Investigating the Citrobacter rodentium NleGs: type III secretion system effector proteins with ubiquitin ligase activity

Agnes Matilde Sågfors

MRC Centre for Molecular Bacteriology and Infection

Imperial College London

Supervisor: Professor Gad Frankel

Thesis submitted to the Department of Life Sciences, Imperial College London, in partial fulfilment of the requirements for the degree of Doctor of Philosophy, September 2018.

Declaration of Originality

I, Agnes Matilde Sågfors, confirm that this thesis represents my own work and that any external contributions are duly acknowledged. The common pronoun "we" is used throughout, but refers to myself unless clearly stated otherwise.

Agnes Matilde Sågfors

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

Enteropathogenic and enterohaemorrhagic *Escherichia coli* (EPEC and EHEC) are important agents of infectious diarrhoea in humans. Together with the murine-restricted pathogen *Citrobacter rodentium*, EPEC and EHEC are attaching and effacing (A/E) pathogens that cause hallmark A/E lesions upon colonisation. Their pathogenesis is underpinned by the injection of effector proteins into the host cytosol via the type III secretion system (T3SS), orchestrating subversion of several eukaryotic pathways. However, the individual roles of many effector proteins are not fully understood. Among these are the NleG family of effectors proteins, which harbour a conserved C-terminus that structurally resembles a eukaryotic E3 ubiquitin ligase domain. *C. rodentium* has three NleG homologs: NleG1, NleG7 and NleG8, which are translocated into eukaryotic cells via the T3SS as demonstrated herein. Biochemical investigations confirmed that NleG1 and NleG8 exhibit E3 ligase activity, reliant on a conserved proline residue P163/P177 that mediates the interaction between E3 and E2 enzymes. As bacteria lack a ubiquitin system, the NleGs are thought to exploit the host ubiquitin code, a notion supported by our findings that NleG7 and NleG8 cause re-localisation of ubiquitin in cultured cells. Both NleG7 and NleG8 bind ubiquitin by yeast two-hybrid assay and the enzymatic activity of NleG8 is required for cytotoxicity in a yeast model. *In vivo*, single Δ*nleG* and double Δ*nleG7* Δ*nleG8* gene deletions did not affect *C. rodentium* colonisation of the murine gut, but a triple *nleG* deletion mutant demonstrated attenuated colonisation, suggesting that NleGs contribute to *C. rodentium* virulence. Reduced faecal Lipocalin 2, an antimicrobial peptide and marker of gut inflammation, was also observed. Significant efforts to find NleG substrates by yeast two-hybrid and mass spectrometry-based screens using transgenic cell lines are outlined, and this work provides a strong foundation for further studies into NleG biology to give insight into A/E pathogenesis.

Acknowledgements

I would like to extend a sincere thank you to Gadi Frankel for the opportunity to undertake my PhD in your research group and for all the encouragement, insight and advice that you have given over the years. I am also grateful to Valérie Crepin-Sevenou, my friend and former mentor in the lab, for teaching me the skills I needed to complete the PhD and for your patience, guidance and camaraderie along the way. Thank you, I hope you know how much your support means to me.

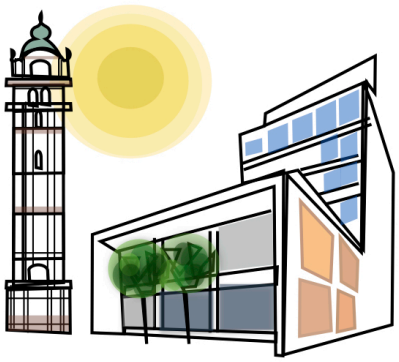
Many thanks to the Biotechnology and Biological Sciences Research Council for funding my PhD project and to all those who directly contributed to it. Maryam Habibzay and the Flow Cytometry Facility at Imperial College London kindly helped with the immunophenotyping. Olga Kotik and Bevin McGeever worked beside me on the NleGs and I am thankful for their efforts and for the tools that we shared. Thank you to Min Goo Lee and Ramanujan Hegde for providing plasmids and to Stacey Schaefer and Randy Hall for the PDZ gel overlay data. I would also like to thank Jyoti Choudhary and her lab for the TMT 10-plex collaboration, and Cedric Berger for his advice and help with data analysis. The work described in the final chapter of this thesis would not have been possible without our collaborators Timurs Maculins, Florian Bonn and Ivan Đikić. A big thank you to the Đikić lab for making me feel welcome in your group, with particular thanks to Evgenij Fiškin for his friendship and support – я не могла бы сделать это без тебя.

I would like to acknowledge the CMBI community and past and present members of the Frankel lab for their scientific- and moral support. Heartfelt thanks to Danielle Carson, for your friendship at work and for the adventures outside. My PhD experience would have been vastly less interesting without Pippa Goddard, and I welcome (almost) all of your ridiculousness in my life. Alexander Byrne and Chris Furniss, thank you for the help you gave and for the brotherly love. Jo and Ernie So, thank you for your wisdom and the shared love of power ballads. Izabela Glegola-Madejska, the *in vivo* experiments I did with you and Valérie are some of my most treasured PhD memories and your generous help was instrumental to my work. Thank you Sabrina Slater for your support and sarcasm and thank you William Pearson for your friendship and encouragement. Thanks also to Rachael Barry, David Ruano-Gallego, Gunnar Schroeder, Matthew Eldridge, Carrie Mullineaux-Sanders, Louise Kerry, Joshua Wong, Elli Mylona, Becky Johnson, Dominic Pollard and Jayne Watson – it was a pleasure working with you.

Importantly, I could not have completed my PhD without the love and support that I received from my dear family. Thank you so much to my parents, my sisters, my brother-in-law and my close friends Ella Taylor, Jessica Berry, Leena Scott and Emily McGovern. I am so lucky to have you. Tusen tack för allt stöd.

Finally, I would like to express my gratitude to Cajsa and Anders Borg. During the second half of my PhD I had the pleasure of regularly babysitting for the Borg family. Having a space to switch off from any science-related worries was invaluable to me, and I feel privileged to have Bea, Nina and Siggy in my life. While I was writing this thesis, Siggy caught an EHEC infection and developed HUS – giving his family a dreadful fright before he started making his recovery.

Siggy, we do this work for children like you.



The Queen's Tower and the Flowers Building,

Imperial College London.

Table of contents

Declaration of originality and copyright declaration ***2***

Abstract ***3***

Acknowledgements ***4***

List of figures ***10***

List of tables ***14***

Abbreviations ***15***

Chapter 1: Introduction ***21***

1.1. Bacterial pathogenesis 22

1.2. *Escherichia coli* 22

1.3. Intestinal pathogenic *E. coli* 23
1.3.1. Enterotoxigenic *E. coli* (ETEC) 24
1.3.2. Diffusely adherent *E. coli* (DAEC) 24
1.3.3. Adherent invasive *E. coli* (AIEC) 25
1.3.4. Enteroaggregative *E. coli* (EAEC) 25
1.3.5. Enteroinvasive *E. coli* (EIEC) 26
1.3.6. Enteropathogenic *E. coli* (EPEC) 26
1.3.7. Enterohaemorrhagic *E. coli* (EHEC) 27
1.3.8. Shiga toxin producing Enteroaggregative *E. coli* (STEAEC) 29

1.4. Attaching and Effacing pathogens 29
1.4.2. Animal models of A/E pathogenesis 30
1.4.2.1. *C. rodentium* infection of mice 31
1.4.2.2. Host immune responses to *C. rodentium* infection 32

1.5. Bacterial secretion systems 35
1.5.1. General secretion mechanisms 35
1.5.2. The type II secretion system 37
1.5.3. The type V secretion system 37
1.5.4. The type I secretion system 38
1.5.5. The type IV secretion system 39
1.5.6. The type VI Secretion System 40
1.5.7. The type VII secretion system 42
1.5.8. The type III secretion system 42
1.5.8.1. The Locus of Enterocyte Effacement-encoded T3SS 44
1.5.8.2. Structure and assembly of the T3SS of A/E pathogens 44
1.5.8.3. T3SS chaperones 46
1.5.8.4. Substrate switch 47
1.5.8.5. Regulation of the LEE 47
1.5.8.6. LEE-encoded T3SS effector proteins 50
1.5.8.7. Non-LEE-encoded T3SS effector proteins 53

1.6. Ubiquitination in health and disease 57
1.6.1. Ubiquitination is a three-step cascade 57
1.6.2. The complexity of the ubiquitin code 58
1.6.3. E3 ligase proteins 60
1.6.4. Bacterial manipulation of the ubiquitination cycle 61
1.6.4.1. E3 ubiquitin ligase mimicry in the host-pathogen arms race 62
1.6.4.2. Novel ubiquitin ligases in mammalian pathogens 63

1.6.4.3. Bacterial HECT-type E3 ligases 64
1.6.4.4. *Legionella* effector proteins hijack host ubiquitination events 65
1.6.4.5. Non-canonical ubiquitination by bacterial pathogens 66
1.6.5. Ubiquitin-rich aggregation formation during bacterial infection 67

1.7. The NleG family of T3SS effector proteins 70
1.7.1. Identification of the NleG family 70
1.7.2. NleG proteins are ubiquitin E3 ligases 71
1.7.3. PDZ-binding motifs in NleG proteins 72

1.8. Aims of this project 72

Chapter 2: Materials and methods 73

2.1. Cells and strains 74
2.1.1. Mammalian cell culture 74
2.1.2. Strains and growth conditions 74
2.1.3. Growth assays 79

2.2. DNA manipulations 79
2.2.1. Plasmids 79
2.2.2. Polymerase chain reaction (PCR) 80
2.2.3. Agarose gel electrophoresis 83
2.2.4. Restriction enzyme digestions and ligations 83
2.2.5. Bacterial transformation by heat shock 83
2.2.6. Bacterial transformation by electroporation 84
2.2.7. Transformation of yeast 84
2.2.8. Chromosomal mutagenesis using pKNG101 84
2.2.9. Chromosomal mutagenesis using pSEVA-612S 85
2.2.10. Generation of stable cell lines 86
2.2.11. *In vitro* transfection 86

2.3. Yeast assays 87
2.3.1. Yeast two-hybrid screening 87
2.3.2. Yeast direct two-hybrid assay 88
2.3.3. Yeast cytotoxicity assay 88

2.4. *In vitro* infection 88

2.5. Immunofluorescent staining 89

2.6. TEM translocation assay 89

2.7. Paraformaldehyde fixation of cultured cells 90

2.8. Methanol fixation of cultured cells 90

2.9. Animal experiments 92
2.9.1. Mouse husbandry 92
2.9.2. Oral infection of mice 92
2.9.3. Lipocalin 2 ELISAs 92
2.9.4. Enterocyte isolation 93
2.9.5. Enterocyte protein extraction 93
2.9.6. Colonic crypt measurements. 93
2.9.7. Immunofluorescent staining of tissue sections 93
2.9.8. Tissue processing for immunophenotyping 94
2.9.9. Multicolour flow cytometry 94

2.10. Protein analysis 95
2.10.1. SDS-PAGE 95
2.10.2. Western blotting 95
2.10.3. Protein BCA assay 96

2.10.4. Protein expression and purification 96
2.10.5. Ubiquitination assays 97
2.10.6. Polyubiquitinated protein enrichment assays 97
2.10.7. FLAG fusion protein immunoprecipitation assays 98
2.10.8. HA fusion protein immunoprecipitation assays 98
2.10.9. Processing of IP samples for mass spectrometry 99
2.10.10. Label free mass spectrometry 100
2.10.11. Enterocyte digestion and TMT labelling for MS 101
2.10.12. Enterocyte proteome MS experiments 101
2.10.13. Analysis of mass spectrometry data 101
2.11. Bioinformatics 102
2.12. Statistical analyses 102

Chapter 3: In vitro characterisation of the C. rodentium NleG proteins 103

3.1. Introduction 104
3.2. NleG proteins are conserved across pathogenic bacteria 104
3.3. *C. rodentium* NleG1, NleG7 and NleG8 are translocated via the T3SS 109
3.4. Assessing the enzymatic activity of *C. rodentium* NleG proteins 111
3.4.1. Purification of recombinant NleG protein for *in vitro* assays 111
3.4.2. *C. rodentium* NleG proteins have E3 ligase activity 114
3.5. The C-terminus of NleG8 mediates yeast cytotoxicity dependent on its E3 ligase activity 116
3.6. The localisation of *C. rodentium* NleG proteins during transfection of cultured cells 118
3.7. Identifying and validating putative interacting partners of NleG proteins by yeast-two-hybrid methods 120
3.7.1. Assessing the suitability of NleGs for yeast two-hybrid screening for interacting partners 122
3.7.2. NleG7 interacts with GRASP55 and WDR61 in yeast 123
3.7.3. *C. rodentium* and EHEC NleG7 interact with an uncharaterised putative membrane protein 127
3.7.4. The interaction of *C. rodentium* NleG7 and NleG8 with NHERF1 and NHERF2 in a yeast two-hybrid assay 129
3.7.5. Full-length NleG7 and NleG8 associate with ubiquitin in yeast 130
3.7.6. The interaction of NleG8 and StoD with the small ubiquitin-like modifier ISG15 131
3.8. Ectopic expression of NleG7 and NleG8 in HeLa cells causes ubiquitin relocalisation 132
3.9. Discussion 133

Chapter 4: Characterising NleG proteins during in vivo C. rodentium infection 139

4.1. Introduction 140
4.2. Making chromosomal *nleG* mutants in C. rodentium 140
4.3. *C. rodentium nleG* deletion mutants are capable of pedestal formation during infection of cultured cells 141
4.4. Immune cell recruitment during *C. rodentium* infection 143
4.5. The role of NleG1 in *C. rodentium* colonisation 148
4.6. The role of NleG1 in immune cell recruitment in response to *C. rodentium* infection 148
4.7. The role of NleG1 in Lcn2 release during *C. rodentium* infection 150

4.8. A TMT 10plex proteomics approach to study the impact of NleG1 on the host proteome during *in vivo* infection 152
4.9. The role of NleG8 in *C. rodentium* colonisation 154
4.10. *C. rodentium* lacking nleG8 induces colonic crypt hyperplasia 156
4.11. The effect of *nleG8* deletion on myeloid cell recruitment 156
4.12. The impact of NleG8 overexpression on *C. rodentium* colonisation 157
4.13. Δ*nleG7* Δ*nleG8 C. rodentium* stimulates less Lcn2 release than WT *C. rodentium* 157
4.14. The effect of *C. rodentium* infection on endogenous ubiquitin 159
4.15. In vivo phenotypes of a triple *nleG* deletion mutant *C. rodentium* 161
4.15.1. Reduced colonisation of a triple *nleG* deletion mutant and lower concentrations of Lcn2 in stools 162
4.16. Discussion 165

Chapter 5: Identifying interactors of C. rodentium NleGs using mass spectrometry..... 168

5.1. Introduction 169
5.2. A mass spectrometry-based strategy to identify NleG interactors 170
5.3. Generation of stably transduced cell lines 171
5.4. Immunoprecipitation of NleG-containing complexes 172
5.5. Mass spectrometry analysis identifies putative interactors of NleG1 and NleG8 174
5.5.1. Novel putative interactors of NleG1 include UBE2O and C1QBP 175
5.5.2. Identifying putative interactors of NleG8 183
5.5.3. UBE2O, C1QBP and CAPER and reproducibly pulled down with NleG1 but not NleG8 in mammalian cells 190
5.6. Validating NleG1 interactions with C1QBP and UBE2O by repeated immunoprecipitations and western blot analysis 194
5.7. The interaction of NleG1 and NleG8 with C1QBP in HeLa cell lysates 196
5.8. Reciprocal IP for the NleG1-UBE2O interaction 197
5.9. Studying the interaction between C1QBP and NleG7 or NleG8 by direct yeast two-hybrid 199
5.10. NleG1 family members can ubiquitinate C1QBP *in vitro* 199
5.11. Ubiquitination of C1QBP and UBE2O by NleG1 in cultured HCT116 cells 203
5.12. Ubiquitination of UBE2O by NleG1 during co-expression in cultured HCT116 cells 205
5.13. Evaluating abundance of C1QBP and UBE2O in the presence and absence of NleG proteins 208
5.14. NleG expression does not affect degradation of IκBα in response to Poly (I:C) transfection 212
5.15. Discussion 212

Chapter 6: Final discussion 218

References 225

List of figures

Figure 1.1:	A/E pathogens induce distinct histopathological lesions in a Type III Secretion System-dependent manner	30
Figure 1.2:	Crypt elongation and immune cell recruitment in response to *C. rodentium infection* in the murine gut	34
Figure 1.3:	The T2SS and T5SS transport proteins across the bacterial outer membrane	36
Figure 1.4:	Bacterial T1SS, T3SS, T4SS and T6SS translocate substrates across both the inner and outer bacterial membranes	40
Figure 1.5:	Schematic representation of the T3SS of *E. coli* and *C. rodentium*	43
Figure 1.6:	The genetic organisation of the Locus of Enterocyte Effacement (LEE)	44
Figure 1.7:	The regulation of the LEE pathogenicity island	49
Figure 1.8:	The genetic organisation of the non-LEE encoded effector genes in *C. rodentium*	54
Figure 1.9:	Ubiquitin transfer is a three-step process	57
Figure 1.10:	Ubiquitin with its eight potential attachment sites for polyubiquitin chain formation	58
Figure 1.11:	The ubiquitin code	59
Figure 1.12:	The nleG genes in *C. rodentium* are located within different genomic islands	70
Figure 1.13:	NleG proteins have a conserved C-terminal motif that structurally resembles an E3 ligase	71
Figure 3.1:	Amino acid sequence alignment of NleG family proteins from EHEC, EPEC, *Citrobacter rodentium* and *Salmonella* spp.	105
Figure 3.2:	NleG family proteins cluster together phylogenetically	107
Figure 3.3:	Amino acid identity matrices for full length NleGs and the respective C- and N-termini	108
Figure 3.4:	NleG-TEM-1 β-lactamase fusion proteins are expressed in *C. rodentium* under IPTG-inducible conditions.	110
Figure 3.5:	*C. rodentium* NleG effectors are translocated in a Type III Secretion System-dependent manner	111
Figure 3.6:	Expression and purification of *C. rodentium* NleG proteins	112
Figure 3.7:	Expression test for GST-tagged NleG proteins in *E. coli*	113
Figure 3.8:	NleG family members are capable of autoubiquitination *in vitro*	114
Figure 3.9:	The C-terminus of NleG1 is has E3 ligase enzymatic activity *in vitro*	115
Figure 3.10:	EHEC and *C. rodentium* NleG8 homologs are cytotoxic in yeast dependent on the ability of their C-terminus to bind E2 ligases	117
Figure 3.11:	*C. rodentium* NleG proteins localise to specific subcellular compartments during transfection	118
Figure 3.12:	*C. rodentium* NleG proteins do not localise to the mitochondrion or to early endosomes during transfection	119
Figure 3.13:	Localisation of *C. rodentium* NleG8 during transfection	120
Figure 3.14:	Schematic representation of yeast two-hybrid methodology for identifying interacting partners	121
Figure 3.15:	*C. rodentium* NleG proteins do not cause autoactivation of reporter genes or a growth defect in in *S. cerevisiae* AH109	123

Figure 3.16:	Direct yeast two-hybrid assays using *C. rodentium* NleG1 did not confirm any new interacting partners	124
Figure 3.17:	Direct yeast two-hybrid assays using *C. rodentium* NleG7^{N} confirmed several interacting partners	125
Figure 3.18:	Structural features of GRASP55 and WDR61, shown to interact with *C. rodentium* NleG7^{N} in a direct two-hybrid assay	126
Figure 3.19:	GRASP55 and NleG7 do not co-localise when co-expressed in HeLa cells	126
Figure 3.20:	NleG8 interacts with GRASP55 by yeast direct two-hybrid	127
Figure 3.21:	Direct yeast two-hybrid assays using N-terminal *C. rodentium* NleG7 confirmed an uncharacterised putative membrane-associated protein as an interacting partner	128
Figure 3.22:	*C. rodentium* NleG7 and NleG8 do not interact with NHERF1 or NHERF2 in a direct two-hybrid yeast assay	130
Figure 3.23:	*C. rodentium* NleG7 and NleG8 associate with ubiquitin in a yeast direct two-hybrid assay	131
Figure 3.24:	*C. rodentium* NleG8 and *S.* typhi StoD associate with ubiquitin but not ISG15 by direct yeast two-hybrid assay	132
Figure 3.25:	*C. rodentium* NleG7 and NleG8 relocalise and co-localise with ubiquitin during transfection	133
Figure 4.1:	Overview of the *in vivo C. rodentium* infections and analyses used in this study	141
Figure 4.2:	A double *nleG* knockout *C. rodentium* strain (Δ*nleG7* Δ*nleG8*) causes pedestal formation during infection of HeLa cells	142
Figure 4.3:	A full *nleG* knockout *C. rodentium* strain (Δ*nleG1* Δ*nleG7* Δ*nleG8*) causes pedestal formation during infection of HeLa cells	142
Figure 4.4:	Δ*nleG1* and a catalytically inactive NleG chromosomal mutant (NleG1$_{P163K}$) is capable of pedestal formation during infection of HeLa cells	143
Figure 4.5:	Colonisation dynamics of WT *C. rodentium* during infection of C57BL/6 mice	144
Figure 4.6:	The flow cytometry-based strategy for analysing myeloid cells collected from mouse intestinal tissue following infection with *C. rodentium*	144
Figure 4.7:	The flow cytometry-based strategy for analysing lymphocytes collected from mouse intestinal tissue following infection with *C. rodentium*	145
Figure 4.8:	Changes in myeloid cell populations in the mouse intestine during *C. rodentium* infection	146
Figure 4.9:	Changes in immune cell populations during *C. rodentium* infection *in vivo*	147
Figure 4.10:	Mice infected with the *C. rodentium nleG1* deletion mutant show similar levels of bacterial colonisation as wild-type (WT) infected mice	148
Figure 4.11:	*In vivo* infection with the *C. rodentium nleG1* knockout mutant does not significantly alter the recruitment of myeloid cells or lymphocytes in the mouse intestine	149
Figure 4.12:	*C. rodentium* Δ*nleG1* and Δ*nleG1* chromosomally complemented with NleG1$_{P163K}$ (catalytically inactive) have no growth defect in liquid minimal media	150
Figure 4.13:	*C. rodentium* lacking NleG1 or NleG1 catalytic activity colonises the murine gut to WT-levels and does exhibit altered stool concentrations of Lcn2, a biomarker of intestinal inflammation	151
Figure 4.14:	A mass spectrometry-based approach to compare proteomes of enterocytes from mice infected with WT, ΔnleG1 or NleG1$_{P163K}$-expressing *C. rodentium*	153

Figure 4.15:	Mice infected with the *C. rodentium nleG8* deletion mutant show similar levels of bacterial colonisation as WT infected mice	155
Figure 4.16:	Mice infected with the *C. rodentium nleG8* deletion mutant develop WT-levels of colonic crypt hyperplasia	155
Figure 4.17:	Intestines of mice infected with the *C. rodentium nleG8* deletion mutant and plasmid-complemented mutant have the same levels of myeloid cell recruitment as WT-infected mice	156
Figure 4.18:	Expressing NleG8 or catalytically inactive NleG8$_{P177K}$ from the pACYC plasmid does not impact the ability of *C. rodentium* to colonise the murine gut	157
Figure 4.19:	A double Δ*nleG7* Δ*nleG8 C. rodentium* displays a decreased level of Lcn2, a biomarker of intestinal inflammation, in stools upon *in vivo* infection of C57BL/6 mice despite no attenuation in bacterial colonisation or colonic crypt hyperplasia	158
Figure 4.20:	Ubiquitin distribution and total ubiquitin in cultured cells infected with *C. rodentium in vitro*	160
Figure 4.21:	Total ubiquitin smear in enterocytes infected with *C. rodentium* WT and Δ*nleG1* and Δ*nleG7* Δ*nleG8* strains	161
Figure 4.22:	A full nleG knockout C. rodentium strain displays no growth defect in liquid media *in vitro*	161
Figure 4.23:	*In vivo* colonisation of a full *nleG* knockout *C. rodentium* strain	162
Figure 4.24:	*In vivo* colonisation of a full *nleG* knockout *C. rodentium*	163
Figure 4.25:	Comparison of colonic crypts from mice infected with WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* at day 8 post-infection	163
Figure 4.26:	*In vivo* infection using triple *nleG* deleted *C. rodentium* induces lower concentrations of Lcn2 in the stool	164
Figure 4.27:	Bacterial colonisation of the mouse intestine by full *nleG* knockout *C. rodentium* and the same strain chromosomally complemented with either WT of catalytically inactive NleG1	165
Figure 5.1:	Schematic overview of the approach used in this study to identify interacting partners of NleGs	170
Figure 5.2:	Doxycycline-inducible expression of the NleGs in cultured cells	172
Figure 5.3:	Immunoprecipitation (IP) of HA-tagged NleG proteins from cell lysates for mass spectrometry analysis	173
Figure 5.4:	Tandem mass spectrometry analysis of immunoprecipitated NleG1$_{P163K}$–containing complexes from mammalian cell lysates identifies putative interacting partners of NleG1	178
Figure 5.5:	Tandem mass spectrometry analysis of immunoprecipitated NleG8$_{P177K}$–containing complexes from mammalian cell lysates identifies putative interacting partners of NleG8	185
Figure 5.6:	A number of proteins are reproducibly enriched in HA-immunoprecipitation samples from lysates of HCT116 cells expressing HA-NleG1$_{P163K}$	190
Figure 5.7:	Overview of the proteins identified as putative interacting partners of NleG1 and NleG8 by affinity purification-coupled mass spectrometry experiments in this study	191
Figure 5.8:	Endogenous UBE2O and C1QBP co-immunoprecipitate with wildtype and catalytically inactive HA-NleG1 in HCT116 cells	194
Figure 5.9:	NleG1 and NleG1$_{P163K}$ expressed in HCT116 cells is not ubiquitinated	195

Figure 5.10: Banding of NleG1 and NleG1$_{P163K}$ from IPs using HA-beads is not due to neddylation or sumoyulation of the NleG 196

Figure 5.11: Recombinant His-NleG1 and His-NleG8 do not interact with C1QBP from HeLa cell lysates 197

Figure 5.12: *In vitro* FLAG pulldown shows no interaction between NleG1 and UBE2O when the proteins are co-expressed in HCT116 cells 198

Figure 5.13: Direct yeast two-hybrid assay did not confirm an interaction between C1QBP and *C. rodentium* NleG proteins 199

Figure 5.14: Preparation of recombinant C1QBP for *in vitro* assays 200

Figure 5.15: *C. rodentium* NleG1 and NleG8 can support ubiquitination of C1QBP *in vitro* 201

Figure 5.16: *C. rodentium* NleG1 can support K48- and K63-linked ubiquitination of itself and of C1QBP *in vitro*. 202

Figure 5.17: The NleG family members NleG1, NleG8 and StoD but not the HECT-type bacterial E3 ligase NleL can ubiquitinate C1QBP *in vitro* 203

Figure 5.18: NleG1 is not ubiquitinated and does not ubiquitinate C1QBP or UBE2O in HCT116 cells 204

Figure 5.19: Immunoprecipitation of ubiquitinated proteins to assess whether UBE2O transiently expressed in NleG1-expressing HCT116 cells ubiquitinates NleG1 206

Figure 5.20: Co-expression of UBE2O (WT or E3 ligase dead) and NleG1 (WT or E3 ligase dead) in HCT116 cells in the presence and absence of a proteasomal inhibitor 207

Figure 5.21: Expressing NleG1 and NleG1$_{P163K}$ in HCT116 cells does not alter C1QBP and UBE2O protein levels 208

Figure 5.22: Endogenous UBE2O is detectable and cytoplasmic in HCT116 cells with and without the induced expression of NleG1 209

Figure 5.23: Endogenous levels of UBE2O and C1QBP in HeLa cells remains unchanged during *C. rodentium* infection, regardless of presence of NleG proteins 209

Figure 5.24: Assessing the level and distribution of endogenous UBE2O during *C. rodentium* infection 210

Figure 5.25: UBE2O but not C1QBP protein abundance in intestinal epithelial cells (IECs) is reduced during *C. rodentium* infection 211

Figure 5.26: NleG1 expression does not affect degradation of IκBα in response to Poly (I:C) transfection in HCT116 cells 212

List of tables

Table 1.1: Manipulation of the host ubiquitin system by bacterial pathogens 68

Table 2.1: Antibiotic stock- and working concentrations 74

Table 2.2: Strains and plasmids used in this study 75

Table 2.3: Primers used in this study 80

Table 2.4: Antibodies used in this study 91

Table 2.5: Antibodies used for flow cytometric analyses in this study 95

Table 5.1: Putative interactors of NleG1 in HCT116 cells as identified by AP-MS/MS, repeat 1 179

Table 5.2: Putative interactors of NleG1 in iBMM cells as identified by AP-MS/MS 182

Table 5.3: Putative interactors of NleG8 in HCT116 cells as identified by AP-MS/MS 186

Table 5.4: Putative interactors of NleG8 in iBMM cells as identified by AP-MS/MS 189

Table 5.5: Putative interactors of NleG1 in HCT116 cells as identified by AP-MS/MS, repeat 2 192

Abbreviations

AAF	Aggregative adherence fimbriae
ABC	ATP-binding cassette
Ade	Adenine
ADIN	Antibody-dependent intracellular neutralisation
ADP	Adenosine diphosphate
A/E	Attaching and effacing
aEPEC	Atypical EPEC
AI-3	Autoinducer-3
AIDA-I	Adhesin involved in diffuse adherence I
AIEC	Adherent invasive *E. coli*
AMPKα2	AMP-activated protein kinase subunit
ANOVA	Analysis of variance
Amp	Ampicillin
AQP2/3	Aquaporin 2/3
Arp2/3	Actin-related protein 2/3
ARRIVE	Animal Research: Reporting of *In Vivo* Experiments
BBS7	Bardet-Biedl syndrome 7 protein
BFP	Bundle-forming pili
BLAST	Basic Local Alignment Search Tool
Bp	Base pair
BSA	Bovine serum albumin
C1QBP	Complement component 1 Q subcomponent-binding protein
CAPER	Co-activator of activating protein-1 and estrogen-receptors
CD55	Complement Decay-accelerating Factor 55
CF	Colonisation factor
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CFU	Colony forming unit
Cif	Cycle inhibiting factor
Cm	Chloramphenicol
cNK cell	Conventional NK cell
DAEC	Diffusely adherent *E. coli*
DAMP	Danger associated molecular pattern
DC	Dendritic cell
DDO	Double dropout
diGly	Diglycine
DMEM	Dulbecco's Modified Eagle Medium
DRiPS	Defective ribosomal products
DTT	1,4-dithiothreitol
Duox2	NAPDH dual oxidase 2
EAEC	Enteroaggregative *E. coli*
EAF	*E. coli* adherence factor
EAST-1	EAEC heat-stable enterotoxin 1
ECP	*E. coli* common pilus

EDTA	Ethylenediamine tetra-acetic acid
EHEC	Enterohaemorrhagic *E. coli*
EIEC	Enteroinvasive *E. coli*
ELISA	Enzyme-linked immunosorbant assay
Ent	Enterobactin
EPEC	Enteropathogenic *E. coli*
Epi	Epinephrine
ER	Endoplasmic reticulum
ETEC	Enterotoxigenic *E. coli*
ExPEC	Extra-intestinal pathogenic *E. coli*
FACS	Fluorescence activated cell sorting
FADD	Fas-associated via death domain
FcR	Fc receptor
FCS	Fecal calf serum
FDR	False discovery rate
FGF7	Fibroblast growth factor 7
FMO	Fluorescence minus one
FRET	Fluorescence resonance energy transfer
GAP	GTPase activating protein
GC-C	Guanylate-cyclase-C
GEF	Guanine-nucleotide exchange factor
GEMS	Global Enteric Multicentre Study
GFP	Green fluorescent protein
GI	Gastrointestinal
GI1-16	Genomic island 1-16
GlcNAc	*N*-acetylglucosamine
GLMN	Glomulin/flagellar-associated protein 68
Gm	Gentamicin
GO	Gene ontology
GRASP55	Golgi reassembly and stacking protein 55
Gsp	General secretin pathway protein
GST	Glutathione S-transferase
H&E	Haematoxylin and eosin
HCP	Haemorrhagic *coli* pilus
HBSS	Hanks' buffered salt solution
HECT	Homologous to the E6-AP carboxy-terminus
HEPES	4-(2-hydroxyethyl)-1- piperazineethanesulphonic acid
His	Histidine
HK2	Hexokinase 2
HlyA	Haemolysin A
IAA	Iodoacetamide
IAMD	Intracellular antibody-mediated degradation
iBMM	Immortalised bone marrow macrophage
ICS	Intracellular cytokine staining
IEC	Intestinal epithelial cell
IFN	Interferon

Ig	Immunoglobulin
IL	Interleukin
ILC3s	Group 3 innate lymphoid cells
ILK	Integrin-linked kinase
IM	Inner membrane
IMP	Inner membrane platform
iNOS	Inducible nitric oxide synthase
InPEC	Intestinal pathogenic *E. coli*
IP	Immunoprecipitation
Ipa	Invasion plasmid antigen
IRF3	Interferon regulatory factor 3
ISG15	Interferon-stimulated gene 15
ISGs	Interferon-stimulated genes
Kan	Kanamycin
KC	Keratinocute chemoattractant
KEGG	The Kyoto Encyclopaedia of Genes and Genomes
KO	Knock out
LA	Localised adherence
LB	Lysogeny broth
LC	Liquid chromatography
LCV	*Legionella*-containing vacuole
Lcn2	Lipocalin-2
LEE	Locus of enterocyte effacement
Leu	Leucine
LFQ	Label-free quantification
LiAc	Lithium acetate
LPS	Lipopolysaccharide
LRR	Leucine-rich repeat
LT	Heat-labile enterotoxin
LTβR	Lymphotoxin β receptor
LUBAC	Linear ubiquitin chain assembly complex
mART	Mono-ADP-ribosyltransferase
MAVS	Mitochondrial antiviral signalling protein
Mbp	Mega basepair
MFP	Membrane fusion protein
MNEC	Meningitis-associated *E. coli*
MOPS	3-(N-morpholino)-propanesulphonic acid
MS	Mass spectrometry
MUSCLE	Multiple sequence comparison by Log-expectation
MyD88	Myeloid differentiation factor 88
NAD	Nicotinamide dinucleotide
Nal	Nalidixic acid
NCBI	National Centre for Biotechnology Information
NDS	Normal donkey serum
NE	Norepinephrine
NEL	Novel E3-like

NEM	*N*-Ethylmaleimide
NEMO	NF-κB essential modulator
NF-κB	Nuclear factor κ-light chain-enhancer of activated B cells
NHE	Sodium-hydrogen exchanger 3
NHERF1/2	NHE regulatory factor 1/2
NK cell	Natural killer cell
Nle	Non-lee encoded
NleG	Non-lee encoded effector protein G
NMR	Nuclear magnetic resonance
Nox1	NADPH oxidase 1
N-WASP	Neural Wiscott-Aldrich syndrome protein
OCRL1	Oculocerebrorenal syndrome of Lowe protein
OM	Outer membrane
OMF	Outer membrane factor
OMP	Outer membrane complex
OMV	Outer membrane vesicle
PAI	Pathogenicity island
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffered saline
PBA	PBS-BSA-sodium azide
PCR	Polymerase chain reaction
PDZ	Psd95, Dlg1, zo-1
PG	Peptidoglycan
PKN1	Protein kinase N1
P/S	Penicillin-streptomycin
Ptcd3	Pentatricopeptide repeat containing protein 2
PVDF	Polyvinylidene fluoride
QDO	Quadruple drop-out
RDEC	Rabbit diarrhoeagenic *E. coli*
REPEC	Rabbit EPEC
Rbm39	Ribosome-binding protein 39
RBR	RING-between-RING
RING	Really interesting new gene
RIPK1	Receptor-interacting serine/threonine protein kinase 1
RNA	Ribonucelic acid
ROS	Reactive oxygen species
RPS3	Ribosomal proteinS3
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription quantitative PCR
SCFA	Short chain fatty acid
scRNA-seq	Single cell RNA sequencing
SD	Synthetic dextrose
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
Sec	General secretory pathway
SEM	Standard error of the mean

SEPEC	Sepsis-associated *E. coli*
SH3	Src homology 3 protein
SILAC	Stable isotope labelling with amino acids in cell culture
Sm	Streptomycin
SNP	Single nucleotide polymorphism
SNX9	Sortin nexin 9
SLC	Solute carrier family
SPATE	Serine protease autotransporter of Enterobacteriaceae
ST	Heat-stable enterotoxin
Stx	Shiga toxin
STEAEC	Shiga toxin-producing Enteroaggregative *E. coli*
T1SS	Type I secretion system
T2SS	Type II secretion system
T3SS	Type III secretion system
T4SS	Type IV secretion system
T5SS	Type V secretion system
T6SS	Type VI secretion system
T7SS	Type VII secretion system
TA	Transit amplifying
TAB2/3	Tak1-binding protein 2/3
Tat	Twin arginine translocation
TCS	Two-component signal transduction system
TCEP	Tris(2-carboxymethyl)phosphine
T-DNA	Transfer DNA
TDO	Triple drop-out
TEAB	Triethylammonium bicarbonate
tEPEC	Typical EPEC
Tet	Tetracycline
TFA	Trifluoroacetic acid
TFBI/II	Transformation buffer I/II
Th1	T helper type 1
Tir	Translocated intimin receptor
TJ	Tight junction
TLR	Toll-like receptor
TMT	Tandem mass tag
TNF	Tumour necrosis factor
TOAt	Tet-ON transactivator
TRADD	TNFRSF1A-associated via death domain
TRE	Tetracycline response element
Trp	Tryptophan
TUBE	Tandem ubiquitin-binding entities
UBAIT	Ubiquitin-activated interaction traps
UBD	Ubiquitin-binding domain
UPEC	Uropathogenic *E. coli*
UTI	Urinary tract infection
WDR61	WD repeat-containing protein 61

Wnt	Wingless
WT	Wild type
x-α-gal	5-Bromo-4-chloro-3-indolyl α-D-galactopyranoside
Y2H	Yeast two-hybrid
YPD	Yeast peptone-dextrose
YPDA	Yeast peptone-dextrose with adenine
ZAP1	Zinc finger CCCH-type antiviral protein 1

Chapter 1:
Introduction

1. Introduction

1.1. Bacterial pathogenesis

Bacterial pathogens have acquired and evolved intricate strategies for colonising their hosts, allowing them to successfully replicate and disseminate in a range of niches. To achieve this, they overcome physiological barriers as well as host cell and systemic responses, setting the scene for a complex host-pathogen interplay. Some bacteria are harmless in immunocompetent individuals and can only cause symptoms following injury or in immunocompromised patients. Others, including several strains of *Escherichia coli,* have acquired specialised molecular armoury that triggers disease even in healthy hosts.

1.2. Escherichia coli

E. coli, a Gram-negative facultative anaerobe of the *Enterobacteriaceae* family, is remarkably versatile and colonises a wide range of niches (Nataro & Kaper, 1998). First described in 1885 as the "common colon bacillus", it is a key constituent of the healthy human intestinal microflora and typically colonises the mammalian gut within hours of birth. *E. coli* is easily grown and genetically manipulated under laboratory conditions and has consequently become central to research in molecular biology and bacterial genetics. It is widely recognised as the world's most comprehensively studied bacterium (Kaper, Nataro & Mobley, 2004; Robins-Browne *et al.*, 2016), with the genome of a laboratory strain of *E. coli* among the first bacterial genomes to be sequenced (Blattner, 1997).

While most *E. coli* are beneficial or harmless to their hosts, some strains can directly cause disease with symptoms ranging from diarrhoea to meningitis, septicaemia and urinary tract infections (UTIs) (Nataro & Kaper, 1998). *E. coli* strains share a 4.1 million base pair (Mbp) genetic backbone and the ability to cause disease is conferred by the insertion of mobile genetic elements, such as large virulence gene clusters (pathogenicity islands, PAIs), plasmids and prophages (Leimbach, Hacker & Dobrindt, 2013). Thus, certain pathogenic *E. coli* have significantly larger genomes compared to commensal strains (Blattner, 1997; Croxen & Finlay, 2010). Fewer than 1,500 genes comprise the core genome conserved across all sequenced *E. coli* strains, while the number of unique genes found across them exceeds 22,000 (Robins-Browne *et al.*, 2016). A comparison between a lab strain of *E. coli* (K-12 strain MG1655), enterohaemorrhagic *E. coli* strain EDL933 and uropathogenic *E. coli* strain CFT077 found that fewer than 40% of genes were shared between the all three strains (Welch *et al.*, 2002).

1.3. Intestinal pathogenic E. coli

The first isolates of *E. coli* shown to be aetiological agents of disease were recovered from children with diarrhoea in the 1940s (Bray, 1945). *E. coli* are now known to contribute to the estimated 1.7 billion cases of acute infectious diarrhoea that occur per annum (leading to over 2 million deaths worldwide) (Snyder & Merson, 1982; Bern *et al.*, 1992; Kosek, Bern & Guerrant, 2003; Hodges & Gill, 2010; Humphries & Linscott, 2015). Infectious diarrhoea thereby poses an immense burden on global health resources. Importantly, it remains within the five most common causes of death in children under five in low- to middle-income countries, accounting for almost 800,000 infant deaths per year (Liu *et al.*, 2015a; Hu & Torres, 2015). Diarrhoea is also a leading cause of infantile malnutrition, which puts patients at further risk of life-threating infections. While many countries have seen a drop in foodborne infections outbreaks, a re-emergence of severe enteric pathogens is accelerated by factors such as globalisation and increased antibiotic resistance (de Nisco, Rivera-Cancel & Orth, 2018).

E. coli that cause gastrointestinal (GI) infections are commonly referred to as intestinal pathogenic *E. coli* (InPEC), separating them from extra-intestinal pathogenic *E. coli* (ExPEC) that comprise uropathogenic *E. coli* (UPEC), meningitis-associated *E. coli* (MNEC) and sepsis-associated *E. coli* (SEPEC) (Robins-Browne *et al.*, 2016). InPEC are grouped into eight different pathotypes based on their clinical symptoms, specific virulence traits and phylogenetic profile. These include the six pathotypes traditionally associated with intestinal *E. coli* infection: enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), diffusely adherent *E. coli* (DAEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC). More recently, two additional pathotypes have been described: adherent invasive *E. coli* (AIEC) and Shiga toxin (Stx) producing enteroaggregative *E. coli* (STEAEC). ETEC, EPEC and EAEC are the most prevalent pathotypes in low- to middle-income countries, where inadequate sanitation and contaminated water sources contribute to the spread of disease. In high-income countries, infections are predominantly associated with outbreaks from contaminated food sources and the most prevalent pathotypes are EHEC, EAEC and STEAEC.

Three pathotypes – EHEC, EPEC and EIEC – utilise a specialized membrane-associated secretion apparatus called the type III secretion system (T3SS) for the translocation of bacterial virulence factors, or effector proteins, straight into host cells to subvert host processes. The T3SS is the major contributor to their virulence mechanisms, while the other pathotypes rely on an array of

colonisation factors and otherwise secreted proteins and toxins to propagate their infection. The following section describes the InPEC pathotypes, with a focus on T3SS-dependent pathotypes.

1.3.1. Enterotoxigenic E. coli (ETEC)

ETEC is the most frequently isolated enteropathogenic bacterium in children under 5 in low- to middle-income countries and is a common cause of traveller's diarrhoea (Lanata *et al.*, 2013). It accounts for millions of diarrhoeal cases per year and tens of thousands of deaths. ETEC-induced diarrhea also occurs in domestic animals (mainly calves and pigs), further contributing to its economic importance (Fairbrother, Nadeau & Gyles, 2005). As its name suggests, ETEC relies on several enterotoxins to cause disease. The cholera-like pathology of ETEC is conferred via a heat-labile (LT) and/or small heat-stable (ST) enterotoxins (STa in human isolates, Qadri *et al.*, 2005). The ETEC LT is an AB_5 toxin (like the cholera toxin) and causes cell death. STa activates the intestinal brush border guanylate-cyclase-C (GC-C) receptor leading to activation of the Cystic Fibrosis Transmbembrane Conductance Regulator (CFTR) ion channel, which results in deregulation of ion absorption and water secretion in the gut. Together, these toxins compromise the structural integrity and absorptive function of the intestinal epithelium, contributing to the severe diarrhoeal phenotype observed following ETEC infection. To deliver the toxins to the intestinal epithelium, ETEC relies on numerous colonisation factors (CFs) that mediate attachment to epithelial cells in the small intestine (Gaastra & Svennerholm, 1996; Nada *et al.*, 2011). CFs differ between human and animal ETEC isolates, and 75% of human ETEC isolates express one of the three major colonisation factors: CFA/I, CFA/II or CFA/IV (Wolf, 1997).

1.3.2. Diffusely adherent E. coli (DAEC)

DAEC strains are traditionally associated with infantile diarrhea and are grouped based on their diffuse adherent histopathology, which involves membrane protrusions of host cells that wrap around infecting bacteria (Scaletsky, Silva & Trabulsi, 1984; Peiffer, Servin & Bernet-Camard, 1998). An autotransporter protein called adhesin involved in diffuse adherence (AIDA-I) was the first determinant of diffuse adherence to be described (Benz & Schmidt, 1992). A second autotransporter, Sat, contributes to the diarrhoeal phenotype by rearranging host tight junctions (TJs, leading to increased paracellular permeability) and by inducing cell death. However, it has also been shown that AIDA-I-expressing *E. coli* often co-express virulence determinants that are characteristic of other InPEC pathotypes (Servin, 2014). Likewise, many AIDA-I-negative DAEC strains in fact cause UTIs, making them ExPEC. Taken together with the inability of two prototypic DAEC strains to cause diarrhoea in human volunteers (following ingestion of 10^{10} colony forming

units, CFUs), there is some doubt regarding a role for this pathotype in diarrhoeal disease (Servin, 2014).

1.3.3. Adherent invasive E. coli (AIEC)

AIEC strains share few virulence factors with other InPEC pathotypes and genetically resemble ExPEC, but remain classified under InPEC since they colonise the intestine. While AIEC does not cause diarrhea, it is associated with the inflammatory gastrointestinal condition Crohn's disease. The aetiology of Crohn's disease is unclear, but likely involves host- as well as environmental factors (Alhagamhmad *et al.*, 2016). Interestingly, the host receptor CEACAM6 required for AIEC colonisation is upregulated in Crohn's patients (Barnich *et al.*, 2007). Following binding of epithelial cells, AIEC produces outer membrane vesicles (OMVs) to deliver bacterial proteins needed to invade host cells. Overexpression of the cellular receptor for the AIEC OMVs has also been associated with AIEC invasion in Crohn's patients (Rolhion, Hofman & Darfeuille-Michaud, 2011). Further, genetic factors linked with Crohn's disease are likely to predispose the gut to AIEC infection (Kaser *et al.*, 2008). Irrespective of whether AIEC colonisation requires a susceptible host or is directly linked to Crohn's disease, immune responses to AIEC invasion appear to exacerbate the inflammation seen in Crohn's patients (Glasser *et al.*, 2001; Carvalho *et al.*, 2009).

1.3.4. Enteroaggregative E. coli (EAEC)

EAEC, another cause of traveller's diarrhoea, are characterised by their autoaggregative attachment to tissue culture cells in a distinctive stacked-brick pattern (Nataro *et al.*, 1987). EAEC pathogenesis relies on the 100-kb virulence plasmid pAA that encodes four Aggregative Adherence Fimbriae (AAF) variants and the virulence regulator AggR (Nataro *et al.*, 1994; Czeczulin *et al.*, 1997; Clements *et al.*, 2012). Together with flagellin, AAF help adherence to the gut mucosa, which occurs via an electrostatic interaction with fibronectin (Berry *et al.*, 2014). AAF also help adherence to salad leaves, which may contribute to bacterial outbreaks from contaminated fresh produce (Berger *et al.*, 2009b; Boisen *et al.*, 2008; Itoh *et al.*, 1997). Furthermore, AAF contribute to biofilm formation, which increases difficulties in clearing the infection (Sheikh *et al.*, 2001). Other virulence factors are found in certain but not all EAEC isolates, such as the pore-forming toxin HlyE and the EAEC heat-stable Enterotoxin 1 (EAST-1; 50% identical to ETEC STa). Some EAEC isolates encode a T3SS (Rasko *et al.*, 2008; Chaudhuri *et al.*, 2010), but their contribution to EAEC virulence has not been shown and EAEC pathogenesis is considered T3SS-independent.

1.3.5. Enteroinvasive E. coli (EIEC)

While EIEC, EPEC and EHEC require a T3SS for their pathogenesis, only EIEC is an intracellular pathogen. EIEC shares its virulence mechanism with *Shigella* spp., the causative agent of bacillary dysentery, and cross-reacts with *Shigella* antisera. Although EIEC may cause bacillary dysentery in rare cases, EIEC infection normally results in watery diarrhea akin other InPEC pathotypes. *E. coli* and *Shigella* spp. are highly related genetically and their separation into different genuses relates to historic factors and the clinical presentation of *Shigella* spp. (Clements *et al.*, 2012). EIEC represents an intermediate between ancestral *E. coli* and *Shigella* spp., as EIEC has acquired the *Shigella* pINV plasmid but retained some *E. coli* biochemical properties (Van Den Beld & Reubsaet, 2012).

For EIEC, like *Shigella*, pathogenesis is characterized by invasion of the colonic epithelium and severe mucosal inflammation (Philpott, Edgeworth & Sansonetti, 2000) and depends on the 210-230 Kb pINV plasmid. pINV plasmids encode numerous virulence factors, including proteins responsible invasion (invasion plasmid antigens; Ipa proteins) and a T3SS. These promote rearrangement of the actin cytoskeleteon to enter the host and to manipulate cellular response (Marteyn, Gazi & Sansonetti, 2012). Upon entry, the bacteria initially reside within a membranous vacuole. The bacteria then lyse the vacuole using IpaC, constituents of the T3SS translocon pore as well as other effector proteins. Following escape into the host cytosol, the bacteria use actin-based motility and spread in a cell-to-cell manner thanks to the surface protein IcsA. Many effector proteins are shared between EIEC, *Shigella* as well as the extracellular pathotypes EPEC and EHEC. Such effector protein conservation reflects similarities in infection strategies and highlights the distinction of EIEC, EHEC and EPEC from T3SS-independent pathotypes.

1.3.6. Enteropathogenic E. coli (EPEC)

EPEC is a significant cause of infant morbidity and mortality in low- to middle-income countries, as recently highlighted by the Global Enteric Multicenter Study (GEMS) that correlated EPEC infection with an elevated risk of mortality in young children with moderate- to severe diarrhea in Sub-Saharan Africa and South Asia (Levine *et al.*, 2012; Kotloff *et al.*, 2013). EPEC is also a cause of infantile diarrhea in England (Sakkejha *et al.*, 2013) and the pathogenic *E. coli* first described in 1945 was in fact EPEC.

EPEC is one of the best understood *E. coli* pathotypes, and thus many aspects of its pathogenesis are well-described. A hallmark of EPEC infection is the formation of distinctive attaching and effacing (A/E) lesions, characterised by the effacement of intestinal microvilli, intimate adherence of the bacterium to the host epithelium and localized actin polymerization (Moon *et al.*, 1983; Tzipori *et al.*, 1985). This histopathology is T3SS-dependent and unites a group of highly adapted pathogens termed A/E pathogens (discussed in **Section 1.4**).

