>ynfG</i>	P0AAJ1	2.26	0.0153	0.043	Probable anaerobic dimethyl sulfoxide reductase chain YnfG
<i>phoB</i>	P0AFJ5	2.24	1.47E-08	3.18E-06	Phosphate regulon transcriptional regulatory protein PhoB
<i>ydfK</i>	P76154	2.23	0.000214	0.00205	Uncharacterized protein YdfK;Uncharacterized protein YnaE
<i>fadD</i>	P69451	2.21	0.00179	0.00878	Long-chain-fatty-acid--CoA ligase
<i>elaA</i>	P0AEH3	2.18	0.0168	0.046	Protein ElaA
<i>dicA</i>	P06966	2.17	0.000237	0.00217	HTH-type transcriptional regulator DicA
<i>birA</i>	P06709	2.14	0.00213	0.00996	Bifunctional ligase/repressor BirA
<i>gatD</i>	P0A9S3	2.13	0.00045	0.00322	Galactitol-1-phosphate 5-dehydrogenase
<i>yoaC</i>	P64490	2.09	0.0041	0.0158	Uncharacterized protein YoaC
<i>yjiU</i>	P39407	2.07	0.00113	0.0061	Uncharacterized protein YjiU
<i>lrhA</i>	P36771	2.06	2.10E-05	0.00044	Probable HTH-type transcriptional regulator LrhA
<i>fucO</i>	P0A9S1	2.03	0.00376	0.0148	Lactaldehyde reductase