EPEC strains are divided into typical or atypical EPEC (aEPEC and tEPEC). tEPEC use bundle-forming pili (BFP) encoded on their *E. coli* adherence factor (EAF) plasmid for their initial attachment to host cells (Girôn, Ho & Schoolnik, 1991). BFP are thin and flexible filaments that mediate inter-bacterial interactions (leading to bacterial microcolonies) and bacteria-host cell interactions (leading to attachment to the host, by binding N-acetyl-lactosamine moieties of host surface glycoproteins). Together, this gives tEPEC a distinctive Localised Adherence (LA) phenotype when infecting a cell monolayer *in vitro*. Although the BFP are the major virulence factor in the initial interaction between the EPEC and the host, bacterial attachment is also aided by adhesins such as the *E. coli* common pilus (ECP) and flagella. Importantly, aEPEC strains are defined by the lack of the BFP-encoding EAF plasmid. A global increase in aEPEC prevalence was recently reported (Hu & Torres, 2015) and while aEPEC are associated with persistent diarrhoea in children (Nguyen *et al.*, 2006; Ochoa *et al.*, 2008), there have also been varied reports regarding their ability to cause acute diarrhoea in adults. Certain aEPEC isolates do not cause diarrhoea in human volunteers (Bieber *et al.*, 1998; Levine *et al.*, 1978), yet many diarrhoeal outbreaks have been attributed to aEPEC (Yatsuyanagi *et al.*, 2002, 2003; Viljanen *et al.*, 1990; Hao *et al.*, 2012). Serotyping EPEC isolates (defining the specific O antigen in their lipopolysaccharide, LPS, and the specific flagellar H antigen) shows that aEPEC belong to a wider variety of serogorups compared to tEPEC (which mainly belong to serogroups O55, O111, O119, O127 or O142; Whale *et al.*, 2007). As the serotypes differ in their repertoire of virulence factors, the capability of isolates to cause diarrhoea likely varies with their serotype (Trabulsi, Keller & Tardelli Gomes, 2002; Tozzoli *et al.*, 2014). Consequently, while the prototype tEPEC strain E2348/69 (serotype O127:H7) is routinely used to study EPEC pathogenesis (Sakkejha *et al.*, 2013), it is sometimes replaced by other isolates to ensure the clinical relevance of studies.

1.3.7. Enterohaemorrhagic E. coli (EHEC)

EHEC is a prominent cause of diarrhea in high-income countries. It was first described in 1977 (Konowalchuk, Speirs & Stavric, 1977) and implicated in human disease soon after (Riley *et al.*,

1983). Genetically related to EPEC, EHEC is also a member of the A/E pathogens (see **section 1.4**). However, while EPEC is human-restricted, EHEC is a zoonotic pathogen with its primary reservoir being the intestines of ruminants (mainly cattle), where it resides asymptomatically (Muniesa *et al.*, 2012). EHEC outbreaks are associated with contaminated produce (traditionally undercooked meat, but also fresh greens) and the incidence of EHEC cases remain constant, perhaps owing to a small number of bacteria required for a successful infection (under 100 CFUs) (Adams *et al.*, 2016; Welinder-Olsson & Kaijser, 2005).

While EPEC colonises the small intestine, EHEC colonises the distal ileum, cecum, large intestine and colon. EHEC causes heamorrhagic colitis characterized by bloody diarrhoea, and in approximately 10% of cases also leads to a potentially life-threatening complication called Haemolytic Uremic Syndrome (HUS), which is characterized by acute renal failure, thrombocytopenia and haemolytic anaemia. In the event of HUS development, mortality rates are high and a third of survivors suffer permanent health consequences (Welinder-Olsson & Kaijser, 2005). HUS development occurs more commonly in children and EHEC is the primary cause of acute pediatric kidney failure in the UK (Lynn *et al.*, 2005; Trachtman, 2013).

The typical EHEC strain (EHEC serotype O157:H7) evolved from a precursor EPEC strain (EPEC O55:H7), and other EHEC strains may also have evolved from this (Ogura *et al.*, 2009). EHEC O157:H7 isolates generally carry a pO157 plasmid that encodes virulence-associated factors, such as the EHEC haemolysin that can be used as a diagnostic marker for EHEC (Feldsine *et al.*, 2016). Related plasmids are found in the main non-O157 serogroups (O26, O45, O103, O111, O121 and O145; Brooks *et al.*, 2005; Ogura *et al.*, 2009).

The defining characteristics of EHEC are the lack of BFP and the presence of *stx* genes that encode Shiga toxins. Instead of relying on BFP for adherence, EHEC uses on a Type IV haemorrhagic *coli* pilus (HCP, Xicohtencatl-Cortes *et al.*, 2009), flagellin (Mahajan *et al.*, 2009), the autotransporter EhaA (Wells *et al.*, 2008), the ECP (Rendon *et al.*, 2007) and long polar fimbria (Fitzhenry *et al.*, 2006). Importantly, the phage-encoded Shiga toxins contribute to a serious pathology and HUS-development. They belong to the AB_5 toxin family and their catalytic A-subunit binds and cleaves rRNA within 28S rRNA of ribosomes, which inhibits protein synthesis and eventually causes cell death. Upon release, Shiga toxins cross the colonic epithelium and access the circulation via colonic blood vessels, causing systemic intoxication (Welinder-Olsson & Kaijser, 2005). The pentameric B-subunit binds cell surface receptors to trigger cell entry by endocytosis, and in accordance with the symptoms of HUS its receptors include a ganglioside

found on kidney epithelial cells. There are two Shiga toxin variants: Stx1 and Stx2. Despite minor structural differences, Stx2 is more commonly expressed in human EHEC isolates and associated with HUS (Fraser *et al.*, 2004). Antibiotic treatment is a risk factor for HUS-development in patients with EHEC, as antibiotic exposure dramatically increases Shiga toxin release in both *in vitro* and *in vivo* infection models (Walterspiel *et al.*, 1992; Kimmitt, Harwood & Barer, 2000; Wong *et al.*, 2000). Alongside the lack of EHEC-specific vaccines, this calls for the development of new treatment options that requires a comprehensive understanding of the mechanisms of EHEC infection.

1.3.8. Shiga toxin producing Enteroaggregative E. coli (STEAEC)

The summer of 2011 saw an outbreak of suspected EHEC in Germany, with many cases of HUS (Krause *et al.*, 2013). The causative strain first thought to be EHEC O104:H4 was, in fact, an Stx2-producing EAEC O104:H4 strain. The emergence of such a hybrid EAEC-EHEC strain warranted the definition of a new pathotype: Shiga-toxin producing enteroaggregative *E. coli* (STEAEC). Approximately 3000 cases were recorded (causing 53 deaths) and over 20% of patients developed HUS (Burger, 2012). The raised rate of HUS was thought to be due to tighter adherence to the epithelium, enhanced inflammation and disruption of cell-to-cell junctions – a consequence of the action of AAF – which could lead to more Stx2 reaching the circulation and causing damage (Boisen *et al.*, 2015). Like the EAEC AAF, the STEAEC AAF has since been shown to mediate adherence to abiotic surfaces and fresh produce (Nagy *et al.*, 2016).

1.4. Attaching and Effacing pathogens

The human pathogens EPEC and EHEC together with rabbit, sheep and pig EPEC, the murine pathogen *Citrobacter rodentium* and the emerging pathogen *Escherichia albertii* form a group of pathogens known as Attaching and Effacing (A/E) pathogens. They are grouped based on the formation of histopathological A/E lesions, characterised by the effacement of brush border microvilli and intimate adherence to the enterocytes, along with the formation of pedestal-like structures directly beneath the site of bacterial attachment (Collins *et al.*, 2014). A/E pathogenesis is underpinned by the expression of genes on the Locus of Enterocyte Effacement (LEE): a 35 Kb chromosomal island encoding the T3SS, transcriptional regulators, the outer membrane adhesin intimin and its receptor Tir, as well as effector proteins and their chaperones. A/E pathogens use their T3SS (discussed in detail in **Section 1.5.8**) to translocate both LEE-encoded and non-LEE encoded effector proteins into host cells to subvert host processes and

ensure bacterial survival. The next section will describe *C. rodentium,* as it is commonly used to model A/E infection *in vivo* and was used in the project presented in this thesis.

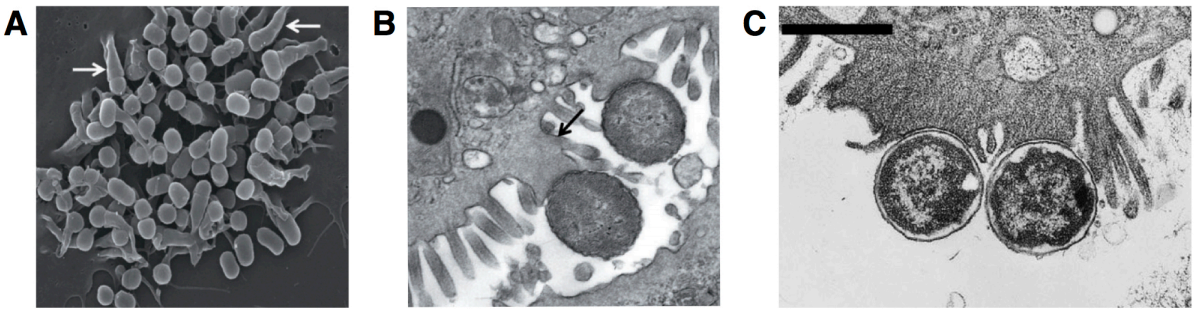


Figure 1.1 | A/E pathogens induce distinct histopathological lesions in a Type III Secretion System-dependent manner. A) Adherent bacteria induce actin pedestals (indicated by white arrows), scanning electron micrograph by Wong *et al.* (2011). **B)** Transmission electron micrograph (Wong *et al.*, 2011) reveals the intestinal attaching and effacing lesions induced by A/E pathogens. Black arrow indicates actin pedestal formation. **C)** *Citrobacter rodentium* causes A/E lesions in the colon of infected mice, transmission electron micrograph by Schauer & Luperchio (2001). Scale bar 1 µm.

1.4.1. Animal models of A/E pathogenesis

Sophisticated manipulation of host responses by A/E pathogens necessitates robust animal infection models to study host-pathogen interactions *in vivo.* Murine models are invaluable thanks to the availability of inbred or genetically manipulated mice housed in pathogen-controlled facilities. However, immunocompetent mice are inherently resistant to EPEC and EHEC infection. The mouse A/E pathogen *C. rodentium,* on the other hand, causes transmissible murine colonic hyperplasia characterised by increased colonic thickness, epithelial crypt elongation and goblet cell loss with A/E lesions indistinguishable from those of EPEC and EHEC (see **Figure 1.1**; Luperchio *et al.*, 2000). The LEE islands of *C. rodentium*, EPEC and EHEC were acquired by horizontal gene transfer and are therefore closely related both in terms of genetic organization and function of gene products (Petty *et al.*, 2010). The shared virulence factors and infection strategy make *C. rodentium* a powerful tool to study A/E pathogenesis in a natural host.

Historically, infant rabbit models using EHEC rabbit diarrhoeagenic *E. coli* (RDEC-1, serotype O15:H-) and rabbit-specific EPEC (REPEC, serotype O103:H2) have also been used to study A/E virulence factors (Ritchie & Waldor, 2005; Agin *et al.*, 1996; Von Moll & Cantey, 1997; Cantey & Moseley, 1991; Cantey, Lushbaugh & In Man, 1981; Inman & Cantey, 1983, 1984) and evaluate

vaccine candidates (Horne *et al.*, 2002). However, their use is declining due to variability in bacterial colonisation, the severity of the infection (Cantey & Blake, 1977) and the lack of genetically modified rabbit strains. Likewise, bovine and ovine models have been used to study EHEC pathogenesis (Hemrajani *et al.*, 2008; Dean-Nystrom *et al.*, 1997; Hoffman *et al.*, 2006) and EHEC vaccine candidates (Allen *et al.*, 2011; Potter *et al.*, 2004), but large animal models suffer from the lack of large-scale facilities, high costs and ethical considerations. A gnotobiotic EPEC piglet model (Tachikawa *et al.*, 1998; Tzipori *et al.*, 1986, 1992, 1985; Dean-Nystrom *et al.*, 2002) on the other hand suffers from not considering the involvement of the microbiota or a developed mucosal immune system during A/E infection: two factors that contribute to the host-pathogen interaction. Similarly, EHEC and EPEC colonisation of the mouse gut can be manufactured using antibiotic pre-treatment, but this does not cause A/E lesion formation and it is unknown whether T3SS effectors are injected into host cells under these conditions (Mundy *et al.*, 2005, 2006). Considering the accumulating evidence of the involvement of the intestinal microbiota during enteric infection (Mullineaux-Sanders *et al.*, 2017), altering the microbiota may compromise the relevance of the data obtained from the use such models. As such, the *C. rodentium* model is our current best *in vivo* model for studying A/E pathogenesis and it has been integral in deciphering roles of specific virulence factors *in vivo* and in characterizing host responses and mucosal immunity to enteric infection.

1.4.1.1. C. rodentium infection of mice

The natural transmission route of *C. rodentium* is the oral-fecal route. In the *in vivo* infection model, where mice are orally gavaged with *C. rodentium*, the pathogen initially colonises the cecum one day post-infection and progresses to the distal colon around 3 days post-infection (Collins *et al.*, 2014). While most of the inoculated bacteria pass through the GI tract, adherent bacteria adapt to the intestinal environment. They undergo a virulence switch that promotes colonic colonisation by ensuring the shedding of hyper-infectious bacteria during the peak of infection (Wiles, Dougan & Frankel, 2005). The genetic background of the mouse determines the severity of the resulting colitis and infection outcome. Certain mouse strains are considered sensitive to *C. rodentium* as they develop severe diarrhoea from day 3 post-infection and succumb to the infection 9-12 days post-infection. Mice considered resistant, such as C57BL/6 mice, clear the infection in 2-3 weeks. At the peak of infection, *C. rodentium* replaces up to 75% of the normal flora in resistant mice and 90% in sensitive mice (Lupp *et al.*, 2007; Berger *et al.*, 2017).

The colonic crypt hyperplasia developed during *C. rodentium* infection results from excessive epithelial regeneration and repair mechanisms (Mundy *et al.*, 2005). In uninfected tissue, intestinal epithelial cells are renewed by $Lgr5^+$ stem cells that can be stimulated by Wingless (Wnt; produced by pericryptal myofibroblasts) to divide asymmetrically. This results in transit-amplifying (TA) cells, which eventually differentiate into luminal intestinal epithelial cells. In *C. rodentium*-infected tissue, however, excessive TA cell-proliferation from $Lgr5^+$ stem cells is triggered by mitogens (Fibroblast Growth Factor 7, FGF7) and cytokines (such as interleukin 11; IL-11) from myofibroblasts, which along with reduced enterocyte shedding results in the characteristic colonic crypt elongation (see **Figure 1.2**). The elongation can be seen by day 7-8 post-infection and peaks at day 12-14 post-infection (Collins *et al.*, 2014).

C. rodentium-induced depletion of goblet cells (which play a role in gut health by secreting mucins that protect against pathogens; Chan *et al.*, 2013; Bergstrom *et al.*, 2008) leads to a loss of epithelial barrier integrity. This allows movement of bacteria to the lamina propria, where they trigger signaling cascades such as NF-κB signallng (Gibson *et al.*, 2008). It is believed that a key difference between *C. rodentium*-resistant and susceptible mice is an enhanced loss of goblet cells in susceptible mice, causing increased exposure of the epithelium to the pathogen, giving more bacterial uptake, antigen presentation and ultimately a hyperactivated T cell response (Kang *et al.*, 2018). The host response to *C. rodentium* infection is discussed next.

1.4.1.2. Host immune responses to C. rodentium infection

During *C. rodentium* infection, immune responses are triggered via pathogen-associated molecular pattern (PAMP)-recognition and Toll-like receptor (TLR)-dependent signalling pathways (mainly TLR2 and TLR4; Kaniuk *et al.*, 2011). In resistant mice, a complex response orchestrated by the innate and adaptive arms of the immune system clears *C. rodentium*. Innate immune cells (neutrophils, macrophages and dendritic cells), intestinal epithelial cells (enterocytes) and helper T cells together with antigen-specific antibody responses promote bacterial clearance. Mice lacking B cells or $CD4^+$ T are more sensitive to *C. rodentium* infection and less able to limit systemic *C. rodentium* dissemination (Vallance *et al.*, 2002; Simmons *et al.*, 2003; Shiomi *et al.*, 2010). The innate immune cells produce and array of pro-inflammatory immunomodulators, including IL-6, IL-23, tumour necrosis factor (TNF), the keratinocyte chemoattractant (KC), IL-1β and IL-18.

The adaptive response to *C. rodentium* is T helper cell type 1 (Th1)-biased and characterised by the production of the immunomodulators interferon-γ (IFNγ), IL-17A and IL-22 by $CD4^+$ T cells (Luperchio & Schauer, 2001; Wiles, Dougan & Frankel, 2005; Collins *et al.*, 2014; Basu *et al.*, 2012; Higgins *et al.*, 1999). IL-17A is pro-inflammatory and promotes neutrophil recruitment while the anti-inflammatory IL-22 (a member of the IL-10 family) stimulates secretion of antimicrobial peptides (RegIIIγ and RegIIIβ) from enterocytes (Collins *et al.*, 2014; Belzer *et al.*, 2011; Higgins *et al.*, 1999; Gallo & Hooper, 2012; Abt & Pamer, 2014; Mukherjee *et al.*, 2014). *C. rodentium* CFU shed in the stool largely correlates with RegIIIγ and RegIIIβ expression levels (Koroleva *et al.*, 2015), which is IL-22 dependent. IL-22-production, in turn, is triggered by IL-23 and mice deficient in either IL-22- or IL-23 succumb to *C. rodentium* infection (Koroleva *et al.*, 2015). However, IL-22-deficient mice are rescued when RegIIIγ is added exogenously (Zheng *et al.*, 2008), emphasizing the role of antimicrobial peptides in this context. The importance of IFNγ lies in enhancing epithelial cell responses and promoting pro-inflammatory lymphocyte phenotypes, supporting the Th1 bias associated with immune responses to A/E pathogens (Luperchio & Schauer, 2001). In line with this, γδ-T cells also contribute to the production of IL-17 and IFNγ (Geddes *et al.*, 2011; Rubino *et al.*, 2013).

Recent work has stressed the role of group 3 innate lymphoid cells (ILC3s) in clearing *C. rodentium* (Ivanov *et al.*, 2009; Fritz *et al.*, 2011; Geddes *et al.*, 2011; Basu *et al.*, 2012; Cella *et al.*, 2009). For instance, mice deficient in the transcription factor RORγt, required for ILC3 development, succumb to the infection (Wang *et al.*, 2010). ILC3s carry the surface lymphotoxin $LT\beta_2LT\alpha_1$, which binds its receptor (LTβR) on dendritic cells, macrophages and enterocytes to enhance their IL-23 production (Wang *et al.*, 2010; Macho-Fernandez *et al.*, 2015; Tumanov *et al.*, 2011). LTβR-deficient mice are sensitive to *C. rodentium* infection (Koroleva *et al.*, 2015) and ILC3s themselves respond to the produced IL-23, increasing their IL-22 expression (Pickert *et al.*, 2009; Grivennikov *et al.*, 2009; Cella *et al.*, 2009; Sugimoto *et al.*, 2008). ILC3s are also able to produce IL-17A. Moreover, LTβR stimulation triggers release of the KC chemokines CXCL1 and CXCL2 that recruit neutrophils to the lamina propria (Wang *et al.*, 2010), thereby contributing to the inflammation.

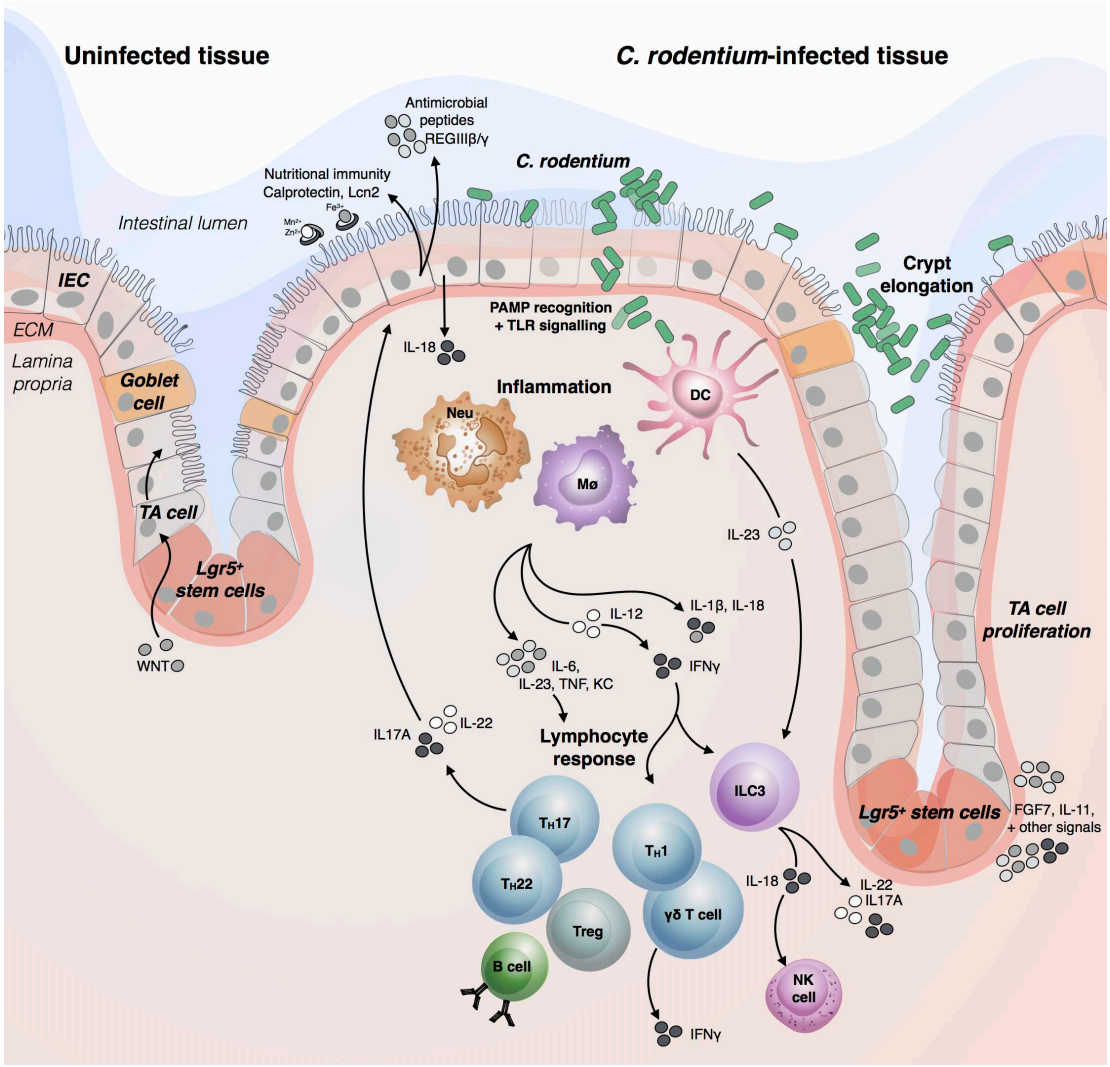


Figure 1.2| Crypt elongation and immune cell recruitment in response to *C. rodentium infection* in the murine gut. Transmissible murine colonic hyperplasia is the pathological hallmark of *C. rodentium* infection. In uninfected tissue intestinal epithelial cells (IECs) are renewed by Lgr5+ stem cells. Wingless (WNT) produced by pericryptal myofibroblasts triggers asymmetric division of Lgr5+ cells, giving transit-amplifying (TA) cells that differentiate into luminal IECs. In infected tissue, excessive TA cell-proliferation is triggered by mitogens (Fibroblast Growth Factor 7, FGF7) and cytokines (such as IL-11) from myofibroblasts. Together with reduced enterocyte shedding, this results in crypt elongation. Upon infection, epithelial barrier integrity is compromised and *C. rodentium* enter the lamina propria. Immune responses are triggered via PAMP recognition and TLR signalling pathways. Dendritic cells (DCs), neutrophils (Neu) and macrophages (Mø) produce an array of pro-inflammatory immunomodulators, including IL-6, IL-23, tumour necrosis factor (TNF) and keratinocyte chemoattractant (KC). Macrophages and DCs secrete IL-1B and IL-18 via NLRP3-dependent inflammasome pathways, contributing to inflammation. Subsequent lymphocyte responses have a Th1-bias. Th17 and Th22 cells secrete IL-17A and IL-22, promoting IEC-production of antimicrobial peptides. Group III innate lymphoid cells (ILC3s) are also implicated in the coordinated response to *C. rodentium* by producing IL-22 and IL-17A and are stimulated by DC-derived IL-23.

Importantly, intestinal epithelial cells play an active role in gut mucosal immunity. Beside forming a physical barrier to microbes (and thereby preventing unnecessary activation of adaptive responses), their secretion of antimicrobial peptides and cross-talk with innate immune cells is needed for eliciting appropriate responses against bacteria such as *C. rodentium* (Kamada *et al.*, 2013; Kruglov *et al.*, 2013; Mortha *et al.*, 2014; Belkaid & Hand, 2014). Further, *C. rodentium*-infected epithelial cells express iNOS (the inducible form of nitric oxide synthase), which can produce reactive nitrogen species and nitric oxide to exert an antimicrobial response, which is known to contribute to clearing the pathogen (Kaniuk *et al.*, 2011; Vallance *et al.*, 2002). Likewise, intestinal epithelial cells respond to IFNγ from immune cells by upregulating expression of NADPH oxidase 1 (Nox1), which stimulates the production reactive oxygen species (ROS) (Kawahara *et al.*, 2004) as an antimicrobial defence.

C. rodentium has developed ways to manipulate many aspects of the host response to favour the infectious process. To appreciate how *C. rodentium* uses the secretion of virulence factors to overcome or hijack host cellular pathways during infection, the following pages will review bacterial secretion systems and their contribution to enteric infection.

1.5. Bacterial secretion systems

The ability to transport proteins, DNA and small molecules across their cell envelope is essential for bacterial physiology, survival and replication. Bacteria employ an array of sophisticated secretion apparatuses to achieve this, including injection of virulence factors into target cells. Six classes of secretion systems are well-characterized in Gram-negative bacteria (T1SS-T6SS) and are broadly classed into those that translocate substrates across both the inner and outer membranes (**Figure 1.4**) and those that translocate substrates across the outer membrane only (**Figure 1.3**) (Costa *et al.*, 2015). The latter rely on inner-membrane transporters such as the general secretory (Sec) pathway or twin arginine translocation (Tat) system to carry substrates from the cytoplasm to the periplasm. This section briefly describes the different secretion systems and their specific contribution to pathogenicity in enteric pathogens. Due to the nature of this project, the T3SS is discussed in more detail.

1.5.1. General secretion mechanisms

The Sec and Tat pathways are universal to eubacteria, archaea and eukaryotes (Natale, Brüser & Driessen, 2008). They deliver substrates to the periplasm in Gram-negative bacteria and to the extracellular space in Gram-positives and archaea. The Sec system comprises three fundamental

parts: a chaperone (SecB) to recognize substrates and maintain them in an unfolded translocation-competent state, a cytoplasmic ATPase (SecA) to power translocation and a membrane-embedded heterotrimeric pore (consisting of SecY, SecE and SecG). SecB recognizes a 20-amino acid signal sequence in the substrate N-terminus that contains three specific regions: an amino-terminal region, a hydrophobic middle region and a polar region (Natale, Brüser & Driessen, 2008). Following translocation, the signal sequence is cleaved by a periplasmic membrane-anchored signal peptidase.

The Tat machinery recognizes a similar cleavable N-terminal signal to the Sec system, but with a consensus basic amino acid-rich motif (S-**R**-**R**-x-F-L-K) prior the hydrophobic region. Moreover, its substrates are folded proteins and secretion is powered using the proton gradient between the periplasm and the cytoplasm. In *E. coli,* the Tat translocase comprises a TatA, TatB and TatC membrane-embedded pore, where TatC recognises substrates. The Tat system is known to be important for the virulence of many pathogens, including EHEC (Pradel *et al.*, 2003). Secretion by the Sec or Tat system may be followed by translocation via a T2SS or a T5SS, which are discussed next.

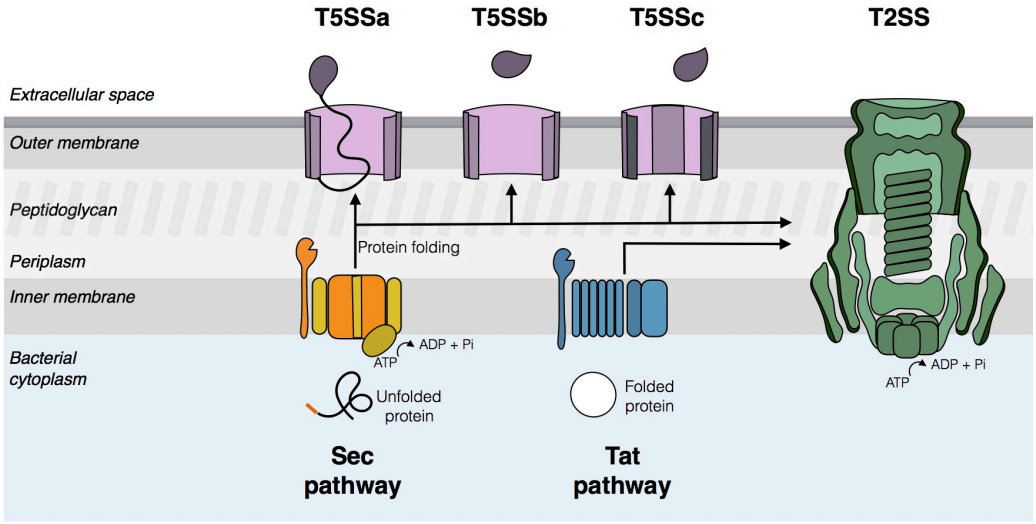


Figure 1.3| The T2SS and T5SS transport proteins across the bacterial outer membrane. Two bacterial secretion systems translocate substrates from the periplasm to the extracellular milieu. Substrates must first be translocated to the periplasm by the Sec- or Tat pathways. The Sec pathway transports unfolded proteins (with N-terminal sequence signal) in an ATP hydrolysis-dependent manner. The Tat pathway transports folded proteins with a conformational secretion signal. The T2SS uses a piston-like motion to transport substrates to the extracellular space. The T5SS can be divided into 3 subgroups, a-c, based on the number of proteins involved. T5SSa autotransporters secrete themselves and comprise just one polypeptide. T5SSb consists of two polypeptides (one per functional domain) while T5SSc requires polypeptides trimerisation to form a channel. See text for details; drawn with help from D. Ruano-Gallego.

1.5.2. The type II secretion system

The T2SS transports proteins from the periplasm across the outer membrane and into the extracellular environment following either the cleavage of a Sec signal sequence and protein folding or the cleavage of a Tat signal sequence. As T2SS substrates must be folded, the T2SS secretion signal is likely conformational rather than linear. T2SS substrates from pathogenic and non-pathogenic bacteria hold a range of biological functions and many are proteases, lipases, phosphatases and carbohydrate-processing enzymes (Green & Mecsas, 2016; Korotkov, Sandkvist & Hol, 2012).

The T2SS is an assembly of 12-15 different proteins, further divided into four subassemblies: (1) the outer-membrane complex (OMP, a channel composed of secretin, general secretin pathway protein, GspD), (2) the inner-membrane platform (IMP, comprising at least GspC, GspE, GspL and GspM), (3) the cytoplasmic ATPase (GspE), and (4) the pseudopilus (comprising GspG, GspH, GspI, GspJ, GspK; Green & Mecsas, 2016). The IMP extends from the inner membrane to secretin in the periplasm and coordinates translocation by making contact with the secretin, the pseudopilus and the ATPase. Substrates interact with secretin, stimulating the ATPase to drive polymerisation of the pseudopilus. The pseudopilus bears structural similarity and evolutionarily relatedness to the bacterial Type IV pilus, and its polymerisation pushes substrates through the outer membrane channel in a piston-like motion (Green & Mecsas, 2016). T2-secreted toxins include the cholera toxin in *Vibrio cholera* and LT in ETEC (Reichow *et al.*, 2010). The T2SS is required for *Legionella pneumophila* and *Yersinia enterolitica* virulence (Rossier, Dao & Cianciotto, 2009; Iwobi *et al.*, 2003). However, there is no evidence of its importance for *C. rodentium* virulence despite the presence of a T2SS cluster within the *C. rodentium* genome.

1.5.3. The type V secretion system

T5SS substrates are unique in that instead of employing a dedicated secretion apparatus for their translocation, they secrete themselves following their Sec-dependent export to the periplasm. T5SSs are divided into three classes based on the number of proteins implicated in the translocation, the predominant class being Type 5a autotransporters. These single polypeptides consist of three or four modular domains, including a 12-stranded β-barrel translocator domain that inserts as a pore into the outer membrane (Green & Mecsas, 2016). They carry a passenger domain (containing the functional part of the autotransporter) that is separated from the translocator domain by a flexible linker domain. Some also carry a protease domain to release the passenger domain following its export through the pore. The second class of T5SS (Type 5b)

is also known as two-partner secretion as the translocator and passenger domain are separate proteins. The third class (Type 5c) requires the trimerisation of three polypeptides to form the β-barrel channel (Green & Mecsas, 2016).

Most known T5SS substrates are virulence factors: either toxins or receptor-binding proteins. For instance, the immunoglobulin A protease of *Neisseria gonorrhoea* cleaves host antibodies, and YadA of *Y. enterolitica* is an important adhesin (El Tahir & Skurnik, 2001; Heise & Dersch, 2006). The *C. rodentium* genome encodes 20 putative autotransporters: 14 were identified owing to homology to AIDA-I, TibA or pertactin – known adhesins that contribute to the virulence of EPEC, ETEC and *Bordetella pertussis*, respectively (Petty *et al.*, 2010; Benz & Schmidt, 1989, 1992; Lindenthal & Elsinghorst, 1999; Elsinghorst & Weitz, 1994; Leininger *et al.*, 1991). Two of the *C. rodentium* autotransporters belong to the serine protease autotransporter of the Enterobacteriaceae (SPATE) family, which are important virulence factors in a number of pathogens (Boisen *et al.*, 2009; Dutta *et al.*, 2002). The *C. rodentium* SPATEs are encoded on the large virulence plasmid pCROD1, together with a TibA-like adhesin (Petty *et al.*, 2010). Furthermore, the adhesin intimin that is required for tight adherence of A/E pathogens to epithelial cells (discussed in **Section 1.5.8.6**) is an autotransporter, although its translocator and passenger domain are in the opposite topology to what is commonly found (Grijpstra *et al.*, 2013).

1.5.4. The type I secretion system

The T1SS delivers substrates of up to 900 kDa across both bacterial membranes in a single step (Thomas *et al*., 2014). It was identified in UPEC, where it secretes the pore-forming toxin Hemolysin A (HlyA; Mackman & Holland, 1984). The HlyA system is the paradigm T1SS and has three indispensable components: (1) an ATP-binding cassette (ABC) transporter (HlyB) to power translocation by adenosine triphosphate (ATP) hydrolysis, (2) a membrane fusion protein (MFP, HlyD) spanning the periplasm and forming a continuous channel to the cell surface, and (3) a pore-forming outer membrane factor (OMF, TolC) connected to the MFP (Thomas, Holland & Schmitt, 2014). Substrates are translocated in an unfolded state and are typically recognised by the ABC transporter via a non-cleaved secretion signal sequence in the substrate C-terminus (Delepelaire, 2004). However, the cytoplasmic N-terminus of the MFP is also believed to contribute to substrate selection (Green & Mecsas, 2016).

Intriguingly, the T1SS bears structural similarity to ABC-type drug efflux pumps (which also have a TolC component) that bacteria employ to export toxic compounds such as antimicrobials (Lubelski, Konings & Driessen, 2007). BapA, a substrate of the SPI-9 T1SS in *Salmonella entercia* serovar Enteriditis, is required for colonisation *in vivo* and for biofilm formation *in vitro* (Latasa *et al.*, 2005). *C. rodentium* encodes four putative T1SSs (with approximately 50% amino acid identity) within separate genomic islands (GI2, GI6, GI8 and GI13; Petty *et al.*, 2010). Each island also encodes a putative substrate: a large repetitive protein of varied size and sequence. Whether the *C. rodentium* T1SSs contribute to virulence remains unknown, but the GI8 T1SS substrate bears sequence similarity to the large repetitive protein of the *S.* Typhi SPI-9 T1SS, which confers adherence to epithelial cells when *S.* Typhi are grown under high osmolarity or low pH (Velásquez *et al.*, 2016). Moreover, the putative GI13 T1SS substrate bears sequence similarity to O island 28 in EHEC EDL933 (Petty *et al.*, 2010).

1.5.5. The type IV secretion system

The T4SS is functionally versatile: it can not only translocate DNA, proteins and DNA-protein complexes into eukaryotic and prokaryotic cells, but also secrete or import genetic material from the extracellular milieu (Wallden, Rivera-Calzada & Waksman, 2010). T4SSs are separated into (1) DNA conjugation systems, (2) effector translocators and (3) contact-independent DNA/protein exchange systems (Christie, 2016). Most T4SSs belong to the first group and the paradigm T4SS is the conjugative *Agrobacterium tumefaciens* VirB/D system, which transports oncogenic transfer DNA (T-DNA) into plant cells. It consists of 12 different proteins (VirB1-VirB11 and VirD4), that form a periplasm-spanning tube for substrate export. Substrates are recognised by VirD4 at the base of the machinery and are transferred to an ATPase complex at the cytoplasmic side of the inner membrane. ATP-hydrolysis powers translocation through the tube. T4SSs that transport DNA use an extracellular pilus for the translocation, but the mechanistic details of this are not fully understood (Green & Mecsas, 2016).

In the context of bacterial infection, a T4SS pilus may serve as an attachment device or an effector-translocating T4SS may deliver virulence factors into competing bacteria or host cells. *L. pneumophila*, for example, requires an effector-translocating T4SS for host cell invasion and subsequent intracellular replication (Segal, Feldman & Zusman, 2005). Only a subset of its 300 known effector proteins have been designated a role during infection, but many of them are known to manipulate host ubiquitin responses or remodel the host cell architecture (Maculins *et al.*, 2016b). In *E. coli*, however, T4SS substrates are restricted to mobile genetic elements

(Christie, 2016; Low *et al.*, 2014). Thus, T4SSs contribute *E. coli* pathogenesis by shaping bacterial genomes and disseminating virulence of antibiotic resistance genes via pili, rather than injecting effector proteins.

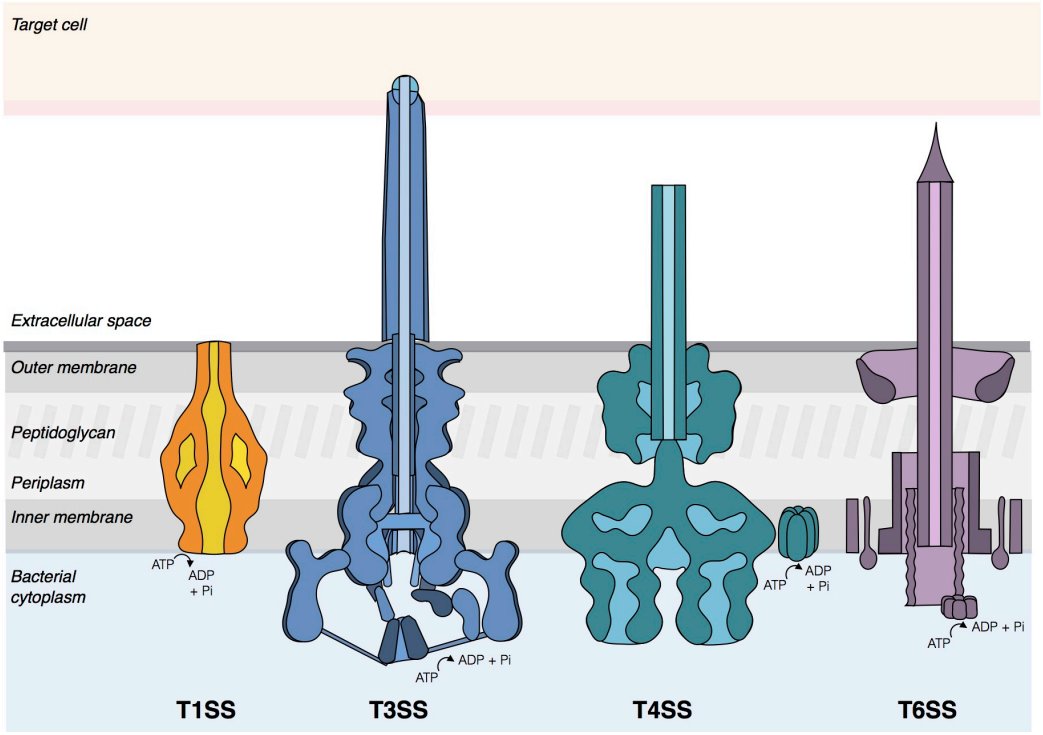


Figure 1.4| Bacterial T1SS, T3SS, T4SS and T6SS translocate substrates across both the inner and outer bacterial membranes. The T1SS and T4SS translocate substrates in an ATP-dependent manner into the extracellular milieu. The T3SS and T6SS can puncture target cells (eukaryotic or bacterial) to deliver effector proteins directly to target cell cytosol. See text for details; figure drawn with helpful input from Dr David Ruano-Gallego.

1.5.6. The type VI Secretion System

The T6SS delivers toxic effector proteins into both eukaryotic and prokaryotic cells, thereby contributing to both pathogenesis and interbacterial competition. Following its *in silico* discovery in *V. cholera*, T6SS genes have been described in a quarter of sequenced Gram-negative genomes (Williams *et al.*, 1996; Das & Chaudhuri, 2003; Mougous *et al.*, 2006). Many T6SS effectors target the bacterial cell wall and membrane (such as peptidoglycan hydrolases), which supports the role of the T6SS bacterial competition. In line with this, many of them are encoded alongside genes that provide immunity to the effector, giving the T6SS^{+} bacterium protection against self-intoxication (Hachani, Wood & Filloux, 2016).

The T6SS is a contractile weapon structurally-, functionally- and evolutionarily related to the puncturing tails of T4 bacteriophages (Leiman *et al.*, 2009). It is formed of a needle-like structure anchored to the cell membrane via a membrane complex (composed of TssJ, TssM and TssL proteins; Silverman *et al.*, 2012). The needle consists of an inner tube (formed by Hcp proteins) enveloped by a contractile sheath (TssB and TssC), capped by a pore-forming spike protein (VgrG) and tethered to the membrane complex via a baseplate complex (TssE, TssE, TssG, TssK). VgrG proteins are often tipped with a PAAR protein, that sharpens the needle, and some VgrGs carry additional C-terminal domains with enzymatic activity that enhance virulence. Following a cell-contact dependent stimulus, the sheath contracts to eject the tail and spike through the target membrane (Basler, 2015; Basler & Mekalanos, 2012; Basler, Ho & Mekalanos, 2013). Contraction provides the energy for the secretion of effector proteins (Russell, Peterson & Mougous, 2014), while a cytoplasmic chaperone with ATPase activity (ClpV) recycles the sheath subunits (Journet & Cascales, 2016).

The *C. rodentium* genome has two T6SS gene clusters: *C. rodentium* Type Six Secretion system cluster 1 (CTS1; 39 genes) and 2 (CTS2, 16 genes) (Petty *et al.*, 2010). They both encode the 13 components necessary for a functional T6SS in other species (Zheng & Leung, 2007). CTS1 is almost identical to an *Enterobacter cancerogenus* cluster. Its *icmF* gene has a frameshift mutation giving a premature stop codon, but this is upstream of a poly-A tract that enables gene regulation by transcriptional slippage – as seen for *mxiE* in *Shigella* spp. (Penno, Sansonetti & Parsot, 2005) and *pyrBI* in *E. coli* (Liu, Heath & Turnbough, 1994). Both gene products are required for a functional T6SS, and it remains unknown what regulates the frameshifting efficiency (Journet & Cascales, 2016). The CTS2 *icmF* also has a frameshift mutation, but Hcp can still be detected in cell culture supernatants (Gueguen & Cascales, 2013), implying that the T6SS remains functional.

In terms of function, T6SS-expressing *C. rodentium* outcompetes co-cultivated *E. coli,* indicating that a *C. rodentium* T6SS plays a role in bacterial competition. T6SS-mediated killing has also been demonstrated between EAEC and *E. coli* K-12 (Journet & Cascales, 2016). Differences between CTS1 and 2 suggest that they have different evolutionary origins, and an EHEC T6SS distinct from CTS1/2 has high homology to the *Shigella sonnei* T6SS. The lack of homology here could reflect that the T6SS is not essential for infecting human cells, but might convey strain-specific adaptations to environmental niches (Petty *et al.*, 2010). $T6SS^{+}$ but not $T6SS^{-}$ EAEC was shown to elicit diarrhoea in human volunteers (Nataro *et al.*, 1995), which might reflect its ability

to help clear competing bacteria (Journet & Cascales, 2016) – an idea also supported by the presence of a T6SS in non-pathogenic *E. coli* strains.

1.5.7. The type VII secretion system

Interestingly, a type VII secretion system (T7SS) has been described in Gram-positive bacteria such as *Mycobacterium* (which have a Gram-negative-like cell envelope) and *Staphylococcus aureus*. *S. aureus* T7SS mutants have attenuated pathogenicity (Burts *et al.*, 2005) and most pathogenic mycobacterial species possess several T7SSs shown to contribute to pathogenicity (Shah & Briken, 2016). For instance, the tubercolusis live vaccine strain *M. bovis* BCG has a genomic deletion that corresponds to its T7SS ESX-1 (Pym *et al.*, 2003).

1.5.8. The type III secretion system

The T3SS resembles a molecular syringe that allows injection of bacterial effector proteins straight into eukaryotic host cells. It is a key virulence factor for many human Gram-negative bacterial pathogens, including *Yersinia* spp., *Salmonella* spp., *Pseudomonas* spp and A/E pathogens (Gaytán *et al.*, 2016). It was first described in *Y. enterolitica* following the discovery of Sec-independent translocation of Yop proteins into host cells (Michiels & Cornelis, 1991). The *Salmonella* T3SS machinery was visualised by negative staining and electron microscopy in 1998 (Kubori *et al.*, 1998) and the EPEC T3SS shortly after (Sekiya *et al.*, 2001; Daniell *et al.*, 2001).

The T3SS consists of a series of concentric rings that span the inner and outer bacterial membranes and form a continuous channel from the bacterial cytosol to the exterior of the cell. It can be divided into four main components: a needle with a translocation pore that inserts into the membrane of target cells, a basal body, an export apparatus and cytoplasmic protein complexes (see **Figure 1.5** for details). In A/E pathogens, key features of the T3SS are encoded within the LEE PAI, described next.

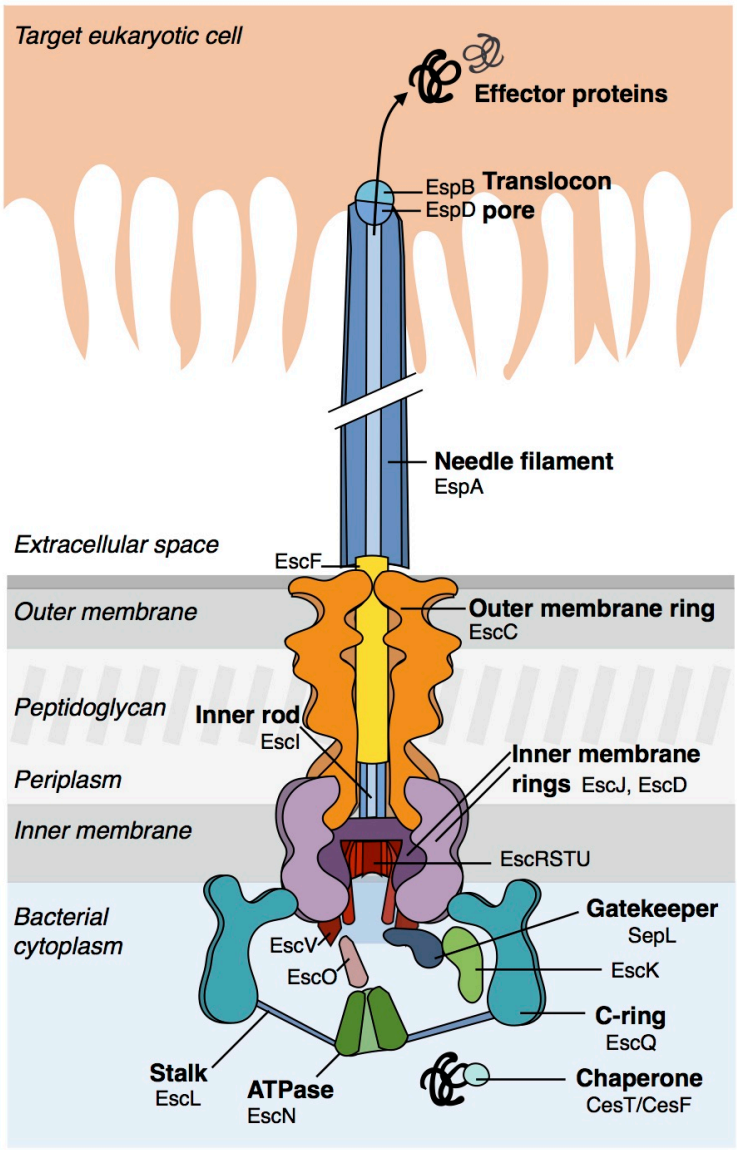


Figure 1.5 | Schematic representation of the T3SS of *E. coli* and *C. rodentium*. It comprises an export apparatus (in red, channel-like structure formed by EscRSTUV within the inner membrane rings), a base complex (in purple and orange), a needle (in yellow), a needle filament (in blue), and cytoplasmic components. EscJ and EscD form the inner and outer MS rings, respectively, while EscC forms the outer membrane ring. EspA polymerisation at the tip of the 10 nm EscF needle generates a filamentous conduit for protein translocation. At the distal end of the channel is an EspB/D pore that inserts into the host membrane. The ATPase EscN powers the translocation of effector proteins through the inner rod (EscI), needle and needle filament and this process is aided by relevant chaperones. The EscN, EscL and Orf15 components are collectively known as the ATPase-complex and are thought to be anchored by the EscQ C-ring. Figure adapted from Slater *et al*., 2018.

1.5.8.1. The Locus of Enterocyte Effacement-encoded T3SS

As mentioned in **Section 1.4**, the T3SS-encoding LEE PAI is the defining genetic feature of A/E pathogens and is quintessential to their virulence. The LEE of A/E pathogens are highly related both in terms of gene repertoire and organisation. A low G+C content (38.3% compared to 50.8% overall for the genome) suggests acquisition by horizontal gene transfer from other species. The LEE comprises 41 conserved open reading frames (ORFs) organised in the 5 main operons and a number of smaller transcriptional units, shown in **Figure 1.6** (Slater *et al.*, 2018). These encode the structural components of a T3SS, the adhesin intimin, core effector proteins, a lytic transglycosylase, chaperones and regulatory proteins (Pallen, Beatson & Bailey, 2005). Apart from the monocistronic EscD, the T3SS structural components are encoded within the operons *LEE1*, *LEE2*, *LEE3* and *LEE4*, and these have been shown to be sufficient for an assembled and functional T3SS (Ruano-Gallego, Álvarez & Fernández, 2015). The following pages will review our current understanding of the T3SS of A/E pathogens in terms of assembly, regulation, function and its contribution to bacterial pathogenesis.

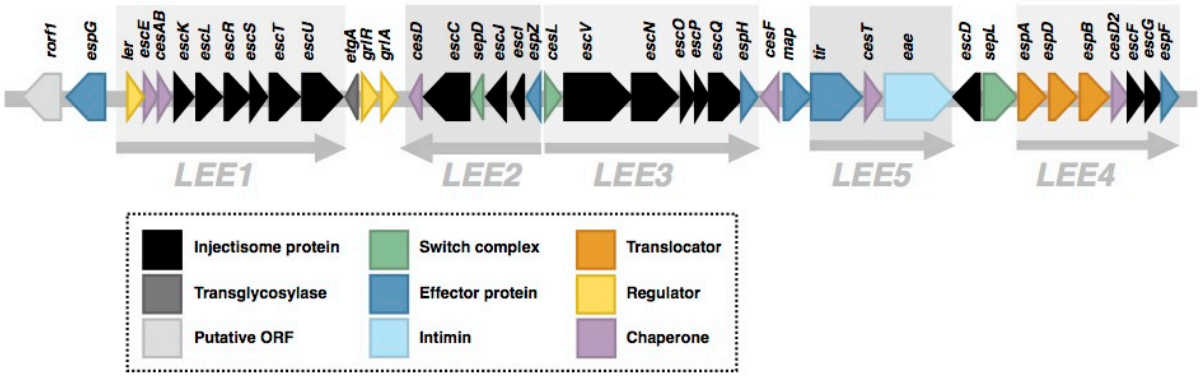


Figure 1.6 | The genetic organisation of the Locus of Enterocyte Effacement (LEE). The LEE genomic island is organised into five operons (*LEE1-LEE5*) containing the genes necessary to assemble a functional T3SS and its associated core effector proteins and relevant chaperones and regulatory proteins. Differences between EPEC, EHEC and *C. rodentium* LEE pathogenicity islands are minor. Figure adapted from Slater *et al.*, 2018.

1.5.8.2. Structure and assembly of the T3SS of A/E pathogens

Signals from transcriptional regulators, chaperones, environmental cues and molecular switches converge to assemble the active T3SS in four stages: (1) basal body and export apparatus assembly; (2) inner rod and needle assembly; (3) filament and translocon assembly and finally (4)

effector protein secretion (reviewed in Slater *et al.*, 2018). The resulting *E. coli* T3SS comprises a 3.5 MDa structure with a width of approximately 8-9 nm (Sekiya *et al.*, 2001). The T3SS of A/E pathogens differ from those of organisms such as *Shigella* and *Yersinia* spp. by having a long proteinaceous filament of polymeric EspA (2.5 Å internal diameter, close to that of the structure it sits upon) attached to the distal end of the needle (schematic representation in **Figure 1.5**). This makes the A/E T3SS around 90 nm long, but able to extend up to 700 nm, which is significantly longer than for example the T3SS of *Yersinia* spp. (average length 58 ± 10 nm) (Sekiya *et al.*, 2001).

An "inside out"-model has been proposed for the assembly of the T3SS of A/E pathogens and *Salmonella* spp. (Wagner *et al*., 2010), meaning that the export apparatus is formed prior to the outer membrane (OM) ring. EscR, EscS and EscT form a stable complex, which recruits the export gate ring (consisting of EscV) and the autoprotease EscU, making the export apparatus EscRSTUV (Diepold, Wiesand & Cornelis, 2011; Diepold *et al.*, 2012). EscU cleavage is a signal for joining the export apparatus to the basal body. The inner membrane (IM) ring is then formed, consisting of an EscD oligomeric ring around an EscJ ring in the inner membrane. These proteins are translocated to periplasm by the Sec pathway before coming together with the export apparatus. On their periplasmic side, they interact with EscC of the OM ring (Creasey *et al.*, 2003; Ogino *et al.*, 2006) and presumably also the OM ring protein EscJ. EscC belongs to a family of secretin proteins, which in many T3SSs are regulated by the lipoprotein pilotin, but in the T3SS of A/E pathogens by other T3SS components (Gauthier & Finlay, 2003). Next, the EscQ cytosoling ring (C-ring) sorting complex (Lara-Tejero *et al.*, 2011) is formed at the foot of the basal body (Pallen, Beatson & Bailey, 2005), further recruiting EscN and EscL that are components of the ATPase complex (Biemans-Oldehinkel *et al.*, 2011).

Binding of EscO to EscN completes the ATPase complex and activates it by inducing a conformational change in EscL (Romo-Castillo *et al.*, 2014). Hexameric EscN is at the core of its function, with a central conserved ATPase domain. Its C-terminus is thought to recognize T3SS substrates and EscN is proposed to serve as a docking site for chaperone-substrate complexes (Gauthier & Finlay, 2003), enabling their ATP hydrolysis-dependent uncoupling and unfolding of proteins to be secreted (as the needle pore is only 2-3 nm wide). Due to structural similarity to the flagellar and F1 ATPase, it is assumed that a combination of ATP hydrolysis and proton motive force powers translocation (Ibuki *et al.*, 2011; Imada *et al.*, 2007; Romo-Castillo *et al.*, 2014; Zarivach *et al.*, 2007). However, despite many well-described T3SS structures (Hodgkinson *et al.*, 2009; Schraidt & Marlovits, 2011; Nans *et al.*, 2015; Kudryashev *et al.*, 2013) and the

observation of effector delivery secretion through the needle conduit (Radics, Königsmaier & Marlovits, 2014), a clear consensus on the exact mechanisms by which secretion occurs and is energised has not been reached (Lee & Rietsch, 2015).

Following docking on the ATPase complex, the structure can secrete the proteins necessary for the construction of the inner rod, needle, tip and translocon. EscI monomers create the inner rod, and EscI is also thought to have a role in substrate regulation. As the structure penetrates the peptidoglycan (PG) layer, a transglucosylase enzyme is encoded within the LEE to help break down PG (García-Gómez *et al.*, 2011). EscF polymers also assemble to make the needle of the T3SS (Gaytán *et al.*, 2016). Needle completion switches the secretion of substrates to mid and late substrates, to give assembly of the filament and translocon (Buttner, 2012; Minamino & Macnab, 2000). Separate models that describe the switch exist (Journet *et al.*, 2003; Wood, Jin & Lloyd, 2008; Creasey *et al.*, 2003; Sal-Man, Deng & Finlay, 2012), but are not discussed here.

The filament, composed of EspA, extends from the EscD structure (Daniell *et al*., 2001). EspA subunits are translocated via the T3SS and self-polymerise. The completed filament has proposed roles in adhesion to epithelial cells and sensing epithelial cells (Cleary *et al.*, 2004). Intriguingly, EspA filaments are absent from A/E lesions, as they are disassembled by an unknown mechanism following intimate adherence to cells during infection (Dahan *et al.*, 2004). At the end of the filament, EspD and EspB form a translocon pore with a pore size of 3-5 nm that can puncture the host cell membrane (Ide *et al.*, 2001; Wachter *et al.*, 1999).

1.5.8.3. T3SS chaperones

Chaperones help maintain T3SS structural components and effector proteins in a translocation-competent state in the bacterial cytoplasm and deliver them to the export gate (Thomas *et al.*, 2012). Further, they are thought to play a role in defining the substrate hierarchy. Within the LEE, eight chaperones have been described. They are structurally very similar – small with very low pI values – and there is no conserved signal sequence between them. Therefore, they are categorised based on their substrate specificity (Tobe *et al.*, 2006). Class IA (CesF and CesL) specifically bind one effector protein, Class IB (CesT) bind multiple effector proteins, Class II (CesAB, CesD and CesD2) bind translocator proteins and Class III (EscE and EscG) bind needle subunits (Gaytán *et al.*, 2016; Izoré, Job & Dessen, 2011; Thomas *et al.*, 2005). In general, the chaperones bind to a region in the N-terminus of these proteins, located within the first 50-100

amino acids (Stebbins & Galán, 2001; Lilic, Vujanac & Stebbins, 2006) after a secretion signal (usually in the first 15 amino acids (Lloyd *et al.*, 2001; Ehrbar *et al.*, 2003).

1.5.8.4. Substrate switch

Following the assembly of the T3SS, a substrate switch occurs. Secretion of translocator proteins is replaced by secretion of effector proteins, which are injected into host cells. This involves a regulatory complex consisting of SepD, SepL and CesL. SepD and SepL are commonly referred to as the gatekeeper proteins, as they promote translocation or secretion whilst preventing secretion of late substrates (Deng *et al.*, 2015; O'Connell *et al.*, 2004; Deng *et al.*, 2005). The mechanisms behind the promotion of translocation is unclear. The prevention of secretion of late substrates, on the other hand, involves the interaction of a SepL-SepD complex with EscV, which represses the affinity of EscV for chaperone-effector pairs (Portaliou *et al.*, 2017). Further, the interaction between SepL and the effector protein Tir (translocated Intimin receptor, the first effector protein to be secreted) also prevents translocation of other effectors (Thomas, Stafford & Hughes, 2004; Wang *et al.*, 2008). The Tir binding site of SepL is shared with EscD (within the inner membrane ring), and this might be involved in relieving the suppression of late effector secretion (Wang *et al.*, 2008). In A/E pathogens, removing calcium from the growth medium switches secretion specificity from translocators to effectors (Deng *et al.*, 2005; Gaytán *et al.*, 2018; Ide *et al.*, 2003; Kenny *et al.*, 1997). Contact between EspA, EspB or EspD with the host membrane was recently shown to not be important in this switch, in contrast to what was previously thought (Gaytán *et al.*, 2018).

1.5.8.5. Regulation of the LEE

Synthesising the T3SS and effectors bears growth rate- and fitness costs to bacteria, necessitating tight regulation of LEE gene expression. As such, A/E pathogens sense and respond to environmental cues derived from the host and from the microbiota (Umanski, Rosenshine & Friedberg, 2002; de Nisco, Rivera-Cancel & Orth, 2018; Carlson-Banning & Sperandio, 2018). Following its ingestion, the pathogen survives the low stomach pH thanks to acid resistance (Nguyen & Sperandio, 2012) and its virulence factors are gradually turned on in response to environmental signals as it travels through the intestine, through the mucus layer towards enterocytes (Connolly, Brett Finlay & Roe, 2015; Furniss & Clements, 2018). The signals include host-derived stress-response hormones epinephrine (Epi) and norepinephrine (NE), the quorum sensing molecule autoinducer-3 (AI-3) produced by the microbiota (Clarke *et al.*, 2006; Hughes *et al.*, 2009) as well as oxygen availability, the presence of fucose in the lumen (Pacheco *et al.*,

2012) and carbon source availability (reliant on metabolic status; Njoroge *et al.*, 2012) and mechanical forces (Alsharif *et al.*, 2015).

Environmental signals are largely integrated by A/E pathogens via two-component signal transduction systems and phosphorylation-independent transcriptional regulators. Over 40 LEE regulators have been described and regulation occurs both on the translation and post-translational level (Bhatt *et al.*, 2016). Many of the systems influence the master LEE regulator, Ler (see **Figure 1.7**), which activates the transcription of operons *LEE2-LEE5* by alleviating repression by the global negative regulator H-NS (Shin, 2017). Ler is encoded within the *LEE1* operon and is constitutively expressed at low levels, but also downregulates *LEE1* transcription, in a negative feedback loop (Bhat *et al.*, 2014).

Simple dietary sugars are absorbed in the small intestine and are unavailable to pathogens that colonise the large intestine, including *C. rodentium* and EHEC. Intestinal pathogenic *E. coli* avoids competition with commensals for these by invading the mucosal layer (covered in glycans) and adhering to enterocytes. A prominent member of the intestinal commensal microbiota, *Bacteroides thetaiotamicron*, cleaves host glycans into the sugar fucose (Fischbach & Sonnenburg, 2011). *C. rodentium* senses and responds to high levels of fucose via their FusKR two-component system, repressing *ler* in the initial stages of infection. Downregulating the T3SS while the pathogen resides in the lumen is an important energy-saving measure to help outcompete commensals (Hughes *et al.*, 2009) and is aided by the two-component system CpxCR (a Cpx envelope stress response), which also represses LEE genes (De la Cruz *et al.*, 2016).

Positive regulators of LEE-genes include Cra, KdpE and QseC. The phosphorylation-independent regulator Cra directly senses metabolic intermediates in the glycolysis pathway. For instance, fructose-1-phosphate and fructose-1,6-biphosphate bind the induces domain of Cra, preventing it from binding its target promoters. This suppresses *ler* expression during glucose-based respiration and promotes *ler* expression in gluconeogenic conditions (Njoroge *et al.*, 2012). Cra acts alongside KdpE, a response regulator that is phosphorylated by the non-cognate sensor histidine kinase QseC in response to host derived hormones Epi and NE as well as microbiota-derived AI-3. Upon phosphorylation, KdpE binds the *ler* promoter to promote efficient transcription initiation (Njoroge *et al.*, 2012).

A further two transcriptional regulators are also phosphorylated by QseC: QseB and QseF. QseB uses the small ribonucleic acid GlyM to indirectly repress *LEE4* and *LEE5*. QseB also activates

flagella and motility genes. QseF indirectly promotes Stx expression in EHEC (Hughes *et al.*, 2009) and expression of the effector Tccp/$EspF_U$, which is essential for A/E lesion formation (Garmendia *et al.*, 2004). They both repress the FusKR two-component system to alleviate fucose-mediated *ler*-repression, and this shifts the regulatory balance and gives expression of the entire LEE (Pacheco *et al.*, 2012). As the pathogen gets closer to the epithelium, Cra also senses the presence of oxygen and activates *ler* transcription (Carlson-Banning & Sperandio, 2018). Alongside *ler* regulation, both Cra and KdpE control the transcription of other genes, including TccP and certain putative EHEC virulence factors. For instance, Cra but not KdpE regulates expression of EspG, a LEE-encoded effector encoded adjacent to *ler* (Njoroge, Gruber & Sperandio, 2013).

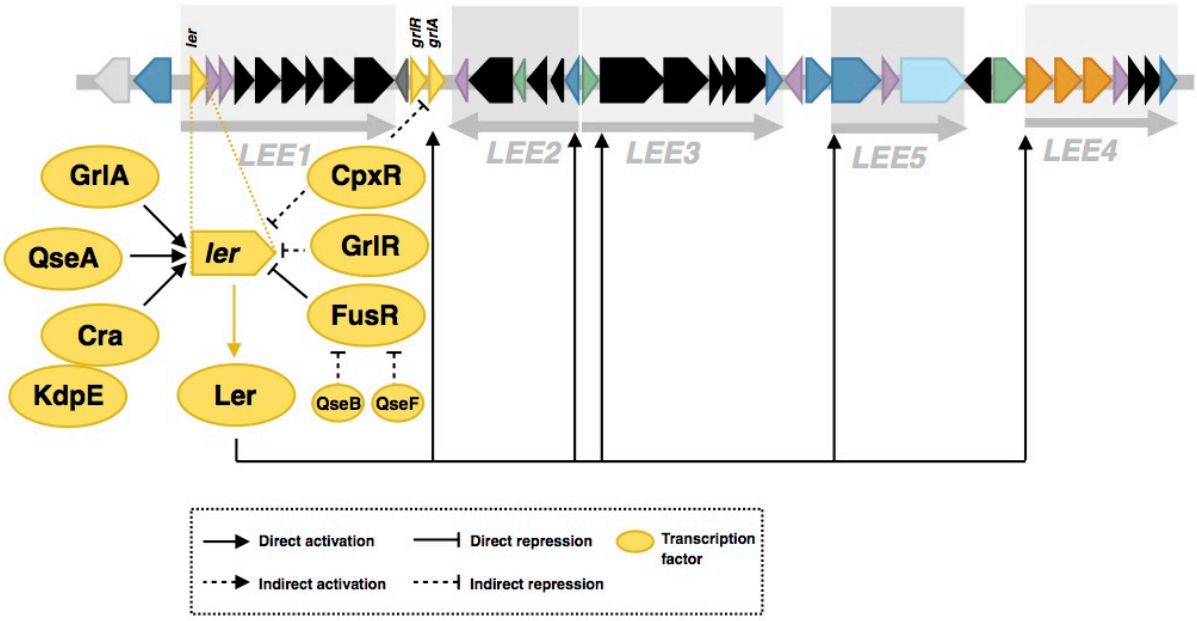


Figure 1.7 | The regulation of the LEE pathogenicity island. Ler is the master regulator of the LEE. External signals from two-component systems (such as FusKR, QseBC, QseE and CdxAR), LysR-type transcription factors (QseA, QseD and LeuO) and response regulators (Cra and KdpeE) regulater ler transcription. Further, Ler-mediated LEE-expression is fine-tuned by the LEE-encoded regulators GrlA and GrlR.

Once two-component system-mediated expression of LEE occurs, transcript levels are fine-tuned by GrlA (activates *ler* transcription) and GrlR (binds antagonisitically to GrlA to repress *ler* transcription) (Deng *et al.*, 2004). Overall, the regulatory mechanisms described here give a high level of control of the expression of T3SS-related genes and unsurprisingly, mutant strains lacking genes encoding the sensor kinase genes FusK, QseC or CpxA show reduced virulence *in vivo* (Pacheco *et al.*, 2012; Clarke *et al.*, 2006; De la Cruz *et al.*, 2016; Moreira *et al.*, 2016). In line with this, *C. rodentium* is less able to colonise mice that do not produce Epi and NE ($Dbh^{-/-}$) (Moreira *et*

al., 2016). Moreover, *cra* knockout strains have reduced pedestal formation on cultured epithelial cells in *vitro* (De la Cruz *et al.*, 2016).

1.5.8.6. LEE-encoded T3SS effector proteins

T3SS effector proteins allow striking yet fine-tuned subversion and manipulation of host responses and pathways upon A/E infection. The following six effector proteins are LEE-encoded: Tir (in *LEE5*), Map (in its monocistronic operon), EspF (in *LEE4*), EspZ (in *LEE2*), EspH (in *LEE3*) and EspG (3' of *LEE1*) (**Figure 1.6**). In addition, the adhesin intimin is encoded within *LEE5*. Each of these contribute to pathogenicity (Ritchie & Waldor, 2005) by one or more function, either on their own or together with other effector proteins. Their respective roles are outlined in this section.

Tir triggers a cascade that results in actin rearrangement in infected host cells. It is essential for A/E lesion formation and is arguably the most studied LEE effector protein. Once translocated, Tir inserts into the host cell OM in a hairpin-loop topology where it acts as the receptor for the bacterial OM adhesin Intimin (Hartland *et al.*, 1999; Kenny, 1999). This allows intimate attachment to the surface of infected cells and leads to the formation of A/E lesions (Wong *et al.*, 2011). The interaction between Tir and Intimin triggers clustering of Tir, which leads to downstream signaling events that result in actin-rich pedestal-like structures directly beneath the adherent bacteria. The signaling cascade involved differs slightly between the A/E pathogens. In EPEC, Tir is phosphorylated on the tyrosine residue Y_{474}, with subsequent recruitment of the host adaptor protein Nck and activation of neural Wiskott-Aldrich syndrome protein (N-WASP) (Schüller *et al.*, 2007). This results in actin-polymerisation by the actin-related protein 2/3 (Arp2/3)-complex. EHEC lacks tyrosine residue Y_{474}, and therefore induces Nck-independent actin polymerization via a conserved asparagine-proline-tyrosine motif (NPY_{458}; NPY_{454} in EPEC). This recruits N-WASP and Arp2/3. However, NPY_{458}-mediated signaling only gives weak actin polymerization *in vitro*. Therefore, EHEC Tir cooperates with the effector TccP/$EspF_U$, to activate Arp2/3 and give robust actin polymerization (Garmendia *et al.*, 2004). *C. rodentium* Tir and EPEC Tir are functionally interchangeable, both requiring tyrosine phosphorylation for actin rearrangement (Deng *et al.*, 2003). However, *C. rodentium* Tir can support actin nucleation in EHEC, while EPEC Tir cannot. Interestingly, although Tir is essential for A/E lesion formation *in vivo*, the tyrosine phosphorylation itself is not essential for *C. rodentium* colonization and A/E lesion formation (Crepin *et al.*, 2015). Thus, the pathways discussed above are not indispensable for recruiting N-WASP or forming A/E lesions, meaning that another host factor must be

important for Tir-mediated actin polymerization and A/E lesion formation. It was found that Tir-mediated signaling regulates the lipid-content of the host outer membrane at the site of bacterial attachment, which might be of importance in this context (Sason *et al.*, 2009; Smith *et al.*, 2010).

Map, which possesses a mitochondrial targeting sequence at its N-terminus, was the first described member of the so called WxxxE family of effector proteins present in A/E pathogens as well as pathogens such as *Shigella* and *Salmonella* spp. (Alto *et al.*, 2006). WxxxE effector were first thought to directly mimic host small GTPases, but in fact function as Guanine-nucleotide exchange factors (GEFs) to promote the activity of host Rho GTPases (Huang *et al.*, 2009; Klink *et al.*, 2010). Map is a GEF for Cdc42, a protein that regulates host actin dynamics (Huang *et al.*, 2009). This results in the formation of transient filopodia around bacteria (Kenny *et al.*, 2002). The precise role of this during A/E infection is unclear. Map-induced filopodia was shown to depend on the interaction between Map and NHE3 regulatory factor 1 (NHERF1) at the membrane (an interaction that requires the C-terminal TRL motif) (Simpson *et al.*, 2006). In EPEC, Map activity and the cessation of filopodia formation is regulated by Tir (Kenny *et al.*, 2002; Berger *et al.*, 2009a). Tir, Map and EspF (discussed below) are implicated in the effacement of microvilli during A/E pathogenesis and the loss of function of the host sodium-glucose transporter SGLT-1 (Dean *et al.*, 2006), which might be important for the diarrhoeal symptoms of EPEC. The Map-NHERF1 interaction was also shown to contribute to the diarrhoeal phenotype (Simpson *et al.*, 2006). In line with his, Map has also been shown to modulate intestinal barrier function, again contributing to diarrhoea. The various functions associated with Map illustrate the multifunctionality of many LEE-encoded effectors, and its regulation by Tir highlights the importance of effector protein cooperation. Further, the various functions attributed to Map may also be reflective of the numerous roles of Cdc42 (affecting actin dynamics, cell cycle progression, apical-basolateral cell polarity and membrane trafficking) (Etienne-Manneville, 2004).

EspF is another LEE-encoded multifunctional effector (Holmes *et al.*, 2010). Its known roles include disruption of the mitochondrion, nucleolus modulation, effacement of microvilli, cell surface receptor disruption and disruption of tight junctions. Its N-terminus harbors both a mitochondrial targeting signal and a nucleolar targeting signal. EspF has multiple proline-rich repeats with Src homoly 3 (SH3) protein scaffolding domains (which mediate binding to many host proteins), an N-WASP binding domain and putative actin-binding motif (Holmes *et al.*, 2010). During infection, EspF is mainly targeted to mitochondria, and induces apoptosis (Nagai,

Abe & Sasakawa, 2005; Nougayrède & Donnenberg, 2004; Dean *et al.*, 2010). Later in the infection (when mitochondrial disruption has already occurred), EspF is targeted to the nucleolus where it redistributes nucleolin to the cytoplasm. This is thought to affect ribosome biogenesis (Dean *et al.*, 2010). Moreover, EspF subverts the Sodium-Hydrogen exchanger 3 (SLC9AE/NHE) (Hodges *et al.*, 2008) in the apical membrane and causes mislocalisation of the channels Aquaporin 2 and 3 (AQP2/3) away from the apical membrane in concert with the effector protein EspG (Guttman *et al.*, 2007). This effect is independent of its effect on mitochondria and apoptosis, and of its binding to sortin nexin-9 (SNX9) that has roles in endocytosis and F-actin organization (Hodges *et al.*, 2008; Guttman *et al.*, 2007). The molecular mechanisms by which EspF affects AQP2/3 is not known, but the roles of EspF in in membrane remodeling and vesicle trafficking could be of importance here (Alto *et al.*, 2007). Taken together with the role of EspF in modulating SGLT-1, themes for the roles of EspF includes the manipulation of ion transport and fluid homeostasis during A/E infection. In line with this, EspF has also been shown to cause disassembly of tight junctions TJs, affecting paracellular diffusion of water and small molecules (McNamara *et al.*, 2001).

EspG is one of the most conserved A/E effector proteins, homologous to VirA in *S. flexneri* (Elliott *et al.*, 2001) and is shown to have pleiotropic effects in infected cells. Initial reports on both VirA and EspG described a role in targeting host microtubules (Shaw *et al.*, 2005; Yoshida *et al.*, 2006; Hardwidge *et al.*, 2005). VirA with tubulin-specific cysteine protease-lie activity was shown to be involved in intra- and inter-cellular spread of *S. flexneri* (Yoshida *et al.*, 2006). Like EspF, EspG was shown to manipulate host TJs (Gujral *et al.*, 2015). Nevertheless, the role of EspG in manipulating microtubules during infection has been contested in the literature (Selyunin *et al.*, 2014; Germane *et al.*, 2008). Recently, EspG/VirA were described as Rab GTPase activating proteins (GAPS), with Rab1 proposed to be the primary target in the cell (Dong, Liu & Shao, 2010; Selyunin *et al.*, 2014) and EspG was also shown to bind ARF GTPases (Selyunin *et al.*, 2011). Interestingly, EspG was found to manipulate host protein trafficking through its interaction with the small GTPases ARF6 and Rab35. Binding ARF6 directly recruits EspG to endosomal structures in EHEC-infected cells, and EspG subsequently interacts with Rab35. This causes intracellular accumulation of the Transferrin Receptor, which recycles host proteins. In this way, EspG causes global manipulation of the host cell surface proteome, as the presence of a range of different host proteins at the cell surface – including AQP2/3 discussed above – is manipulated in an EspG-dependent manner (Furniss *et al.*, 2016).

EspZ is only found in A/E pathogens. It regulates the injection of T3SS effector proteins into the host. It is suggested that EspZ thus ensures that cells only become infected in one wave of bacteria, rather than being continuously infected in multiple waves (Berger *et al.*, 2012). In line with this, EspZ was shown to inhibit cytotoxicity due to excessive or disordered effector delivery. Ectopic delivery of EspZ resulted in protection of staurosporin-induced apoptosis of cultured eukaryotic cells, demonstrating a direct anti-apoptotic function for EspZ (Roxas *et al.*, 2012). EspZ is required for the establishment of A/E infection; as shown in an *in vivo* REPEC model, rabbits infected with *espZ* knockout REPEC did not become colonised and also had increased epithelial cell death one day post-infection (Wilbur *et al.*, 2015). Thus, in the absence of EspZ, infected epithelial cells die and come off, preventing normal colonisation dynamics in the gut.

Lastly, EspH is also involved in the formation of filopodia and pedestal dynamics. It is required for effective pedestal formation *in vitro* (Tu *et al.*, 2003; Wong *et al.*, 2012b). In the absence of EspH, extended filopodia are formed and there is a delay in pedestal formation. Over-expressing EspH on the other hand gives elongated pedestals directly beneath adherent bacteria (Tu *et al.*, 2003). EspH is also known to promote cell detachment via focal-adhesion disassembly (Wong *et al.*, 2012b). It functions by competitively inactivating host Rho GEFs, preventing Rho GTPase activation (Dong, Liu & Shao, 2010). Importantly, it does not target or affect bacterial T3SS-translocated Rho GEFs, and bacterial Rho GEFs may thus replace host Rho GEFs to control Rho GTPase activation (Wong *et al.*, 2012a, 2011).

1.5.8.7. Non-LEE-encoded T3SS effector proteins

A number of effector proteins are encoded outside of the LEE island, termed Non-LEE encoded (Nle)-effectors (Tobe *et al.*, 2006). They are found across the genome, but are mainly organized in 16 genetic islands (GIs, GI1 to GI16; **Figure 1.8**). Recently, all effectors – including 17 Nle effector proteins – were sequentially deleted from EPEC to generate a toolkit of a full effector-mutant and a number of intermediates (Cepeda-Molero *et al.*, 2017). These strains were used to study the intricate relationships between effector proteins and confirmed that Nle protein are essential to A/E lesion formation on mucosal surfaces. This section briefly describes a number of studied Nle effector proteins and their contribution to the infection.

The first prophage-encoded effector protein to be identified was cycle inhibiting factor (Cif; Marchès *et al.*, 2003; Charpentier & Oswald, 2004). While *cif* is absent in *C. rodentium* or key EHEC strains, it is found in a subset of EPEC strains. It acts as a cyclomodulin, which is a family of

bacterial toxins that modulate the cell cycle (Taieb *et al.*, 2006; Hsu *et al.*, 2008). It causes cell cycle arrest, giving rise to an accumulation of inactive phosphorylated Cdk1. Moreover, Cif also induces delayed apoptosis in intestinal epithelial cells (Samba-Louaka *et al.*, 2009).

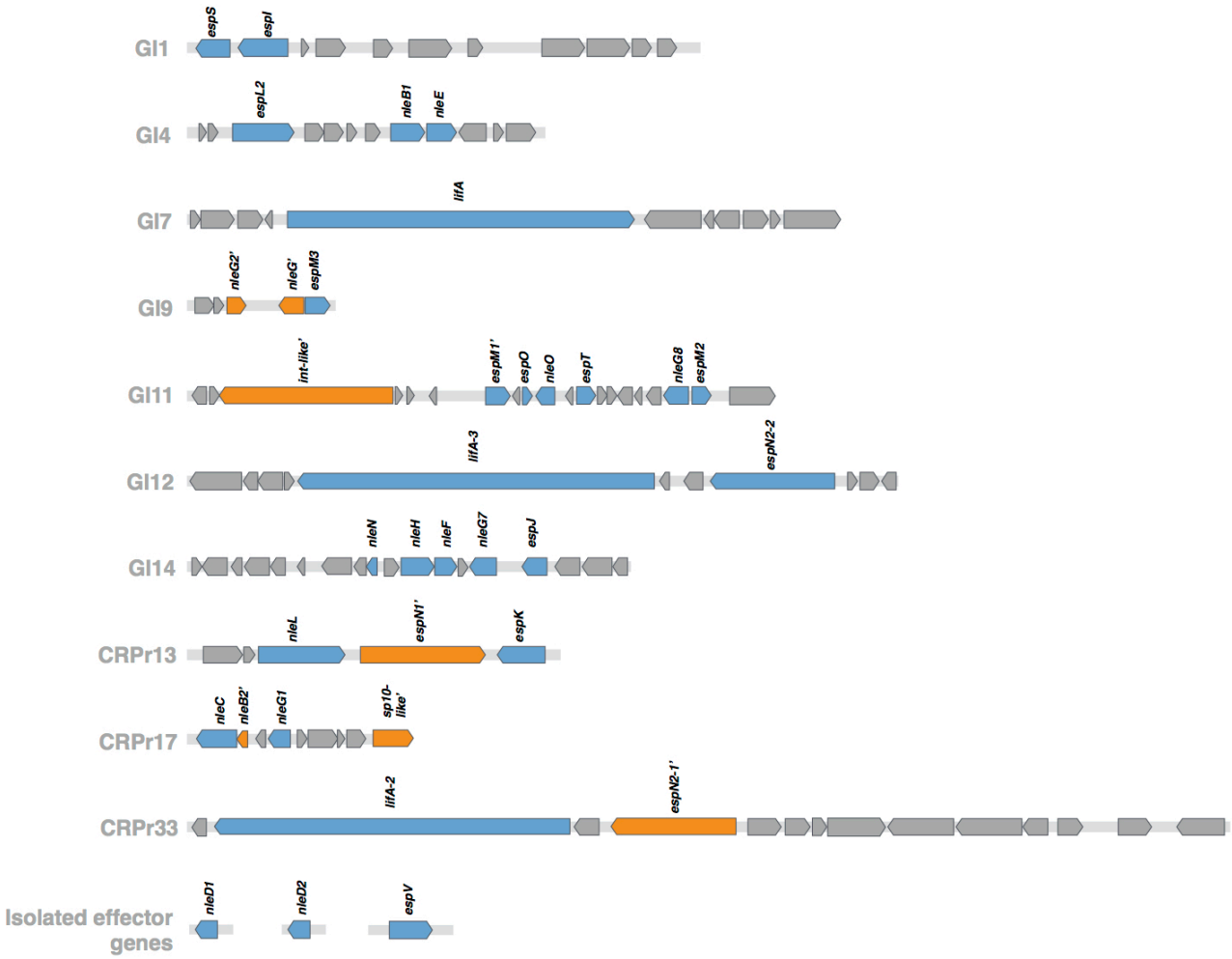


Figure 1.8 | The genetic organisation of the non-LEE encoded effector genes in *C. rodentium*. Non-LEE encoded effector genes are found within a number of genomic islands. Effector genes are shown in blue; pesudogenes are shown in orange.

NleA is encoded within a prophage-associated PAI and translocates to the Golgi apparatus during transfection of cultured cells. NleA is conserved in LEE-encoding pathogens and is absent in non-pathogenic strains of *E. coli*. Strikingly, NleA is essential for bacterial colonisation of the murine gut; *C. rodentium* lacking *nleA* is unable to exert its full virulence *in vivo*. Infection of *C. rodentium*-susceptible mice is lethal, while Δ*nleA C. rodentium*-infected susceptible mice do not succumb to the infection (Gruenheid *et al.*, 2004). NleA was shown to compromise the Sec23/24 complex, which is part of the COPII protein coat that shapes intracellular protein transport

vesicles (Kim *et al.*, 2007). Thus, NleA inhibits COPII-dependent protein export from the endoplasmic reticulum (ER). Moreover, NleA has also been shown to suppress NLRP3 inflammasome activity (Yen, Sugimoto & Tobe, 2015).

NleB is also essential for colonisation. It is an arginine lycosyltransferase protein that modifies conserved arginine residues in death domain-containing host proteins with *N*-acetylglucosamine (GlcNAc) (Scott *et al.*, 2017). Thus, it can block extrinsic apoptosis signaling (Gao *et al.*, 2013). Moreover, ectopic NleB expression mediates modification of Fas-associated via death domain (FADD), TNFRSF1A-associated via death domain (TRADD), and receptor-interacting serine/threonine protein kinase 1 (RIPK1) (Pearson *et al.*, 2013). Recently, an affinity proteomics approach enriching for arginine-GlcNAcylated glycopeptides looked at the global dynamics of arginine GlcNAcylation during NleB-overexpression and during *C. rodentium* infection *in vivo*. This verified that FADD is rapidly and stably modified during infection, suggesting that NleB antagonizes death receptor-induced apoptosis via irreversible FADD modificiation (Scott *et al.*, 2017).

EspJ inhibits autophosphorylation of tyrosine Src family kinases and phosphorylation of its substrates Tir and FcγRIIa, thereby inhibiting opsono-phagocytosis via FcγRIIa. EspJ is a unique adenosine diphosphate (ADP) ribosyltransferase, able to directly inhibit Src kinase by amidation and ADP ribolsylation of a conserved reside (E310) required for its kinase activity (Young *et al.*, 2014). Downstream effects of this involve the inhibition of actin polymerization in response to EPEC infection. This was shown to bear importance during infection *in vivo*, combining mass spectrometry based methodologies with *C. rodentium* infection to demonstrate that EspJ-induced ADP ribosylation controls immune responses by targeting multiple non-receptor tyrosine kinases (Pollard *et al.*, 2018).

Certain Nle effectors, just like some LEE-encoded effectors, target small GTPases. Two of these, EspT and EspM, also belong to the WxxxE family of effector proteins (Bulgin *et al.*, 2010). EspT modulates the host cell cytoskeleton by activating Rac1 and Cdc42 (Bulgin *et al.*, 2009a, 2009b). This allows the subset of aEPEC strains that carry EspT (such as aEPEC E110019) to display an intracellular phenotype not normally associated with EPEC. EspM is found in EPEC, EHEC and *C. rodentium* (duplicated in the latter two) and activates RhoA, giving it roles in stress fibre formation, repression of actin pedestal formation and TJ manipulation (Arbeloa *et al.*, 2009, 2008; Simovitch *et al.*, 2010). The presence of EspM proteins is associated with severe infections (Arbeloa *et al.*, 2009).

Interestingly, it has been shown that the host can sense and responds to the puncturing by a T3SS, as contact with an EPEC T3SS activates NF-κB signalling (Litvak *et al.*, 2017). Unsurprisingly, A/E pathogens dedicate a proportion of their effector protein repertoire to down-regulating or inhibiting the host immune responses and to limit excessive host cell death. Indeed, NleE, NleC, NleD and NleH1/2 all inhibit TNF- and IL-1β-induced NF-κB responses during *C. rodentium* infection. During infection of cultured cells, an EPEC mutant deleted of *nleB*, *nleE*, *nleC* and *nleD* exhibits strong NF-κB signaling (Litvak *et al.*, 2017). NleF on the other hand is pro-inflammatory and activates NF-κB early during infection of cultured cells epithelial cells (Echtenkamp *et al.*, 2008; Pallett *et al.*, 2014). NleF also has another function directly inhibiting caspase proteins to inhibit cell death pathways (Blasche *et al.*, 2013; Song *et al.*, 2017). NleH1/2 are bacterial kinases that have also been shown to target cell death (García-Angulo *et al.*, 2008; Grishin *et al.*, 2014). They target a cofactor of NF-κB (ribosomal protein S3, RPS3) to repress NF-κB translocation to the nucleus. NleE exerts its anti-inflammatory functions by methylating Tab2/3, while NleD and NleC are metalloproteases and the latter was shown to directly cleave degrade the p65 subunit of NF-κB (Nadler *et al.*, 2010; Yao *et al.*, 2014; Yen *et al.*, 2010; Giogha *et al.*, 2015).

The Nle effector protein LifA has also been studied (Cassady-Cain *et al.*, 2016). Also called lymphostatin, LifA has putative glycosyltransferase activity and is important for EHEC colonisation of cattle (by serogroups O5, O111 and O26), as an *efa1* deletion mutant was attenuated in its ability to associate to the mucosa and colonise the intestine (Klapproth *et al.*, 2000, 2005; Abu-Median *et al.*, 2006). However, the mechanism behind this remains to be determined.

EspO binds integrin-linked kinase (ILK), a protein involved in pro-inflammatory signaling following LPS stimulation. The role of EspO during *C. rodentium* murine infection was recently investigated (Ale *et al.*, 2017), showing that an *espO* deletion mutant triggered less recruitment of $CD4^+$ T cells producing IFNγ, IL-17A or IL-22. While bacterial colonisation was increased for the mutant compared to WT *C. rodentium*, the lower recruitment of $CD4^+$ T cells coincided with reduced levels of hyperplasia.

A number of Nle effectors, such as NleG and NleL, exploit host ubiquitin pathways. As the focus of this project is the NleG family of effectors, the ubiquitin system will be described in detail before a review of our knowledge regarding bacterial effector proteins that modulate host ubiquitination (**Secion 1.6.4**) and of the NleG proteins (**Section 1.7**).

1.6. Ubiquitination in health and disease

Post-translational modification is a fundamental regulatory tool in the eukaryotic cell, where most proteins encounter at least one form of post-translational modification in their lifespan (Komander, 2009). One key post-translational modification is ubiquitination, which regulates virtually every process in the eukaryotic cell, including cell division, differentiation, signal transduction, protein trafficking, DNA repair and protein regulation.

1.6.1. Ubiquitination is a three-step cascade

Ubiquitination is the reversible covalent attachment of the highly conserved 76-amino acid (8.5 kDa) ubiquitin to the ε-amino group of lysine residues of diverse intracellular target proteins. The process is a multi-enzyme cascade (outlined in **Figure 1.9**) that requires the concerted action of 3 classes of enzymes: E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugation enzymes and E3 ubiquitin-ligase enzymes (Komander & Rape, 2012).

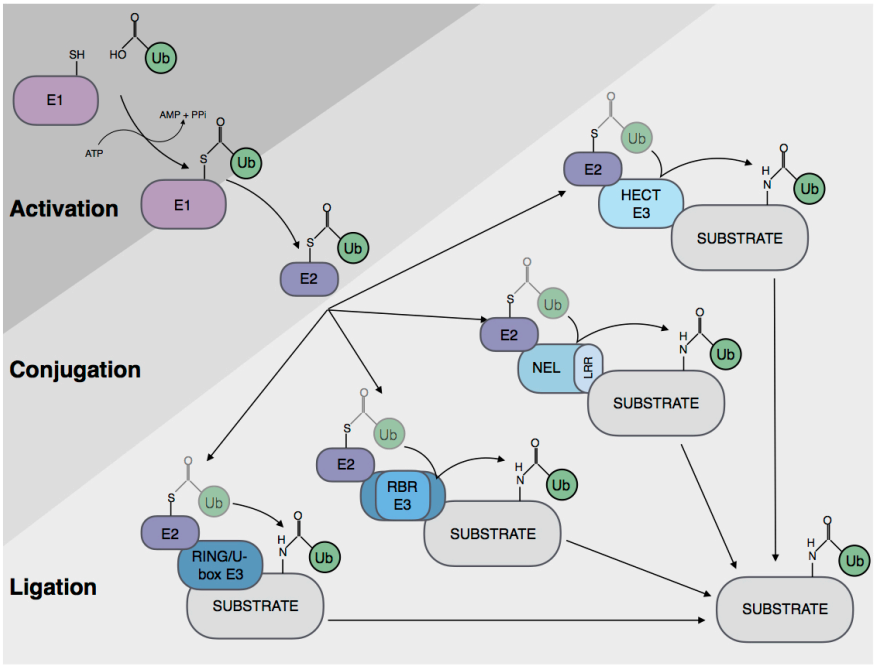


Figure 1.9 | Ubiquitin transfer is a three-step process. A ubiquitin activating (E1) enzyme mediates ATP-dependent adenylation of the C-terminus of ubiquitin and its subsequent attachment to the E1 active cysteine by thioesterification. Ubiquitin is then transferred to the catalytic cysteine of a ubiquitin-conjugating (E2) enzyme, before being tranfsferred to its substrate by a ubiquitin ligase (E3) enzyme. RING/U-box E3 ligases mediate direct transfer of ubiquitin from the E2 to a substrate, while the other classes (HECT, NEL and RBR-type E3 ligases) mediate indirect transfer where ubiquitin is transferred first to the E3 itself and then onto its substrate.

In humans, one of two E1 enzymes activates ubiquitin by attaching ubiquitin to itself via a thioester bond between the C-terminal glycine of ubiquitin and an E1 cysteine residue in an ATP-dependent manner. Ubiquitin is then transferred from the E1 to a cysteine residue within an E2

ubiquitin-conjugating enzyme (37 encoded in the human genome), forming yet another thioester bond. The resulting E2-ubiquitin complex is brought into contact with a substrate protein by an E3 ubiquitin ligase (over 600 putative E3s encoded in the human genome; Hershko & Ciechanover, 1998), that acts to directly or indirectly transfer ubiquitin from the E2 to the substrate. As a result, an isopeptide bond is formed between the C-terminal glycine of ubiquitin and the ε-amino group of a lysine within the substrate. Thus, the E3 ligase determines substrate specificity as well as catalysing the final ubiquitin transfer step of the cascade.

1.6.2. The complexity of the ubiquitin code

One round of ubiquitination gives a mono-ubiquitinated substrate and is classically associated with cell surface receptor endocytosis and lysosomal degradation. Although the attachment of ubiquitin normally occurs on lysine residues, conjugation to other substrate side chains such as threonine, serine, tyrosine and cysteine has also been demonstrated in certain cases (McDowell & Philpott, 2013; Komander, 2009; Kulathu & Komander, 2012). Typically, successive rounds of ubiquitination generate poly-ubiquitin chains where one of the seven lysine resides within ubiquitin itself (positions 6, 11, 27, 29, 33, 48 and 63) serve as a recipient for additional attachment of ubiquitin (**Figure 1.10**). Ubiquitin can also form linear chains, where the C-terminal glycine of ubiquitin forms an isopeptide bond with the N-terminal methionine of the next ubiquitin. Polyubiquitin chains can be homotypic (same linkage type throughout chains) or heterotypic (multiple linkage types per chain). The latter may also be branched or unbranched, adding another level of complexity to this system. The structural range of resulting chains means that although they share the ubiquitin moiety as their basic building block, the different chains could perhaps be regarded as distinct post-translational modifications.

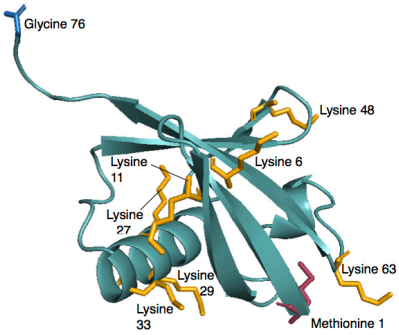


Figure 1.10 | Ubiquitin with its eight potential attachment sites for polyubiquitin chain formation. The solved structure of ubiquitin (shown in teal; PDB code 1UQB; Vijay-kumar, Bugg & Cook, 1987) shows that the seven lysine residues (shown in yellow) are exposed on different surfaces of the ubiquitin molecule. The N-terminal methionine (shown in pink) is used as a ubiquitin-ubiquitin linkage point for linear chain formation. The C-terminal glycine (shown in blue) required for the isopeptide bond-formation is highlighted.

The recognition of ubiquitin modification by proteins containing specialised ubiquitin-binding domains (UBD) couples the modification to downstream events. 20 UBDs have been described and many are specific to certain ubiquitin linkage types. Thus, the type of chain linkage that occurs determines the fate of a substrate *in vivo* (**Figure 1.11**) and this is commonly referred to as the ubiquitin code (Komander & Rape, 2012). K48-linked chains, for instance, label proteins for proteasomal degradation and represent the most common and best-characterised outcome of ubiquitination. K11-linked chains also serve as a signal for proteasomal degradation and as a whole ubiquitin-dependent proteosomal degradation is a fundamental contributor to post-translational re-organisation of the proteome. K63-linked ubiquitin chains have non-degradative scaffolding roles during the assembly of signalling complexes and regulate for instance NFκB-dependent pathways.

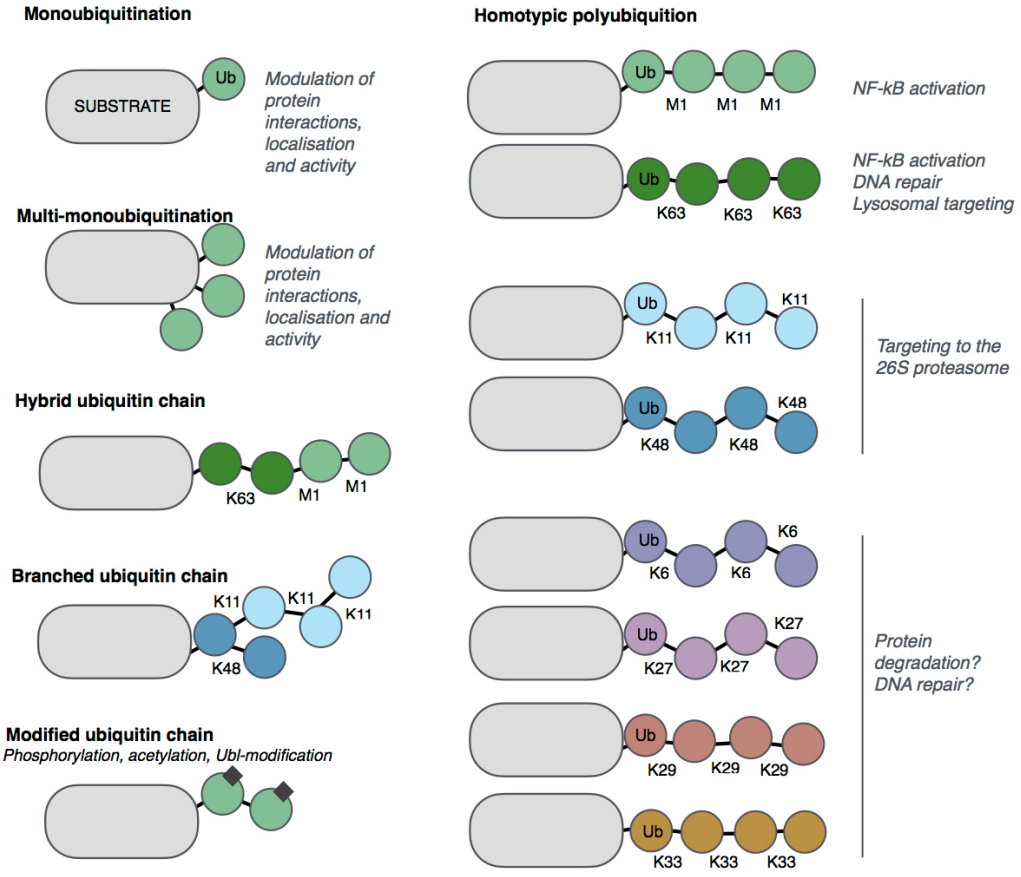


Figure 1.11 | The ubiquitin code. Ubiquitin can be conjugated to a substrate (in grey) in different ways, determining the fate of the substrate. Monoubiquitination and multi-monoubiquitination are known to modulate protein interactions, protein localisation and protein activity. A number of structurally distinct homotypic polyubiquitin chains can be formed depending on the specific residues involved in the ubiquitin-ubiquitin linkage (shown on the right), with different downstream consequences of cellular processes. Hybrid (or heterotypic) ubiquitin chains containing more than one type of linkage may also be formed, as well as branched or forked ubiquitin chains. Chains can also be further modified by post-translational modification, such as phosphorylation, acetylation or modification by small ubiquitin-like (Ubl) modifiers. Adapted from Komander, 2009.

1.6.3. E3 ligase proteins

Four classes of E3 ligases have been described to date, the two main classes being the Really Interesting New Gene (RING)/U-box-type E3 ligase and the Homologous to the E6-AP Carboxyl-terminus (HECT)-type (Zheng & Shabek, 2017). RING/U-box-type E3 ligases bridge ubiquitin-bound E2s and target proteins to mediate direct transfer of ubiquitin from the E2 to the target. The difference between RING domains and the U-box is that the RING domain requires coordination of two zinc ions and eight conserved cysteine/histidine residues to adopt the required structure, while the U-box does this without zinc ions (O'Connor & Huibregtse, 2017). HECT-type E3 ligases mediate indirect transfer of ubiquitin, where ubiquitin is first transferred to a cysteine residue within the E3 ligase before being transferred onto the target protein (Bernassola *et al.*, 2008; Rotin & Kumar, 2009). More recently, two additional classes mediating indirect transfer of ubiquitin have been described. New E3-like (NEL)-type E3s were identified from T3SS effectors of pathogenic bacteria and harbour a leucine-rich repeat (LRR) domain in their N-termini for substrate recognition (Quezada *et al.*, 2009). Their C-termini have a conserved ligase domain, with a single conserved cysteine residue required to form the E3-ubiquitin active intermediate before transfer of ubiquitin to the substrate. In the absence of a substrate, the LRR domain binds to the C-terminus, which masks and sequesters the catalytic cysteine residue, giving autoinhibition that is relieved upon substrate binding and subsequent rearrangement of the LRR domain (Keszei & Sicheri, 2017). The last described class of E3 enzymes is the RING-Between-RING (RBR)-type, which consists of two RING domains: one to recognize and bind substrates and one to catalyse the transfer of ubiquitin to an active cysteine of the E3 ligase prior to transferring the ubiquitin onto a substrate (Spratt, Walden & Shaw, 2014; Smit & Sixma, 2014).

Moreover, E3 ligases may be modular, single protein or multi-protein complexes, such as the SCF or APC complexes (Ardley & Robinson, 2005). The majority of human E3 ligases are RING-type E3 ligases (over 600 annotated in human genome), while there are 26 human HECTs and 14 RBRs (O'Connor & Huibregtse, 2017). The biggest family of RING E3 ligases is the Cullin-RING Ligase (CRL) family that exist as large multi-subunit complexes (Duda *et al.*, 2011).

Ubiquitination is reversible by specialised proteases termed de-ubiquitinases (DUBs, approximately 85 encoded in the human genome) that release ubiquitin moieties from their chains or substrates for recycling in the cell. DUBs are classified into 5 classes based on their

catalytic mode of action and may be chain linkage-specific or promiscuous (Herhaus & Dikic, 2017). Alongside the ubiquitination/deubiquitination system, a number of small ubiquitin-like modifiers, including SUMO-1, SUMO-2, SUMO-3, NEDD8, ISG15 and LC3, modify substrates in a similar manner to ubiquitin (Lu & Yeh, 2013). These modifiers utilize components of the ubiquitination cascade (E1, E2 and E3 enzymes) and hold important roles in cell signaling in the eukaryotic cell.

Ubiquitination has a profound effect on the cell proteome, inducing changes in protein stability, function, localization and interaction with other proteins. As such, perturbations in this system are associated with a wide range of cancers and diseases such as metabolic disorders. Further, the ubiquitination system is an important component in the response to bacterial pathogens, including triggering innate responses that recruit inflammatory cells to the site of infection (Ashida & Sasakawa, 2017). Ubiquitination can mark bacteria for killing via proteasome-, phagosome- or autophagosome pathways (Herhaus & Đikić, 2017). For instance, invading cytosolic *Salmonella* become poly-ubiquitinated by M1- and K-63-linked chains that are recognized by ubiquitin-binding autophagy receptors mediating autophagosomal degradation of bacteria (Herhaus & Dikic, 2017; Fiskin *et al.*, 2016; Thurston *et al.*, 2012).

1.6.4. Bacterial manipulation of the ubiquitination cycle

Although some bacteria have systems that functionally resemble the degradative arm of the ubiquitination system – such as the pupylation system in *Mycobacterium tuberculosis* (Pearce *et al.*, 2008) – ubiquitin itself remains exclusive to eukaryotic cells. Given the contribution of ubiquitination to innate immune signaling, protein homeostasis and cell development, it is unsurprising that microbial pathogens have evolved strategies that target or exploit the host ubiquitination system. Global proteomics analyses reveal dynamic changes in the pool of ubiquitinated proteins of *S.* Typhimurium-infected cultured cells, including changes in the actin cytoskeleton, NF-κB pathway and autophagy pathways (Fiskin *et al.*, 2016). A growing pool of bacterial effector proteins are known to interfere with or exploit host ubiquitination (summarised in **Table 1.1**). A wide range of mechanisms is used by these effectors to accomplish their goal, including mimicking host factors (E3 ligases and DUBs), catalysing E1/E2-independent ubiquitination and modifying components of the ubiquitination machinery. For instance, Cif in EPEC (homologous to CHBP in *Burkholderia pseudomallei*), is a deamidase that directly targets the eukaryotic SCF complex to interfere with regulation of the cell cycle (Morikawa *et al.*, 2010). The deamidase OspI in *S. flexneri*, on the other hand, directly modifies the E2 protein Ubc13 and

prevent it from functioning as the E2 enzyme for TRAF6, which is important for activation of immune responses (Sanada *et al.*, 2012). NleE from A/E pathogens methylates Tak1-binding protein 2 (TAB2) and TAB3 during infection to prevent their binding to K63-linked ubiquitin chains, which inhibits downstream NF-κB-activation (Zhang *et al.*, 2012). Δ*nleE C. rodentium* causes a lower bacterial burden in the mouse gut and reduced colonic crypt hyperplasia (Wickham *et al.*, 2007). Further, NleE methylates ZRANB3 (has a Npl4-like Zinc Finger domain akin TAB2/3), which is involved in DNA repair pathways, and NleE may therefore perhaps provide a link between A/E pathogen infection, virulence proteins and genomic integrity (Yao *et al.*, 2014).

Many pathogens are now appreciated to synthesise E3 ubiquitin ligase mimics to exert virulence strategies (reviewed in Maculins *et al.*, 2016), and much research has been done to identify their host substrates to elucidate which host signalling cascades the pathogen modulates in this manner for its own survival or dissemination. The following pages describe bacterial E3 ligase proteins.

1.6.4.1. E3 ubiquitin ligase mimicry in the host-pathogen arms race

The secreted effector protein AvrPtoB of *Pseudomonas syringae* (a pathogen of tomato and *Arabidopsis thaliana*) structurally and functionally mimics a RING-type E3 ligase via its C-terminus (Abramovitch *et al.*, 2003; Rosebrock *et al.*, 2007). It is a striking example of a bacterium hijacking host ubiquitin-dependent proteasomal degradation, while also telling the story of an evolutionary arms race between the pathogen and its host. The C-terminus of AvrPtoB is essential for avoidance of plant immunity pathway recognition upon infection, as its loss results in localised immunity-associated programmed cell death in the host (a means by which plants attempt to inhibit infection spread) (De Torres *et al.*, 2006). The N-terminus of AvrPtoB, however, is recognised and phosphorylated by the host Fen kinase, which activates a process known as effector-triggered immunity (resulting in immunity-associated programmed cell death). *P. syringae* strikes back by ubiquitinating the same Fen kinase, leading to its proteasomal degradation and avoidance of the host immunity response. To counteract this, yet another host kinase (Pto kinase) phosphorylates a regulatory threonine-residue within AvrPtoB, reducing its ability to ubiquitinate Fen. As such, the outcome of the infection depends on a competition between the phosphorylation and ubiquitination activities of the host kinase and the E3 ligase effector protein of the pathogen.

1.6.4.2. Novel ubiquitin ligases in mammalian pathogens

A recently described family of effector proteins constitute the novel class of ubiquitin ligases (NELs) that is found in human pathogens of significant clinical relevance, including *Shigella* and *Salmonella* spp. They include at least 15 bacterial enzymes, spanning across 6 pathogenic genera. They are structurally distinct from eukaryotic E3 ligases but functionally mimic eukaryotic HECT-type E3 ligases, employing an indirect ubiquitin transfer process to modify its substrates. NELs contain three important features: (1) E2-interacting surface residues involved in hijacking host E2s charged with ubiquitin, (2) a catalytic cysteine reside required for ubiquitin transfer (despite the lack of a characteristic HECT domain bi-lobed structure) and (3) LRRs of variable length for substrate recognition.

The IpaH T3SS effector proteins in *S. flexneri* were the first NELs described (reviewed in Ashida & Sasakawa, 2016). *S. flexneri* harbours three pINV-encoded IpaH effectors and seven gDNA-encoded IpaH effectors. In an *in vivo* model of infection, mice infected with *Shigella* expressing a catalytically inactive form of IpaH9.8 suffered a lower bacterial burden and exhibited a smaller inflammatory response, illustrating the importance of E3 ligase mimicry to *Shigella* virulence (Ashida *et al.*, 2010). IpaH9.8 was shown to inhibit ubiquitin-dependent NF-κB-signalling by targeting the NF-κB essential modulator (NEMO) for ubiquitin-dependent proteasomal degradation (Ashida *et al.*, 2010). Moreover, IpaH9.8 also ubiquitinates the splicing factor U2AF35, preventing it from taking part in an mRNA splicing reaction and thereby down-regulating U2AF35-dependent pro-inflammatory gene expression (Okuda *et al.*, 2005; Seyedarabi *et al.*, 2010; Seyedarabi, 2010). IpaH7.8, on the other hand, contributes to the infectious process by facilitating bacterial escape from endocytic vacuoles in mouse and human macrophages (Fernandez-Prada *et al.*, 2000). During mouse infection, *Shigella* mutants either lacking IpaH7.8 or expressing ligase-dead IpaH7.8 were attenuated in their ability to multiply within the host. IpaH7.8 was also shown to activate inflammasome pathways by ubiquitinating glomulin/flagellar-associated protein 68 (GLMN), a negative regulator of NLRP3/NLRC4 inflammasome activation, leading to its proteasomal degradation and enhanced macrophage cell death (Suzuki *et al.*, 2014).

IpaH0722 is one of many bacterial effector proteins that target NF-kB activation, which is pivotal in inflammatory and immune responses (see **Table 1.1**). IpaH0722 ubiquitinates and ultimately degrades TFN-associated regulatory factor 2 (TRAF2), resulting in inhibition of protein kinase C-mediated NF-κB activation (Ashida, Nakano & Sasakawa, 2013). IpaH4.5 targets TNFa-mediated

NF-κB activation *in vitro* from yet another angle, by direct ubiquitination of the p65 subunit of NF-κB (Zheng *et al.*, 2016; Wang *et al.*, 2013). Unsurprisingly, the mammalian E3 ligase linear ubiquitin chain assembly complex (LUBAC; Tokunaga & Iwai, 2012) that generates linear (Met-1 linked) ubiquitin chains in response to invading pathogens to trigger NF-κB signalling is also targeted. IpaH1.4 and IpaH2.5 ubiquitinate HOIP, a subunit of LUBAC, targeting it for degradation to down-regulate NF-κB responses (De Jong *et al.*, 2016).

In *Salmonella* spp., three known NELs - SspH1, SspH2 and Slrp – all contribute to *Salmonella* virulence in animal models of infection and all require a catalytic cysteine residue for their activity (Rohde *et al.*, 2007; Quezada *et al.*, 2009). SspH1 localises to the nucleus and inhibits NF-κB-dependent gene expression. Moreover, SspH1 was shown to interact with serine/threonine protein kinase N1 (PKN1). SspH1-mediated PKN1 ubiquitination and subsequent degradation inhibits the androgen receptor response (Keszei *et al.*, 2014). A mass spectrometry-based approach identified the NOD receptor co-chaperone Sgt1 as a partner for SspH2, and it was shown that SspH2 enhances NOD-1 mediated responses in cells. Slrp is a homologue of YopM in *Yersinia pestis* and is required for *S.* Typhimurium colonisation of Peyer's patches during mouse infection (Tsolis *et al.*, 1999; Wei *et al.*, 2016). Slrp interacts with the reduction/oxidation regulatory protein thioredoxin-1 (implicated in mediating oxidative stress-induced signalling processes and cell death; Bernal-Bayard & Ramos-Morales, 2009) and the chaperone ERdj3 (Bernal-Bayard, Cardenal-Muñoz & Ramos-Morales, 2010). Importantly, Slrp ubiquitinates thioredoxin-1 *in vitro* and decreases its enzymatic activity (Zouhir *et al.*, 2014), making Slpr an example of an NEL that catalyses the formation of non-degradative ubiquitin chains. The SlrP-ERdj3 interaction changes the localisation of SlrP to the ER and decreases the ability of ERdj3 to bind unfolded proteins, but independently on Slrp E3 ligase activity (Bernal-Bayard, Cardenal-Muñoz & Ramos-Morales, 2010). Recently, a proteomics-based approach revealed that upon stable expression in cultured epithelial cells, Slrp decreases protein levels of CKB and S1004A (elevated in many pathological conditions, Tarabykina *et al.*, 2007), which have both been implicated in clearance of bacteria (Cordero-Alba *et al.*, 2016). Whether this is due to Slrp-mediated ubiquitination and proteasomal degradation is yet to be confirmed.

1.6.4.3. Bacterial HECT-type E3 ligases

SopA from *Salmonella* spp. was the first T3SS effector shown to structurally and functionally mimic HECT-type E3 ligases (Zhang *et al.*, 2006). Its homologue NleL in A/E pathogens (based primarily on C-terminal sequence homology) catalyses the formation of unanchored

polyubiquitin chains via K6- and K48 linkage *in vitro* (Lin *et al.*, 2011). Its activity depends on a conserved cysteine residue. Although precise NleL-mediated mechanisms during infection remain undefined, its ligase activity was shown to down-regulate the rate of pedestal formation during *in vitro* EHEC-infection (Piscatelli *et al.*, 2011). *C. rodentium* NleL is required for infection of enterocytes in an *in vivo* mouse model (Piscatelli *et al.*, 2011), while SopA is required for polymorphonuclear neutrophil (PMN) transepithelial migration (a hallmark of *S*. Typhimurium pathogenesis) and fluid secretion in the intestinal lumen of ligated ileal calf loops (Wood *et al.*, 2000).

Recently, SopA was found to ubiquitinate (by K11- and K48-linkage) the eukaryotic RING-type E3 ligases TRIM56 and TRIM65. TRIM56 and TRIM65 play roles in initiating innate immune responses by interacting with host RNA sensors RIG-1 and MDA-5, respectively, and are important for type I interferon signalling (Fiskin *et al.*, 2017). One study proposed that SopA-mediated ubiquitination of TRIM proteins was non-degradative, but enhanced their catalytic activity towards RIG-I and MDA5, resulting in increased transcription of IFN-B mRNA (Kamanova *et al.*, 2016). However, Fiskin and colleagues structurally analysed TRIM56 and SopA binding, and found that the binding occluded the E2-binding site of TRIM56. This means that bound TRIM56 cannot ubiquitinate targets or SopA (Fiskin *et al.*, 2017). Thus, SopA controls both the cellular concentration and the activity of its target to impair host defences against an invading pathogen.

1.6.4.4. Legionella effector proteins hijack host ubiquitination events

The Dot/Icm T4SS of *L. pneumophila* translocates a number of effector proteins that manipulate the host ubiquitin system. For instance, AnkB is anchored to the cytosolic face of *Legionella*-containing vacuoles (LCVs) where it mediates K48-linked ubiquitination of proteins via the SCF complex. This causes protein degradation *en masse*, which is thought to satisfy the high demand of amino acids required by *L. pneumophila* during its replication within the host cell (Lomma *et al.*, 2010). The contribution of AnkB to the infection strategy is highlighted by the attenuated virulence of AnkB knockout strains in a mouse model of infection (Lomma *et al.*, 2010). *L. pneumophila* also secretes the U-box-type E3 ligase effector protein LubX. LubX harbours two distinct U-box domains, where one is one is crucial for E3 activity and the second is required for substrate binding (Kubori, Hyakutake & Nagai, 2008; Quaile *et al.*, 2015). It ubiquitinates the host kinase Cdh1, and, at later stages of infection, LubX acts as a "meta-effector", regulating the function of other effectors. For instance, LubX targets the effector SidH for proteasomal degradation. Its importance during virulence of *Legionella* was established in a *Drosophila*

melanogaster infection model (Nagai & Kubori, 2013; Kubori *et al.*, 2010) and sheds light on how bacteria have evolved intricate ways of hijacking the host ubiquitin system to optimise timing for terminating the function of other effectors.

1.6.4.5. Non-canonical ubiquitination by bacterial pathogens

L. pneumophila SidE T4SS effector family members (SidE, SdeA, SdeB, SdeC) are capable of non-canonical ubiquitination in the host. This type of ubiquitination is independent of E1 and E2 enzymes and of ATP hydrolysis (Qiu *et al.*, 2016; Hsu *et al.*, 2014; Qiu & Luo, 2017). A full deletion SidE family knockout mutant displayed attenuated virulence against *Dictyostelium discoideium* (Qiu *et al.*, 2016). This phenotype was fully complemented by the expression of mono-ADP-ribosyltransferase (mART)-containing SdeA, as the family are characterised by their mART motifs. They ubiquitinate ER-associated Rab GTPases, and the full SidE family knockout mutant was unable to recruit ER markers to the LCVs–a hallmark of *L. pneumophila* infection – in the *D. discodeium* model. Once again, this could be complemented by mART-containing SdeA.

Instead of using E1 and E2 enzymes, SidE family members use nicotinamide dinucleotide (NAD) as a cofactor and add ADP-ribose to ubiquitin (arginine residue 42) to activate ubiquitin. This results in an ADP-ribosylated ubiquitin intermediate. The phosphodiester bond within this intermediate is cleaved by the nucleotidase-phosphohydrolase (NP) domain of SidE, giving phospho-ribosylated ubiquitin that is transferred onto a Rab GTPase substrate in a reaction that releases AMP (Bhogaraju *et al.*, 2016). This "all-in-one" ubiquitination is an efficient way of hijacking host ubiquitin pathways and the discovery of this system might give clues to the increasing complexity of the ubiquitin system in eukaryotes and in host-pathogen interactions.

The *L. pneumophila* E3 ligase RavN was recently described (Lin *et al.*, 2018). It has a U-box E3 domain in its N-terminus followed by a domain of 3 alpha-helices and its activity depended on the interaction with a subset of E2 enzymes. They solved the crystal structure of RavN and found that the U-box was only remotely similar to previously described U-box domains. Using an *in silico* fold homology search, they also identified another four bacterial ligases in *L. pneumophila* that were not known to possess E3 ligase activity. Three of them were capable of catalysing the formation of K48-linked chains *in vitro*. The fourth effector mediated K-11 and K-33 ubiquitin linkages, similar to another effector SidC, which possesses E3 ligase activity via an N-terminal catalytic triad (Cys-His-Asp) with no tertiary similarity to NELS, HECT-type- or RING-type E3

ligases (Hsu *et al.*, 2014). This effector was renamed SdcB, a new paralog of SidC, as it also contained the SidC catalytic triad.

Taken together, these discoveries indicate that a significant proportion of the *L. pneumophila* effector battery is dedicated to manipulating the host ubiquitin system, and more mechanisms by which pathogens do this are rapidly being uncovered.

1.6.5. Ubiquitin-rich aggregation formation during bacterial infection

Cytoplasmic ubiquitinated protein aggregates are known to form in response to certain pathogens, as first implied by their formation in LPS-challenged dendritic cells (Lelouard *et al.*, 2002, 2004). The aggregates are believed to contain defective ribosomal products (DRiPS), that are normally proteasomally degraded (Kaniuk *et al.*, 2011). While this appears to be a cellular response to infection, its role is not fully understood.

One theory is that DRiPS are storage compartments for pathogen-derived MHC Class I antigens (Yewdell & Nicchitta, 2006), which could explain why Listeriolysin O, the major antigen required for clearing *Listeria monocytogenes,* co-localises with DRiPS (Viala *et al.*, 2008). Intriguingly, *L. pneumophila* was recently shown to disrupt such ubiquitin aggregation formation using its T4SS (Ivanov & Roy, 2009). In contrast, they are present in the cytosol of *C. rodentium*-infected colonic cells in mice at day six post-infection (starting to resolve at day eight and being undetectable by day 12) (Kaniuk *et al.*, 2011), which could indicate that *C. rodentium* does not target this cellular response. Further, ubiquitin-rich aggregates were not observed in MyD88$^{-/-}$ *C. rodentium*-infected mice and the same observation was true in a *Helicobacter pylori* infection model in the stomach. For *H. pylori*, it was also shown that ubiquitin aggregate formation *in vitro* required Nox1 subunits. As both MyD88-dependent TLR stimulation and oxidative stress are known to increase translational infidelity and as LPS-challenge increases translational rates in dendritic cells, it might be that the sheer stress of an enteric infection is causing the protein aggregation within the DRiPs. Whether these aggregates serve a purpose (antigen storage depots) or are a consequence of other cellular immune responses (translational burst or ROS production) is currently not known.

Table 1.1 | Manipulation of the host ubiquitin system by bacterial pathogens.

RING/U-box E3 ligase	Species	Target	Function	Reference
NleG	EHEC, EPEC, *C. rodentium*	Unknown	Unkown	Wu *et al.*, 2010
StoD	*Salmonella* spp.	Unknown	Unknown	Wu *et al.*, 2010
AvrPtoB	*Pseudomonas syringae*	Fen kinase	Inhibit immunity-associated programmed cell death	Rosebrock *et al.*, 2007
LubX	*Legionella pneumophila*	CLK1, SidH.	Promotion of growth in macrophages and regulating the function of other effectors	Kubori, Hyakutake & Nagai, 2008; Nagai & Kubori, 2013; Lin *et al.*, 2015; Quaile *et al.*, 2015
GobX	*L. pneumophila*	Unknown	Unknown	Lin *et al.*, 2015
RavN	*L. pneumophila*	Unknown	Unknown	Lin *et al.*, 2018
MavM	*L. pneumophila*	Unknown	Unknown	Lin *et al.*, 2018
Lpg2370	*L. pneumophila*	Unknown	Unknown	Lin *et al.*, 2018
HECT E3 ligase	**Species**	**Target**	**Function**	**Reference**
SopA	*Salmonella* spp.	TRIM56, TRIM65	Regulation of host inflammation	Diao *et al.*, 2008; Lin, Diao & Chen, 2012; Kamanova *et al.*, 2016; Fiskin *et al.*, 2017; Lin *et al.*, 2011
NleL	EHEC, EPEC, *C. rodentium*	Unknown	Pedestal formation	Lin *et al.*, 2011; Hospenthal, Freund & Komander, 2013; Piscatelli *et al.*, 2011; Sheng *et al.*, 2017
MavJ	*L. pneumophila*	Unkown	Unknown	Lin *et al.*, 2018
NEL E3 ligase	**Species**	**Target**	**Function**	**Reference**
SlrP	*Salmonella* spp.	TXN, ERdj3	Regulation of host cell death	Bernal-Bayard, Cardenal-Muñoz & Ramos-Morales, 2010; Bernal-Bayard & Ramos-Morales, 2009; Zouhir *et al.*, 2014
SspH1	*Salmonella* spp.	PKN1	Inhibition of androgen receptor and NFkB signalling	Keszei *et al.*, 2014; Haraga & Miller, 2006; Miao *et al.*, 1999
SspH2	*Salmonella* spp.	SGT1, NOD1	Activation of NOD-mediated NFkB signalling	Bhavsar *et al.*, 2013; Quezada *et al.*, 2009
IpaH9.8	*Shigella flexneri*	NEMO, U2AF53	Inhibition of NFkB signalling and inhibition of splicing	Ashida *et al.*, 2010; Takagi *et al.*, 2016; Okuda *et al.*, 2005
IpaH7.8	*S. flexneri*	GLMN	Inflammasome activation (pyroptosis)	Suzuki *et al.*, 2014; Fernandez-Prada *et al.*, 2000
IpaH4.5	*S. flexneri*	p65, TBK1	Inhibition of TNFa-mediated NFkB signalling	Wang *et al.*, 2013; Zheng *et al.*, 2016
IpaH3	*S. flexneri*	Unknown	Unknown	Zhu *et al.*, 2008
IpaH1.4	*S. flexneri*	HOIP	Inhibition of NFkB activation	De Jong *et al.*, 2016
IpaH2.5	*S. flexneri*	HOIP	Inhibition of NFkB activation	De Jong *et al.*, 2016
IpaH0722	*S. flexneri*	TRAF2	Inhibition of PKC-mediated NFkB signalling	Ashida, Nakano & Sasakawa, 2013
NopM	*Rhizobium* spp.	Unknown	Dampen ROS burst, induction of plant defence genes	Xin *et al.*, 2012
F-box E3 ligase	**Species**	**Target**	**Function**	**Reference**
AnkB	*L. pneumophila*	Parvin B	Nutrient acquisition	Lomma *et al.*, 2010
LegU1	*L. pneumophila*	BAT3	Unknown	Ensminger & Isberg, 2010
LicA	*L. pneumophila*	Unknown	Unknown	Qiu & Luo, 2017; Ensminger & Isberg, 2010
GALA	*Ralstonia solanacearum*	Unknown	Unknown	Angot *et al.*, 2006; Remigi *et al.*, 2011; Chen *et al.*, 2018

VirF	*Agrobacterium tumefaciens*	Unknown	Proteolysis of the T complex?	Anand *et al.*, 2012; Zaltsman *et al.*, 2010; Magori & Citovsky, 2011
XL-box E3 ligase	**Species**	**Target**	**Function**	**Reference**
XopL	*Xanthomonas campestris* pv. vesicatoria	Unknown	Induces plastid clustering; inhibits stromule formation	Erickson *et al.*, 2018; Soni & Mondal, 2018; Singer *et al.*, 2013
Other E3 ligases	**Species**	**Target**	**Function**	**Reference**
SidE family	*L. pneumophila*	Rab1, Rab6A, Rab30, Rab33b	E1/E2 independent ubiquitination	Qiu *et al.*, 2016; Lin *et al.*, 2018
Sid C and SdcB	*L. pneumophila*	Unknown	Unknown	Hsu *et al.*, 2014
DUBs/deNEDDs/ deSUMOs	**Species**	**Target**	**Function**	**Reference**
TssM	*Burkholderia pseudomallei*	TRAF-6, TRAF-3, IkBa	Inhibition of NFkB signalling	Shanks *et al.*, 2009; Tan *et al.*, 2010
AvrA	*Salmonella* spp.	IkBa, B-catenin, MAPKK	Inhibition of NFkB and JNK signalling. Stabilisation of tight junctions. Anti-apoptotic	Ye *et al.*, 2007; Collier-Hyams *et al.*, 2002; Liao *et al.*, 2008
SseL	*Salmonella* spp.	IkBa, K63-ubiquitinated aggregates, S100A6, HNRPK	Inhibition of NFkB signalling and autophagy	Rytkonen *et al.*, 2007; Mesquita *et al.*, 2012; Arena *et al.*, 2011; Le Negrate *et al.*, 2008
ChlaDub1	*Chlamydia trachomatis*	IkBa	Inhibition of NFkB signalling	Le Negrate *et al.*, 2008b
ChlaDub2	*C. trachomatis*	Unknown	Inhibition of NFkB signalling	Misaghi *et al.*, 2006; Claessen *et al.*, 2013
ChlaOTU	*C. trachomatis, C. pneumoiae*	NDP52	Inhibition of autophagy?	Furtado *et al.*, 2013
XopD	*X.campestris* pv. vesicatoria	SIERF4	Suppression of plant ethylene production to inhibit anti Xcv-immunity	Kim *et al.*, 2008; Kim, Stork & Mudgett, 2013; Chosed *et al.*, 2007
Deamidases	**Species**	**Target**	**Function**	**Reference**
CHBP	*B. pseudomallei*	SCF complex	Arrests cell cycle	Yao *et al.*, 2012; Cui *et al.*, 2010
Cif	EHEC, EPEC, *C. rodentium*	SCF complex	Arrests cell cycle	Marchès *et al.*, 2003; Charpentier & Oswald, 2004; Loukiadis *et al.*, 2008; Morikawa *et al.*, 2010
OspI	*Shigella flexneri*	Ubc13	Inhibits TRAF6 activation	Sanada *et al.*, 2012; Fu *et al.*, 2013
Kinases	**Species**	**Target**	**Function**	**Reference**
OspG/NleH	*S. flexneri,* EHEC, EPEC, *C. rodentium*	Host E2s	Inhibit host E2s	Kim *et al.*, 2005; Zhou *et al.*, 2013; Pruneda *et al.*, 2014
Methyl-transferases	**Species**	**Target**	**Function**	**Reference**
OspZ/NleE	EHEC, EPEC, *C. rodentium, Shigella* spp.	TAB2/3, ZRANB3	Inhibits NFkB activation. Polymorphonuclear leukocyte migration. DNA repair pathways?	Nadler *et al.*, 2010; Yao *et al.*, 2014; Zhang *et al.*, 2012; Zurawski *et al.*, 2008

1.7. The NleG family of T3SS effector proteins

Although similar, the T3SS effector pool within A/E pathogens is not identical. The effector repertoire is expanded in EHEC, with around 50 effector proteins compared to 21 found in EPEC (Tobe *et al.*, 2006; Iguchi *et al.*, 2009) and 31 in *C. rodentium*. *C. rodentium* has a large number of genomic islands within its chromosome that show evidence of integration, and most virulence factors of *C. rodentium* are encoded within these. Among Nle effectors are the NleG proteins. NleG1, NleG7 and NleG8 are described in *C. rodentium* (**Figure 1.12**) along with two putative pseudogenes, *nleG* and *nleG2*. These are the focus of this project.

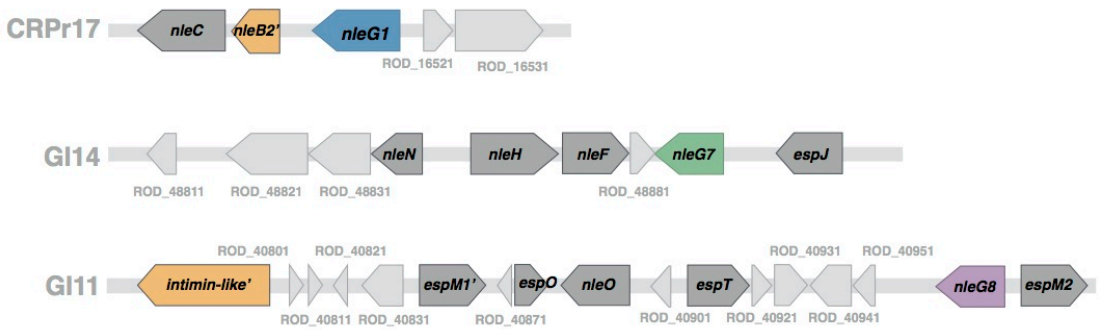


Figure 1.12 | The *nleG* genes in *C. rodentium* are located within different genomic islands. Effector genes are shown in dark grey, pesudogenes in orange. *nleG1*, *nleG7* and *nleG8* are highlighted in blue, green and purple, respectively.

1.7.1. Identification of the NleG family

NleG was first described in EPEC E2348/69 in 2006; the NleG family has since become the largest known family of non-LEE encoded effector proteins (Li *et al.*, 2006). Bioinformatic analysis identifies at least 20 NleG homologues in A/E pathogens, including 14 in EHEC strain Sakai and five in *C. rodentium* (Tobe *et al.*, 2006). Likewise, NleG homologs also exist in *Salmonella* spp., including StoD in *Salmonella enterica* serovar Typhi. Many homologs were confirmed to be T3SS effectors via immunofluorescence or CyA- /β-lactamase translocation assays (Tobe *et al.*, 2006). A recent study that sought to identify gene clusters associated with lethality in EPEC cases found that several *nleG* alleles in EPEC associated with the severity of diarrhoeal symptoms, motivating further efforts to understand contributions of NleGs to A/E pathogenesis.

The conservation of NleG family members in enteropathogenic bacteria suggests that they hold a role in infection. NleG proteins have variable N-termini but a conserved C-terminal motif of approximately 100 amino acids with no sequence homology to proteins of known function. The structure of the C-terminus of NleG2.3$_{EHEC}$ was solved by nuclear magnetic resonance (NMR) spectroscopy and revealed striking structural homology to the RING/U-box domain of eukaryotic E3 ubiquitin ligases (**Figure 1.13**) (Wu *et al.*, 2010). This prompted the hypothesis that NleG proteins functionally mimic eukaryotic E3 ligases to subvert host ubiquitin-mediated pathways upon infection. If the C-terminal domain of NleG proteins confers enzymatic activity, it is tempting to speculate that the variable N-termini of NleG proteins reflect specificity to different binding partners or ubiquitination substrates.

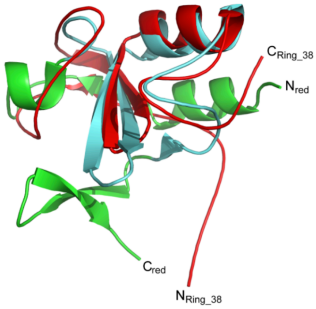


Figure 1.13 | NleG proteins have a conserved C-terminal motif that structurally resembles an E3 ligase. A RING finger domain (in red) superimposed onto the structure of EHEC NleG 2.3 C-terminus. The RING/U-box motif (3 β-sheets, 1 α-helix) of the NleG is shown in cyan, while the remainder of the conserved C-terminus is shown in green. Data from Wu *et al.*, 2010.

1.7.2. NleG proteins are ubiquitin E3 ligases

Importantly, it was revealed that purified full length NleG2.3, NleG5.1 and NleG9 from EHEC strain Sakai confer E3 ubiquitin ligase activity in the presence of ubiquitin, an E1 enzyme and an E2 enzyme (UBE2D2) in an *in vitro* auto-ubiquitination assay (Wu *et al.*, 2010). Auto-ubiquitination was detected when the conserved C-terminal domains only of NleG2.3 and NleG5.1 were tested in this manner. In contrast, the N-terminus of NleG5.1 did not support auto-ubiquitination, suggesting that the C-terminal domain conserved in NleG proteins confers E3 ubiquitin ligase activity across the family. The interaction between EHEC NleG2.3 and UBE2D2 was analysed further. Three single point mutations (I121K, L152K and P160K) within the RING/U-box motif of NleG2.3 abolished the interaction with UBE2D2 and subsequent auto-ubiquitination, supporting the idea that NleG proteins evolved to manipulate ubiquitin-dependent processes in the host (Wu *et al.*, 2010).

1.7.3. PDZ-binding motifs in NleG proteins

EHEC NleG7, EHEC NleG8.1 and *C. rodentium* NleG8 also harbour C-terminal PDZ-binding motifs known to recognise PDZ-domains: ubiquitous 70-90 residue protein-protein interaction modules (Beuming *et al.*, 2005). The acronym PDZ comes from the proteins where the domains were first described, namely post-synaptic density protein (PSD95), *Drosophila* disc large tumour suppressor (Dlg1) and zonula occludens-1 protein (zo-1). PDZ-containing proteins play key roles in orchestrating complex scaffolding formations at cellular membranes, such as anchoring receptors or ion channels to cytoskeletal components. Examples include scaffolding proteins NHERF1 and NHERF2 that anchor Na^+/H^+ exchanger 3 (NHE3) to actin via ezrin (Cha & Donowitz, 2008). Intriguingly, NHE3 is downregulated in a T3SS-dependent manner during EPEC-infection and the T3SS effector protein Map binds NHERF1 via its PDZ-binding motif to subvert actin pathways and trigger filopodia formation during early EPEC-infection (Berger *et al.*, 2009a; Martinez *et al.*, 2010). PDZ-binding motifs clearly play a role in the context of T3SS-mediated A/E pathogenesis but the significance of NleG- and PDZ-protein interactions remain unknown.

1.8. Aims of this project

Studying the impact of effector secretion on host cells has revealed fascinating complexity and sophistication to the T3SS system. However, our understanding of its complex effector protein repertoire during *in vivo* infection is far from complete. The contribution of effector proteins that exploit the ubiquitination system is not well characterized in *C. rodentium*, and this project sought to elucidate the roles of the NleG family of effector proteins to A/E pathogenesis.

The specific aims of this project are threefold:

1) Examine the *C. rodentium* NleG proteins *in vitro,* assessing their putative enzymatic activity and assessing their intracellular localisation during transfection.
2) Identify interacting partners of *C. rodentium* NleG1, NleG7 and NleG8 and characterise putative interactions.
3) Examine the *in vivo* importance of NleG proteins using the *C. rodentium* model for A/E pathogenesis.

Chapter 2:
Materials and methods

2. Materials and methods

2.1. Cells and strains

2.1.1. Mammalian cell culture

HeLa, Swiss 3T3 and HCT116 cells (obtained from ATCC) and murine immortalised bone marrow macrophages (iBMMs, courtesy of Dr Timurs Maculins (Đikić Lab, Institute of Biochemistry II, Goethe University Frankfurt - Medical Faculty, University Hospital) were cultured as a monolayer in a humidified 37°C atmosphere (5% CO_2) in Dulbecco's Modified Eagle's Medium (DMEM; Sigma) containing 4500 mg/l glucose (1000 mg/l glucose for HeLa cells). DMEM was supplemented with 10% heat-inactivated foetal calf serum (FCS) and 2 mM glutamine (Invitrogen). Cell lines were passaged at a 2-3-day interval (1 in 10 dilution) by 5-minute treatment in 0.02% ethylenediamine tetra-acetic acid (EDTA) solution (Sigma) supplemented with 0.1% trypsin (Gibco). Mycoplasma screenings were carried out routinely using the MycoAlert™ kit (Lonza).

Table 2.1| Antibiotic stock- and working concentrations.

Antibiotic	Supplier	Stock concentration	Working concentration
Ampicillin (Amp)	Sigma-Aldrich	100 mg/ml [diluted in water]	100 µg/ml
Chloramphenicol (Cm)	Sigma-Aldrich	30 mg/ml [diluted in ethanol]	8 µg/ml (Cm^{Low}), 30 µg/ml or 90 µg/mL (Cm^{High})
Doxycyline (Dox)	Sigma-Aldrich	1 mg/ml [diluted in water]	1 µg/ml
Gentamycin (Gm)	Sigma-Aldrich	10 mg/ml [diluted in water]	10 µg/ml
Kanamycin (Kan)	Sigma-Aldrich	50 mg/ml [diluted in water]	50 µg/ml
Nalidixic acid (Nal)	Alfa Aesar	50 mg/ml [diluted in water]	50 µg/ml
Streptomycin (Sm)	Sigma-Aldrich	100 mg/ml [diluted in water]	50 µg/ml or 2 mg/ml (Sm^{High})
Tetracycline (Tet)	Sigma-Aldrich	15 mg/ml [diluted in 50% v/v ethanol in water]	15 µg/ml

2.1.2. Strains and growth conditions

Bacterial and yeast strains used in this study are listed in **Table 2.2**. Bacteria were cultured in Lysogeny broth (LB broth; Merck Millipore) at 37°C shaking at 200 rpm or statically at 37°C on LB agar (Merck Millipore) plates. If applicable, the LB was supplemented with antibiotics (listed in **Table 2.1**). Yeast were cultured at 30°C in yeast peptone dextrose medium (YPD; 1% yeast extract, 2% peptone, 2% glucose 20 mg/mL adenine hemisulfate) or synthetic dextrose medium (SD; 6.7 g/L yeast Nitrogen base, Difco) supplemented with 2% glucose and appropriate amino acids for plasmid selection.

Table 2.2 | Strains and plasmids used in this study.

Strain	Relevant characteristics	Resistance	Source
C. rodentium ICC169	Wildtype *C. rodentium*, Nal^R derivative of ICC168, O152 serotype	Nal^{50}	Wiles *et al.*, 2004
C. rodentium Δ*nleG1*	*nleG1* deletion mutant	Nal^{50}	This study
C. rodentium Δ*nleG7*	*nleG7::cat* (*nleG7* replaced with chloramphenicol resistance cassette)	Nal^{50}, Cm^8	Laboratory collection
C. rodentium Δ*nleG8*	*nleG8*::*aphT* (*nleG8* replaced with kanamycin resistance cassette)	Nal^{50}, Kan^{50}	Laboratory collection
C. rodentium Δ*nleG7* Δ*nleG8*	*nleG7::cat nleG8::aphT*	Nal^{50}, Kan^{50}, Cm^8	Laboratory collection
C. rodentium Δ*nleG1* Δ*nleG7* Δ*nleG8*	Δ*nleG1 nleG7::cat nleG8::aphT*	Nal^{50}, Kan^{50}, Cm^8	This study
C. rodentium $NleG1_{P163K}$	WT *nleG1* gene replaced with *nleG1* gene with P163K mutation	Nal^{50}	This study
C. rodentium Δ*nleG1* + NleG1	*nleG1* deletion mutant with the *nleG1* gene re-introduced in its native site	Nal^{50}	This study
C. rodentium Δ*nleG1* Δ*nleG7* Δ*nleG8* + NleG1	Δ*nleG1* Δ*nleG7* Δ*nleG8* mutant with the *nleG1* gene re-introduced in its native site	Nal^{50}, Kan^{50}, Cm^8	This study
C. rodentium Δ*nleG1* Δ*nleG7* Δ*nleG8* + $NleG1_{P163K}$	Δ*nleG1* Δ*nleG7* Δ*nleG8* mutant with the *nleG1* gene with P163K mutation introduced into the *nleG1* site	Nal^{50}, Kan^{50}, Cm^8	This study
C. rodentium Δ*espA*	*espA::cat* (*espA* replaced with kanamycin resistance cassette)	Nal^{50}, Kan^{50}	Petty *et al.*, 2010
C. rodentium EspH-TEM	*C. rodentium* transformed with pCX340 expressing EspH-TEM-1	Tet^{15}	Petty *et al.*, 2010
C. rodentium Δ*espA* EspH-TEM	*C. rodentium* Δ*espA* transformed with pCX340 expressing EspH-TEM-1	Tet^{15}	Petty *et al.*, 2010
E. coli 1047 helper strain	*E. coli* containing pRK2013 that conveys mobilisation functions *in trans*	Sm^{50}	Kaniga, Delor & Cornelis, 1991
E. coli CC118-λpir	Δ*(ara-leu) araD ΔlacX74 galE galK phoA20 thi- rpsE rpoB argE(Am) recA1, λpir*	-	Herrero, De Lorenzo & Timmis, 1990
E. coli TOP10	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*E *gal*K *rps*L (Str^R) *end*A1 λ	-	Invitrogen
E. coli BL21 Star	F^-*ompT hsd*S_B (r_B^-, m_B^-) *galdcmrne*131 (DE3)	-	Thermo Fisher Scientific
E. coli BL21 pLysS	F^-*ompT hsd*SB (r_B^-, m_B^-) *galdcmrne*131 (DE3) pLysS (Cam^R)	Cm^{50}	Thermo Fisher Scientific
S. cerevisae AH109	ATa, *trp*1-901, *leu*2-3, 112, *ura*3-52, *his3*-200, *gal4*Δ, *gal80*Δ, LYS2::GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3::MEL1UAS-MEL1TATA-*lacZ*, MEL1) and Y187 (MATa, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ gal80Δ, met-, URA3::GAL1UAS-GAL1TATA-lacZ, MEL1)	-	Clontech
S. cerevisae BY4741	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	-	Brachmann *et al.*, 1998

Plasmid	Relevant characteristics	Resistance	Source
pKNG101	Suicide vector, *sacB*	Sm^{50}	Kaniga, Delor & Cornelis, 1991
pKNG101-$NleG1^{FLANK}$	pKNG101 derivative for triparental mating; encodes 500 bp upstream- and downstream of *nleG1*	Sm^{50}	This study
pSEVA-612S	GmR; R6K ori, oriT, I-SceI restriction sites flanking multicloning site	Gm^{50}	Martínez-García *et al.*, 2015

pSEVA-612S-NleG1^{FLANK}	pSEVA612S-derivative for triparental mating; encodes 500 bp upstream- and downstream of *nleG1*	Gm^{50}	This study
pSEVA-612S-NleG1^{FLANK}-NleG1	pSEVA612S-derivative for triparental mating; encodes NleG1 and 500 bp upstream- and downstream of *nleG1*	Gm^{50}	This study
pSEVA-612S-NleG1^{FLANK}-NleG1$_{P163K}$	pSEVA612S-derivative for triparental mating; encodes NleG1$_{P163K}$ and 500 bp upstream- and downstream of *nleG1*	Gm^{50}	This study
pACBSR	p15A ori, PBAD, I-SceI endonuclease and λred genes	Cm^{50}	Ruano-Gallego, Álvarez & Fernández, 2015
pCX340	pBR322 derivative; cloning vector used to fuse effector proteins to the mature form of TEM-1 β-lactamase	Tet^{15}	Charpentier & Oswald, 2004
pCX340-NleG1	Encodes NleG1-TEM-1 fusion	Tet^{15}	This study
pCX340-NleG2	Encodes NleG2-TEM-1 fusion	Tet^{15}	This study
pCX340-NleG7	Encodes NleG7-TEM-1 fusion	Tet^{15}	This study
pCX340-NleG8	Encodes NleG8-TEM-1 fusion	Tet^{15}	This study
pACYC184	Medium copy-number bacterial expression vector	Tet^{15}, Cm^{50}	New England Biolabs
pACYC-NleG8	pACYC184 derivative; encodes NleG8	Tet^{15}	This study
pACYC-NleG8$_{P177K}$	pACYC184 derivative; encodes NleG8 with a P177K point mutation	Tet^{15}	This study
pRETRO-X-Tight-Pur	Transfection vector for generating stable cell lines with doxycycline-inducible HA-tagged transgene expression	Pur^{R}	Clonetech
pRETRO-X-Tight-Pur-NleG1	Transfection vector for generating stable cell lines with doxycycline-inducible HA-NleG1 expression	Pur^{R}	Dr Tim Maculins
pRETRO-X-Tight-Pur-NleG1$_{P163K}$	Transfection vector for generating stable cell lines with doxycycline-inducible HA-NleG1$_{P163K}$ expression	Pur^{R}	Dr Tim Maculins
pRETRO-X-Tight-Pur-NleG8	Transfection vector for generating stable cell lines with doxycycline-inducible HA-NleG8 expression	Pur^{R}	Dr Tim Maculins
pRETRO-X-Tight-Pur-NleG8$_{P177K}$	Transfection vector for generating stable cell lines with doxycycline-inducible HA-NleG1$_{P163K}$ expression	Pur^{R}	Dr Tim Maculins
pETM11	Bacterial expression vector for purification of proteins with an N-terminal TEV-cleavable 6His-tag	Kan^{50}	Dümmler, Lawrence & de Marco, 2005
pETM11-NleG1	pETM11 derivative encoding 6His-NleG1	Kan^{50}	This study
pETM11-NleG1$_{P163K}$	pETM11 derivative encoding 6His-NleG1$_{P163K}$	Kan^{50}	This study
pETM11-NleG7	pETM11 derivative encoding 6His-NleG7	Kan^{50}	This study
pETM11-NleG8	pETM11 derivative encoding 6His-NleG8	Kan^{50}	This study
pETM11-NleG8$_{P177K}$	pETM11 derivative encoding 6His-NleG8$_{P177K}$	Kan^{50}	This study
pGEX-6P-2	Bacterial expression vector for protein purification	Amp^{100}	GE Healthcare
pGEX-6P-2-NleG1	pGEX-6P-2 derivative encoding GST-NleG1	Amp^{100}	This study
pGEX-6P-2-NleG1$_{P163K}$	pGEX-6P-2 derivative encoding GST-NleG1$_{P163K}$	Amp^{100}	This study
pGEX-6P-2-NleG7	pGEX-6P-2 derivative encoding GST-NleG7	Amp^{100}	This study
pGEX-6P-2-NleG7$_{P177K}$	pGEX-6P-2 derivative encoding GST-NleG7$_{P177K}$	Amp^{100}	This study
pGEX-6P-2-NleG8	pGEX-6P-2 derivative encoding GST-NleG8	Amp^{100}	This study
pGEX-6P-2-NleG8$_{P177K}$	pGEX-6P-2 derivative encoding GST-NleG8$_{P177K}$	Amp^{100}	This study
pGEX-6P-2-NleG7	Encodes GST-NleG7	Amp^{100}	This study
pGEX-6P-2-	Encodes GST-NleG7 with a P177K point mutation	Amp^{100}	This study

NleG7$_{P177K}$			
pGEX-6P-2-NleG8	Encodes GST-NleG8	Amp^{100}	This study
pGEX-6P-2-NleG8$_{P177K}$	Encodes GST-NleG8 with a P177K point mutation	Amp^{100}	This study
pRK5	Plasmid for transfection of mammalian cells (no tag)	Amp^{100}	Laboratory collection
pRK5-HA	Transfection vector; HA-tag upstream of MCS	Amp^{100}	Laboratory collection
pRK5-Myc	Transfection vector; Myc-tag upstream of MCS	Amp^{100}	Laboratory collection
pRK5-Myc-NleG1	Encodes NleG1 with an N-terminal Myc-tag	Amp^{100}	This study
pRK5-HA-NleG7	Encodes NleG7 with an N-terminal HA-tag	Amp^{100}	This study
pRK5-Myc-NleG7	Encodes NleG7 with an N-terminal Myc-tag	Amp^{100}	This study
pRK5-Myc-NleG8	Encodes NleG8 with an N-terminal Myc-tag	Amp^{100}	This study
pRK5-Myc-NleG8$_{P177K}$	Encodes NleG8$_{P177K}$ with an N-terminal Myc-tag	Amp^{100}	This study
pRK5-Myc-NleG8$_{\Delta DTRL}$	Encodes NleG8 with an N-terminal Myc-tag, PDZ-binding motif deleted	Amp^{100}	This study
pRK5-Myc-GFP	Encodes GFP with N-terminal Myc-tag	Amp^{100}	Laboratory collection
pRK5-HA-mCherry	Encodes mCherry with N-terminal HA-tag	Amp^{100}	Laboratory collection
pCMV-GRASP55-Myc	Mammalian transfection vector encoding C-terminally Myc-tagged human GRASP55		Gee *et al.*, 2011
pGBKT7	Plasmid for expression of proteins fused with GAL4A binding domain and N-terminal Myc-tag in yeast	Kan^{50}	Clontech
pGBKT7-NleG1	Derivative of pGBKT7, encodes NleG1-GAL4A-DNA-binding domain fusion	Kan^{50}	This study
pGBKT7-NleG1^{N}	Derivative of pGBKT7, encodes NleG1 N-terminus fused with GAL4A-DNA-binding domain	Kan^{50}	This study
pGBKT7-NleG7	Derivative of pGBKT7, encodes NleG7-GAL4A-DNA-binding domain fusion	Kan^{50}	This study
pGBKT7-NleG7^{N}	Derivative of pGBKT7, encodes NleG7 N-terminus (amino acids 1-97) fused with GAL4A-DNA-binding domain	Kan^{50}	This study
pGBKT7-NleG8	Derivative of pGBKT7, encodes NleG8-GAL4A-DNA binding domain fusion	Kan^{50}	This study
pGBKT7-NleG8^{N}	Derivative of pGBKT7, encodes NleG8 N-terminus (amino acids 1-109) fused with GAL4A-DNA-binding	Kan^{50}	This study
pGBKT7-NleG8.1$_{EHEC}$	Derivative of pGBKT7, encodes EHEC NleG8.1-GAL4A-DNA-binding domain fusion	Kan^{50}	Laboratory collection
pGBKT7-StoD	Derivative of pGBKT7, encodes StoD-GAL4A-DNA binding domain fusion	Kan^{50}	Laboratory collection
pGBKT7-StoD$_{P204K}$	Derivative of pGBKT7, encodes StoD$_{P204K}$-GAL4A-DNA binding domain fusion	Kan^{50}	Laboratory collection
pGADT7	Plasmid for expression of proteins fused with GAL4A activation domain and N-terminal HA-tag in yeast	Amp^{100}	Clontech
pGADT7-PA28	Derivative of pGADT7, encodes PA28 subunit fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-Cdc20	Derivative of pGADT7, encodes human CDC20 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-Merp1	Derivative of pGADT7, encodes Merp1 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-WDR77	Derivative of pGADT7, encodes WDR77 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-AD	Derivative of pGADT7, encodes alcohol dehydrogenase fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-A6NEOI	Derivative of pGADT7, encodes A6NEO1 fused with GAL4A activation domain	Amp^{100}	This study

pGADT7-MTK	Derivative of pGADT7, encodes mitochondrial thymidine kinase fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-Marc2	Derivative of pGADT7, encodes Marc2 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-Nidogen1	Derivative of pGADT7, encodes Nidogen-1 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-GRASP55	Derivative of pGADT7, encodes GRASP55 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-BBS7	Derivative of pGADT7, encodes BBS7 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-OCRL1	Derivative of pGADT7, encodes OCRL1 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-WDR61	Derivative of pGADT7, encodes WDR61 fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-HAX1	Derivative of pGADT7, encodes HAX1 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-uPAR	Derivative of pGADT7, encodes uPAR fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-PRKRIR	Derivative of pGADT7, encodes PRKRIR fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-Calpain1	Derivative of pGADT7, encodes Calpain-1 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-RPD3	Derivative of pGADT7, encodes RPD3 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-STAT5	Derivative of pGADT7, encodes STAT5 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-H3	Derivative of pGADT7, encodes histone H3 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-LOC401397	Derivative of pGADT7, encodes LOC401397 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-$LOC401397^{Coding}$	Derivative of pGADT7, encodes LOC401397 (putative coding region only) fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-PRMT5	Derivative of pGADT7, encodes PRMT5 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGAD424-NHERF1	Derivative of pGAD424, encodes human NHERF1 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGAD424-NHERF2	Derivative of pGAD424, encodes human NHERF2 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-ubiquitin	Derivative of pGADT7, encodes ubiquitin fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-ISG15	Derivative of pGADT7, encodes ISG15 fused with GAL4A activation domain	Amp^{100}	Laboratory collection
pGADT7-C1QBP	Derivative of pGADT7, encodes full-length C1QBP fused with GAL4A activation domain	Amp^{100}	This study
pGADT7-$C1QBP_{Mature}$	Derivative of pGADT7, encodes C1QBP (mature form, amino acids 74-282) fused with GAL4A activation domain	Amp^{100}	This study
pTT3-C1QBP-bio-His	Transfection vector with pCMV-promoter encoding 6his-tagged C1QBP	Amp^{100}	Addgene
pETM11-C1QBP	pETM11 derivative encoding 6His-C1QBP	Kan^{50}	This study
pETM11-$C1QBP_{74-282}$	pETM11 derivative encoding 6His-tagged mature C1QBP	Kan^{50}	This study
pYES2	Glucose-repressible galactose-inducible yeast expression vector. *URA* gene for uracil biosynthesis, for selection in yeast.	Amp^{100}	Invitrogen
pYES2-NleG1	Encodes *C. rodentium* NleG1	Amp^{100}	This study
pYES2-NleG7	Encodes *C. rodentium* NleG7	Amp^{100}	This study
pYES2-$NleG7^{N}$	Encodes *C. rodentium* NleG7 N-terminus (amino acids 1-97)	Amp^{100}	This study
pYES2-$NleG7^{C}$	Encodes *C. rodentium* NleG7 C-terminus (amino	Amp^{100}	This study

	acids 98-209)		
pYES2-NleG8	Encodes *C. rodentium* NleG8	Amp^{100}	This study
pYES2-$NleG8^{N}$	Encodes *C. rodentium* NleG8 N-terminus (amino acids 1-109)		
pYES2-$NleG8^{C}$	Encodes *C. rodentium* NleG8 C-terminus (amino acids109-213)	Amp^{100}	This study
pYES2-$NleG8_{P177K}$	Encodes *C. rodentium* NleG8 with P177K point mutation	Amp^{100}	This study
pYES2-$NleG8_{DARA}$	Encodes *C. rodentium* NleG8 with mutated PDZ-binding motif (DTRL to DARA)	Amp^{100}	This study
pYES2-$NleG7^{N}$-$NleG8^{C}$	Encodes *C. rodentium* $NleG7^{1-97}$-$NleG8^{109-213}$ chimeric fusion protein	Amp^{100}	This study
pYES2-$NleG8^{N}$-$NleG7^{C}$	Encodes *C. rodentium* $NleG8^{1-109}$-$NleG7^{98-209}$ chimeric fusion protein	Amp^{100}	This study
pYES2-$NleG7_{EHEC}$	Encodes EHEC NleG7	Amp^{100}	Laboratory collection
pYES2$NleG8.1_{EHEC}$	Encodes EHEC NleG8.1	Amp^{100}	Laboratory collection
pcDNA3.1-3xFLAG-TEV-UBE2O	Transfection vector with pCMV-promoter encoding FLAG-tagged UBE2O	Amp^{100}	Yanagitani, Juszkiewicz & Hegde, 2017
pcDNA3.1-3xFLAG-TEV-$UBE2O_{C1040S}$	Transfection vector with pCMV-promoter encoding FLAG-tagged UBE2O with C1040S mutation	Amp^{100}	Yanagitani, Juszkiewicz & Hegde, 2017
pcDNA3.1-3xFLAG-GFP	Transfection vector with pCMV-promoter encoding FLAG-tagged GFP	Amp^{100}	Yanagitani, Juszkiewicz & Hegde, 2017

R=resistance.

2.1.3. Growth assays

For growth curves, bacterial strains were grown in a 96-well plate in 200 µl of LB, DMEM or minimal M9 media at 37°C with agitation (200 rpm). Yeast strains were grown similarly in SD broth. Growth was monitored by regularly measuring the optical density at 600 nm (OD_{600}) using a FLUOstar OPTIMA Microplate reader (BMG LABTECH), starting at an OD_{600} of 0.1.

2.2. DNA manipulations

2.2.1. Plasmids

Plasmids used in this study are listed in **Table 2.2**. A Miniprep kit was used as recommended by the manufacturer (Peqlab) for small-scale plasmid extraction from 5 ml overnight bacterial cultures. For transfection experiments, endotoxin-free plasmid DNA was extracted from 150 ml cultures using a QIAPrep Spin Midiprep kit as recommended by the manufacturer (QIAGEN). DNA was eluted in sterile de-ionised water and DNA concentrations measured using a NanoDrop NS-1000 (Thermo Scientific).

2.2.2. Polymerase chain reaction (PCR)

Genomic DNA was extracted using a DNA extraction kit (QIAGEN). Genes were PCR-amplified from genomic and plasmid DNA or inverse PCR-amplified from plasmid DNA using KOD Hot Start polymerase kit (Novagen) and primers listed in **Table 2.3**. Colony screening PCRs were performed with pre-mixed One*Taq* 2X Master Mix with Standard Buffer (New England Biolabs, NEB) in 15 µl reactions containing 0.5 µM primers.

Table 2.3 | Primers used in this study. *With silent mutation to disrupt EcoRI restriction site; RBS = ribosome binding site, inv. = inverse.

Primer	Sequence	Enzyme		Plasmid	Insert/amplicon
AS01	CGCGGATCCATGCAGCCTGTAGGACTATC	BamHI	FWD	pRK5-HA/Myc, pYES2	NleG1
AS02	CGCCGGAATTCTTAACTATTTTTTATAACGAAGTTAC	EcoRI	REV	pRK5-HA/Myc, pYES2	NleG1
AS05	CGCGGATCCATGCCTATTATGTTAAATTTCTCAAATG	BamHI	FWD	pRK5-HA/Myc, pYES2	NleG8
AS06	CGCCGGAATTCTCAAAGACGCGTATCCGACC	EcoRI	REV	pRK5-HA/Myc, pYES2	NleG8
AS07	CGCCGGAATTCAAGAAGGAGATATACCTACGTAATGCCTATTATGTTAAATTTCTCAAATG	EcoRI	FWD	pACYC	RBS + NleG8
AS08	CGCAGTACTTCAAAGACGCGTATCCGACC	ScaI	REV	pACYC	NleG8
AS09	CGCGGATCCATGCCTATCTTTTTAAACTTCTCC	BamHI	FWD	pRK5-HA/Myc, pYES2	NleG7
AS10	CGCCGGAATTCTCAGTCTGTATCCAGTATCCTGAAATGCTGATTCATAATATC	EcoRI	REV	pRK5-HA/Myc, pYES2	NleG7*
AS23	ATGTGAAGCATTGCGCCGG	-	REV	pRK5	NleG8 P177K inv. PCR fragment
AS24	GCATTGAGCCGTGAAGCTTTTACACC	-	FWD	pRK5	NleG8 P177K inv. PCR fragment
AS27	CGCGATCGATGCCACCATGGCGATGCAGCCTGTAGGACTATC	ClaI	FWD	pRK5 no tag	Kozac + NleG1
AS28	CGCCGGAATTCTCAAGCGTAATCTGGAACATCGTATGGGTAACTATTTTTTATAACGAAGTTACCC	EcoRI	REV	pRK5 no tag	NleG1 + HA
AS29	CGCCGGAATTCTCACAGGTCCTCCTCGGAGATCAGCTTCTGCTCACTATTTTTTATAACGAAGTTACCC	EcoRI	REV	pRK5 no tag	NleG1 + Myc
AS30	CGCCGGAATTCGCCACCATGGCGATGCCTATCTTTTTAAACTTCTCC	EcoRI	FWD	pRK5 no tag	Kozac + NleG7
AS31	CGCAAGCTTTCAAGCGTAATCTGGAACATCGTATGGGTAGTCTGTATCCAGTATCCTGAAATGCTGATTCATAATATC	HindIII	REV	pRK5 no tag	NleG7* + HA
AS32	CGCAAGCTTTCACAGGTCCTCCTCGGAGATCAGCTTCTGCTCGTCTGTATCCAGTATCCTGAAATGCTGATTCATAATATC	HindIII	REV	pRK5 no tag	NleG7* + Myc
AS33	CGCGATCGATGCCACCATGGCGATGCCTATTATGTTAAATTTCTCAAATG	ClaI	FWD	pRK5 no tag	Kozac + NleG8
AS34	CGCCGGAATTCTCAAGCGTAATCTGGAACATCGTATGGGTAAAGACGCGTATCCGACCTGG	EcoRI	REV	pRK5 no tag	NleG8 + HA
AS35	CGCCGGAATTCTCACAGGTCCTCCTCGGAGATCAGCTTCTGCTCAAGACGCGTATCCGACCTGG	EcoRI	REV	pRK5 no tag	NleG8 + Myc
AS48	CGCCATATGATGCAGCCTGTAGGACTATC	NdeI	FWD	pCX340	NleG1
AS49	CGCCATATGATGCCTATCTTTTTAAACTTCTCC	NdeI	FWD	pCX340	NleG7
AS50	CGCCATATGATGCCTATTATGTTAAATTTCTCAAATG	NdeI	FWD	pCX340	NleG8
AS51	CGCCGGAATTCCGACTATTTTTTATAACGAAGTTACCC	EcoRI	REV	pCX340	NleG1

AS52	CGCCGGAATTCCGGTCTGTATCCAGTATCCTGAAATGCTGATTCATAATATC	EcoRI	REV	pCX340	NleG7*
AS53	CGCCGGAATTCCGAAGACGCGTATCCGACCTG	EcoRI	REV	pCX340	NleG8
AS54	CGCCATATGATGTTAAGTATGCTTTCAGCCT	NdeI	FWD	pCX340	NleG2
AS55	CGCCGGAATTCCGATTATTTTTTATAACGAAGTTCCCC	EcoRI	REV	pCX340	NleG2
AS69	CGTGAGCCAATAAAGGCATCAATG	-	FWD	Any with NleG1	NleG1 P163K inv. PCR fragment
AS70	CGTCAGTTTGTGAACTCCACCGTCACGG	-	REV	Any with NleG1	NleG1 P163K inv. PCR fragment
AS71	CTTAGCAGGGAAAAGATAGCAA	-	FWD	Any with NleG7	NleG7 P177K inv. PCR fragment
AS72	CTTATGGGGGTGGTGCTTTCTG	-	REV	Any with NleG7	NleG7 P177K inv. PCR fragment
AS76	TCCTGTAATGGCTGCTTCCTA	-	FWD	-	LOC401397
AS77	AAGAAATGGTATATATAGAATTCTAT	-	REV	-	LOC401397
AS78	CGCGGATCCCGTCCTGTAATGGCTGCTTCCTA	BamHI	FWD	pGADT7	LOC401397
AS79	CGCGCTCGAGTCAAAGAAATGGTATATATAGAATTCTAT	XhoI	REV	pGADT7	LOC401397
AS80	CAATTGAAGCATAAAATCTCGAG	EcoRI	REV	pYES2	$NleG7^{N}$
AS81	GAGATTTTATGCTTCAATTGGAAGTTCAGAGTAGAGTTAACTC	BamHI	FWD	pYES2	$NleG7^{N}$ (20 last bp) + $NleG8^{C}$
AS82	ATTTCTGACACTTTCCTGCAAAG	EcoRI	REV	pYES2	$NleG8^{N}$
AS83	GCAGGAAAGTGTCAGAAATTCTTTAAATATAGGATGGCAAAAAG	BamHI	FWD	pYES2	$NleG8^{N}$ (20 last bp) + $NleG7^{C}$
AS84	CGCCGGAATTCTCACAATTGAAGCATAAAATCTCGAG	EcoRI	REV	pYES2	$NleG7^{N}$ (stop added)
AS85	CGCGGATCCATGTCTTTAAATATAGGATGGCAAAAAG	BamHI	FWD	pYES2	$NleG7^{C}$ (start added)
AS86	CGCGGATCCATGGAAGTTCAGAGTAGAGTTAACTC	BamHI	FWD	pYES2	$NleG8^{C}$ (start added)
AS87	CGCCGGAATTCTCACGACCTGGGTAATATGCAGAA	EcoRI	REV	pYES2	$NleG8^{C}$ ΔDTRL
AS88	CGCCGGAATTCTCATGCACGTGCATCCGACCTGGGTAATATGCAGAA	EcoRI	REV	pYES2	$NleG8^{C}$ DARA mutant
AS106	ATCAGCATGCCCTTACTTCGTTGAGTGTCCAACTTTTGGG	SphI	FWD	pSEVA	NleG1upstream region
AS107	ACTAGGATCCCCGGAGGTGTATTCCGGCTTTTGTGAAC	BamHI	REV	pSEVA	NleG1 downstream region
AS108	CGCCGGAATTCATGTCTTTAAATATAGGATGGCAAAAAG	EcoRI	FWD	pGBKT7	$NleG7^{C}$ (start added)
AS109	CGCCGGAATTCATGGAAGTTCAGAGTAGAGTTAACTC	EcoRI	FWD	pGBKT7	$NleG8^{C}$ (start added)
AS110	CGCCGGAATTCATGTCATCATTACAGACAAGAATATCA	EcoRI	FWD	pGBKT7	$NleG1^{C}$ (start added)
AS145	CGCCGGAATTC ATGCTGCCTCTGCTGCGC	EcoRI	FWD	pGADT7	C1QBP
AS146	CGC GGATCC CTACTGGCTCTTGACAAAACTC	BamHI	REV	pGADT7	C1QBP
BM73	CTATTCGATGATGAAGATACCCCACCAAACCC	-	FWD	pGADT7	Sequencing amplicon
BM74	GTGAACTTGCGGGGTTTTTCAGTATCTACGATT	-	REV	pGADT7	Sequencing amplicon
Upkn	CCCTGGATTTCACTGATGAG	-	REV	pKNG101	Sequencing amplicon
Rpkn	CATATCACAACGTGCGTGGA	-	FWD	pKNG101	Sequencing amplicon
OK01	CATCCATGGGACCAGCTAGCTCATTATTGAATG	NcoI	FWD	pETM11	NleG1 [10-194]
OK02	GGCCTCGAGTTAACTATTTTTTATAACGAAGTTACC	XhoI	REV	pETM11	NleG1 [10-194]
OK03	CATCCATGGGACCTGAAAATGAACTAAACTCTTTAC	NcoI	FWD	pETM11	NleG7 [14-209]
OK04	CCCAAGCTTTCAGTCTGTATCCAGAATTCTG	HindIII	REV	pETM11	NleG7 [14-209]
OK05	CATCCATGGGAAGAATATTGCCAGAAAGTGAA	NcoI	FWD	pETM11	NleG8 [11-210]

	GT				
OK06	CCGCTCGAGTCAAAGACGCGTATCCGAC	XhoI	REV	pETM11	NleG8 [11-210]
OK07	CATCCATGTCATCATTACAGACAAGAATATCA	NcoI	FWD	pETM11	NleG1^{C}
OK08	CATGCCATGGCAATGGTCCCTCTGCTGAGATG	NcoI	FWD	pETM11	C1QBP [1-282]
OK09	CATGCCATGGCAATGCACACCGACGGCGA	NcoI	FWD	pETM11	C1QBP [74-282]
OK10	CCGCTCGAGTTACTGGGACTTCACGAAGGAC	XhoI	REV	pETM11	C1QBP

2.2.3. Agarose gel electrophoresis

DNA fragments (PCR products or digested plasmids) were resolved via agarose gel electrophoresis. Gels were made with 0.8-1.2% agarose (w/v) in 1X TAE buffer (40 mM Tris-acetate and 1 mM EDTA) and using SYBRSafe™ DNA Stain (Invitrogen; 1:10000 dilution). Samples were mixed with 6X DNA Loading Buffer (NEB), separated at 150V in 1X TAE buffer and visualised with a SafeImage™ blue light trans-illuminator (Invitrogen). Following verification on agarose gels, PCR products were purified using the Qiaquick PCR purification kit (QIAGEN) and vectors purified using the Qiaquick Gel Extraction kit (QIAGEN).

2.2.4. Restriction enzyme digestions and ligations

Restriction endonucleases were used as recommended by the manufacturer (New England Biolabs) and the digested DNA fragments were purified using the Qiaquick PCR purification kit (for digested PCR products) or the Qiaquick Gel Extraction kit (for digested vectors). Ligation reactions were performed using T4 DNA ligase (NEB) as per the manufacturer's instructions.

2.2.5. Bacterial transformation by heat shock

Purified recombinant plasmid DNA and ligation products were transformed into chemically competent *E. coli* TOP10 (Invitrogen). To make *E. coli* chemically competent, cells were grown in liquid LB to OD_{600}=0.5 and harvested by centrifugation (4000 x *g*) for 10 minutes at 4°C. Cells were resuspended in ice-cold transformation buffer I (TFBI, pH 5.8) containing 30 mM potassium acetate, 10 mM $CaCl_2$, 50 mM $MnCl_2$ and 15% v/v glycerol and left on ice for 5 minutes before harvesting as above. Cells were then resuspended in TFBII buffer (pH 6.5) containing 10 mM MOPS (3-(N-morpholino)-propanesulphonic acid), 75 mM $CaCl_2$, 10 mM $RbCl_2$ and 15% v/v glycerol, in one tenth of the original culture volume and stored at -80°C until use. For one transformation, 80 µl competent cells were incubated on ice for 30 minutes with 5-10 µl ligation reaction or circa 100 ng purified plasmid DNA prior to a 42°C-heat shock for 60 seconds, a 5-minute incubation on ice, recovery for 1-2 hours in 500 µl S.O.C. media (containing 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose; Thermo Fisher Scientific) in at 37°C with shaking. The transformation mix was plated on selective antibiotic LB plates and incubated at 37°C overnight.

2.2.6. Bacterial transformation by electroporation

Purified recombinant plasmid DNA was transformed into *C. rodentium* strains by electroporation. To make electrocompetent *C. rodentium*, cells were grown in 50 ml LB broth to OD_{600}=0.5 and cooled on ice for 45 minutes prior to harvesting by centrifugation at 4000 x *g* (10 minutes, 4°C). Cells were washed twice in 25 ml ice-cold 15% v/v glycerol and pelleted by centrifugation as above. After the second wash, cells were resuspended 15% v/v glycerol in one tenth of the original culture volume and aliquoted into 80 µl aliquots in. One aliquot was used per transformation and 3-5 µl plasmid DNA was added to the cells and mixed in an electroporation cuvette (0.2 cm width, BioRad). The GenePulser™ electroporation machine (BioRad) was used for the electroporation 300 Ω resistance, 2.5 kV voltage and 25 µF capacity. 1 ml LB broth was immediately added to the transformation mixture and left to recover at 37°C (with shaking) for 1-2 hours before plating on selective antibiotic plates and incubated at 37°C overnight.

2.2.7. Transformation of yeast

The lithium acetate (LiAc) method was used to transform *Saccharomyces cerevisiae* with pGBKT7, pGADT7 or pYES2 constructs. Yeast were cultured for 24 hours with 200 rpm shaking in 10 ml YPD media, diluted 1:10 in 50 mL YPD and grown for 4 hours at 30°C (200 rpm shaking). Cells were pelleted (15 seconds at 14000 rpm in a table-top microcentrifuge), washed with sterile water and re-suspended in 100 mM LiAc to generate competent yeast. A transformation mix containing 5 µg plasmid DNA, 50% w/v polyethylene glycol, 10 mM LiAc and 50 µg boiled herring sperm DNA was added on top of 50 µl competent yeast. Yeast were resuspended in the transformation mix and incubated at 30°C for 30 minutes and then at 42°C for 20 minutes. Yeast cells were pelleted as above, re-suspended in sterile water and plated on selective SD media.

2.2.8. Chromosomal mutagenesis using pKNG101

C. rodentium nleG1 was knocked out from the genome of WT *C. rodentium* (ICC169) via triparental conjugation using a pKNG101 suicide plasmid as previously described (Vasseur *et al.*, 2005). Briefly, *nleG1*-flanking regions were fused by overlapping PCR and cloned into pKNG101 (primers in **Table 2.3**). The resulting plasmid was maintained in *E. coli* CC118λpir to create the donor parental strain. The plasmid was mobilised into WT *C.*

rodentium (the receiver strain) aided by *E. coli* 1047 carrying the conjugative plasmid pRK2013 (the helper strain) (Herrero, De Lorenzo & Timmis, 1990; Kaniga, Delor & Cornelis, 1991). For this to occur, the strains were spotted on top of one-another on LB agar plates and incubated at 37°C for four hours. Spots were collected using sterile loops, resuspended in LB broth and plated on selective LB agar plates. Clones where recombination had occurred were selected using streptomycin-supplemented LB plates (2000 µg/ml). Mutants where a second recombination event occurred, resulting in gene deletion, were selected for by streaking onto LB plates supplemented with 20% sucrose. Successful *nleG1* gene deletion was verified via PCR and sequencing.

2.2.9. Chromosomal mutagenesis using pSEVA-612S

To generate a triple (full) *nleG* knockout mutant in *C. rodentium*, *nleG1* was knocked out from the genome of Δ*nleG7* Δ*nleG8 C. rodentium* (a strain previously generated in the lab) using a second triparental conjugation method that uses the pSEVA suicide vector. Homology regions flanking the *nleG1* gene were fused by overlapping PCR as above and cloned into pSEVA-612S (Martínez-García *et al.*, 2015). The resulting plasmid, pSEVA-612S-NleG1^{FLANK}, was maintained in *E. coli* CC118λpir to generate the donor parental strain. 20 µl of the donor strain was spotted on top of 20 µl helper *E. coli* 1047 pRK2013 on LB agar and incubated for two hours at 37°C. Next, 40 µl of the receiver strain (Δ*nleG7* Δ*nleG8 C. rodentium* carrying a pACBSR plasmid) was spotted on top and the plate incubated at 37°C for a further four hours. Bacteria were then collected with a sterile loop, resuspended in LB broth and plated on selective media. Conjugants were selected for on LB agar supplemented with Gm and Cm. Individual colonies were picked and grown in LB broth supplemented with Sm and 0.4% w/v L-arabinose (Sigma) for 8 hours (37°C shaking incubation) to induce expression of the I-SceI restriction endonuclease from pACBSR. I-SceI recognises and cleaves a DNA sequence present within pSEVA but not within the WT *C. rodentium* genome, so colonies able to grow on Cm but not Gm following L-arabinose treatment must have undergone homologous recombination to eject pSEVA from the genome. Depending on the recombination site, this will either generate the desired mutant or give reversion to WT. Successful deletion of *nleG1* was verified by PCR (see **Table 2.3** for primers) and sequencing. The strain was then passaged in liquid LB several times to remove pACBSR and Cm-sensitive mutants selected for.

2.2.10. Generation of stable cell lines

Cells with doxycycline-inducible expression of HA-tagged NleG proteins were generated by Dr Timurs Maculins and Gabor Varga (Đikić Lab, Institute of Biochemistry II, Goethe University Frankfurt - Medical Faculty, University Hospital) using the retrovirus-based pRetroX Tet-On Advanced transfection system. Briefly, *nleG* genes were PCR-amplified from *C. rodentium* genomic DNA and cloned into the pRETRO-X-Tight-Pur vector (Clontech) using standard restriction cloning methods (EcoRI and BamHI sites). The constructs were transformed into Phoenix GP cells (that express the viral proteins Gag and Pol) along with a pVSV-G plasmid encoding the Env protein required for virus assembly. Resulting virions were collected and used to infect two parental cell lines: HCT116 cells (ATCC) and immortalised murine macrophages (courtesy of Dr Timurs Maculins), both stably expressing the Tet-On transactivator (TOAt) gene in the presence of doxycycline. Stably transduced cells were selected using puromycin (Sigma) and resistant cell colonies were expanded and tested for doxycycline-inducibility of the transgene.

2.2.11. In vitro transfection

HeLa ATCC cells were seeded onto coverslips in a 24-well plate at a density of 4-5 x 10^5 cells per ml to achieve 40-60% confluency after 24 hours. 24-hour transfections were performed using the GeneJuice Transfection Reagent (Novagen) in accordance with the manufacturer's instructions. A 3:1 ratio of GeneJuice reagent to DNA (0.25 µg) was adopted. Transfected cells were fixed with 4% paraformaldehyde (PFA) in PBS for 15 minutes and washed in PBS prior to immunofluorescent staining.

Transfections of plasmid DNA into HCT116 cells were carried out using Lipofectamine® 2000 (Thermo Fischer Scientific) in accordance with the manufacturer's recommendations. Cells were seeded in 6-well plates (5 x 10^5 cells per well) or 10 cm dishes (1.5-2.0 x 10^6 cells per well) and grown for 24 hours. Transfections where carried out using 2.5 µg DNA with 7.5 µl Lipofectamine reagent (6-well format) or 4 µg DNA with 30 µg Lipofectamine reagent (10 cm dish format) for up to 24 hours. If doxycycline-inducible cell lines were used and the expression of the relevant protein was desired, doxycycline (1 µg/ml) was added to the cell medium for up to 48 hours. Where

necessary, 10 µM MG-132 treatment to inhibit the proteasome was introduced 5.5 hours prior to cell lysis and further analysis. Likewise, 5 µM PR-619 treatment to inhibit DUBs was introduced 4 hours prior to cell lysis and further analysis where necessary. Transfected cells were lysed in 1X RIPA buffer (50 mM Tris-HCL pH 7.5, 150 mM NaCl, 1% NP-40, 1 mM EDTA pH 8, and 0.1% sodium deoxycholate) supplemented with appropriate inhibitors. One proteasome inhibitor cocktail tablet (Roche) was used per 10 ml lysis buffer. In lysis buffer, PR-619 was used at a concentration of 50 µM. A BCA assay was performed (see **Section 2.9.3**) prior to further analysis by western blot or IP experiments.

Poly (I:C) transfections were also carried out using Lipofectamine® 2000 in accordance with the manufacturer's recommendations. A 6-well plate format was used with HCT cells seeded at a density of 5 x 10^5 cells per well 24 hours prior to transfection. 200-2500 ng Poly (I:C) (Sigma) resuspended in sterile water was added per well for 3 or 6 hours. Cells were lysed in RIPA buffer as outlined above and a BCA assay was performed prior to Western blot analysis.

2.3. Yeast assays

2.3.1. Yeast two-hybrid screening

Yeast-two-hybrid screens were performed with one of two bait strains: *S. cerevisiae* AH109 transformed with pGBKT7 expressing either the N-terminus of NleG7 ($NleG7^N$) or $NleG1^N$. These strains were grown in 50 ml SD medium with 2% glucose (v/v) but lacking tryptophan (-Trp) (30°C, 200 rpm shaking) until OD_{600}=0.8. Cells were harvested by centrifugation and re-suspended in 5 ml SD -Trp media for mating with the Universal Mouse (Normalized) Mate & Plate™ Library strain containing pGADT7 prey plasmids (Clontech). 45 ml 2X YPDA was added and the culture shaken at 30 rpm for 25 hours (30°C). The cells were harvested and re-suspended in 0.5X YPDA, 200 µl aliquots of which was plated on SD quadruple drop-out (QDO) plates (lacking leucine, tryptophan, histidine and adenine) supplemented with glucose (2% v/v) and X-α-gal (5-Bromo-4-chloro-3-indolyl α-D-galactopyranoside; Clontech) to select for interactions between bait and prey proteins. Plates were incubated for up to three weeks at 30°C. Resulting colonies were re-streaked thrice on QDO and pGADT7 primers used to amplify the library plasmid gene. Amplicons were sent for sequencing (GATC).

2.3.2. Yeast direct two-hybrid assay

Co-transformation of 0.1-10 µg pGBKT7 and pGADT7 plasmids into AH109 was achieved via the LiAc method (see **Section 2.2.7**). Tranformants were selected using double drop-out (DDO) SD medium lacking leucine and tryptophan with 2% glucose (v/v). Colonies were picked and re-suspended in 200 µl sterile water and 10 µl were dropped onto DDO and QDO plates for incubation at 30°C for a minimum of three days.

2.3.3. Yeast cytotoxicity assay

Transformation of 0.1-10 µg pYES2 plasmid constructs into *S. cerevisae* BY4741 was achieved via the LiAc method as already described (**Section 2.2.7**). Tranformants were selected using SD medium lacking uracil. Single colonies were cultured in 5 ml SD –uracil supplemented with 2% glucose until OD_{600}=2 and 1 ml was harvested by centrifugation (1 minute, 14000 rpm), washed twice in 1 ml sterile water and re-suspended in 1 ml sterile water. The yeast suspension was serially diluted (up to 10^{-8} dilution) in sterile water. 20 µl of each dilution was dropped onto SD plates lacking uracil, supplemented with either 2% glucose (which represses expression from pYES2) or 2% galactose (which induces expression from pYES2) for a 3 day 30°C-incubation.

2.4. In vitro infection

Glass coverslips were seeded with 1 x 10^5 Swiss cells, HeLa cells or HCT116 cells 48h prior to infection. To prime bacteria, *C. rodentium* cultures (grown for 8 hours in 2 ml LB broth) were diluted 1:500 in DMEM (1000 mg/l glucose) and incubated overnight in a humidified 37°C atmosphere (5% CO_2). On the day of infection, 100 µl DMEM-grown bacteria was added to the coverslips in 1 ml fresh media. The bacteria were centrifuged onto the cells for 5 minutes (200 x *g*). Cells were washed after 2 hours of infection and at 0.5, 3 or 5 hours-post-inoculation they were washed and fixed in 4% PFA.

For analysis of cell lysates following infections, cells were either seeded and infected as above or scaled-up up by seeding in a 6-well plate format (seeding density 1.5 x 10^5 cells per well in 3 ml media, 48 hour seed). For infections in the 6-well format, 300 µl primed DMEM-grown bacteria was added to the coverslips in 3 ml fresh media. Cells were washed in PBS and lysed in 500 µl (for 24-well format) or 200-300 µl (for 6-well format) PBS with 1% Triton X-1000 supplemented with 25 µM PR619, 25 U benzonase and 1 protease inhibitor cocktail tablet per 10 ml lysis buffer for 10 minutes on ice and cells

were scraped and collected. Lysates were cleared by centrifugation (14000 rpm in microcentrifuge at 4°C) and 5X SDS loading buffer was added to lysates. Samples were boiled for 5 minutes and analysed by SDS-PAGE and western blot. In some cases, the lysis buffer was substituted to 1X RIPA buffer supplemented with appropriate inhibitors.

2.5. Immunofluorescent staining

Cells were washed in PBS and quenched with 50 mM ammonium chloride in PBS for 10 minutes prior to permeabilisation using Triton-X100 (0.2% v/v in PBS; Sigma) for 4 minutes. A 5-minute incubation in BSA (0.2% w/v; Sigma) was used to saturate the cells prior to the addition of the primary antibody (listed in **Table 2.4**; diluted in PBS-BSA) and 60-minute incubation at room temperature in a humidified chamber. Cells were washed in PBS and re-saturated before the addition of the secondary antibody (**Table 2.4**) and a 45-minute incubation in a dark humid environment. DNA was counterstained with Hoechst 33342 (1 in 1000 dilution) and actin detected using AlexaFluor647 or AlexaFluor633; all added together with secondary antibodies. Coverslips were washed again in PBS and in water, mounted using Prolong Gold Antifade Reagent (Invitrogen) and analysed either using a Zeiss Axio Imager M1 microscope (Carl Zeiss Microscopy) or a Zeiss AxioVision Z3 microscope (Carl Zeiss Microscopy). Image processing was achieved using Zen 2.3 (Blue version) software (Carl Zeiss Microscopy) or FiJi software (Schindelin *et al.*, 2012).

2.6. TEM translocation assay

The T3SS translocation assay was performed using the TEM-1 β-lactamase kit (Invitrogen) as previously described (Charpentier & Oswald, 2004; Petty *et al.*, 2010). Briefly, 6 x 10^4 Swiss or HeLa cells were seeded in a black-wall clear-bottom 96-well plate and incubated until 100% confluent. *C. rodentium* strains were primed as above and 30 µl was added per well. Infection was carried out in a 37°C 5% CO_2 environment for two hours, cells were then washed and expression of TEM-1 fusion proteins induced by adding 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) in DMEM for 90 minutes. When the proteasome inhibitor MG-132 (Sigma) was required, it was added to the medium (final concentration 10 µM) 30 minutes before IPTG. Cells were washed with Hanks' buffered salt solution (HBSS) and incubated for 90 minutes at room temperature in HBSS-20 mM HEPES (4-(2-hydroxyethyl)-1- piperazineethanesulphonic acid; Sigma) buffer supplemented with 20 µl TEM-1 substrate CFF2/AM (Invitrogen). Finally, cells

were washed thrice in HBSS and the fluorescence measured using a FLUOstar Optima plate reader. If the substrate has not been cleaved by β-lactamase, it will emit green light upon 410 nm excitation (emission detected via 520 nm filter, 10 nm band-pass). If the substrate has been cleaved, it will emit blue fluorescence (emission detected via 450 nm filter). Translocation can thus be quantified and expressed as the 450/520 emission ratio. The assay produces a higher ratio in HeLa cells compared to Swiss cells. HeLa cells were therefore preferred if differences in response ratios are subtle.

2.7. Paraformaldehyde fixation of cultured cells

Washed cells on coverslips were fixed with 300 µl 4% PFA (Agar Scientific) in PBS for 15 minutes at room temperature. Fixed cells were then washed quickly three times in 500 µl PBS and stored at 4°C until required.

2.8. Methanol fixation of cultured cells

For methanol fixation, washed cells on coverslips were fixed with 500 µl 100% ice cold methanol at -20°C for 20 minutes, followed by three 5-minute washes in PBS and immediate immunostaining.

Table 2.4 | Antibodies used in this study. Preference for BSA or milk in blocking solution indicated when important.

Primary	Species	Dilution	Source
α-ubiquitin clone FK2	Mouse	1 in 100 (IF) 1 in 1000 (WB, milk)	Enzo Life Sciences
α-GAPDH	Rabbit	1 in 2000 (WB)	Abcam
α-giantin	Rabbit	1 in 750 (IF)	Abcam
α-GFP	Mouse	1 in 200 (WB)	Abcam
α-GST (clone 3G10/1B3)	Mouse	1 in 2000 (WB)	Abcam
α-*Citrobacter*	Rabbit	1 in 200 (IF)	Sigma
α-DnaK (clone 8E2/2)	Mouse	1 in 10000 (WB)	Enzo Life Sciences
α-HA.11 (clone 16B12)	Mouse	1 in 1000 (IF)	Cambridge Bioscience
α-IκBα	Rabbit	1 in 1000 (WB, BSA)	Cell Signaling Technologies
α-ISG15	Rabbit	1in 1000 (WB, BSA)	Cell Signaling Technologies
α-Myc	Mouse	1 in 200 (IF)	Millipore
α-NEDD8	Rabbit	1 in 1000 (WB)	Cell Signaling Technologies
α-TEM	Mouse	1 in 1000 (WB)	QED Biosciences
α-SUMO1	Rabbit	1 in 1000 (WB)	Cell Signaling Technologies
α-SUMO2/3	Rabbit	1 in 1000 (WB)	Cell Signaling Technologies
α-TRIM25	Rabbit	1 in 2500 (IF)	Sigma
α-Tubulin	Mouse	1 in 2000 (WB)	Sigma
α-C1QBP (clone D7H12)	Rabbit	1:1000 (WB, milk)	Cell Signaling Technologies
α-UBE2O	Rabbit	1:1000 (WB)	Thermo Fischer Scientific
α-intimin280β	Chicken	1:50 (IF)	Mundy *et al.*, 2007
α-Ki67 (clone SP6)	Rabbit	1:20 (IF)	Thermo Fischer Scientific
Secondary	**Species**	**Dilution**	**Source**
α-mouse IgG (H+L) AF-488	Donkey	1 in 200 (IF)	Jackson Immunoresearch
α-mouse IgG (H+L) RRX	Donkey	1 in 200 (IF)	Jackson Immunoresearch
α-chicken IgY Cy™3	Goat	1 in 200 (IF)	Jackson Immunoresearch
α-rabbit IgG (H+L) AF-488	Donkey	1 in 200 (IF)	Jackson Immunoresearch
α-rabbit IgG (H+L) RRX	Donkey	1 in 200 (IF)	Jackson Immunoresearch
α-mouse IgG (Fc) HRP	Goat	1 in 10000 (WB)	Jackson Immunoresearch
α-rabbit IgG (Fcγ) HRP	Goat	1 in 25000 (WB)	Jackson Immunoresearch
Conjugate	**Species**	**Dilution**	**Source**
α -FLAG-HRP	Mouse	1 in 1000 (WB)	Sigma
α -HA-TRITC	Mouse	1 in 100 (IF)	Sigma
α -HA-HRP	Rabbit	1 in 4000 (WB, BSA)	Sigma
α -His-HRP	Mouse	1 in 5000 (WB)	Sigma
α-ubiquitin(K48-linked)-HRP	Mouse	1 in 2000 (WB)	Cell Signaling Technologies
α-ubiquitin(K63-linked)-HRP	Mouse	1 in 2000 (WB)	Cell Signaling Technologies
α-ubiquitin(P4D1)-HRP	Mouse	1 in 1000 (WB)	Santa Cruz Laboratories
α -Myc-FITC	Mouse	1 in 100 (IF)	Sigma

IF = immunofluorescence, WB = Western blot, AF = Alexa Fluor®

2.9. Animal experiments

2.9.1. Mouse husbandry

All animal experiments were conducted in compliance with the Animals (Scientific Procedures) Act 1986 following approval by the local Ethical Review Committee to meet the UK Government Home Office guidelines for animal welfare. Pathogen-free female C57Bl/6 mice (18-20 g; purchased from Charles River Ltd., UK) were housed in individually ventilated HEPA-filtered cages (containing sterile bedding and nesting) with sterilized food and water *ad libitum*. Independent single infections experiments were performed with 4-6 mice per group. Mice were terminated by cervical dislocation. Animal Research: Reporting of *In Vivo* Experiments (ARRIVE)-guidelines for executing and reporting animal experiments were followed (Kilkenny *et al.*, 2010).

2.9.2. Oral infection of mice

Mice were infected via oral gavage of 200 µl *C. rodentium* grown overnight in 15 ml LB and re-suspended in 1.5 ml PBS to obtain an inoculum of approximately 5×10^9 CFUs. The number of viable bacteria was also assessed retrospectively by plating serial dilutions of the inoculum onto LB agar plates with antibiotics (see **Table 2.1**). Strains were tested in parallel to WT- and mock-infected mice. Following infection, stool samples were recovered at regular intervals and the viable number of bacteria per gram of stool determined by plating onto selective LB agar plates. Mice were culled at day 8 post-inoculation and colonic tissues were collected for further analyses.

2.9.3. Lipocalin 2 ELISAs

The concentration of Lipocalin 2 (Lcn2) per gram stool was determined by Enzyme-linked Immunosorbent Assay (ELISA). To this end, stool samples were homogenised in PBS with 0.1% v/v Tween-20, centrifuged at 12000 rpm for 10 minutes and the supernatant stored at -80°C. The Lcn2 concentrations in the supernatant was then determined using the DuoSet Mouse Lipocalin-2/NGAL ELISA kit (R&D Systems) in accordance with the manufacturer's instructions. Stool supernatants from uninfected mice were diluted 1 in 50 in assay diluent, while stool supernatants from infected mice were diluted 1 in 1000 (day 4 post-infection), 1 in 1500 up to 1 in 3000 (day 6 post-infection) or 1 in 2000 to 1 in 4000 (day 8 post-infection). 100 µl of diluted sample was added per well of a 96-well ELISA plate in technical duplicate.

2.9.4. Enterocyte isolation

Enterocytes were isolated from the distal 4 cm of mouse colonic tissue. Following tissue harvest and rinsing of contents, colons were incubated at 37°C (shaking at 200 rpm) for 40 minutes in 5 ml freshly prepared enterocyte isolation buffer composed of 10 mM HEPES, 1 mM EDTA (Life Technologies) and 0.5% v/v β-mercaptoethanol in HBSS lacking $MgCl_2$ and $CaCl_2$; Thermo Fischer Scientific) to lift the enterocytes from the tissue. The tissue was removed and the lifted enterocytes centrifuged at 4°C for 10 minutes at 2000 x *g*. Following two washes in PBS (in 10 ml and 1 ml, respectively), the pellet was stored at -80°C until digestion and tandem mass tag (TMT) labelling or until enterocyte protein purification for analysis by Western blot.

2.9.5. Enterocyte protein extraction

Isolated enterocytes were lysed in 180 µl lysis buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA pH 8.0, 1% v/v NP-40, 6.25 mM NaPPi, 1 mM NaF (Sigma), 1 mM Na_3VO_4 (Sigma), 50 mM DUB-inhibitor PR-619 (Sigma) and protease inhibitor (Roche; 1 tablet per 10 ml buffer). Next, 20 µl 20% SDS was added for a final SDS concentration of 2%. Samples were passed through 0.8 ml centrifuge columns (Pierce) by centrifuging at 14000 rpm for 5 minutes and a BCA assay was used to determine the resulting protein concentrations (see **section 2.9.3**). Samples were stored at -80°C until analysis by Western blot.

2.9.6. Colonic crypt measurements.

The most distal 0.5 cm of colon was harvested and fixed in 10% v/v buffered formalin solution (Sigma-Aldrich) prior to sectioning and staining with haematoxylin and eosin (H&E) using standard techniques (Crepin *et al.*, 2016). Colonic crypt hyperplasia was determined by measuring the lengths of all well-oriented colonic crypts from every mouse.

2.9.7. Immunofluorescent staining of tissue sections

Paraffin-embedded sections were also subjected to immunofluorescence analysis, performed following paraffin removal, heat-induced epitope retrieval and staining. Briefly, paraffin was removed by two 10-minute incubations in Histoclear (National Diagnostics), followed by rehydration in 100% ethanol (two 10-minute washes), 95% ethanol (two three-minute washes), 80% ethanol (3-minute wash) and two three-minute washes in PBS supplemented with 0.1% v/v Tween-20 and 0.1% w/v saponin (Sigma) (PBS-TS). Epitope retrieval was then achieved by placing slides in a canopic jar filled with demasking solution (10 mM trisodium citrate, 0.05% v/v

Tween-20, pH 6.0) and boiling for 20 minutes in a rice cooker (Breville). Following cooling, slides were washed in PBS-TS and blocked for 20 minutes in a humidified chamber using PBS-TS supplemented with 10% v/v normal donkey serum (PBS-TS-NDS; 60 µl per slide). Samples were treated for 1 hour with primary antibodies (listed in **Table 2.4**) diluted in PBS-TS-NDS (60 µl per slide), washed twice in PBS-TS (10 minutes per wash), treated for 1 hour with secondary antibodies (listed in **Table 2.4**) and Hoescht 33342 (1 in 1000) in PBS-TS-NDS (60 µl per slide) and washed twice as before. All antibody incubations were performed in a humidified chamber. Slides were then dipped in distilled water and mounted using Prolong Gold Antifade Reagent. Immunofluorescent images were acquired using a Zeiss Axio Vision Z3 microscope and processed using Zen 2.3 software (Blue).

2.9.8. Tissue processing for immunophenotyping

For analysis of immune cells in the intestine, a 4 cm segment of the distal colon was opened longitudinally and rinsed of content. The tissue was placed in 1 ml RPMI-1640 medium (Sigma-Aldrich) supplemented with 10% FCS and penicillin-streptomycin (P/S, Sigma-Aldrich) and prepared for flow cytometry. Another 3 ml RMPI-1640 supplemented with liberase (1:37.5 dilution, Roche) and DNaseI (1:2000 dilution, Promega) was added per colon and the tissue was mechanically dissociated using a Gentle MACS instrument prior to a 30-minute shaking incubation at 37°C. 40 µl 0.5 M EDTA was then added to limit enzymatic activity. A single cell suspension was achieved by passage through a 70 µm nylon sieve. Cells were re-suspended in 800 µl RPMI-1640 with 10% FCS and P/S.

2.9.9. Multicolour flow cytometry

For surface marker staining, 2-4 × 10^6 cells were pelleted onto V-bottomed 96-well plates (2000 rpm, 2 minutes, 4°C) and washed in PBS. Dead cell exclusion was implemented by incubating cells in 25 µl Live/Dead Fixable Near IR stain (1:1000 in PBS; Life Technologies; 30 minute 4°C-incubation). Cells were washed and incubated for 10 minutes at 4°C in Fc receptor (FcR) blocking reagent (1:10 dilution in PBS; Miltenuyi Biotec) to minimize non-specific antibody binding. They were washed in PBA (PBS supplemented with 1% BSA and 0.01% sodium azide) and stained with relevant antibodies (see **Table 2.5**; diluted in PBA) for 30 minutes at 4°C. Stained cells were fixed using Fixation Buffer (Biolegend), washed and re-suspended in PBA. Stained cells were run on a FACS Fortessa III instrument (BD Biosciences). Flow cytometric data was analysed using FlowJo software (Tree Star Inc.).

Table 2.5| Antibodies used for flow cytometric analyses in this study.

Marker	Fluorochrome	Dilution	Clone	Source
B220 (CD45R)	PE	1:200	RA3-682	eBioscience
CD3e	PE-Cy7	1:200	145-2C11	eBioscience
CD4	PerCP-Cy5.5	1:200	RM4-5	eBioscience
CD8	FITC	1:100	53-6.7	eBioscience
CD11b	PerCP-Cy5.5	1:400	M1/70	eBioscience
CD11c	PE-Cy7	1:200	N418	BD Biosciences
F4/80	APC	1:200	BM8	eBioscience
γδT	FITC	1:100	eBioGL3	eBioscience
Ly6G	FITC	1:100	1A8	BD Biosciences
NK1.1	APC	1:200	PK136	eBioscience

2.10. Protein analysis

2.10.1. SDS-PAGE

Prior to protein analysis, samples were denatured at 95°C for 5 minutes in Laemmli buffer (60 mM Tris-HCL pH 6.8, 1% SDS, 5% v/v glycerol, 5% v/v β-mercaptoethanol and 0.01% w/v bromophenol blue). For ubiquitination samples, denaturing was performed at 70°C for 10 minutes. Samples were loaded onto protein gels where the resolving gel portions were cast using 8-15% acrylamide in 0.375 M Tris-HCl pH 8.8, 0.1% SDS) and stacking gel portions were cast using 4% acrylamide in 0.125 M Tris-HCl pH 6.8, 0.1% SDS. For both resolving and stacking gels, 0.1% (w/v) ammonium persfulfate and 0.1% (v/v) TEMED were added to induce acrylamide polymerisation. Samples were separated by polyacrylamide gel electrophoresis (SDS-PAGE) at 140V in Tris-Glycine-SDS buffer (GeneFlow). The molecular weights of bands were estimated by comparison to a pre-stained broad range protein standard (NEB). For mass spectrometry work involving in gel-trypsinisation, pre-cast graduated gels (4-20% acrylamide) were used (BioRad). Gels were either Coomassie-stained using Instant*Blue* Protein Stain (Expedion) or transferred to Polyvinylidene fluoride (PVDF) membranes (GE Healthcare) for Western blotting.

2.10.2. Western blotting

PVDF membranes were activated in methanol for 1 minute prior to gel transfer. For autoubiquitination experiments, wet transfers of protein onto membranes were performed for 90 minutes at 400 mA in Tris-Glycine buffer (GeneFlow). For other experiments, semi-dry transfers were performed using the Trans-blot Turbo Blotting System (Bio-Rad) in accordance with the manufacturer's instructions. Membranes were blocked overnight in TBS or PBS with

0.1% Tween-20 (TBS-T or PBS-T) with 5% w/v skimmed milk (Merck) or 3-5% or w/v BSA. Membranes were probed with primary antibodies (listed with appropriate dilutions in **Table 2.4**) for one hour at room temperature or overnight at 4°C in TBS-T or PBS-T with 3% BSA or PBS-T 5% milk. Following three 10-minute washes in PBS-T, membranes were incubated in HRP-conjugated secondary antibodies (see **Table 2.4**) in PBS-T or TBST-T with 3% BSA 5% milk for 1 hour. Membranes were washed again three times and ECL plus reagent (GE Healthcre) was added as per manufacturer's instructions. Chemiluminescence was detected using a LAS-3000 Fuji Imager.

2.10.3. Protein BCA assay

Protein concentrations were determined using the Pierce^{TM} BCA Protein Assay Kit (ThermoFischer Scientific) as recommended by the manufacturer.

2.10.4. Protein expression and purification

6His-tagged protein expression and purification was performed by Dr Olga Kotik. To produce 6His-tagged NleG proteins or 6His-tagged C1QBP for functional assays, *E. coli* BL21 (Star) strains containing pETM11-*nleG* constructs were grown at 37°C (200 rpm, in 100 ml LB with appropriate antibiotics) until an OD_{600} of 0.6-0.8. Protein expression was induced with 1 mM IPTG and incubated at 20°C overnight (at 4°C, with 200 rpm agitation). Expression was tested by sampling of cultures, SDS-PAGE and Coomassie analysis. Cells expressing the relevant 6X-His-tagged proteins were then harvested by centrifugation at 4550 x *g* for 15 minutes at 4°C and stored at -80°C until lysis. Cells were lysed by sonication on ice in 15 ml lysis buffer consisting of 20 mM Tris-HCL pH 8.0 with 250 mM NaCl. Soluble and insoluble fractions were separated by centrifugation at 29000 x *g* for 20 minutes at 4°C. The soluble fractions were applied to pre-equilibrated TALON His-Tag Purification Resin and purified per the manufacturer's recommendations (GE Healthcare). Briefly, resins were incubated with the soluble fractions for one hour at 4°C with end-over-end rotation before being applied to gravity-flow columns. Unbound proteins were washed with 20 volumes of lysis buffer. Bound fractions were eluted from the resin with 1 ml 100 mM imidazole in lysis buffer. Protein concentrations were determined using a NanoDrop 1000 (Thermo Fischer Scientific).

When TEV-protease cleavage of the N-terminal tag was desired, the purified protein was incubated with TEV protease (1:100 w/w ratio of protease to protein) in TEV buffer (20 mM Tris-

HCl pH 8.0, 100 mM NaCl and 5 mM β-mercaptoethanol) for one hour at room temperature then re-applied to TALON resin as before to bind the 6His-tagged TEV protease. The unbound fraction was analysed for full cleavage by SDS-PAGE and Coomassie staining.

To express Glutathione S-Transferase (GST)-tagged NleG proteins, pGEX-6P-2-NleG vectors were transformed in *E. coli* BL21 pLysS and *E. coli* BL21 Star. 5 ml LB overnight cultures were diluted 1 in 100 in 50 ml LB with appropriate antibiotics and grown at 37°C with 200 rpm agitation until an OD_{600} of 0.4-0.6. A pre-induction sample (1 ml) was collected (bacteria pelleted by centrifugation and resuspended in 5X SDS loading buffer. To induce protein expression, 1 mM IPTG was added to each culture and proteins were expressed either at 30°C for 3 hours, 37°C for 3 hours or 18°C overnight (always with 200 rpm agitation). The OD_{600} was measured and two 1 ml samples were collected and bacteria pelleted as above. The first sample was resuspended in 5X SDS loading buffer. The second sample was resuspended in 1 ml GST lysis buffer (50 mM Tris pH 8, 2 mM EDTA, 10% v/v glycerol, 150 mM sodium chloride and 5 mM DTT with 1 tablet cOmplete Mini protease inhibitor cocktail per 10 ml buffer). Bacteria were then lysed by sonication on ice and centrifuged to separate soluble and insoluble fractions before being resuspended in 5X SDS loading buffer according to the OD_{600}. Samples were then analysed by SDS-PAGE followed by Coomassie staining or western blot to assess expression of GST-tagged NleG proteins.

2.10.5. Ubiquitination assays

Cell-free ubiquitination assays to test for E3 ligase activity were performed using a protocol adapted from the E2 Scan Kit (Ubiquigent 2013). Reactions (40 µl total volume in 50 mM HEPES, 5 mM 1.4-dithiothreitol (DTT; Sigma) and 5M $MgCl_2$ assay buffer) were assembled on ice up with 0.1 µM 6-His-tagged Ube1 (Ubiquigent), 0.05 nM UBE2D3 (Ubiquigent), 100 µM ubiquitin (Ubiquigent) and 1 µM putative E3 ligase with or without 1 mM ATP (Thermo Fischer Scientific). After a 1-hour incubation at 30°C, the reaction was stopped by adding 5X SDS loading buffer and boiling for 5 min at 95°C. Samples were then separated by SDS-PAGE and transferred to PVDF membranes as previously described and probed for ubiquitin by western blot.

2.10.6. Polyubiquitinated protein enrichment assays

For polyubiquitinated protein enrichment assays, HCT116 cells were seeded at a density of 5 x 10^6 cells per 10 cm diameter dish (with or without 1 µg ml^{-1} doxycycline) and grown for 24 hours to induce expression HA, HA-NleG1 or HA-$NleG1_{P163K}$. If transient transfection of FLAG-UBE2O or FLAG-GFP was also required, the cells were seeded at 2.5 x 10^6 cells per 10 cm diameter dish and

grown for 24 hours prior to transfection (as described in **Section 2.2.11**) for 24 hours. When indicated, the medium was supplemented with 10 µM MG132 5.5 hours prior to lysis and with 5 µM PR-619 4 hours prior to lysis. Cells were washed in 2 ml cold PBS and lysed on ice in 500 µl RIPA buffer supplemented with protease inhibitor cocktail, 50 µM PR-619 and 10 mM *N*-Ethylmaleimide (NEM; Sigma). BCA assays were performed on cleared lysates to determine protein concentrations and samples were diluted to 1000 µg protein per ml lysis buffer. 40 µl samples were collected for input controls and the rest was added to polyubiquitin affinity beads (Calbiochem) or negative control beads (Calbiochem). Samples were incubated with the beads for 4 hours at 4°C with end-over-end rotation. Beads were pelleted by centrifugation (3000 rpm, 1 minute, 4°C) and 40 µl were collected for unbound controls. Beads were then washed 4 times in 1 ml lysis buffer. Samples were boiled for 5 minutes in 50 µl 5X SDS Loading Buffer to elute peptides. Beads were pelleted by centrifugation (13000 rpm, 1 minute) and the supernatants collected, separated by SDS-PAGE and analysed by western blot.

2.10.7. FLAG fusion protein immunoprecipitation assays

For FLAG immunoprecipitation (IP) assays, HCT116 cells were seeded at 2.5 x 10^6 cells per 10 cm diameter dish with 1 µg ml^{-1} doxycycline to induce expression of the transgene (HA, HA-NleG1 or HA-$NleG1_{P163K}$) and grown for 24 hours before 24-hour transfection of FLAG-tagged constructs (FLAG-UBE2O or FLAG-GFP). Cells were washed twice with 2 ml ice cold PBS and subsequently lysed with 500 µl ice cold RIPA buffer supplemented with a protease inhibitor cocktail tablet (1 tablet per 10 ml buffer) and 50 µM PR-619 DUB inhibitor. Lysates were cleared by centrifugation (13000 rpm for 10 minutes at 4°C) and the buffer was exchanged to TBS pH 7.4 using Slide-A-Lyzer® 3.5K Dialysis Cassettes (Thermo Scientific). 40 µl was retained as an input control and the remainder was incubated at 4°C with 40 µl ANTI-FLAG® M2 Affinity Gel (Sigma; equilibrated by 3 washes in 1 ml TBS) for 2 hours with end-over-end rotation. Beads were pelleted by centrifugation (3000 rpm, 1 minute, 4°C) and 40 µl supernatant was collected as an unbound control. Beads were washed 3 times in TBS and incubated at 4°C for 30 minutes in 100 µl TBS with 150 ng $µl^{-1}$ 3X FLAG peptide (Sigma) with gentle shaking to elute bound peptides. Beads were pelleted by centrifugation (as above) and the supernatants were collected using narrow-ended pipette tips and separated by SDS-PAGE for analysis by western blot.

2.10.8. HA fusion protein immunoprecipitation assays

For HA immunoprecipitation assays, HCT116 cells were seeded at a density of 1 x 10^7 cells per 15 cm diameter dish (with or without 1 µg ml^{-1} doxycycline to induce expression of the transgenes

HA, HA-NleG1 or HA-NleG1$_{P163K}$) and grown for 24 hours. Cells were washed with 5 ml ice cold PBS and subsequently lysed with 1.5 ml ice cold RIPA buffer supplemented with a protease inhibitor cocktail tablet (1 per 10 ml buffer) and if required 50 µM PR-619 DUB inhibitor. Lysates were cleared by centrifugation (13000 rpm for 10 minutes at 4°C). Of the collected supernatants, 25 µl were retained as an input control and the remainder was incubated at 4°C with monoclonal anti-HA agarose beads (Sigma; equilibrated by 3 washes in 1 ml lysis buffer) for 2 hours (for analysis by western blot only) or overnight (for analysis by mass spectrometry) with end-over-end rotation. Beads were pelleted by centrifugation (3000 rpm, 1 minute, 4°C) and 25 µl supernatant was collected as an unbound control. Beads were washed 3 times in lysis buffer and incubated at room temperature for 30 minutes in 40 µl PBS with 250 µg ml^{-1} purified HA peptide (Sigma) with 1000 rpm agitation to elute bound peptides. Beads were pelleted by centrifugation (as above) and the supernatants were collected using narrow-ended pipette tips. Immunoprecipitated proteins were separated by SDS-PAGE and analysed by western blot or mass spectrometry.

2.10.9. Processing of IP samples for mass spectrometry

For label-free mass spectrometry (MS) interactome analysis, IP eluates were subjected to in-gel trypsin digest. First, the proteins from IP eluate were separated as technical triplicates by Tris-glycine SDS–PAGE (4-20% gradient gels; BioRad) and each lane was cut into four pieces. Each quarter-lane piece was further cut into 1 mm^3 pieces to allow processing in a 96-well plate (1 well per quarter-lane). Pieces were covered in sterile water (200 µl per well) and stored at 4°C overnight.

On the next day, pieces were washed three times in 50 mM ammonium bicarbonate solution with 50% ethanol (100 µl per well) for 15 minutes at room temperature. Unless otherwise indicated, solutions were carefully removed using a multichannel pipette and discarded between all steps. Pieces were dehydrated by two consecutive 5-minute incubations in absolute ethanol (100 µl per well) and dried by SpeedVac for 5-10 minutes. Next, proteins were reduced with 10 mM DTT in 50 mM ammonium bicarbonate solution (100 µl per well) for 45 minutes at 56°C and then alkylated with 55 mM iodoacetamide (Sigma) in 50 mM ammonium bicarbonate solution (100 µl per well) for 30 minutes at room temperature, protected from light by aluminium foil. Pieces were then washed with 50 mM ammonium bicarbonate solution (100 µl per well) for 15 minutes at room temperature before being dehydrated in absolute ethanol (100 µl per well) for 15 minutes at room temperature. This washing-dehydration was repeated

twice, followed by an additional 15-minute ethanol dehydration step at room temperature and SpeedVac for 5-10 minutes to ensure pieces were dry.

Proteins were then digested with 12 ng μl^{-1} trypsin (Sigma) in 50 mM ammonium bicarbonate solution (40 µl per well), first for 15 minutes at 4°C to rehydrate the gel pieces and then overnight at 37°C. The resulting peptides were sequentially extracted from gel pieces and during the extraction steps, solutions were not discarded between steps but sequentially collected into a clean 96-well plate. Pieces were incubated first in 30% acetonitrile in 3% trifluoroacetic acid (TFA, 100 µl per well) for 20 minutes at room temperature, then twice in 70% acetonitrile (100 µl per well) for 20 minutes at room temperature and twice in 100% acetonitrile (100 µl per well) for 20 minutes at room temperature. The collected peptide samples (approximately 600 µl per well) were transferred to clean Eppendorfs and the volume reduced to 80 µl by SpeedVac (30°C) for around 2 hours. During this time, C18 resin stage tips were conditioned. 30 µl methanol was added to each stage tip and stage tips were centrifuged at 1000 rpm for 20-30 seconds to leave 5 mm of liquid over the resin. This was repeated using an elution buffer (80% acetonitrile and 0.1% formic acid in sterile water), and then twice with a wash buffer (0.1% formic acid in sterile water). 80 µl of a binding buffer (containing 5% acetonitrile and 1% TFA) was added to each peptide sample (for final concentrations of 2.5% acetonitrile and 0.5% TFA) and added to the conditioned stage tips for desalting (centrifuged at 2600 rpm for 5 minutes). Following the addition of peptides to the resin, the resin was washed once more with 30 µl wash buffer (centrifuged at 2600 rpm for 2 minutes). A syringe was used blow air through the resin to ensure it was dry, and stage tips were then stored at 4°C until elution.

To elute the peptides, 30 µl elution buffer was added per stage tip. Stage tips were centrifuged for 40s at 1000 rpm and incubated for 15 minutes at room temperature to soak the resin. Samples were eluted into a 96-well plate by centrifugation for 5 minutes at 1500 rpm. The volume was reduced to 2-3 µl by Speedvac (30°C) for 10-12 minutes and topped up to 8 µl with wash buffer. Thus, de-salted eluted IP-captured peptides were reconstituted in 0.5% acetonitrile and 0.1% TFA. Samples were stored at -20°C protected from light until analysis by MS.

2.10.10. Label free mass spectrometry

Label-free mass spectrometry was kindly performed by Dr Florian Bonn in the Ivan Đikić Laboratory (Institute of Biochemistry II, Goethe University Frankfurt - Medical Faculty,

University Hospital, Frankfurt am Main, Germany). The peptides were separated using a gradient of 5%–33% acetonitrile in 0.5% acetic acid over 226 min and detected on an Orbitrap Elite mass spectrometer (Thermo Scientific) as previously described (Bhogaraju *et al.*, 2016).

2.10.11. Enterocyte digestion and TMT labelling for MS

Enterocytes (extracted as described in **Section 2.10.4**) were kindly digested and TMT labelled for MS by Lu Yu and Kate Hadavizadeh in the Jyoti Choudhary Laboratory (Functional Proteomics Group, Chester Beatty Laboratories, Institute of Cancer Research, London, United Kingdom). Enterocytes were dissolved in 100 µl 0.1 M triethylammonium bicarbonate (TEAB) solution with 0.1% SDS and lysed by pulse probe sonication. For each sample, 80 µg protein was dried by SpeedVac. Next, the proteins were resuspended in 100 µl 0.1 M TEAB solution with 4% SDS and 15 mM tris(2-carboxymethyl)phosphine (TCEP) in an ultrasonic bath. Samples were reduced for 20 minutes at 56°C, cooled to 25°C and then alkylated with iodoacetamide (IAA) for 30 minutes. Proteins were then purified by precipitation in 20% trichloroacetic acid precipitation, washed in cold acetone, resuspended in 0.1 M TEAB and digested with 3 µg MS-grade trypsin (Pierce) for 18 hours at 37°C. Of the digested protein, 40 µg were labelled with 0.4 mg TMT10plex following the manufacturer's recommendations (Thermo Fischer Scientific).

2.10.12. Enterocyte proteome MS experiments

Samples were fractionated with a U3000 high-performance liquid chromatography (LC) system (Thermo Fisher Scientific) coupled to an XBridge ethylene-bridged hybrid C_{18} column (Waters). Peptides were injected onto the Orbitrap Fusion Tribrid mass spectrometer coupled to a U3000 RSLCnano ultrahigh-performance LC system (Thermo Fisher Scientific) loading onto a PepMap C_{18} trap and separated on a PepMap C_{18} column. Data acquisition was done by the SPS10-MS3 method. Further analysis was performed using Perseus (version 1.5.6.9) and only peptides with a FDR-corrected *P* value <0.01 were retained (Pollard *et al*., 2018).

2.10.13. Analysis of mass spectrometry data

MS data was processed with Perseus (version 1.5.6.9). Absolute label free quantification (LFQ) intensities were logartimized (log_2) and peptides identified by one or more one unique peptide sequence were retained. The false-discovery rate (FDR) was set to 0.05 and at least two valid values across the test technical repeats were required for further comparisons. Theoretical LFQ values were then imputed for remaining missing values (width set to 0.3; down-shift set to 1.8).

Ratios between test and control groups were created and significant enrichment proteins (over 1.5-fold changed) as assessed by an FRD-corrected *P* value>0.05 were analysed further. The data was exported to Excel to make volcano plots.

2.11. Bioinformatics

Basic Local Alignment Search Tool (BLAST) searches were performed using standard settings (Boratyn *et al.*, 2013) to search nucleotide or protein sequences from the National Centre for Biotechnology Information (NCBI) website (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence protein alignments were generated Clustal Omega tool with standard settings, accessible online (Sievers *et al.*, 2011). The Kyoto Encyclopaedia of Genes and Genomes (KEGG) was also used to retrieve DNA sequences for *Salmonella* genes (Kanehisa *et al.*, 2016).

2.12. Statistical analyses

Statistical analysis was performed using GraphPad Prism 5.0 (GraphPad Software Inc.) and data shown represent means ± standard error of the mean (SEM) for all groups of interest unless otherwise indicated. Where appropriate, analysis of variance (ANOVA) followed by Tukey's multiple comparison post-test or Bonferroni's multiple comparison post-test was used as indicated to determine statistical significance. Difference denoted as significant fall below the p-value 0.05.

Chapter 3:
In vitro characterisation of the C. rodentium NleG proteins

3. In vitro characterisation of the C. rodentium NleG proteins

3.1. Introduction

Defining the activity of T3SS effector proteins is vital for deciphering A/E pathogenesis. This chapter focuses on characterising the NleG family of T3SS effector proteins *in vitro*: describing the family bioinformatically, testing *C. rodentium* NleG localisation within the host cell, defining their enzymatic activities and attempting to identify interacting partners that they target during A/E infection.

3.2. NleG proteins are conserved across pathogenic bacteria

BLAST analyses (Boratyn *et al.*, 2013) and previous published reports (Deng *et al.*, 2004; Tobe *et al.*, 2006; Wu *et al.*, 2010; Fookes *et al.*, 2011) identifies 24 NleG homologs (including pseudogenes) across A/E pathogens and *Salmonella* spp., with 14 homologs in EHEC alone and 5 homologs in *C. rodentium.* In *C. rodentium,* two of the NleGs are predicted to be pseudogenes. To begin examining the NleG family of bacterial effector proteins, a multiple sequence alignment and sequence comparison of the NleG protein homologs from EHEC, EPEC, *C. rodentium*, *S. bongori* and *S. enterica* was performed (**Figure 3.1**). This confirmed that NleG family proteins indeed vary in both length and sequence, but they all share a conserved C-terminus of approximately 100 amino acids that was shown in EHEC NleG2.3 to comprise an E3 ubiquitin ligase domain (Wu *et al.*, 2010). The N-termini of NleG family proteins display lower conservation than the C-terminus, with amino acid sequence identities as low as 2.86% (EHEC NleG2.2 compared to *S.* Paratyphi B StoD, see **Figure 3.3C**). The C-termini of NleG proteins, on the other hand, are more homologous (see **Figure 3.3B**), with homologs from different bacterial strains being up to 80% identical (*C. rodentium* NleG7 compared to EHEC NleG7). While *C. rodentium* NleG7 and NleG8 are only 36.78% identical in their N-termini, they are 51.52% identical in their C-termini. Likewise, *C. rodentium* NleG1 is 42.71% identical to *S.* Typhi StoD, despite an N-terminal identity of only 10.11%.

The C-terminus of NleG proteins encompasses the RING/U-box domain of E3 ligases, with core residues conserved between all NleG homologs (highlighted in **Figure 3.1**), including three residues (V165, L167 and P204) previously shown to be essential for the interaction between the EHEC NleG2.3 homolog and the E2 ligase enzyme UBE2D3 (Wu *et al.*, 2010; indicated by arrows in **Figure 3.1**). Furthermore, *C. rodentium* NleG8 and EHEC NleG7 and NleG8.1 have a predicted PDZ-binding motif in their C-termini (highlighted in purple in **Figure 3.1**). Such motifs mediate

binding to proteins containing PDZ domains that are highly abundant in eukaryotic proteins (Lee & Zheng, 2010).

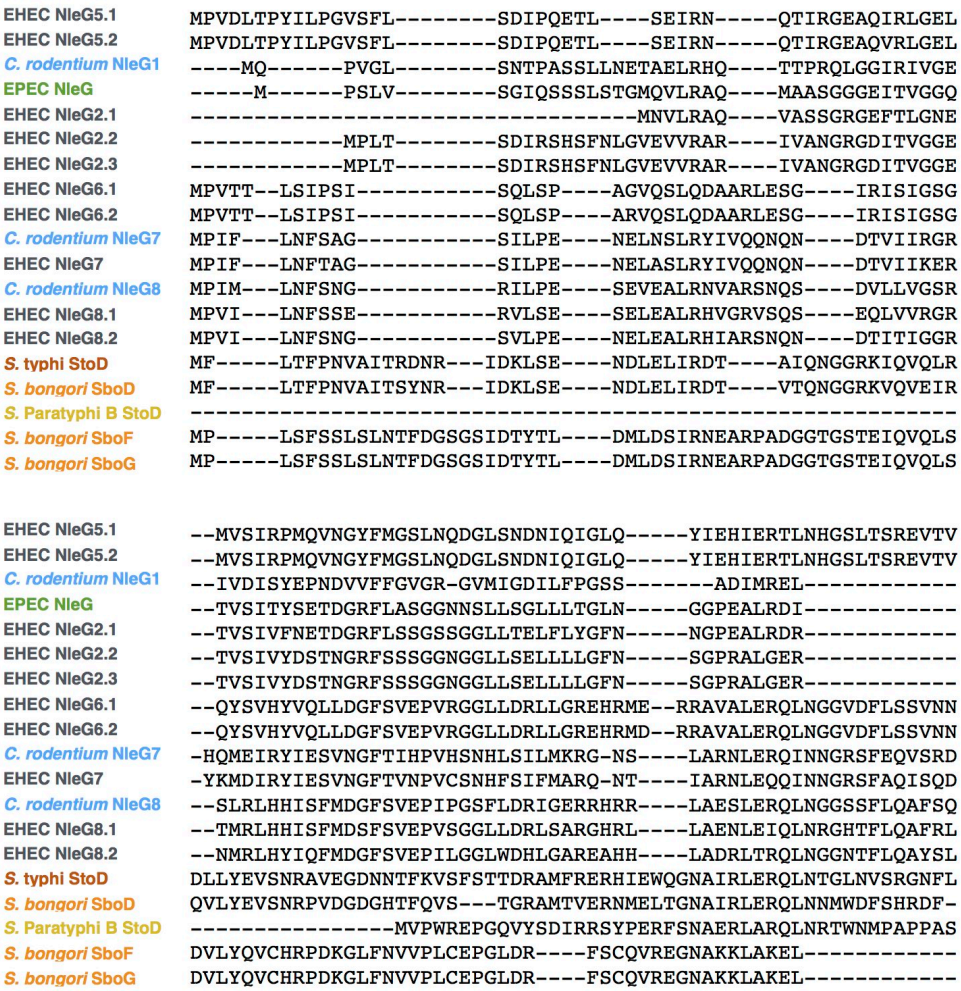
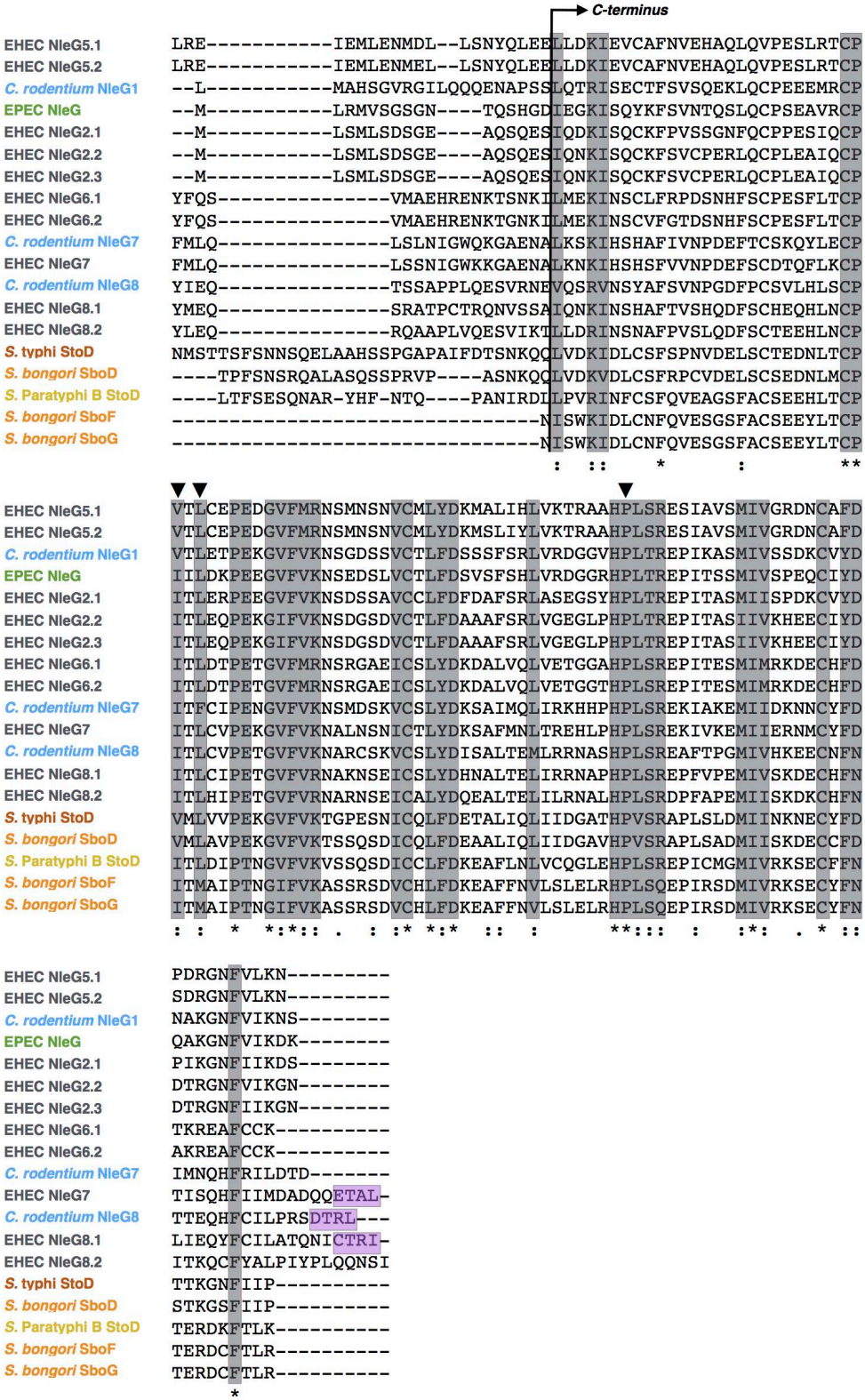


Figure 3.1 (continued over to the next page) | Amino acid sequence alignment of NleG family proteins from EHEC, EPEC, *Citrobacter rodentium* and *Salmonella* spp. NleG homologs have a conserved C-terminal domain that structurally resembles an E3 ligase and residues known to be structurally important for the RING-type domain are shaded in grey. The start of the C-terminus is indicated by the black arrow. Arrowheads indicate the residues found to be required for the interaction between EHEC NleG2.3 and the E2 enzyme UBE2D2 in an auto-ubiquitination assay (Wu *et al.*, 2010). Asterisks indicate positions that have a single, fully conserved residue. Colons indicate conservation of amino acids with highly similar properties, while full stops indicate conservation of amino acids with weakly similar properties. Predicted PDZ-binding motifs are shaded in purple. Alignment performed in Clustal Omega (Sievers *et al.*, 2011).


C-terminus
EHEC NleG5.1
EHEC NleG5.2
C. rodentium NleG1
EPEC NleG
EHEC NleG2.1
EHEC NleG2.2
EHEC NleG2.3
EHEC NleG6.1
EHEC NleG6.2
C. rodentium NleG7
EHEC NleG7
C. rodentium NleG8
EHEC NleG8.1
EHEC NleG8.2
S. typhi StoD
S. bongori SboD
S. Paratyphi B StoD
S. bongori SboF
S. bongori SboG

Overall, NleG lengths range from 163 amino acids (*S*. Paratyphi B StoD) to 233 amino acids (*S*. typhi StoD). *S. bongori* SboF and SboG are 100% identical (also on the DNA level)(Fookes *et al.*, 2011). Similarly, EHEC NleG5.1 and NleG5.2 are 97.65% identical and NleG2.2 and NleG2.3 are 99.46% identical on the amino acid level (see **Figure 3.3A**). The lowest overall conservation within the NleG family with 23.20% amino acid identity is between *C. rodentium* NleG1 and *S. bongori* SboF/SboG. A cladogram constructed based on these percentages revealed that the NleG homologs cluster together phylogenetically across the strains (**Figure 3.2**). Indeed, *C. rodentium* NleG7 is 74.16% identical to EHEC NleG7, while only being 43.75% identical to *C. rodentium* NleG8 or 29.17% identical to *C. rodentium* NleG1 (**Figure 3.3A**).

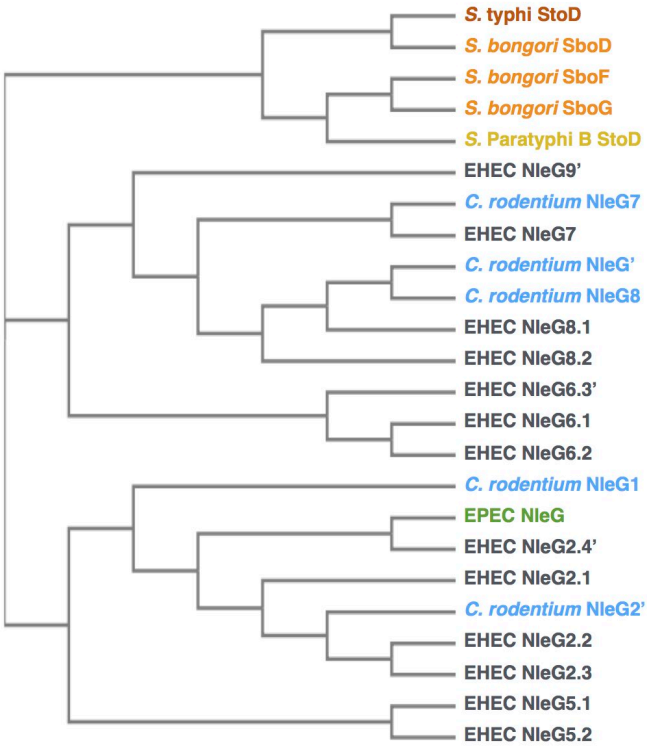


Figure 3.2 | NleG family proteins cluster together phylogenetically. Pseudogenes are denoted by '. Cladogram created in Clustal Omega (Sievers *et al.*, 2011) based on amino acid identity of *C. rodentium* NleG proteins (shown in blue), EPEC NleG (in green), EHEC NleG proteins (in grey), *S. enterica* serovar Typhi StoD (in red), *S*. Paratyphi B StoD (in yellow) and *S. bongori* SboD, SboF and SboG (in orange).

A **Full length NleG family member sequence comparison (percentage amino acid identity)**

	StoD	SboD	SboFG	ParaStoD	CRNleG1	EPECNleG	NleG2.2	NleG2.3	NleG5.1	NleG5.2	NleG6.1	NleG6.2	CRNleG7	NleG7	CRNleG8	NleG8.2	NleG8.1
StoD	**100**	73.61	32.43	31.65	25	23.53	25.27	25.81	28.36	27.86	31.07	30.58	29.56	32.51	29.56	31.03	27.09
SboD	73.61	**100**	32.43	33.77	27.66	24.06	26.88	27.42	28.43	28.43	32.52	31.55	29.56	31.03	29.06	33	29.06
SboFG	32.43	32.43	**100**	50.68	23.2	27.32	28.02	28.02	29.41	28.88	27.81	27.27	27.81	25.13	24.19	23.66	25.81
ParaStoD	31.65	33.77	50.68	**100**	29.68	33.33	34.64	34.64	31.37	30.72	36.94	36.31	31.21	31.85	32.48	33.76	32.48
CRNleG1	25	27.66	23.2	29.68	**100**	44.92	44.09	43.55	33.51	33.51	29.32	29.32	29.17	29.69	25	30.73	28.12
EPECNleG	23.53	24.06	27.32	33.33	44.92	**100**	59.16	58.64	31.58	32.11	31.58	31.05	29.69	28.65	27.6	27.08	28.65
NleG2.2	25.27	26.88	28.02	34.64	44.09	59.16	**100**	99.48	34.92	34.92	33.33	32.8	28.8	29.84	31.41	28.8	30.89
NleG2.3	25.81	27.42	28.02	34.64	43.55	58.64	99.48	**100**	34.39	34.39	33.33	32.8	28.8	30.37	31.41	28.8	30.89
NleG5.1	28.36	28.43	29.41	31.37	33.51	31.58	34.92	34.39	**100**	97.65	34.87	34.87	30.41	28.87	25.77	29.38	28.87
NleG5.2	27.86	28.43	28.88	30.72	33.51	32.11	34.92	34.39	97.65	**100**	33.85	33.85	29.38	27.84	25.77	28.35	28.87
NleG6.1	31.07	32.52	27.81	36.94	29.32	31.58	33.33	33.33	34.87	33.85	**100**	96.17	33.17	32.2	38.83	43.2	42.72
NleG6.2	30.58	31.55	27.27	36.31	29.32	31.05	32.8	32.8	34.87	33.85	96.17	**100**	33.17	31.71	38.35	43.2	42.72
CRNleG7	29.56	29.56	27.81	31.21	29.17	29.69	28.8	28.8	30.41	29.38	33.17	33.17	**100**	74.16	43.75	43.27	43.75
NleG7	32.51	31.03	25.13	31.85	29.69	28.65	29.84	30.37	28.87	27.84	32.2	31.71	74.16	**100**	40.57	39.25	40.65
CRNleG8	29.56	29.06	24.19	32.48	25	27.6	31.41	31.41	25.77	25.77	38.83	38.35	43.75	40.57	**100**	56.81	59.15
NleG8.2	31.03	33	23.66	33.76	30.73	27.08	28.8	28.8	29.38	28.35	43.2	43.2	43.27	39.25	56.81	**100**	61.4
NleG8.1	27.09	29.06	25.81	32.48	28.12	28.65	30.89	30.89	28.87	28.87	42.72	42.72	43.75	40.65	59.15	61.4	**100**

B **NleG family C-terminal sequence comparison (percentage amino acid identity)**

	NleG2.2	NleG2.3	CRNleG1	EPECNleG	StoD	SboD	ParaStoD	SboFG	NleG5.1	NleG5.2	CRNleG7	NleG7	NleG6.1	NleG6.2	CRNleG8	NleG8.1	NleG8.2
NleG2.2	**100**	98.98	62.24	63.27	39.58	41.67	44.79	42.71	44.33	43.3	37.76	42.86	42.71	42.71	38.78	38.78	31.63
NleG2.3	98.98	**100**	61.22	62.24	40.62	42.71	44.79	42.71	43.3	42.27	37.76	43.88	42.71	42.71	38.78	38.78	31.63
CRNleG1	62.24	61.22	**100**	66.33	42.71	44.79	43.75	35.42	48.45	49.48	38.78	39.8	42.71	42.71	36.73	38.78	38.78
EPECNleG	63.27	62.24	66.33	**100**	39.58	37.5	41.67	38.54	43.3	43.3	39.8	39.8	40.62	40.62	37.76	37.76	33.67
StoD	39.58	40.62	42.71	39.58	**100**	81.25	41.67	38.54	41.67	40.62	42.71	48.96	43.75	42.71	38.54	40.62	43.75
SboD	41.67	42.71	44.79	37.5	81.25	**100**	43.75	39.58	39.58	39.58	40.62	42.71	43.75	41.67	37.5	41.67	43.75
ParaStoD	44.79	44.79	43.75	41.67	41.67	43.75	**100**	68.75	42.71	41.67	42.71	44.79	51.04	50	44.79	44.79	45.83
SboFG	42.71	42.71	35.42	38.54	38.54	39.58	68.75	**100**	37.5	36.46	42.71	39.58	41.67	40.62	39.58	39.58	37.5
NleG5.1	44.33	43.3	48.45	43.3	41.67	39.58	42.71	37.5	**100**	96.91	43.3	41.24	52.08	51.04	36.08	43.3	42.27
NleG5.2	43.3	42.27	49.48	43.3	40.62	39.58	41.67	36.46	96.91	**100**	42.27	40.21	51.04	50	35.05	42.27	41.24
CRNleG7	37.76	37.76	38.78	39.8	42.71	40.62	42.71	42.71	43.3	42.27	**100**	67.68	43.75	43.75	51.52	52.53	46.46
NleG7	42.86	43.88	39.8	39.8	48.96	42.71	44.79	39.58	41.24	40.21	67.68	**100**	44.79	43.75	45.63	51.43	41.9
NleG6.1	42.71	42.71	42.71	40.62	43.75	43.75	51.04	41.67	52.08	51.04	43.75	44.79	**100**	94.79	43.75	52.08	48.96
NleG6.2	42.71	42.71	42.71	40.62	42.71	41.67	50	40.62	51.04	50	43.75	43.75	94.79	**100**	42.71	52.08	48.96
CRNleG8	38.78	38.78	36.73	37.76	38.54	37.5	44.79	39.58	36.08	35.05	51.52	45.63	43.75	42.71	**100**	60.19	54.37
NleG8.1	38.78	38.78	38.78	37.76	40.62	41.67	44.79	39.58	43.3	42.27	52.53	51.43	52.08	52.08	60.19	**100**	66.67
NleG8.2	31.63	31.63	38.78	33.67	43.75	43.75	45.83	37.5	42.27	41.24	46.46	41.9	48.96	48.96	54.37	66.67	**100**

C **NleG family N-terminal sequence comparison (percentage amino acid identity)**

	ParaStoD	SboFG	StoD	SboD	NleG5.1	NleG5.2	NleG6.1	NleG6.2	CRNleG7	NleG7	NleG8.1	CRNleG8	NleG8.2	CRNleG1	EPECNleG	NleG2.2	NleG2.3
ParaStoD	**100**	39.29	23.88	25.4	10.87	10.87	19.23	17.31	14.29	12.24	18	20	18	5.56	5.71	2.86	2.86
SboFG	39.29	**100**	31.82	28.24	16.05	14.81	24.05	24.05	14.47	17.11	25	23.68	19.74	10.77	18.18	10.77	10.77
StoD	23.88	31.82	**100**	66.4	19.05	18.1	24.3	22.43	21.36	18.45	22.12	19.23	22.12	10.11	15.91	12.64	12.64
SboD	25.4	28.24	66.4	**100**	13.4	14.43	27.27	25.25	24.21	21.05	21.88	17.71	20.83	13.58	12.5	10.13	10.13
NleG5.1	10.87	16.05	19.05	13.4	**100**	98.28	16.04	16.04	16.5	13.59	17.31	15.38	19.23	25	26.37	23.33	23.33
NleG5.2	10.87	14.81	18.1	14.43	98.28	**100**	15.09	15.09	17.48	14.56	17.31	15.38	18.27	23.86	25.27	22.22	22.22
NleG6.1	19.23	24.05	24.3	27.27	16.04	15.09	**100**	97.35	23.85	21.1	34.55	33.64	38.18	14.13	24.47	24.73	24.73
NleG6.2	17.31	24.05	22.43	25.25	16.04	15.09	97.35	**100**	23.85	21.1	34.55	33.64	38.18	14.13	23.4	23.66	23.66
CRNleG7	14.29	14.47	21.36	24.21	16.5	17.48	23.85	23.85	**100**	80	35.78	36.7	40.37	21.98	21.51	18.48	18.48
NleG7	12.24	17.11	18.45	21.05	13.59	14.56	21.1	21.1	80	**100**	30.28	33.94	36.7	21.98	18.28	15.22	15.22
NleG8.1	18	25	22.12	21.88	17.31	17.31	34.55	34.55	35.78	30.28	**100**	56.36	55.45	16.48	21.51	22.83	22.83
CRNleG8	20	23.68	19.23	17.71	15.38	15.38	33.64	33.64	36.7	33.94	56.36	**100**	59.09	13.19	17.2	17.39	17.39
NleG8.2	18	19.74	22.12	20.83	19.23	18.27	38.18	38.18	40.37	36.7	55.45	59.09	**100**	18.68	20.43	20.65	20.65
CRNleG1	5.56	10.77	10.11	13.58	25	23.86	14.13	14.13	21.98	21.98	16.48	13.19	18.68	**100**	24.18	22.22	22.22
EPECNleG	5.71	18.18	15.91	12.5	26.37	25.27	24.47	23.4	21.51	18.28	21.51	17.2	20.43	24.18	**100**	53.76	53.76
NleG2.2	2.86	10.77	12.64	10.13	23.33	22.22	24.73	23.66	18.48	15.22	22.83	17.39	20.65	22.22	53.76	**100**	100
NleG2.3	2.86	10.77	12.64	10.13	23.33	22.22	24.73	23.66	18.48	15.22	22.83	17.39	20.65	22.22	53.76	100	**100**

Figure 3.3 | Amino acid identity matrices for full length NleGs and the respective C- and N-termini. Matrices created in Clustal Omega (Sievers *et al.*, 2011) based on amino acid identity of *C. rodentium* NleG proteins (shown in blue), EPEC NleG (in green), EHEC NleG proteins (in grey), *S.* Typhi StoD (in red), *S.* Paratyphi B StoD (in yellow) and *S. bongori* SboD, SboF and SboG (in orange). **A)** Percentage amino acid identity matrix of full-length NleG family proteins**. B)** Percentage amino acid identity matrix of the more conserved C-terminus of NleG family proteins. **C)** Percentage amino acid identity matrix of the more variable N-terminus of NleG family proteins.

These analyses strengthen the idea that the C-termini of NleG family proteins contain the functional enzymatic domain, and that the N-terminus is much more variable, perhaps to confer specificity in substrate recognition.

3.3. C. rodentium NleG1, NleG7 and NleG8 are translocated via the T3SS

To verify T3SS-dependent translocation of NleG proteins into *C. rodentium*-infected eukaryotic cells, a TEM-translocation assay was employed. In this assay (schematic in **Figure 3.4A**), effector proteins are fused to TEM-1 β-lactamase and expressed in bacteria with a functionally competent or incompetent T3SS. Eukaryotic cells are then infected in the presence of a TEM-1 β-lactamase substrate that generates green fluorescent emission upon 409 nm excitation. If an effector-TEM fusion protein is translocated into the eukaryotic cell, it will cleave the TEM-1 β-lactamase substrate, which disrupts fluorescence resonance energy transfer (FRET) and results in blue fluorescent emission upon 409 nm excitation. By comparing response ratios (emissions at 450/520 nm) in T3SS-expressing and T3SS-null strains, it is possible to determine whether an effector-TEM fusion protein is translocated in a T3SS-dependent manner.

Here, *nleG* genes were fused to the *blaM* gene (encoding TEM-1 β-lactamase) in the IPTG-inducible plasmid pCX340 and transformed into WT and T3SS-null (Δ*espA) C. rodentium.* A known T3SS effector protein (EspH) fused to TEM-1 β-lactamase served as a positive control. The resulting protein constructs with their expected molecular weights are shown in **Figure 3.4B**. Expression of the TEM-1 β-lactamase-fused proteins was assessed by sampling primed cultures of each *C. rodentium* strain (with and without IPTG induction) and analysing them by SDS-PAGE and western blot for TEM-1 β-lactamase. All constructs expressed well both in the wildtype and Δ*espA C. rodentium* backgrounds (**Figure 3.4C**). These strains were therefore deemed suitable for use in a TEM-1 β-lactamase translocation assays.

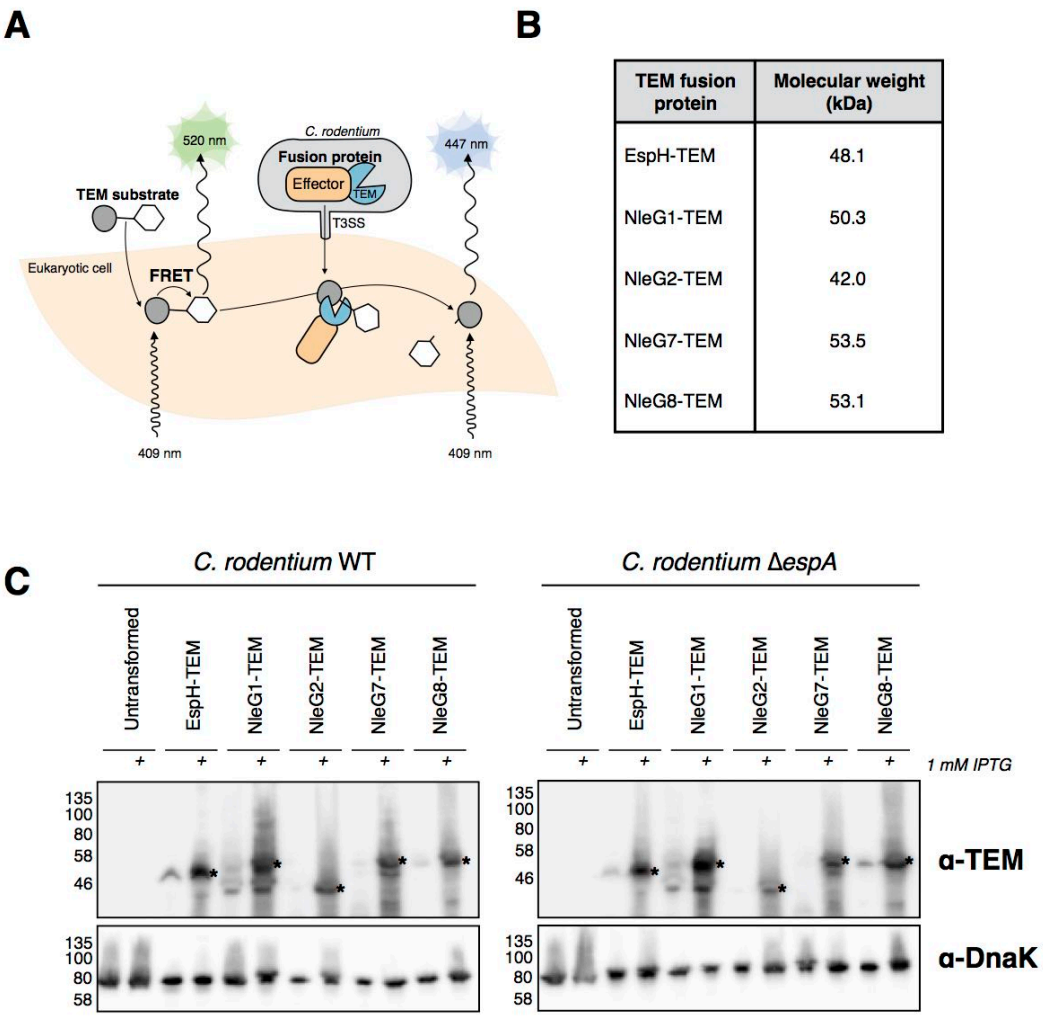


Figure 3.4 | NleG-TEM-1 β-lactamase fusion proteins are expressed in *C. rodentium* under IPTG-inducible conditions. A) Schematic of the TEM-1 β-lactamase T3SS translocation assay. T3SS-translocated TEM-1 β-lactamase effector fusion proteins cleave the TEM-1 β-lactamase substrate, disrupting fluorescence resonance energy transfer (FRET) to give blue rather than green fluorescence emission upon 409 nm excitation. Translocation can be expressed as the 450/520 nm emission ratio. **B)** TEM-1 β-lactamase effector fusion proteins used in this study, with the predicted molecular weight for each construct. **C)** Expression of TEM-1 β-lactamase effector fusion proteins in *C. rodentium* was verified by SDS-PAGE and western blot-analysis of primed *C. rodentium* cultures. An anti TEM-1 antibody was used to detect TEM-1 tagged proteins and an anti-DnaK antibody was used as a loading control. Fusion proteins of the correct molecular weights are indicated by an asterisk.

NleG7 and NleG8 were translocated into Swiss cells via the T3SS, indicated by a significantly higher response ratio when the fusions were expressed in WT- compared to Δ*espA C. rodentium* (**Figure 3.5A**). Translocation of NleG1 via the T3SS could only be observed in HeLa cells (which gives a better readout compared to Swiss cells in this assay) and in the presence on the proteasome inhibitor MG-132 (**Figure 3.5B**), suggesting that NleG1-TEM-1 might be unstable in the host cell or toxic to the host cell and therefore targeted for degradation.

C. rodentium has two additional *nleG* sequences annotated as pesudogenes: *nleG'* and *nleG2'*. The former is considered a pseudogene due to several frame-shift mutations and premature stop codons, while *nleG2'* has an alternative short open reading frame that includes the conserved E3 ligase domain. We used the translocation assay to test if the resulting gene product was translocated into HeLa cells but confirmed that *nleG2'* is a pseudogene (**Fig. 3.5B**). NleG1, NleG7 and NleG8 are henceforth considered the full NleG protein repertoire in *C. rodentium*.

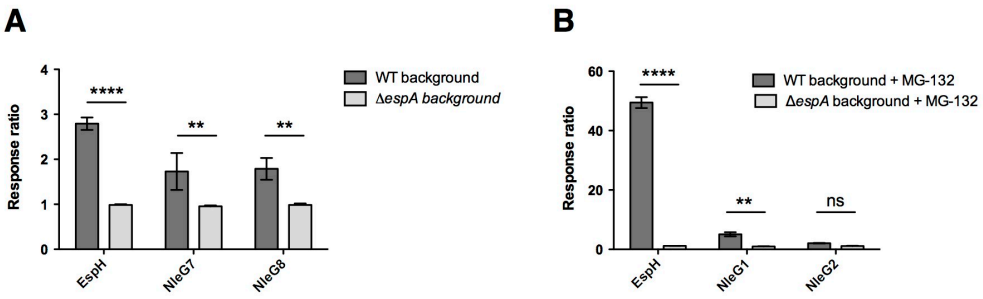


Figure 3.5 | *C. rodentium* NleG effectors are translocated in a T3SS-dependent manner. A) Translocation assay for NleG7-TEM-1 and NleG8-TEM-1 fusions expressed in wildtype and T3SS-null (Δ*espA*) *C. rodentium.* EspH-TEM-1 served as a positive control. Translocation is expressed as the 450/520 nm emission ratio. NleG7 and -8 were translocated into Swiss cells when expressed from the wild-type but not the Δ*espA* background. **B)** Translocation assay for NleG1-TEM1 and NleG2-TEM1 fusions expressed in wildtype and Δ*espA C. rodentium.* The proteasome inhibitor MG-132 was added 30 minutes prior to induction of fusion protein expression to prevent protein degradation. EspH-TEM1 was used as a positive control. NleG1 was translocated into HeLa cells via the T3SS, while NleG2 was not. Error bars represent ±SEM; **p ≤ 0.01, ****p ≤ 0.0001, ns = non-significant.

3.4. Assessing the enzymatic activity of C. rodentium NleG proteins

A common feature of E3 ligases is their ability to self-ubiquitinate, and *in vitro* auto-ubiquitination assays are therefore commonly used to assess whether putative E3 ligases harbour E3 ligase activity (Komander & Rape, 2012). This is particularly useful when there are no known substrates for the E3 ligase. In this section, we sought to verify the predicted E3 ligase enzymatic activity of *C. rodentium* NleG proteins by autoubiquitination assays.

3.4.1. Purification of recombinant NleG protein for in vitro assays

Autoubiquitination assays require purified recombinant putative E3 ligase protein, and the first step was therefore to purify NleG proteins of interest. We set out to purify recombinant WT NleG proteins as well as NleG proteins harbouring point mutations that we predicted would abolish their ability to autoubiquitinate and ubiquitinate their substrates. This would allow us to incorporate catalytically inactive NleGs in future *in vitro* and *in vivo* experiments, which would be hugely beneficial to the NleG project. We decided to mutate a conserved proline reside within the RING-domain of the NleG proteins to lysine, as this specific mutation was previously shown

to completely abolish the interaction between EHEC NleG2.3 and the E2 enzyme UBE2D2, thereby rendering NleG2.3 incapable of autoubiquitination in vitro (Wu *et al*., 2010). Genes for NleG1, NleG1$_{P163K}$, NleG7, NleG8 and NleG8$_{P177K}$ were cloned with a TEV protease-cleavable 6His-tag in the pETM11 inducible bacterial expression vector. Protein expression was tested by sampling pre-induced and IPTG-induced *E. coli* BL21 Star cultures transformed with pETM11-constructs and analysis by SDS-PAGE and Coomassie-staining. The *E. coli* BL21 Star strain was used as it contains a mutation within the RNase E gene, stabilising mRNA produced by the T7 RNA polymerase (Lopez *et al*., 1999; Studier & Moffatt, 1986). The expression test (**Figure 3.6A**) showed that while NleG1 and NleG1$_{P163K}$ expressed well in this system, expression was low for NleG8 and NleG8$_{P177K}$ and undetectable for NleG7.

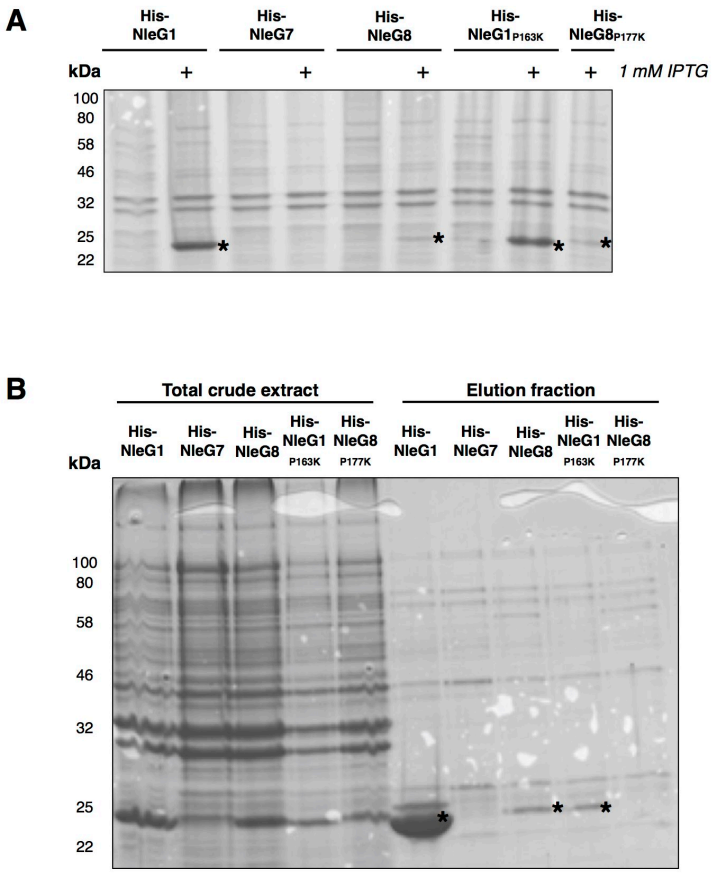


Figure 3.6 | Expression and purification of *C. rodentium* NleG proteins. A) Expression test to assess expression of N-terminally 6His-tagged NleG proteins from a pETM11 vector in *E. coli* BL21 Star. Cultures were grown at 37°C to an OD_{600} of 0.6-0.8. For induced samples, 1 mM IPTG was added and protein expressed overnight at 20°C. Pre-induced and induced samples were compared by SDS-PAGE and Coomassie. Asterisks indicate NleG expression at the predicted molecular weight. There was good expression of NleG1 and NleG1$_{P163K}$, NleG8 and NleG8$_{P177K}$ expressed poorly, and there was no expression of NleG7 in this system. **B)** Purification of NleG proteins using His-TALON affinity chromatography gave soluble His-NleG1 as well as His-NleG1$_{P163K}$ and NleG8. No soluble NleG7 or NleG8$_{P177K}$ was purified by this method. Analysis by SDS-PAGE and Coomassie stain. Asterisks indicate NleG elution at the predicted molecular weight. *These tests were conducted by Dr Olga Kotik at Imperial College London.*

The recombinant proteins were then purified by affinity chromatography using a His-TALON resin, which generated large quantities of soluble recombinant 6His-NleG1, and smaller quantities of both 6His-NleG1$_{P163K}$ and 6His-NleG8 (as analysed by SDS-PAGE and Coomassie staining, **Figure 3.6B**). However, no purified NleG7 or NleG8$_{P177K}$ was recovered using this method. The expression tests and protein purifications were kindly performed by Dr Olga Kotik (Frankel Lab, Imperial College London).

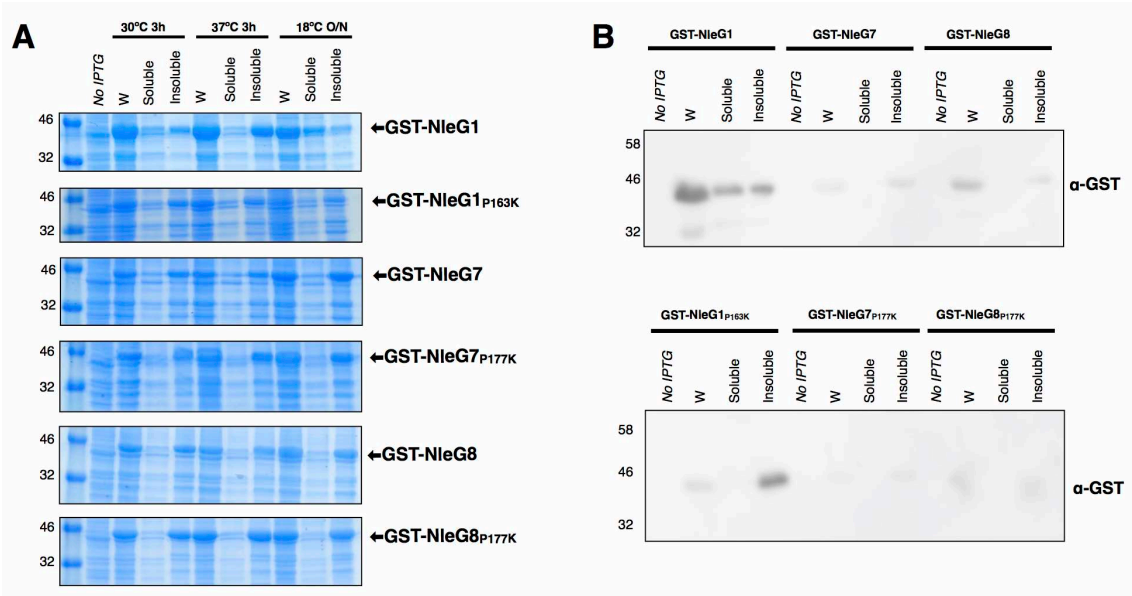


Figure 3.7 | Expression test for GST-tagged NleG proteins in *E. coli*. A) Expression test to assess expression of N-terminally GST-tagged NleGs from the inducible pGEX-6P-2 vector in *E. coli* BL21 Star or pLysS. Cultures were grown at 37°C to OD_{600}=0.4-0.6. To induce expression, 1mM IPTG was added and proteins expressed overnight at 18°C or for 3 hours at 30°C or 37°C. Pre-induced (no IPTG)-, whole cell lysate (W)-, soluble- and insoluble samples were analysed in parallel by SDS-PAGE and Coomassie staining. **B)** Expression of NleGs in selected samples from (A) were also analysed by western blot using an anti-GST antibody. Samples from the 37°C 3h condition were analysed for GST-NleG1$_{P163K}$. For all other NleGs, the 18°C overnight expression samples were analysed.

In an attempt to generate a complete set of purified *C. rodentium* NleG proteins, pGEX-6P-2 vectors were also generated for expressing recombinant GST-NleG proteins. The expression was tested under several conditions (overnight growth at 18°C, 3h at 30°C and 3h at 37°C) in *E. coli* BL21 (**Figure 3.7**). However, NleG7 and NleG7$_{P177K}$ were still not solubly expressed, as seen by Coomassie (**Figure 3.7A**; no bands at the right molecular weight observed for the soluble fraction) and likewise by western blot (**Figure 3.7B**). Moreover, there was not more soluble protein for the other NleGs here compared to the previous expression method, and therefore further *in vitro* assays were performed using the purified 6His-tagged NleG proteins. No further protein work was conducted for NleG7. As a result, NleG7 could not be included in the autoubiquitination assays, discussed next.

3.4.2. C. rodentium NleG proteins have E3 ligase activity

The E3 ligase activity of the purified recombinant NleG1, NleG1$_{P163K}$, NleG8 alongside purified *S*. Typhi StoD (kindly provided by Dr Alexander Byrne; Frankel Lab) was tested by autoubiquitination assay. The putative E3 ligases were combined with an E1 enzyme (UBE1-His), an E2 enzyme (UBE2D3), ubiquitin and ATP. Autoubiquitination was determined by western blot against ubiquitin. In case of autoubiquitination, a sample contained numerous species of the E3-ubiquitin ligase (with different numbers of conjugated ubiquitin moieties) with a range molecular weights, resulting in a smear. Such smears were clearly detected for NleG1, NleG8 and StoD in the presence of ATP (**Figure 3.8**), confirming that they indeed have E3 ligase activity that akin eukaryotic RING-type E3 ligases (but not all bacterial E3 ligases) requires ATP hydrolysis within the ubiquitination cycle. As shown in the representative blot in **Figure 3.8**, separate bands corresponding to different species of ubiquitinated NleG1 (mono-ubiquitinated, di-ubiquitinated., and so on) can be detected. The absence of smearing seen for the control reaction lacking a putative E3 ligase confirmed that the E2 enzyme used was not able to exhibit any E3 ligase activity in the absence of another E3, as can be seen for some E2 ligases (such as UBE2E1). Importantly, no smear was detected for NleG1$_{P163K}$ in the presence or absence of ATP, demonstrating that mutating a residue predicted to be essential for binding of *C. rodentium* NleG1 to an E2 enzyme is sufficient to completely abolish its ability to ubiquitinate itself. Within our research group, this result was replicated for the equivalent StoD mutant (StoD$_{P204K}$). In further assays in this project, NleG proteins containing the proline to lysine mutation are considered catalytically inactive.

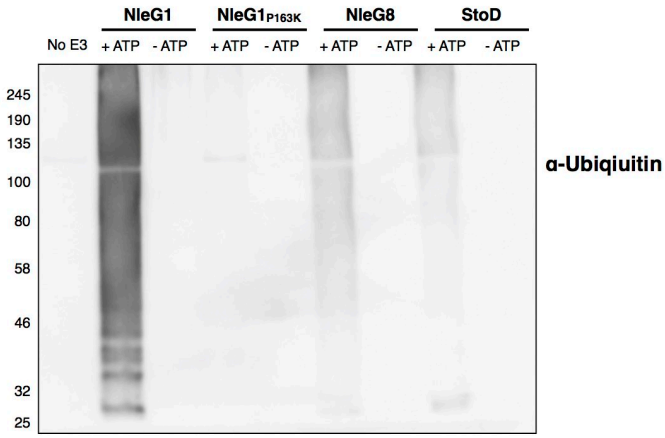


Figure 3.8 | NleG family members are capable of autoubiquitination *in vitro*. Autoubiquitination assays were performed using *C. rodentium* NleG1 [10-194], NleG1$_{P163K}$ [10-194], NleG8 [11-210] and *S*. Typhi StoD [full length]. Reactions were analysed by SDS-page and western blot using an anti-ubiquitin antibody (F2K) to detect conjugated to NleG family members. All reactions contained ubiquitin, an E1 enzyme (UBE1-His) and an E2 enzyme (UBE2D3). ATP was added as indicated in the figure, and the No E3 control contained ATP. Image representative of two independent repeats.

We also tested whether the C-terminus of NleG on its own could autoubiquitinate. The NleG1 C-terminus (comprising amino acids 97-104) was expressed, purified and tested by the *in vitro* auboubiquitination assay as done for the full-length NleG proteins, which revealed a clear ubiquitin smear upon western blotting using the anti-ubiquitin antibody (**Figure 3.9**). An anti-His antibody was also used to confirm that the observed smears consisted of NleG-His proteins of varying molecular weight (clear laddering in **Figure 3.9**). This was also seen for the *S*. Typhi StoD C-terminus (assay performed by Dr Alexander Byrne, data not shown), and our data taken together with published reports of EHEC NleG C-terminal E3 ligase activity strengthen the hypothesis that C-terminal E3 ligase activity is conserved for all NleG family members.

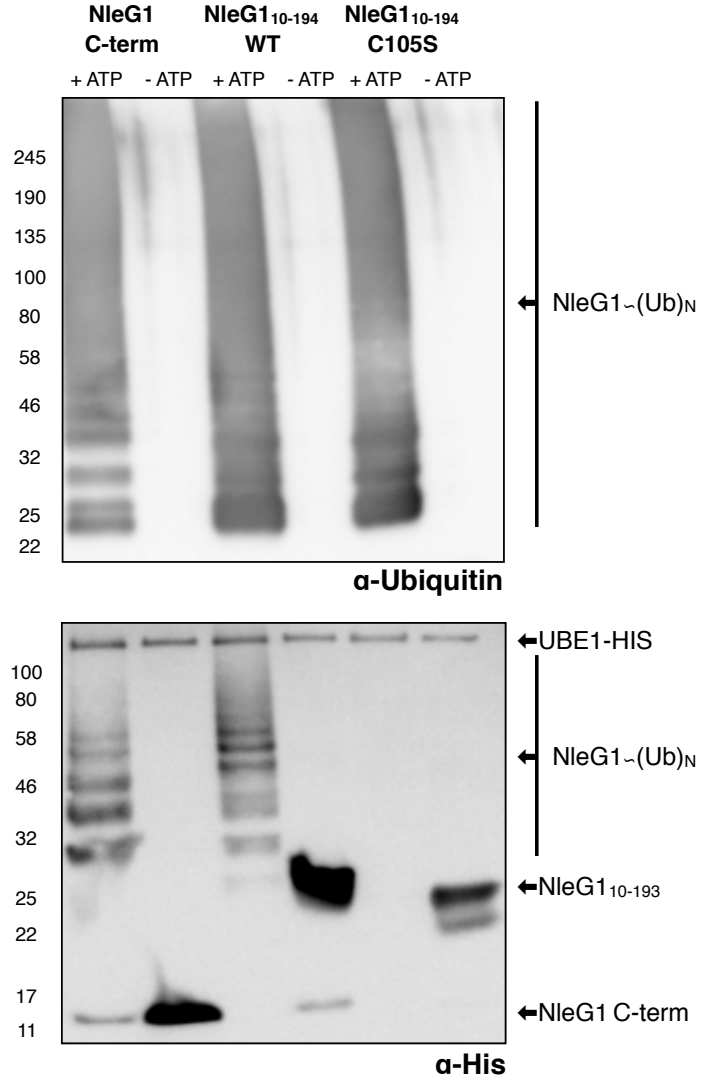


Figure 3.9 | The C-terminus of NleG1 is has E3 ligase enzymatic activity *in vitro*. Autoubiquitination assays were performed using full-length NleG1 [10-194], $NleG1_{C105S}$ [10-194] mutant, and C-terminal NleG1 [97-194]. Reactions were analysed by SDS-page and western blot using an anti-ubiquitin antibody (F2K) to detect conjugated to NleG family members and an anti-His antibody for His-tagged proteins. Reactions contained ubiquitin, UBE1-His and an UBE2D3; ATP as indicated. Molecular weight (kDa) shown on the left-hand side.

Moreover, we mutated a cysteine residue (C105) within the C-terminus of NleG1 to serine, in an attempt to reduce protein aggregation to increase the purity of protein samples for better structural analyses of NleG1. This did not affect its enzymatic activity in the *in vitro* assay (as shown in **Figure 3.9**). In parallel to the work described here, substantial efforts were made by Dr Olga Kotik towards solving the structure of NleG1 (with N-terminal truncations of varying lengths) by X-ray crystallography and NMR spectroscopy. Unfortunately, these efforts will not be discussed further in this thesis as they are beyond the scope of this PhD project.

3.5. The C-terminus of NleG8 mediates yeast cytotoxicity dependent on its E3 ligase activity

To test if *C. rodentium* NleG proteins are cytotoxic in yeast, they were expressed from pYES2 in *S. cerevisiae* BY474 under a glucose-repressible and galactose-inducible promoter. Yeast growth could then be compared between repressible and inducible conditions, revealing whether NleG expression affected yeast survival. The pYES2 method has previously shown that the enzymatic activity of other bacterial effector proteins can confer yeast cytotoxicity (Lesser & Miller, 2001; Arbeloa *et al.*, 2011), giving insight into whether they target crucial cellular pathways. Of importance to this project, EHEC NleG8.1 was previously shown to be cytotoxic in this manner (Dr Bevin McGeever, Frankel Lab). Here, yeast cultures transformed with pYES2-*nleG* constructs were serially diluted and plated on selective media containing either galactose or glucose. As expected, all strains exhibited growth on glucose-containing media, where nothing is expressed from the pYES2 plasmid (**Fig 3.10A**). Strikingly, while yeast expressing *C. rodentium* NleG1 and NleG7 as well as EHEC NleG7 grew on galactose-supplemented media, yeast expressing *C. rodentium* NleG8 and EHEC NleG8.1 did not (**Fig3.10A**, black arrowheads). Disrupting the putative PDZ–binding motif at the C-terminus of NleG8 – a motif that is absent in NleG1 and NleG7 – by mutating it from DTRL to DARA did not reduce cytotoxicity. On the other hand, the cytotoxicity was relieved when the P177K mutation (shown by auto-ubiquitination assays to render NleGs enzymatically inactive) was introduced, confirming that the observed cytotoxicity is dependent on the E3 ligase activity of NleG8 and EHEC NleG8.1.

NleG7 and NleG8 are 44% identical on the amino acid level. The C-terminus, which is conserved within the NleG family, is 51% identical between NleG7 and NleG8, while the N-terminus is 39% identical. The difference in cytotoxicity between NleG7 and NleG8 raised questions about the contribution of the respective N- and C-termini to the observed phenotype. First, we asked if the individual termini mediated cytotoxicity and found that no N- or C-terminus was cytotoxic when

expressed alone (**Fig 3.10B**). Next, we wondered if the cytotoxicity of NleG8 required the N-terminus of NleG8 specifically, as the variable N-termini of NleG family proteins were hypothesised to reflect binding specificity to different cellular targets or targeting to different subcellular localisations. Two chimeric constructs were thus created: one where the NleG7^{N} was fused to NleG8^{C}, the other where the NleG8^{N} was fused to NleG7^{C}. Strikingly, only the former chimera was cytotoxic in yeast (**Fig. 3.10B**, black arrowhead). Collectively, these data suggest that NleG8^{C} mediates yeast cytotoxicity dependent on its E3 ligase activity and requires an NleG^{N} domain, but not specifically the N-terminus of NleG8.

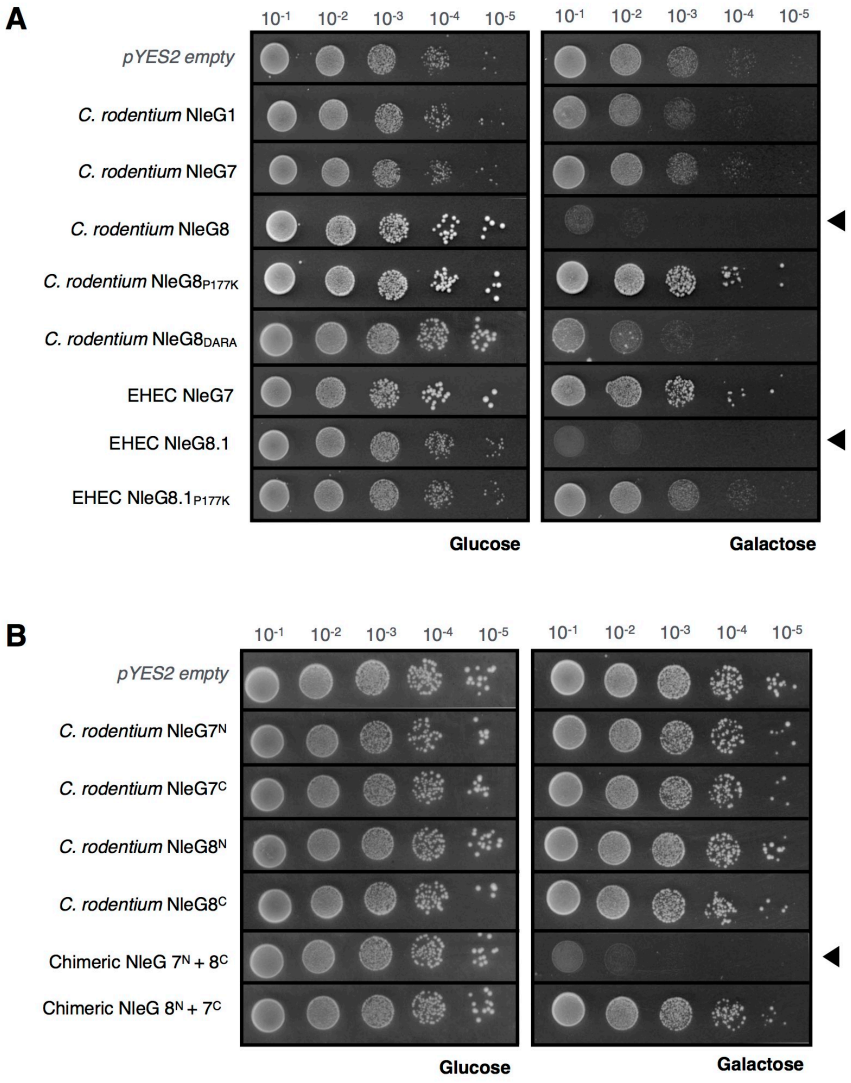


Figure 3.10 | EHEC and *C. rodentium* NleG8 homologs are cytotoxic in yeast dependent on the ability of their C-terminus to bind E2 ligases. pYES2-NleG-constructs were transformed into *S. cerevisiae* BY474. NleG expression was repressed on glucose-containing media and induced on galactose-containing media. Yeast was plated in serial dilutions (10^{-1} to 10^{-5}). **A)** EHEC and *C. rodentium* NleG8 are cytotoxic, indicated by lack of growth galactose-containing medium (black arrows). NleG8$_{P177K}$ was not cytotoxic, but mutating the PDZ-binding motif of NleG8$_{CR}$ (DTRL to DARA) failed to reduce cytotoxicity. NleG1 and NleG7 were not cytotoxic. **B)** The N- and C-termini of *C. rodentium* NleG7 and NleG8, respectively are not cytotoxic. A chimeric protein consisting of NleG7^{N} and NleG8^{C} is cytotoxic (black arrow), while an NleG8^{N} + NleG7^{C} chimera is not.

3.6. The localisation of C. rodentium NleG proteins during transfection of cultured cells

To ascertain the subcellular localisation of *C. rodentium* NleG proteins in eukaryotic cells, epitope-tagged NleGs were expressed in HeLa cells and their localisation determined by immunofluorescence microscopy. The pRK5 transfection vector was used in these experiments to transiently express Myc-NleG proteins alongside a Myc-GFP control in HeLa cells. As shown in **Figure 3.11**, Myc-tagged NleG1 was cytoplasmic and diffuse, while Myc-NleG7 displayed perinuclear vesicular staining for approximately 65% of transfected cells and diffuse staining for remaining cells. In contrast, the majority (circa 80%) of Myc-NleG8-transfected cells displayed punctate staining with large puncta concentrated around or inside the nucleus (examples in **Figure 3.11**). We assessed whether perinuclear staining was indicative of localisation to the Golgi apparatus, but found that the NleGs did not co-localise with the Golgi marker giantin (**Figure 3.11**) or trigger Golgi fragmentation.

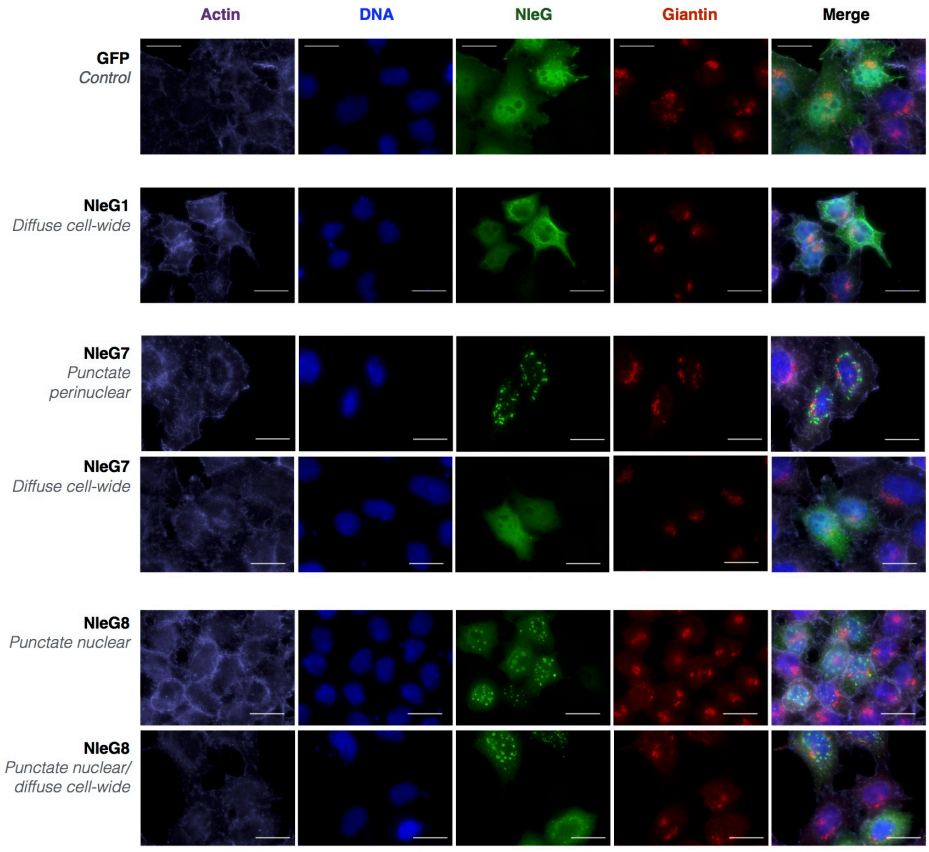


Figure 3.11 | *C. rodentium* NleG proteins localise to specific subcellular compartments during transfection. HeLa cells were transfected with pRK5 constructs expressing Myc-NleG1, Myc-NleG7 or Myc-NleG8. Actin is shown in purple, DNA in blue**.** NleG localisation was assessed by immunofluorescence an anti-Myc antibody (green). NleG1 was diffuse, NleG7 punctate perinuclear or diffuse and NleG8 diffuse or punctate (in nucleus or around the nucleus). NleG did not co-localise with giantin staining (in red). Scale bar 20 µm.

To define the precise subcellular localisation of NleGs during transfection, we stained transfected cells for markers of specific compartments. This revealed that the NleG proteins did not co-localise with the mitochondrial marker TOMM22 (**Figure 3.12A**) or with the early endosome marker EEA1 (**Figure 3.12B**). There was no difference in localisation between wild-type NleG8, the NleG8$_{P177K}$ point mutant incapable of interacting with E2 enzymes and NleG8 ΔDTRL (lacking the predicted C-terminal PDZ-binding motif; **Figure 3.12**).

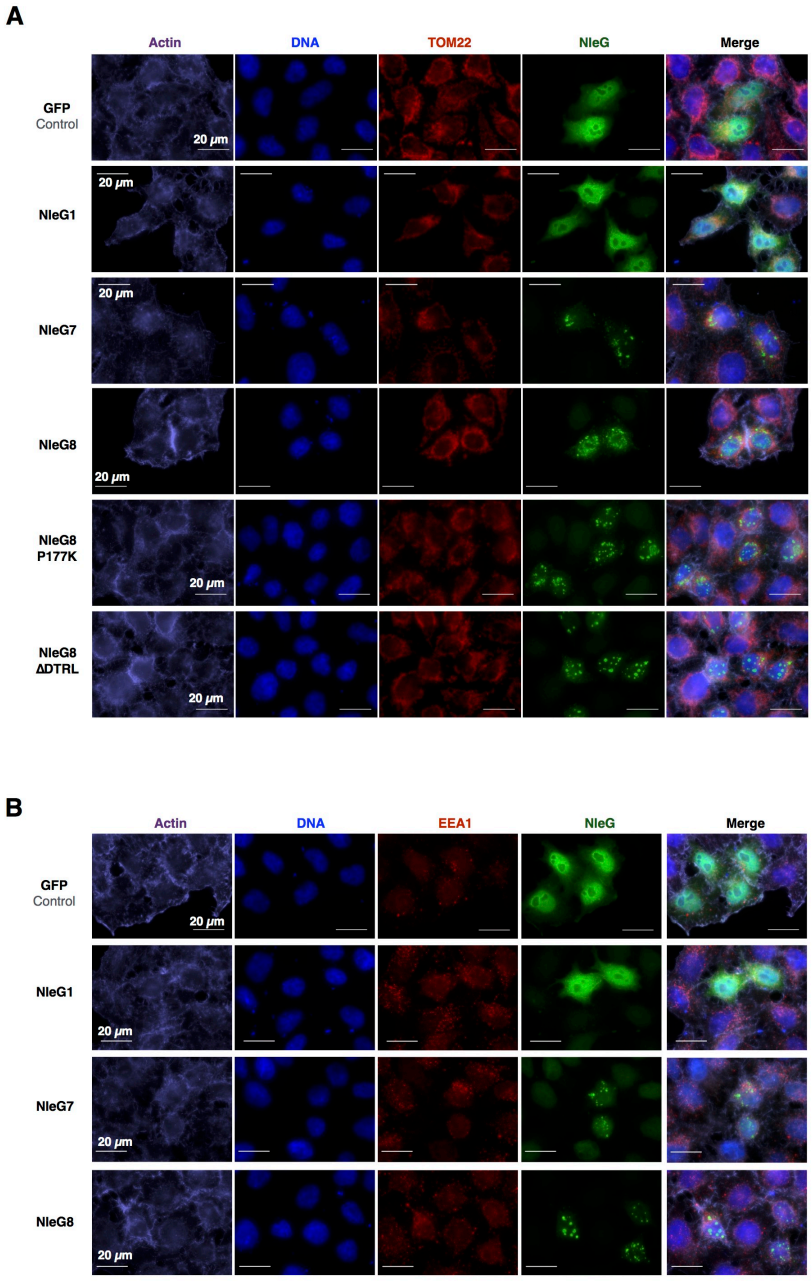


Figure 3.12 | *C. rodentium* NleG proteins do not localise to the mitochondrion or to early endosomes during transfection. HeLa cells were transfected with pRK5 constructs expressing Myc-NleG1, Myc-NleG7, Myc-NleG8 or Myc-GFP. Actin is shown in purple, DNA in blue. NleG localisation was assessed by immunofluorescence using an anti-Myc antibody (green). **A)** NleG-staining did not co-localise with TOMM22 (red; mitochondrial marker). NleG8-localisation was punctate around/in the nucleus for NleG8$_{WT}$, NleG8$_{\Delta DTRL}$ and NleG8$_{P177K}$. **B)** NleG-staining did not co-localise with EEA1 (red; marker of early endosomes). Scale bar 20 μm.

Moreover, to determine if NleG8 localised to the nuclear membrane in Myc-NleG8-transfected cells, we stained the transfected cell for the nuclear membrane marker Lamin A. This revealed that some puncta were clearly outside the nucleus, but appeared closely associated with the nuclear membrane (**Figure 3.13**). Analysis of epi-fluorescence microscopy Z-stacks of transfected cells showed that the puncta were in focus when the Lamin A staining was not, suggesting that they were localised over and under the nucleus.

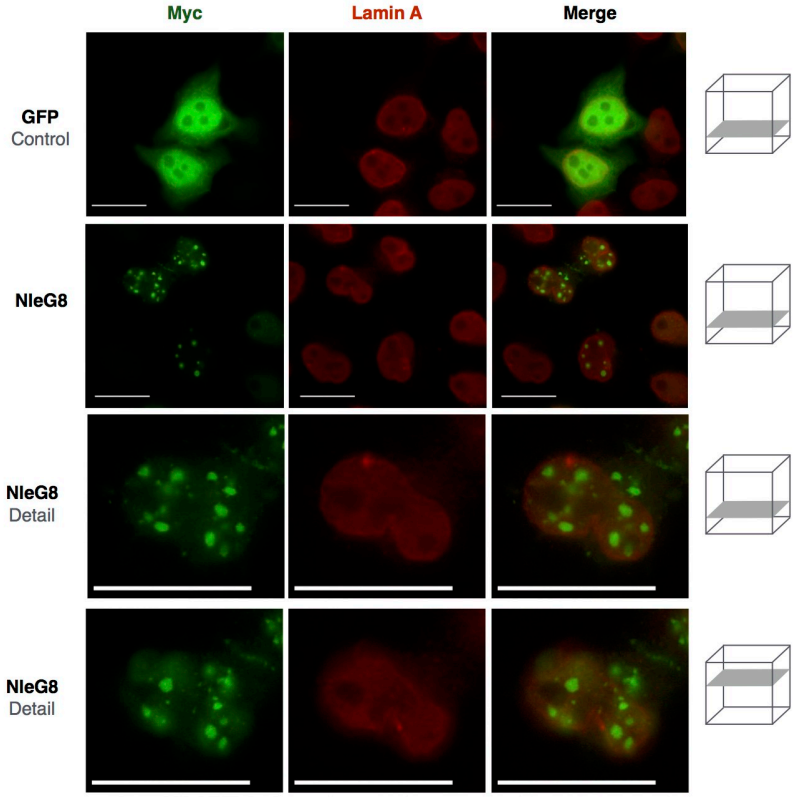


Figure 3.13 | Localisation of *C. rodentium* NleG8 during transfection. HeLa cells were transfected with pRK5 expressing Myc-NleG8 or Myc-GFP. Localisation in relation to the nucleus was assessed using an anti-Myc antibody (green) and Lamin A staining (nuclear membrane marker; red). Two Z-planes represented, scale bar 20 µm.

3.7. Identifying and validating putative interacting partners of NleG proteins by yeast two-hybrid methods

An important clue to understanding the specific function of E3 ubiquitin ligases in eukaryotic cells comes from identifying their physiologically relevant interacting partners. We hypothesised that the variable N-termini of NleG proteins reflect different interacting partners or substrates in the host upon translocation via the T3SS, and we therefore sought to identify such interacting proteins. We performed yeast two-hybrid screens (methodology outlined in **Figure 3.14**), as this

has given promising results for other bacterial E3 ligase proteins in the literature (Lomma *et al.*, 2010).

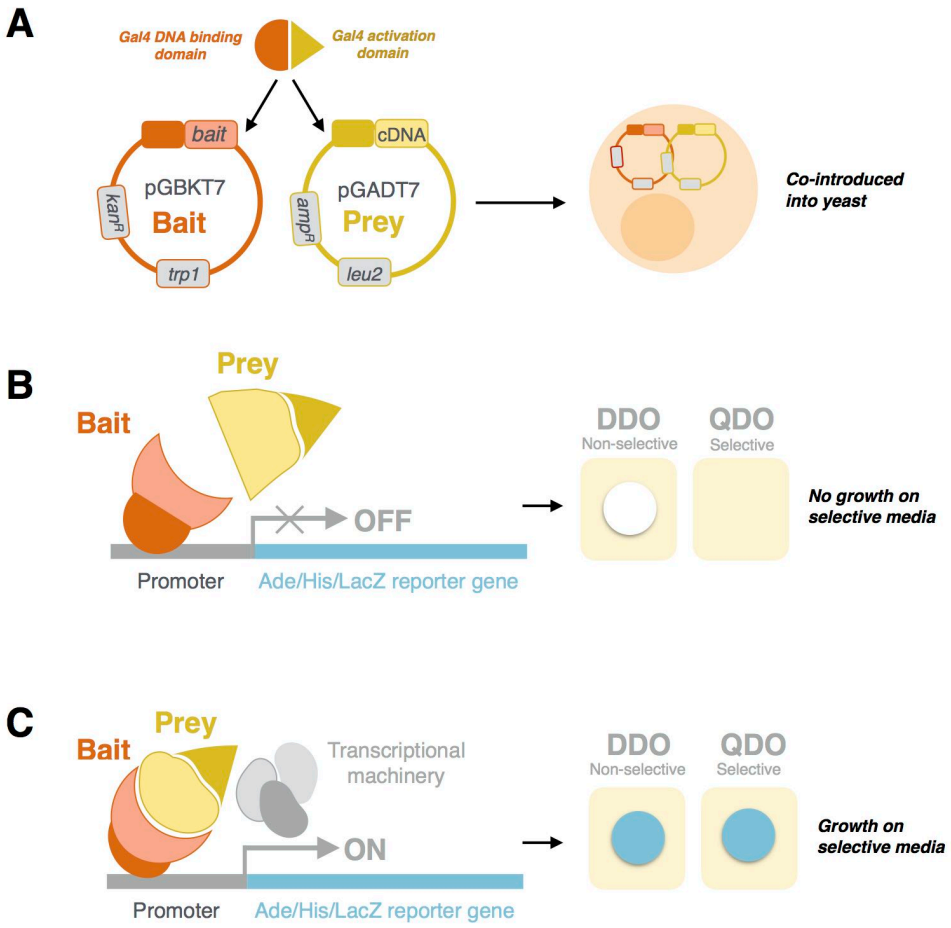


Figure 3.14 | Schematic representation of yeast two-hybrid methodology for identifying interacting partners. A) The Gal4a transcription factor is split into DNA-binding- and Gal4a activating domains and fused to bait and prey proteins in pGBKT7 and pGADT7 plasmids, respectively. pGBKT7 encodes *TRP1* for tryptophan biosynthesis; pGADT7 encodes *LEU2* for leucine biosynthesis. Bait and prey plasmids are co-introduced into *S. cerevisiae* AH109. **B**) If bait and prey proteins do not interact, reporter genes are not expressed and growth occurs on media lacking tryptophan and leucine (double drop-out media, DDO) but not on media lacking lacking histidine, adenine, tryptophan and leucine (quadruple drop-out media, QDO). **C**) If bait and prey proteins interact, a Gal4a transcriptional activator is formed, resulting in histidine and adenine biosynthesis. This permits growth on QDO. A *lacZ* reporter also gives expression of a β-galactosidase, which metabolises X-α-gal and colours the yeast blue.

Briefly, the screen works by separating the Gal4a transcription factor into its DNA-binding- and Gal4a activating domains and expressing them from separate plasmids. Within these vectors, the domains are fused to bait proteins (in pGBKT7) or a library of prey proteins (in pGADT7) (**Figure 3.14A**; Giniger, Varnum & Ptashne, 1985; West, Yocum & Ptashne, 1984). pGBKT7 confers the ability to synthesise tryptophan, while pGADT7 encodes *LEU2* for leucine biosynthesis. *S. cerevisiae* AH109 transformed with both plasmids can grow on media lacking tryptophan and leucine (double drop-out media, DDO). If bait and prey proteins interact, the Gal4a-binding and

–activating domains are brought into close proximity and form a functional transcriptional activator, recruiting the necessary translational machinery that promotes histidine and adenine biosynthesis. This allows growth on media lacking histidine, adenine, tryptophan and leucine (quadruple drop-out media, QDO; see **Figure 3.14C**). Additionally, a *lacZ* reporter gene downstream of the histidine and adenine biosynthesis genes encodes a β-galactosidase enzyme that metabolises x-α-gal, which gives a blue coloration of the resulting yeast colonies. If the bait and prey proteins do not interact, there is no expression of the reporter genes and yeast growth is observed on DDO but not on QDO (**Figure 3.14B**). This makes yeas two-hybrid methodology a powerful tool for identifying protein-protein interactions.

3.7.1. Assessing the suitability of NleG proteins for yeast two-hybrid screening for interacting partners

Prior to starting yeast two-hybrid screening, control experiments were performed to determine the suitability of pGBKT7-*nleG* constructs for use in the screen. The ability of NleG proteins to activate expression of the reporter genes *HIS3*, *ADE2* and *lacZ* in the absence of an interacting partner (thereby considered autoactivating) was assessed by transformation of pGBKT7-*nleG* constructs (or empty vector) into *S. cerevisiae* AH109 and plating on selective media (**Figure 3.15A**). Growth on SD media lacking tryptophan was used to demonstrate successful transformation of the yeast. Growth on media lacking tryptophan, histidine and adenine (triple drop out media, TDO) was used to assess autoactivation of reporter genes. Promisingly, full length and N-terminal NleG1, NleG7 or NleG8 from *C. rodentium* exhibited no yeast growth on TDO media, and were deemed suitable for further testing. We carried out further suitability tests for two potential baits: the NleG1 N-terminus and the NleG7 N-terminus. The N-termini were chosen as they are hypothesised to confer binding to the interacting partners and that the absence of the catalytic C-terminus might stabilise the interaction. Proceeding with a full length enzymatically active NleG as a bait would plausibly complicate partner identification as E3 ligase-partner interactions tend to be transient in nature and often result in proteasomal degradation of the partner.

Whether the chosen baits were cytotoxic to yeast or caused a growth defect in yeast (which might affect yeast mating efficiency in two-hybrid screens) was assessed by monitoring growth in SD media (lacking tryptophan to ensure maintenance of the bait plasmid). As shown in **Figure 3.15B**, both baits tested gave growth curves similar to yeast transformed with the empty vector.

The N-terminal NleG1 and NleG7 baits were therefore deemed suitable for use in yeast two-hybrid screens to identify potential interacting partners of NleG1 and NleG7.

A

pGBKT7 plasmid in AH109	Growth on -Trp media	Growth on TDO media
No plasmid	No	No
Empty plasmid	Yes	No
$NleG1_{CR}$ full length	Yes	No
$NleG1_{CR}$ N-terminus	Yes	No
$NleG7_{CR}$ full length	Yes	No
$NleG7_{CR}$ N-terminus	Yes	No
$NleG8_{CR}$ full length	Yes	No
$NleG8_{CR}$ N-terminus	Yes	No

B

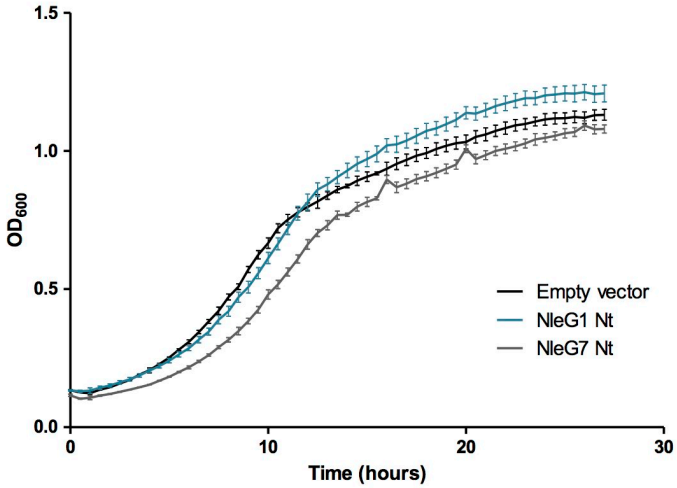


Figure 3.15 | *C. rodentium* NleG proteins do not cause autoactivation of reporter genes or a growth defect in in *S. cerevisiae* AH109. A) The yeast strain *S. cerevisiae* AH109 used in yeast-two-hybrid screening was transformed with pGBKT7-NleG constructs or empty vector and the growth of resulting colonies on selective media was documented. Transformants were plated onto SD media lacking tryptophan (-Trp; growth expected in transformed with pGBKT7) and SD media lacking tryptophan, adenine and histidine (triple drop-out media, TDO; growth only seen in genes within pGBKT7 constructs auto-activate *HIS* and *ADE* reporter genes). **B)** *S. cerevisiae* AH109 transformed with empty pGBKT7 or pGBKT7 expressing the N-terminus of NleG1 or NleG7 were grown in SD broth lacking tryptophan. OD_{600} measurements were recorded using a FluoStar Optima plate reader to assess whether NleG expression affected yeast growth. The graph shows mean OD_{600} readings ± SD.

3.7.2. NleG7 interacts with GRASP55 and WDR61 in yeast

In the first yeast two-hybrid screen performed in this project, yeast expressing the N-terminus of NleG1 (henceforth $NleG1^{N}$) was mated with a murine universal library of prey proteins expressed in pGADT7. QDO media was then used to select for yeast where bait and prey proteins interact.

Eleven resulting clones on QDO media were sequenced and the validity of eight hits of interest (including two plasmids expressing the PA28 β-subunit of the proteasome complex) was tested in a so called direct-two-hybrid assay (**Figure 3.16**), where known bait and prey constructs are transformed into yeast and the growth on QDO assessed. This assay did not confirm the putative interaction between NleG1 (full length or NleG1^{N}) and PA28-β, indicated by the lack of growth on QDO (**Figure 3.16A**). For the remaining seven hits of interest, all the prey clones that gave growth on QDO upon co-transfection with the bait construct were autoactivating, as growth was also observed upon co-transformation with empty pGBKT7 (**Figure 3.16B**). For this reason, these hits were deemed false positives and were not analysed further.

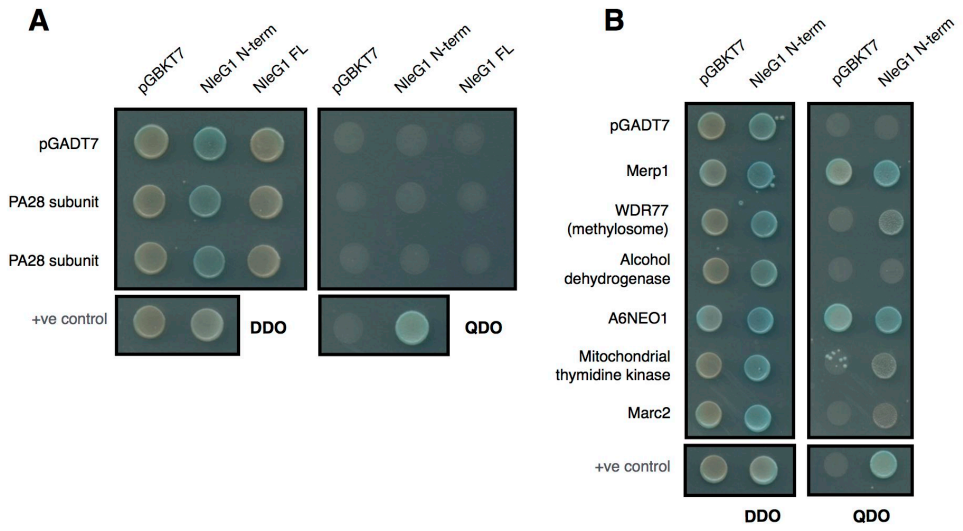


Figure 3.16 | Direct yeast two-hybrid assays using *C. rodentium* NleG1 did not confirm any new interacting partners. A) Two PA28 subunit-expressing clones identified in a yeast two-hybrid screen using *C. rodentium* NleG1^{N} and a commercial mouse library did not interact with NleG1 (full length, FL, or N-terminal) in a direct yeast two-hybrid assay. **B)** Another seven hits identified in the yeast two-hybrid screen did not interact with N-terminal NleG1 in a direct two-hybrid assay. The positive control (+) for the assay is Cdc20 interacting with NleG8$_{EHEC}$.

The second yeast-two-hybrid screen using pGBKT7-NleG7^{N} bait against a mouse library yielded 19 clones on QDO. Only seven of them were protein-coding (listed in **Figure 3.17A**) and five were tested in a direct-two-hybrid assay: Nidogen-1 (a basal membrane protein), Golgi Reassembly and Stacking Protein 55 (GRASP55, involved in stacking Golgi cisternae), Bardet-Biedl Syndrom 7 protein (BBS7; involved in ciliogenesis), WD repeat-containing protein 61 (WDR61; also known as Ski8, involved in transcriptional regulation) and Oculocerebrorenal syndrome of Lowe protein (OCRL1; involved in inositol signalling) (**Figure 3.17B**). Growth on QDO was observed when prey plasmids encoding Nidogen-1, BBS7 or OCRL1 were co-transformed with empty pGBKT7, indicative of autoactivation. The assay did, however, confirm that NleG7^{N} interacts with

GRASP55 and with WDR61 as growth appeared on QDO when yeast was co-transformed with pGBKT7-NleG7^{N} but not with the empty plasmid (indicated in **Figure 3.17B** by black arrowheads).

A

Protein	Associated function
Nidogen-1	Basement membrane protein involved in cell-cell adhesion
GORASP2 (GRASP55)	Golgi stacking protein, implicated in autophagy
BBS7	Ciliogenesis, microtubule-based transport.
OCRL1	Inositol signalling, binds Rac and Rho GTPases
WDR61	WD40 repeat-containing protein
GNB4	Guanine nucleotide-binding protein
NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10	Electron transport

B

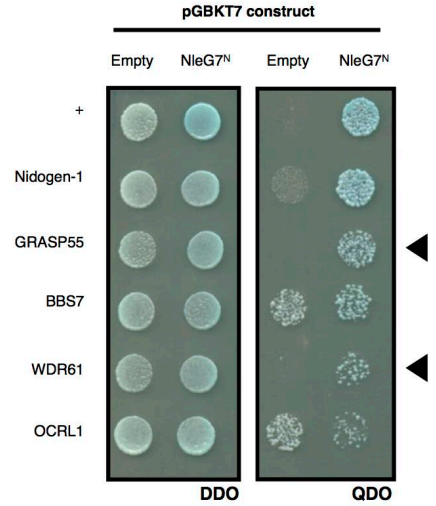


Figure 3.17 | Direct yeast two-hybrid assays using *C. rodentium* NleG7^{N} confirmed several interacting partners. A) Hits from a yeast-two-hybrid screen with *C. rodentium* NleG7^{N} against a universal mouse library. **B**) A direct yeast two-hybrid assay where pGADT7 hits from (A) were co-transformed in yeast with empty pGBKT7 or pGBKT7-NleG7^{N}. Black arrows indicate interactions with GRASP55 and with WDR61. The positive control (+) is Cdc20 interacting with NleG8$_{EHEC}$. **C)** A direct two-hybrid assay where pGBKT7-NleG8 and pGBKT7-NleG8^{N} were co-transformed in yeast with pGADT7-GRASP55 gave growth on QDO (black arrowhead).

The key structural domains of both GRASP55 and WDR61 are shown in **Figure 3.18**. The overlap between the full-length protein and bait recovered in the yeast screens is also indicated for both GRASP55 and WDR61. At the time that the above yeast two-hybrids assays were performed, we had immunofluorescence microscopy data showing that NleG7 was predominantly perinuclear when expressed in HeLa cells. This prompted the hypothesis that NleG7 localised to the Golgi apparatus in HeLa cells, making the Golgi protein GRASP55 an interesting putative interacting partner to consider in further validation experiments. WDR61 on the other hand is predominantly nuclear (Jaehning, 2010), unlike NleG7, bringing into question the biological relevance of the WDR61-NleG7 interaction in yeast. As such, subsequent validation experiments focused solely on GRASP55.

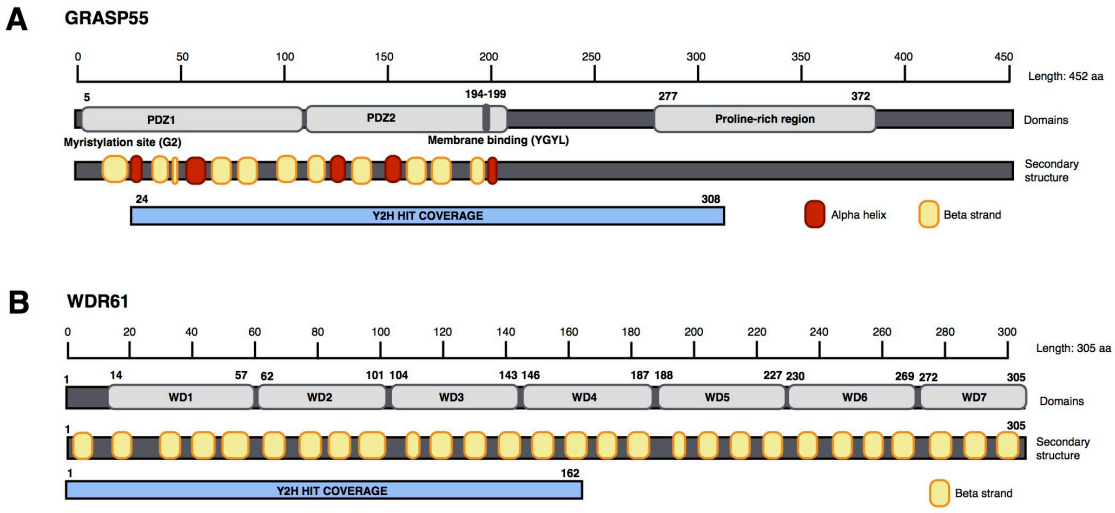


Figure 3.18 | Structural features of GRASP55 and WDR61, shown to interact with *C. rodentium* NleG7^{N} in a direct two-hybrid assay. A) Structural overview of GRASP55 of *Homo sapiens* (sharing 91% amino acid identity with the murine homolog), based on the solved crystal structure. The yeast hit (in blue) overlaps with the N-terminal half of the protein, covering three quarters of the PDZ1 domain and the entire PDZ2 domain. **B)** Structural overview of the WDR61 in *Homo sapiens* (88% amino acid identity shared with murine WDR61). The solved crystal structure was solved and revealed 7 tandem-WD40 repeats arranged in a β-propeller architecture (Xu & Min, 2011). 162 residues were covered by the hit in our yeast screen.

To test if GRASP55 and NleG7 co-localised during transient transfection of HeLa cells, co-transfection experiments were performed using pRK5 expressing HA-NleG7 and a construct expressing GRASP55-Myc kindly provided by Professor Min Goo Lee (Yonsei University College of Medicine, South Korea). We expected a true target or interacting partner of NleG7 to co-localise with NleG7 when co-expressed in cultured cells. However, immunofluorescence microscopy showed that GRASP55 localisation did not co-localise with NleG7 (**Figure 3.19**). Further experiments to determine the precise subcellular localisation of NleG7 (described previously in this chapter, see **Section 3.6**, **Figure 3.11**) later confirmed that NleG7 does not co-localise with giantin, a Golgi marker that GRASP55 co-localises with in published reports (Xiang & Wang, 2010).

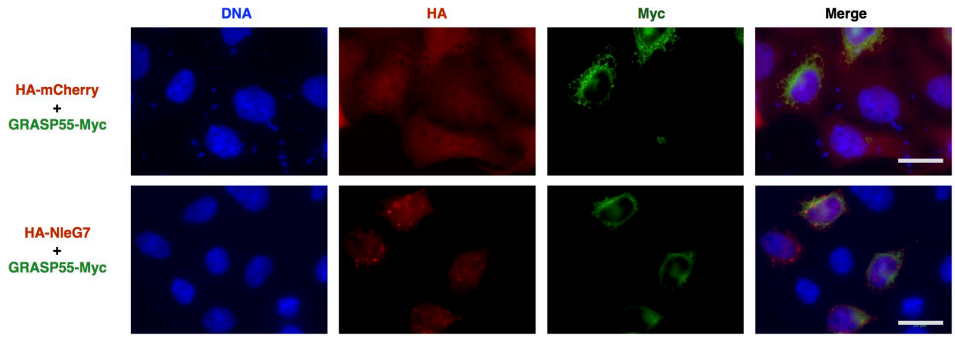


Figure 3.19 | GRASP55 and NleG7 do not co-localise when co-expressed in HeLa cells. HeLa cells were transfected with constructs expressing C-terminally Myc-tagged GRASP55 and N-terminally HA-tagged NleG7 and analysed by immunofluorescence microscopy. GRASP55 is shown in green, NleG7 in red and DNA in blue. Scale bar 20 µm.

The final experiment following up on GRASP55 as a putative interacting partner of NleG7 tested the specificity of the interaction in yeast. We hypothesised that different NleG proteins have different cellular targets in the host, and therefore performed a yeast direct two-hybrid assay to ascertain if another NleG family member, NleG8, could interact with GRASP55. As shown in **Figure 3.20**, an interaction with GRASP55 was indeed observed both for full-length and N-terminal NleG8. GRASP55 harbours several PDZ motifs (see **Figure 3.18**), but the interaction between GRASP55 and NleG8 did not rely on the C-terminal PDZ-binding motif of NleG8, as N-terminal NleG8 also bound GRASP55.

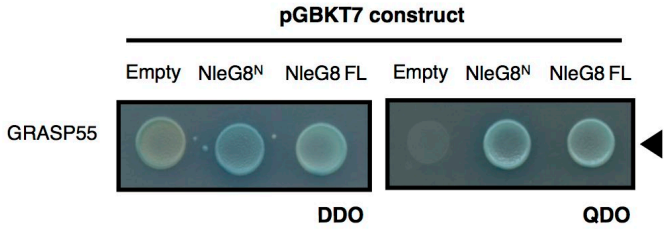


Figure 3.20 | NleG8 interacts with GRASP55 by yeast direct two-hybrid. A direct two-hybrid assay where pGBKT7-NleG8 FL (full length) and pGBKT7-NleG8^{N} (N-terminus) co-transformed in *S. cerevisiae* AH109 with pGADT7-GRASP55 gave growth on QDO (black arrowhead).

Ultimately, the lack of specificity of the GRASP55-NleG7 interaction in yeast taken together with the absence of co-localisation between NleG7 and GRASP55 in eukaryotic cells brought into question the validity of this putative interaction.

3.7.3. C. rodentium and EHEC NleG7 interact with an uncharaterised putative membrane protein

NleG7 from *C. rodentium* and NleG7 EHEC share 74% amino acid identity and might therefore share the same or similar interacting partners in their hosts. Before the start of this project, a yeast two-hybrid screen had been performed by Dr Bevin McGeever (Frankel Lab) using the EHEC NleG7 N-terminus as a bait against a library of HeLa cell prey proteins. This generated a number of interesting hits, which are listed in **Figure 3.21A**. In the project presented here, we tested to see if the hits for EHEC NleG7 also interacted with *C. rodentium* NleG7 by direct yeast two-hybrid methodology. This showed that *C. rodentium* NleG7^{N} interacts with LOC401397, an uncharacterised predicted transmembrane protein with no sequence homology to proteins of known function (**Figure 3.21B**). In the original yeast-two-hybrid screen with EHEC NleG7^{N}, 11 hits had been identified as LOC401397, highlighting its potential as a true target of NleG7. To clean

up the plasmid, only the coding region of the hit was subcloned into pGADT7 and tested in another direct yeast two-hybrid assay against *C. rodentium* NleG7 and NleG7^{N}. This confirmed a weak interaction between the uncharacterised protein and NleG7 (**Figure 3.21C**), although the plates had to be incubated for an additional 3 days before any growth was observed. Moreover, when all 11 hit sequences were analysed further it became clear that LOC401397 was not always in frame with the Gal4-domain. Finally, we were unable to detect LOC401397 mRNA in cultured HeLa cells using reverse transcription quantitative PCR (RT-qPCR; data not shown), even though the original screen that identified LOC401397 had been performed using a commercial HeLa library. Taken together, this brought into question the likelihood of LOC401397 as a true interacting partner of NleG7, and no further efforts were made to validate this putative interaction.

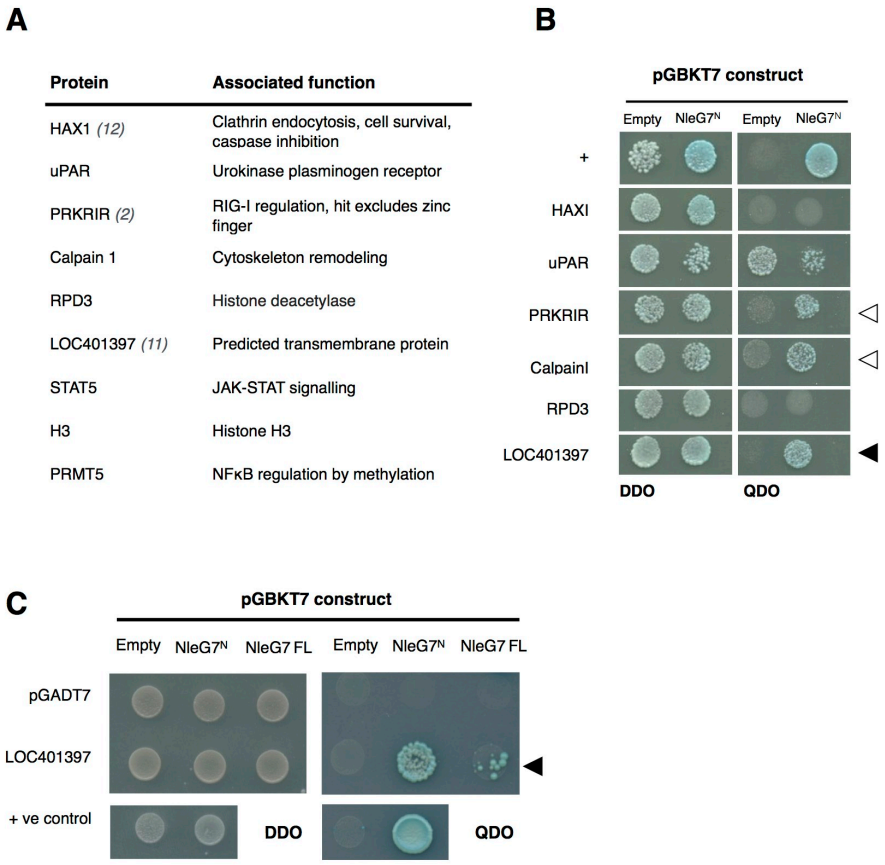


Figure 3.21 | Direct yeast two-hybrid assays using N-terminal *C. rodentium* NleG7 confirmed an uncharacterised putative membrane-associated protein as an interacting partner. A) The hits from a previously performed yeast-two-hybrid screen with EHEC NleG7 N-terminus against a HeLa library. The number of hits per protein is indicated in brackets if more than one. Data courtesy of Dr Bevin McGeever. **B)** A direct yeast two-hybrid assay, where plasmids of interest from (A) were co-transformed with the bait plasmid (encoding N-terminal NleG7, NleG7^{N}) or empty pGBKT7 into yeast. **C)** A direct yeast two-hybrid assay, where empty pGADT7 or the prey plasmid (pGADT7-LOC401397) were co-transformed into yeast with the bait plasmid (encoding NleG7^{N} or full length NleG7) or empty pGBKT7. The positive control (+) is Cdc20 interacting with NleG8$_{EHEC}$. Black arrowheads indicate the positive interaction with LOC401397. Open arrowheads highlight examples of more growth with the bait plasmid than with the empty plasmid, but still showing some growth with the empty plasmid, indicative of autoactivation.

3.7.4. The interaction of C. rodentium NleG7 and NleG8 with NHERF1 and NHERF2 in a yeast two-hybrid assay

C. rodentium NleG8 and EHEC NleG7 and NleG8.1 contain C-terminal PDZ-binding motifs, as indicated in grey in **Figure 3.22A**. They recognise PDZ-domains, which are highly abundant modules that mediate protein-protein interactions (Lee & Zheng, 2010). As part of a previous global PDZ array screen (courtesy of Professor Randy Hall, Emory University School of Medicine, United States of America), it was revealed that the 50 most C-terminal amino acids of *C. rodentium* NleG8 and EHEC NleG8.1 fused with GST specifically bind PDZ-1 and (albeit to a lesser extent) PDZ-2 of both NHERF1 and NHERF2 (**Figure 3.22B**). NHERF1 and NHERF2 are involved in trafficking, membrane retention and regulation of membrane proteins (including ion channels) and are highly expressed in enterocytes (Martinez *et al.*, 2010). We therefore tested the interaction between NHERF1/2 and *C. rodentium* NleG8 or NleG7 using the direct yeast two-hybrid approach. However, no growth was observed on QDO medium following co-transformation of yeast with NleG8- or NleG7-expressing pGBKT7 and NHERF1/2-expressing pGAD424 (**Fig. 3.22C**).

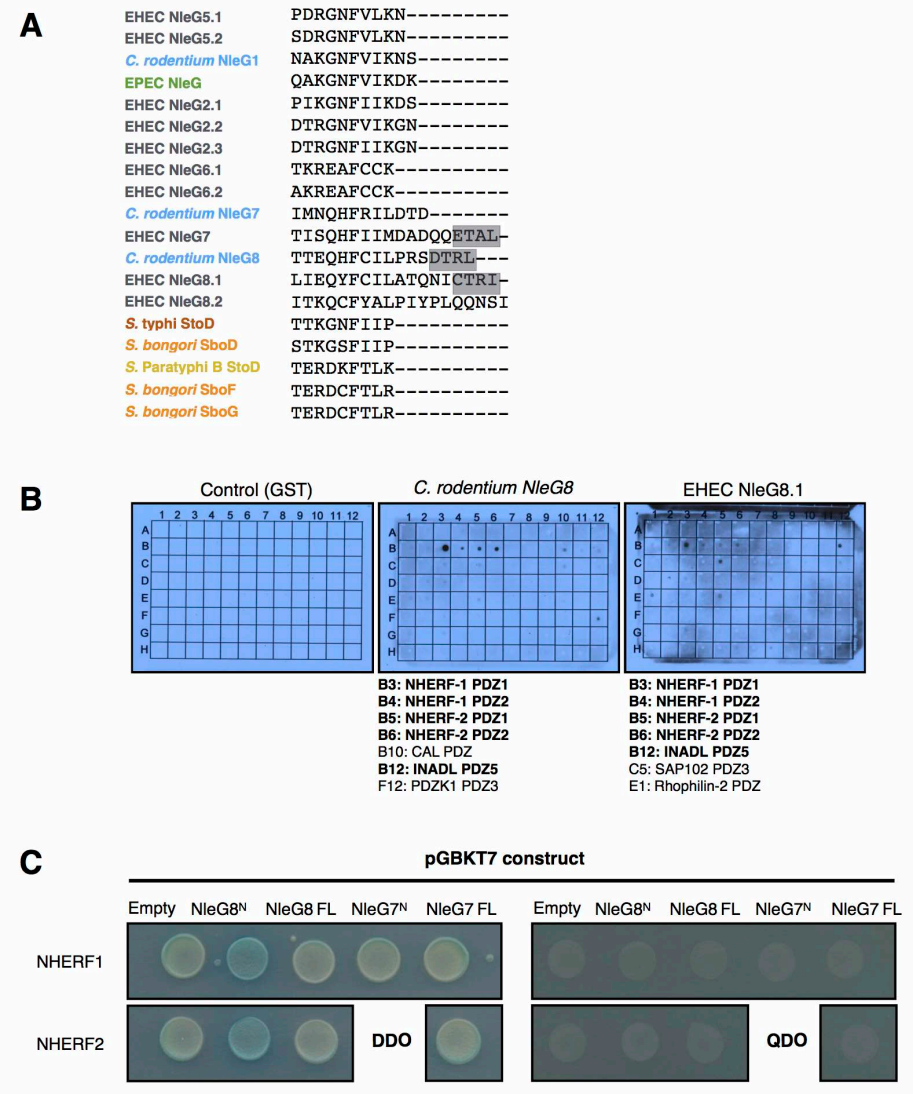


Figure 3.22 | *C. rodentium* NleG7 and NleG8 do not interact with NHERF1 or NHERF2 in a direct two-hybrid yeast assay. A) Amino acid alignment of the very C-terminus of NleG family proteins. The PDZ-binding motifs found in EHEC NleG7., *C. rodentium* NleG8 and EHEC NleG8.1 are highlighted in grey. **B)** Gel overlay assay using the last 50 amino acids of *C. rodentium* NleG8 or EHEC NleG8, alongside a GST control, against PDZ-domains from proteins of known function. Positive interactions indicated by a spot. PDZ domains implicated in putative interactions are indicated below the blot, with stronger interactions in bold. Experiment kindly performed by Dr. Stacey Schaefer, data courtesy of Professor Randy Hall. **C)** pGBKT7-NleG constructs were co-transformed into *S. cerevisiae* AH109 and plated on double and quadruple drop out medium (DDO and QDO). Positive interactions (indicated by grow on QDO but not DDO upon co-transformation with with pGAD424 constructs) were not observed.

3.7.5. Full-length NleG7 and NleG8 associate with ubiquitin in yeast

Knowing that NleGs are E3 ligases enzymes, we investigated whether they could directly associate with ubiquitin in a direct yeast-two hybrid. Ubiquitin was fused with the Gal4a-DNA activating domain in pGADT7 and autoactivation tests were performed as described earlier in this chapter. The ubiquitin construct was found suitable for use in further experiments (**Figure 3.23A**). Interestingly, yeast co-transformed with pGBKT7 expressing full-length NleG7 or NleG8

and pGADT7-ubiquitin but not empty pGADT7 grew on QDO (**Fig. 3.23B**), suggesting that full length NleG8 and NleG8 but not the N-termini associate with ubiquitin in a direct yeast-two-hybrid system. From the data presented here, it cannot be determined if NleG7 and NleG8 are interacting with ubiquitin or are covalently bound to ubiquitin (ie have undergone ubiquitination), as both would result in growth on QDO. If the latter is true, this would be the first indication that the NleGs are ubiquitinated in eukaryotic cells. No interaction was observed between NleG1 and ubiquitin, but it should be noted that no expression tests were performed to verify that full-length NleG1 was being expressed from this plasmid.

A

pGADT7 plasmid in AH109	Growth on -Leu media	Growth on TDO media
No plasmid	No	No
Empty plasmid	Yes	No
Ubiquitin	Yes	No

B

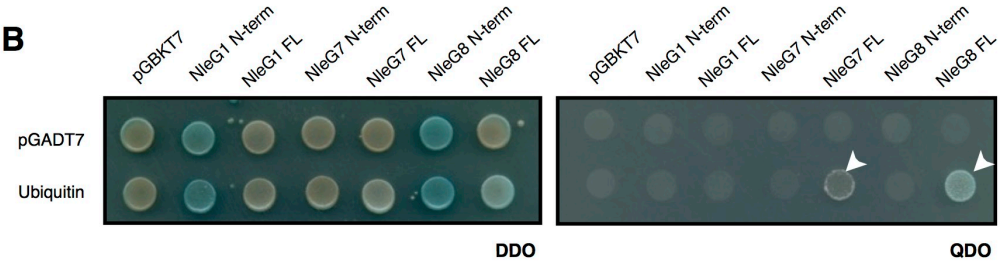


Figure 3.23 | *C. rodentium* NleG7 and NleG8 associate with ubiquitin in a yeast direct two-hybrid assay. A) Autoactivation testing showed that ubiquitin expressed from pGADT7 does not autoactivate *HIS* and *ADE* genes in *S. cerevisiae* AH109, as indicated by lack of growth on TDO media (lacking leucine, histidine and adenine). Growth on SD media lacking leucine confirmed that the yeast was successfully transformed with the pGADT7 plasmid. **B)** A direct yeast-two-hybrid assay where pGBKT7 plasmids encoding *C. rodentium* NleGs (N-terminus or full length, FL) were co-transformed with empty pGADT7 or pGADT7 encoding ubiquitin into yeast. Interactions are indicated by growth on selective QDO media following co-transformation with the NleG-encoding plasmid but not with empty plasmid (white arrows).

3.7.6. The interaction of NleG8 and StoD with the small ubiquitin-like modifier ISG15 in yeast experiments

In parallel to this NleG project, the interaction between ubiquitin and StoD was studied by NMR spectroscopy by Dr Melanie McDowell and Professor Susan Lea (Sir William Dunn School of Pathology, University of Oxford). They found that both the N- and C-terminus of StoD could bind to ubiquitin and that StoD bound K48- or K63-linked di-ubiquitin moieties with greater affinity than for mono-ubiquitin (unpublished data). The ability to bind two ubiquitin moieties simultaneously prompted the idea that StoD and other NleG family members might bind

Interferon-stimulated Gene 15 (ISG15), which is a small ubiquitin-like modifier structurally reminiscent of a di-ubiquitin molecule (Blomstrom *et al.*, 1986; Kunzi & Pitha, 2003). To test this idea, we performed a yeast direct two-hybrid assay, where the ability of NleG8 and StoD (catalytically active or inactive) to interact with ISG15 or ubiquitin was assessed (**Figure 3.24**). While growth was observed for yeast co-transformed with ubiquitin and NleG8/StoD constructs, co-transformation with ISG15 and NleG8/StoD constructs did not give growth on QDO. Therefore, it was concluded that NleG8 and StoD do not interact with ISG15 in this system. We could distinguish between an interaction with ISG15 and covalent modification by ISG15 (ISGylation) here, as the yeast strain that the assay was performed in does have the E2 enzyme necessary for ISGylation.

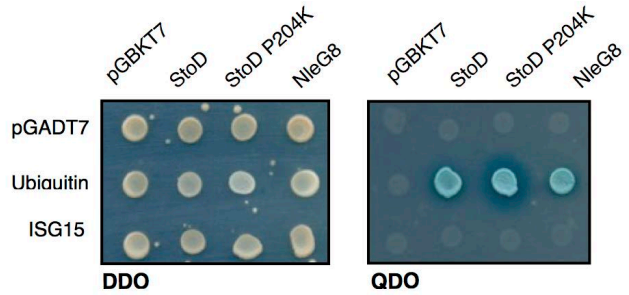


Figure 3.24 | *C. rodentium* NleG8 and *S.* typhi StoD associate with ubiquitin but not ISG15 by direct yeast two-hybrid assay. A direct yeast-two-hybrid assay where pGBKT7 plasmids encoding *C. rodentium* NleG8/StoD were co-transformed with empty pGADT7 or pGADT7 encoding ubiquitin or ISG15 into yeast. Interactions are indicated by growth on selective QDO media.

3.8. Ectopic expression of C. rodentium NleG7 and NleG8 in HeLa cells causes ubiquitin relocalisation

Having established that NleG7 and NleG8 associate with ubiquitin in yeast two-hybrid assays, we sought to determine the effect of their expression on host ubiquitin localisation in cultured cells. HeLa cells were transfected with pRK5 constructs encoding Myc-GFP, Myc-NleG1, Myc-NleG7 or Myc-NleG8 and visualised by immunofluorescence microscopy to determine whether they co-localised or altered ubiquitin staining in the cell (**Figure 3.25**). Transfection of the Myc-GFP control in Hela cells gave a diffuse ubiquitin signal across the cell, as expected for a protein that is not involved in ubiquitin signalling. This was also observed for Myc-NleG1. Transfection of Myc-NleG7 and Myc-NleG8 in HeLa cells, on the other hand, re-localised endogenous ubiquitin: a punctate ubiquitin signal that co-localised with NleG7 or NleG8 was observed, as indicated by white arrowheads in **Figure 3.25**. This suggests that either overexpression of NleG7 and NleG8 re-organises the pool of ubiquitinated proteins in the cell, or that NleG7 and NleG8 heavily ubiquitinate themselves or their cellular targets under these conditions. Irrespective of which,

this supports the involvement of the NleG family members in hijacking with the host ubiquitination system.

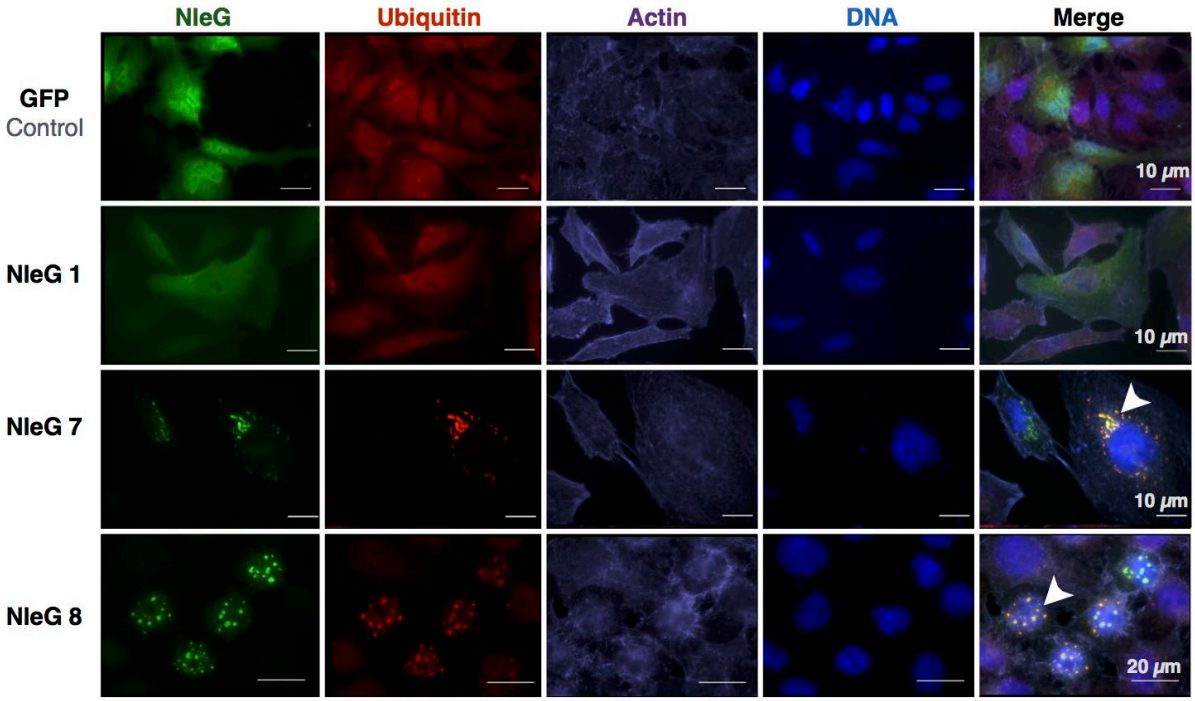


Figure 3.25 | *C. rodentium* NleG7 and NleG8 relocalise and co-localise with ubiquitin during transfection. HeLa cells were transfected with pRK5 constructs expressing Myc-NleG1, Myc-NleG7, Myc-NleG8 or Myc-GFP. NleG localisation was assessed an anti-Myc antibody (green). Ubiquitin is shown in red, actin in purple and DNA in blue. Co-localisation with ubiquitin was seen for NleG7- and NleG8-transfected cells and is indicated by white arrowheads.

3.9. Discussion

E3 ubiquitin ligase enzymes, by catalysing the final step of the ubiquitination cascade, control virtually every aspect of eukaryotic cell biology (Zheng & Shabek, 2017) and a growing pool of evidence suggests that viral and bacterial proteins modulate or mimic E3 ligase proteins. While the NleG family of T3SS effector proteins are known to structurally mimic RING/U-box-type E3 ubiquitin ligases via their conserved C-termini (Wu *et al.*, 2010), their specific contributions to A/E pathogenesis remain undetermined. The N-termini, with more pronounced variation in sequence length and similarity, might reflect diversification of NleG homologs in pathogenic bacteria to target a range of cellular functions during infection. NleG family members are also found in *S. bongori* as well as in *S. enterica*, but only in the typhoidal serovars *S.* Typhi and *S.* Paratyphi B. As shown in this chapter, *C. rodentium* can translocate three NleG homologs NleG1, NleG7 and NleG8 into eukaryotic cells via its T3SS and the overarching aims of this project are to characterise these proteins and investigate their function and importance during A/E pathogenesis.

The idea that N-terminal amino acid sequence variation directs the NleG proteins to different host interacting partners or substrates is not exclusive to the NleG family. For instance, the IpaH family of NEL-type bacterial E3 ligase mimics in *Shigella* and *Salmonella* spp. share an overall topology with a ubiquitin ligase domain and an LRR domain. However, differences within the LRR confer substrate specificity, and a range of different substrates and binding partners have been associated with specific IpaH homologs (refer to **Table 1.1** for details). Moreover, different binding partners may explain why certain pathogens express numerous NleG proteins, as seen for *C. rodentium*.

Finding substrates for E3 ubiquitin ligases is, however, challenging for several reasons. The first reason is the low-affinity transient substrate-enzyme interactions (generally in the high nM to µM range). Secondly, many targets are quickly subjected to proteasomal degradation and their detection is therefore difficult. Thirdly, the human ubiquitiniome and ubiquitin code is extremely complex and, finally, if substrate proteins are localised to poorly soluble cellular compartments, their identification by standard IP-based methods becomes substantially more difficult (Coyaud *et al.*, 2015). As such, many bacterial E3 ligases have been described but only few of their substrates identified. One method used to identify host proteins bound by T3SS effector proteins is the yeast two-hybrid screening method (Popa *et al.*, 2016; Simpson *et al.*, 2006; Marchès *et al.*, 2006), and this has also been used for the identification of host proteins bound by the *Legionella* E3 ligase AnkB (Lomma *et al.*, 2010). *S. cerevisiae* itself harbours one E1 enzyme, 11 E2 enzymes and at least 51 E3 enzymes (Finley *et al.*, 2012).

In this study, we conducted several yeast two-hybrid screens and identified three putative interacting partners of *C. rodentium* NleG7^{N}. The first was WDR61, a WD40-domain-containing histone methyl-lysine binder within the nuclear PAF1 complex that associates with RNA polymerase II and confers transcriptional regulation (Chen *et al.*, 2017; Van Oss, Cucinotta & Arndt, 2017). This complex recruits an E3 ubiquitin ligase complex and an E2 enzyme to trigger monoubiquitination of histone H2B, which in turn promotes histone H3 trimethylation (Cole, Clifton-Bligh & Marsh, 2015; Shema *et al.*, 2008; Chu *et al.*, 2013; Vermeulen & Timmers, 2010; Kim & Gibson, 2010). Previous reports of bacterial effector proteins interfering with host histone modification, such as the *Legionella* RomA methyltransferase, lends some supports to the idea that *C. rodentium* effectors could have evolved to interfere with this process (Rolando *et al.*, 2013). However, considering that upon transfection of HeLa cells using Myc- and HA-tagged NleG7, NleG7 predominantly localises around the nucleus (rather than inside the nucleus, where the PAF1 complex is found), the interaction between NleG7 and WDR61 might not be of

relevance during *C. rodentium* infection. Therefore, we did not lend further efforts to establish whether this interaction was detectable or relevant outside the yeast system.

The second binding partner was GRASP55, which is associated with the stacking of Golgi cisternae (Shorter *et al.*, 1999; Giuliani, Grieve & Rabouille, 2011). GRASP55 was previously shown to contribute to unconventional secretion of proteins, which is primarily stress-induced. Extracellular delivery of IL-1β is classified as unconventional secretion and Dupont and colleagues showed that it required autophagy-associated proteins as well as GRASP proteins (Dupont *et al.*, 2011; Gee *et al.*, 2011; Levi & Glick, 2007). Autophagy regulates numerous inflammatory transcriptional responses and studies suggest that autophagy in intestinal epithelium controls *C. rodentium* infection (Deretic, Jiang & Dupont, 2012; Inoue *et al.*, 2012). GRASP55 localises to medial- to trans cisterna and displayed perinuclear staining and co-localisation with giantin in published reports (Xiang & Wang, 2010; Kuo *et al.*, 2000; Roghi *et al.*, 2010; Linstedt & Hauri, 1993). Our data, however, show that NleG7 does not co-localise with giantin during transfection of HeLa cells. Further, co-transfections of NleG7 and GRASP55 in HeLa cells did not reveal any co-localisaiton between these proteins. An appreciated drawback of the yeast screening methodology for identifying interacting partners is the high incidence of false positives (Bruckner *et al.*, 2009; Rajagopala *et al.*, 2012). In our laboratory, we have previously observed non-specific interactions between NleG proteins and targets identified in yeast-two-hybrid screens; we hypothesise that this is also the case for GRASP55. Ultimately, this stresses that yeast-two-hybrid screening may not be the most appropriate method for identifying substrates of NleG family members, which may be translocated into host cells in very low concentrations during infection and only partake in transient, unstable interactions. Although we started characterising the interaction between $NleG7^{N}$ and an uncharacterised predicted membrane protein that was originally a hit from an EHEC NleG7 screen conducted in parallel to this project, we soon realised that this interaction was not specific to NleG. As the protein bound NleG8, we decided to turn our attention to putative hits identified by non-yeast methods.

One such way in which the binding partners of NleG8 had previously been studied was via gel overlay assay to detect interactions between $NleG8^{C}$ (which contains a PDZ-binding motif) and an array of known PDZ-domains. This revealed interactions between *C. rodentium* NleG8 and the NH3 regulatory factors NHERF1 and NHERF2, respectively, which was also detected for EHEC NleG8.1. NHERF proteins are highly expressed in intestinal and colonic tissues and orchestrate intracellular protein trafficking, localisation and signalling pathways, best illustrated by their

regulation of ion transporters such as NHE3 in the apical membrane (Donowitz *et al.*, 2005). This putative interaction is noteworthy considering the implication of NHERF-proteins and NHE3 in (dys-)regulation of intestinal ion transport and consequently diarrhoea (Lamprecht & Seidler, 2006). NHE3 activity is in fact known to be down-regulated during EPEC-infection dependent on the T3SS effector protein EspF (Berger *et al.*, 2009a; Hodges *et al.*, 2008; Li *et al.*, 2007). NHERF1 and NHERF2 both harbour two PDZ domains and exhibit overlapping interactomes due to high sequence homology (Martinez *et al.*, 2010; Morales *et al.*, 2007). Over 30 binding partners have previously been reported including T3SS effectors of A/E pathogens, such as Map, EspI and NleH1 (Martinez *et al.*, 2010; Shenolikar *et al.*, 2004). Martinez and colleagues confirmed interactions between NHERF2 and Map, EspI and NleH1, respectively, by gel overlay, co-immunoprecipitation and co-localisation experiments (Martinez *et al.*, 2010). They hypothesised that NHERFs may function as a distribution hub at the apical membrane to regulate trafficking of effector proteins. We propose that NleG8 distribution and function could be regulated in a similar manner and that the PDZ-binding motif could help direct or anchor NleG proteins and their targets to the membrane. Although we could not confirm the interaction between *C. rodentium* NleG8 or NleG7 and NHERF1/2 by yeast methodology, further efforts could perhaps determine whether NHERF1 or NHERF2 are degraded upon *C. rodentium* infection.

An interesting yeast result from this project concerns the interaction of NleG proteins with ubiquitin. We show that full-length but not N-terminal NleG7 and NleG8 interact with ubiquitin by direct yeast two-hybrid, suggesting that NleG proteins require their C-termini to bind ubiquitin. However, we cannot rule out that the ubiquitin association here is reflective of ubiquitination of NleG proteins in the yeast cells, as many E3 ligases self-ubiquitinate. A way to discriminate between the two hypotheses would be to test the catalytically inactive NleG proteins in this assay, as it should not be able it self-ubiquitinate. As for full-length StoD, the catalytically inactive mutant still binds ubiquitin, presumably via its N-terminal domain as this is supported by structural studies. Since NleG7 and NleG8 do not bear high homology to StoD in their N-termini, we propose that it should be tested whether they associate with ubiquitin in yeast assays when the catalytically inactive mutants are used.

To the best of our knowledge, we demonstrate for the first time that NleG proteins in *C. rodentium* harbour catalytic E3 ligase activity, as determined by *in vitro* autoubiquitination assays. We also show that mutating a conserved proline residue (residue P163 in NleG1) to lysine inhibits the E3 ligase activity, as this residue is needed for interacting with E2 enzymes.

This finding was essential in directing other arms of this PhD project, as the catalytic mutant was extensively used in further experiments to find interacting partners by mass spectrometry-based methods, described and discussed in **Chapter 5**. Further support that NleG proteins evolved to modulate or hijack ubiquitin-dependent cellular pathways comes from the co-localisation of NleG7 and NleG8 with ubiquitin in transfected cells.

Moreover, the cytotoxicity of NleG8 shown in this work suggest that NleG8 is biochemically active. This is comparable to reported findings for the *L. pneumophila* effector protein RavN, which was recently identified as a bona fide E3 ligase (Lin *et al.*, 2018). RavN also causes no detectable cytotoxicity when expressed in mammalian cells, but strongly attenuates yeast growth on media plates as seen for both EHEC NleG8.1 and *C. rodentium* NleG8 in the yeast cytotoxicity assay in this project. Like RavN, the cytotoxcitiy of NleG8 depended on a conserved residue (P177) located within the E2-E3 interface, as the P177K mutation restored growth (Lin *et al.*, 2018). Intriguingly, the cytotoxic phenotype required a full length NleG8, but the N-terminus could be substituted with that of NleG7. This could be an indication that the NleG termini should not be studied on their own, and that further experiments in cell-based systems should involve full-length NleG proteins (despite our data that the C-terminus of NleG1 on its own is capable of autoubiquitination).

We report different subcellular localisation patterns for NleG family members during transfection. NleG1 displays cytosolic cell-wide localisation similar to its closest homologue, EPEC NleG (Li *et al.*, 2006). We only observed NleG1-translocation into HeLa cells in a β-lactamase translocation assay when protein degradation was inhibited, as NleG1 may have a short half-life in eukaryotic cells in the absence of MG-132. The issue of instability due to rapid proteasome-dependent degradation following autoubiquitination is a recognised obstacle in the bacterial E3 ligase field of research (Suzuki *et al.*, 2014). Here, it may have implications for future *in vitro* studies concerning NleG1 (including transfections) and emphasises the need for including MG-132 or using the NleG1 point mutant incapable of supporting auto-ubiquitination in further experiments.

Although NleG7 predominantly displayed perinuclear staining, it did not co-localise with the early endosome marker EEA1, the cis- and medial Golgi marker giantin or the mitochondrial marker TOMM22 during transfection. The precise compartment or vesicle that NleG7 localises to thereby remains unknown. NleG8 puncta were observed around the nucleus but did to localise

within the nuclear membrane marker Lamin A. As useful as the *in vitro* transfection method can be, it must however be remembered that drawbacks of this over-expression system include the risk of localisation artefacts that are not relevant under A/E infection conditions. On the other hand, investigating the localisation of NleG proteins during infection is also problematic due to the typically very low number of translocated effector proteins alongside the possibility that NleG proteins undergo autoubiquitination and subsequent proteasomal degradation. Ultimately, further transfection experiments might be more informative once cellular binding partners of NleG proteins have been identified or a role for NleG proteins during infection is suggested.

The work outlined in this chapter contributes novel information regarding *C. rodentium* NleG E3 ligase proteins, furthering our understanding of the NleG family as a whole. As an important aspect of understanding E3 ligase biology is identifying their substrates and characterising E3-substrate interactions, we conclude that other methods than yeast screening should be employed to find the true targets of the *C. rodentium* NleG proteins.

Chapter 4: Characterising NleG proteins during in vivo C. rodentium infection

4. Characterising NleG proteins during in vivo C. rodentium infection

4.1. Introduction

In vitro studies using cell-line models (such as HeLa and Swiss 3T3 cells) have been invaluable in studying the functions of many T3SS effector proteins. However, *in vitro* findings do not always reflect the physiological function of effector proteins *in vivo*. Effector proteins may mediate interactions with mucosal surfaces or subvert host responses in enterocytes in the context of commensal bacteria, calling for robust *in vivo* laboratory models to study their roles. *C. rodentium* is the gold standard animal infection model to study virulence of the related A/E pathogens EPEC and EHEC, as mice are inherently resistant to EHEC and EPEC infection. The *C. rodentium* model has been used extensively to study specific virulence factors during colonisation, host responses to infection, mucosal immune responses and the role of the microbiota in host-pathogen interactions. The model has also been key to a myriad of important discoveries concerning immunity to enteric pathogens, including the secretion of IL-22 and IL-17 by T cells during enteric infection (Geddes *et al.*, 2011; Zheng *et al.*, 2008; Rubino, Geddes & Girardin, 2012), the involvement of the noncanonical inflammasome pathway (Kayagaki *et al.*, 2011, 2013; Lupfer *et al.*, 2014) and the involvement of NLRP6 inflammasomes (Wlodarska *et al.*, 2014). Recently, regulation of specific host immune responses by the single T3SS effector protein EspJ was demonstrated *in vivo* by combining an established and robust *C. rodentium* infection protocol with state-of-the-art mass spectrometry techniques (Pollard *et al*., 2018).

The main aim of this chapter was to employ the *C. rodentium in vivo* infection model to investigate the contribution of the NleG proteins to the kinetics and histopathological features of *C. rodentium* infection.

4.2. Making chromosomal nleG mutants in C. rodentium

Assessing the contribution of NleG proteins to *C. rodentium* infection *in vivo* required numerous *nleG* knockout mutants. Prior to the start of this project, Δ*nleG7* and Δ*nleG8* alongside a double knockout (Δ*nleG7* Δ*nleG8*) strains had been generated by Dr Valerie Crepin-Sevenou. Here, a Δ*nleG1 C. rodentium* strain and a full *nleG* knockout *C. rodentium* strain (Δ*nleG1* Δ*nleG7* Δ*nleG8)* were generated (by three-partner-conjugation with the pKNG101 suicide plasmid and the pSEVA-615S suicide plasmid, respectively – see **Sections 2.2.8** and **2.2.9** for details). Further, the new single *nleG1* deletion mutant and the triple *nleG* deletion mutant were complemented on the genome by introducing the *nleG* gene back into its native site using the pSEVA-615S tri-

parental cloning method. Complementation was performed both with the wildtype *nleG* gene and with the mutated gene encoding the $NleG1_{P163K}$ single point mutant that was shown in **Chapter 3** to be enzymatically inactive. With a catalogue of *nleG* mutant *C. rodentium* strains, it was possible to perform a range of murine infections to assess the contribution of NleG proteins to the infectious process *in vivo*. A graphical overview of the *in vivo* infection protocol and subsequent analyses performed in this part of the project is given in **Figure 4.1**.

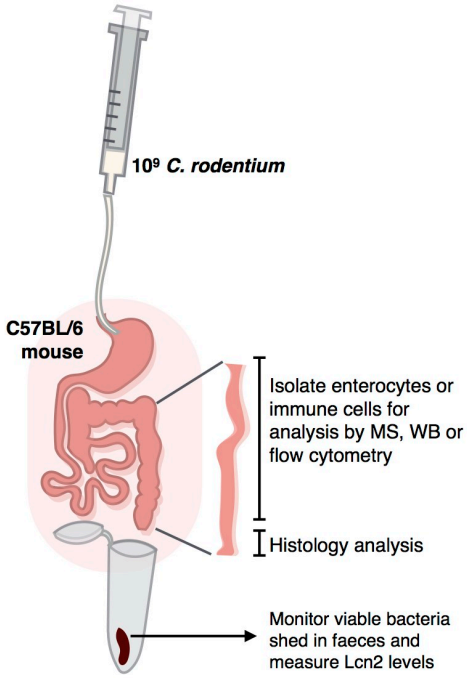


Figure 4.1| Overview of the *in vivo C. rodentium* infections and analyses used in this study. Female C57BL/6 mice were orally gavaged with 10^9 *C. rodentium* in 200 µl PBS. During the infection, viable bacteria shed in faeces were enumerated by resuspending faeces in PBS and plating serial dilutions on selective media. Faecal Lcn2, a biomarker of intestinal inflammation, was measured by ELISA. Mice were terminated by cervical dislocation and their colons collected for analysis, the distal 0.5 cm for histology analysis and a further 3-4 cm for isolation of enterocytes (for analysis by mass spectrometry, MS, or western blot, WB) or immune cells (for analysis by flow cytometry).

4.3. C. rodentium nleG deletion mutants are capable of pedestal formation during infection of cultured cells

Prior to testing new chromosomal mutants *in vivo*, their ability to grow in liquid media and to infect cultured cells *in vitro* must always be assessed. The strains generated here exhibited no growth defects in LB or M9 minimal media (growth curves shown later in chapter; **Figures 4.12** and **4.22**). To establish whether knocking out *nleG* genes had an impact on the ability of *C. rodentium* to adhere to cells and translocate effector proteins into infected cells, HeLa cells were infected with the new mutants. Infections were carried out for 2.5 to 3 hours and the ability of the mutants to trigger the formation of actin-rich pedestals, which requires a functional T3SS, was assessed by immunofluorescence microscopy.

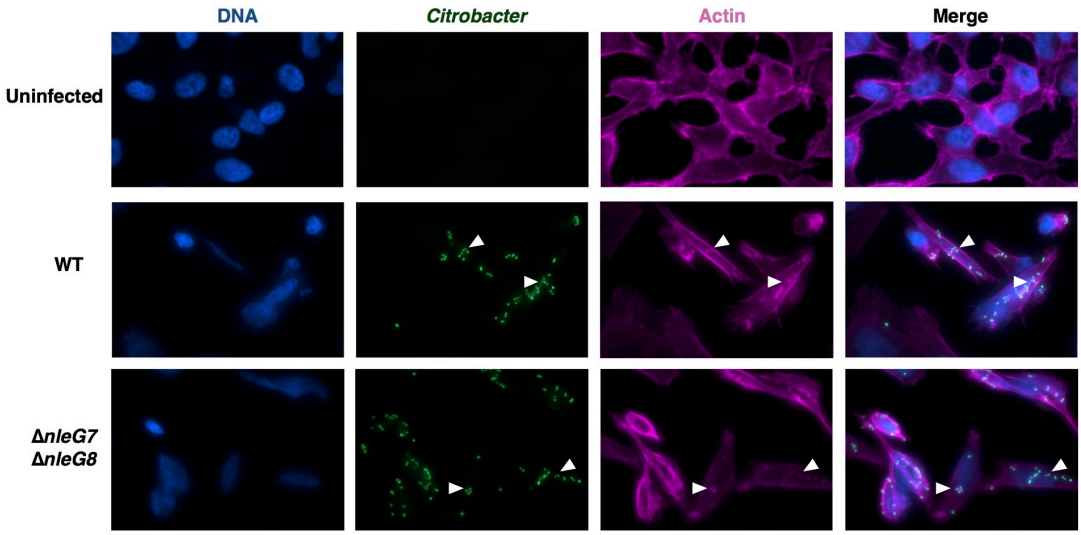


Figure 4.2 | A double *nleG* knockout *C. rodentium* strain (Δ*nleG7* Δ*nleG8*) causes pedestal formation during infection of HeLa cells. HeLa cells were infected for 2.5 hours with the WT and Δ*nleG7* Δ*nleG8 C. rodentium*. Cells were analysed by immunofluorescence. DNA in blue, *C. rodentium* in green and actin in pink. Actin polymerisation is observed under bacteria for WT and double *nleG* knockout *C. rodentium* (examples indicated by white arrowheads).

As shown in **Figure 4.2**, the double deletion mutation (Δ*nleG7* Δ*nleG8*) – the parental strain from which the triple deletion mutant was generated – was capable of pedestal formation during infection of HeLa cells. No other phenotypes (such as increased cell death or changes in cell morphology) were observed during the infections. When the triple *nleG* deletion mutant was used to infect HeLa cells, it also behaved like the WT strain in terms of pedestal formation and infection level (**Figure 4.3**). This indicated that the Δ*nleG1* Δ*nleG7* Δ*nleG8* mutant had a functional T3SS that was not blocked or otherwise compromised by the deletion of *nleG* genes.

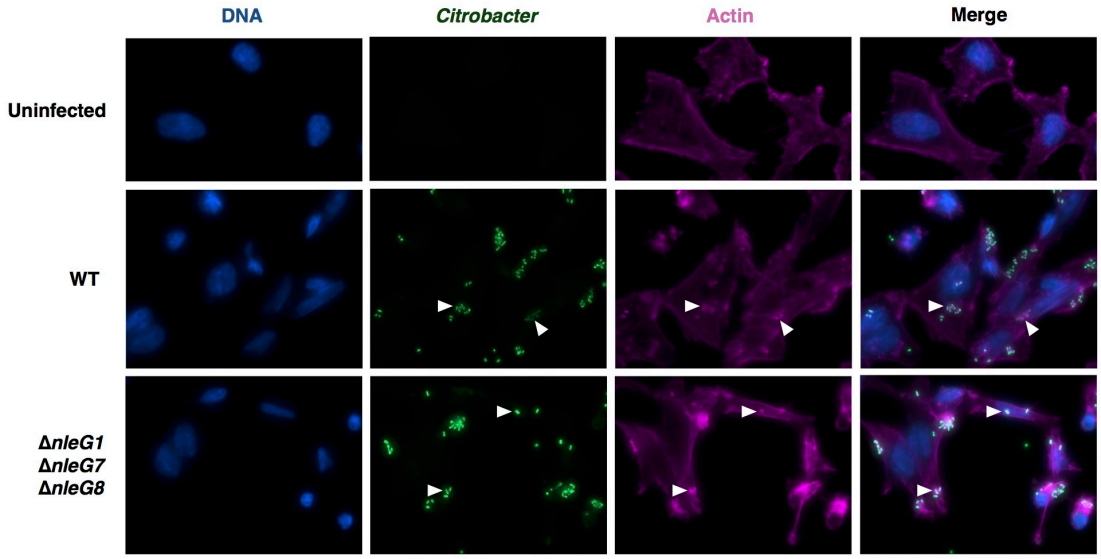


Figure 4.3 | A full *nleG* knockout *C. rodentium* strain (Δ*nleG1* Δ*nleG7* Δ*nleG8*) causes pedestal formation during infection of HeLa cells. HeLa cells were infected for 3 hours with the WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium*. Cells were analysed by immunofluorescence. DNA in blue, *C. rodentium* in green and actin in pink. Actin polymerisation is observed under bacteria for WT and *nleG* knockout *C. rodentium* (examples indicated by white arrowheads).

The newly generated single *nleG1* gene deletion strain was tested in the same way and displayed no obvious infection defect or inability to form pedestals (**Figure 4.4**). Similarly, we ensured that complementing the strain by re-introducing the *nleG* gene (but with a mutation to render it enzymatically inactive) into its native site had no impact on the infection *in vitro* (**Figure 4.4**). As expected, the infections caused by Δ*nleG1* and NleG1$_{P163K}$ strains were indistinguishable and we concluded that all *nleG* mutant strains generated in this study were suitable for use in subsequent *in vivo* infection experiments.

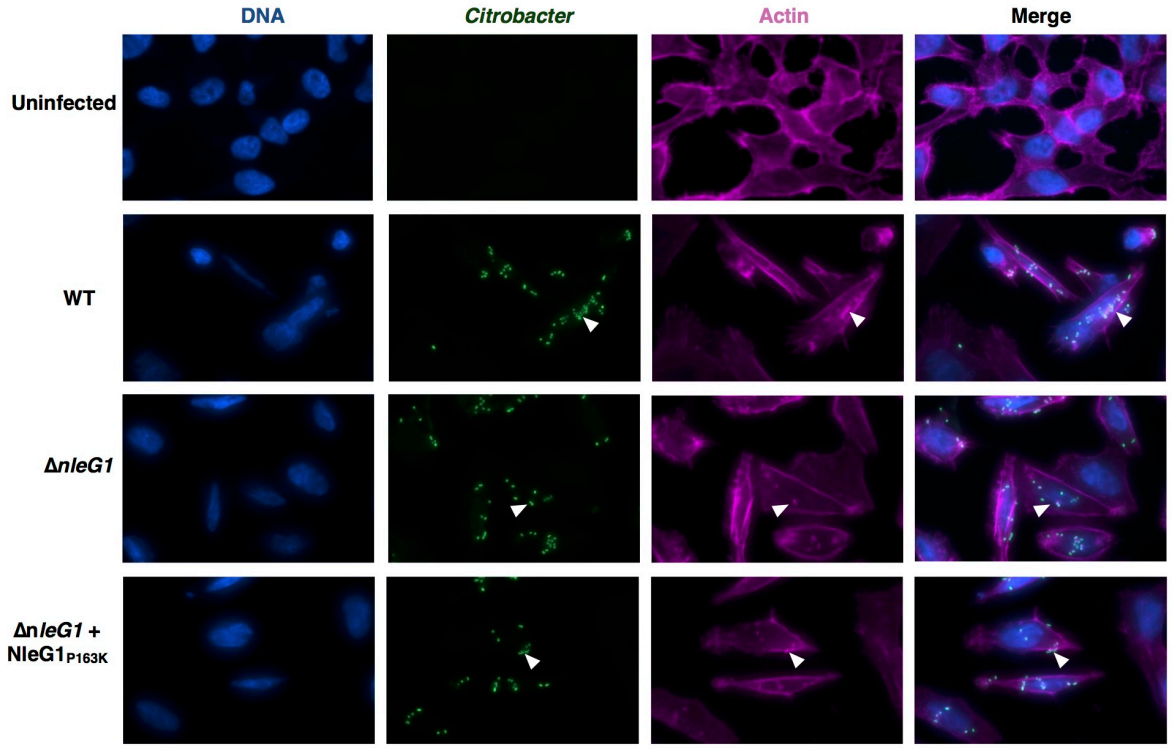


Figure 4.4 | Δ*nleG1* and a catalytically inactive NleG chromosomal mutant (NleG1$_{P163K}$) is capable of pedestal formation during infection of HeLa cells. HeLa cells were infected for 2.5 hours with the following *C. rodentium* strains: WT, Δ*nleG1* and Δ*nleG1* chromosomally complemented with the catalytically inactive NleG1$_{P163K}$. Cells were analysed by immunofluorescence. DNA in blue, *C. rodentium* in green and actin in pink. Actin polymerisation is observed under adherent bacteria for WT, Δ*nleG1* and Δ*nleG1* + NleG1$_{P163K}$ (examples indicated by white arrowheads).

4.4. Immune cell recruitment during C. rodentium infection

The *C. rodentium* model has been used to demonstrate that T3SS effector proteins can affect the recruitment of specific immune cells *in vivo* at defined time points post-inoculation (Crepin *et al.*, 2015). To assess whether bacterial colonisation or immune cell recruitment was altered in the absence of NleGs, we first had to characterise the WT *C. rodentium* infection and set up suitable staining panels and gating strategies for flow cytometry analysis. Therefore, an *in vivo* infection was performed and the colons of WT- or mock (PBS)-infected mice were collected for immunophenotyping at days 4, 8, 14 and 21 post-infection. The infection dynamics (**Figure 4.6**)

were as expected for WT *C. rodentium*: the numbers of viable bacteria shed in the stool peaked at day 8-10 post-infection, the infection started to clear at day 14 post-infection and viable *C. rodentium* were no longer recovered in the stool by day 21 post-infection.

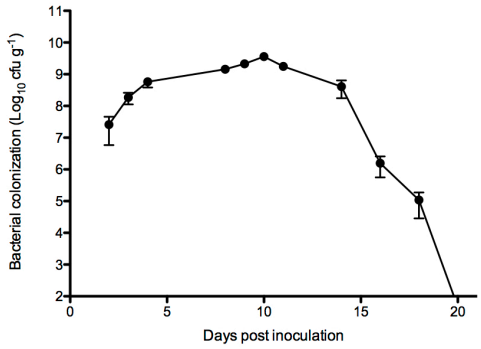


Figure 4.5 | Colonisation dynamics of WT *C. rodentium* during infection of C57BL/6 mice. C57BL/6 mice (n=20) were infected with WT *C. rodentium* ICC169 by oral gavage. Five animals were sacrificed for further analyses at day 4-, day 8-, day 14- and day 21 post-infection. Error bars represent ±SEM.

The collected colons were homogenised and the presence of specific cell surface markers of individual cells was evaluated by antibody staining and flow cytometry. The gating strategy for distinguishing between macrophages, dendritic cells and neutrophils is shown in **Figure 4.6**. Gating strategies for lymphocyte populations were also designed, shown in **Figure 4.7A-B**.

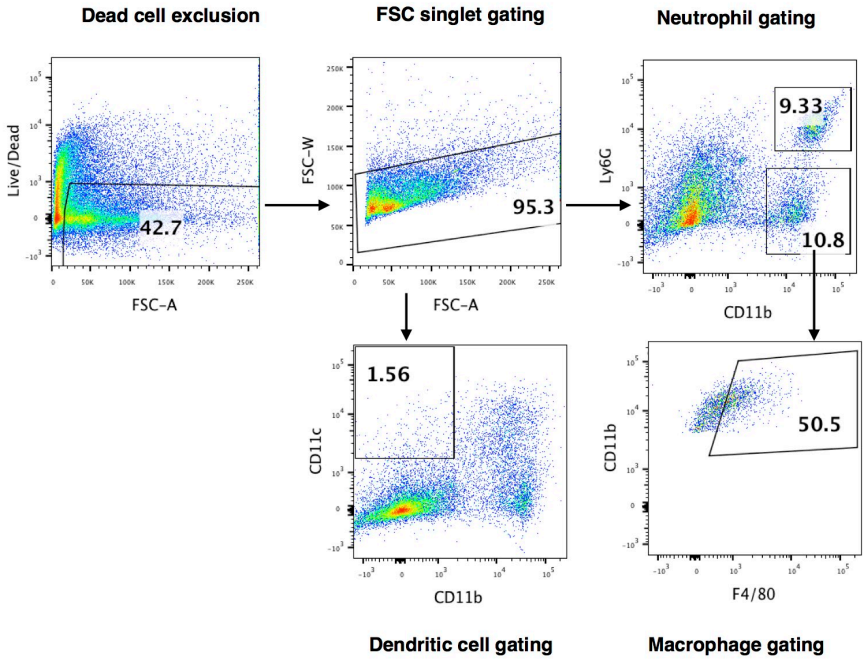


Figure 4.6 | The flow cytometry-based strategy for analysing myeloid cells collected from mouse intestinal tissue following infection with *C. rodentium*. The gating strategy for myeloid cells using representative contour plots to illustrate dead cell exclusion by Live/Dead viability staining, exclusion of doublets, and gating based on the expression of cell surface markers CD11c, CD11b, Ly6G and F4/80 to distinguish between dendritic cells ($CD11c^+$ $CD11b^-$), neutrophils ($Ly6G^+$ $CD11b^+$) and macrophages ($Ly6G^-$ $CD11b^+$ $F4/80^+$). Numbers represent percentages of total cells within the parent gate.

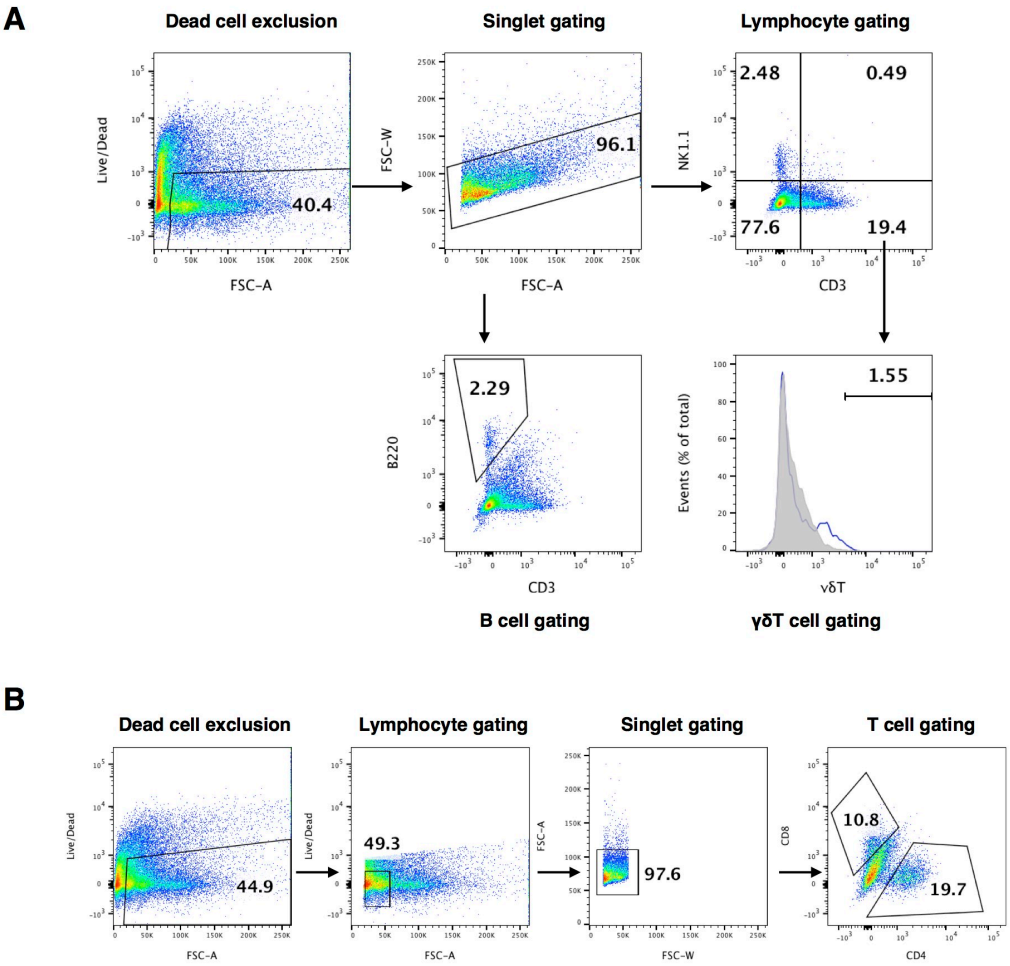


Figure 4.7 | The flow cytometry-based strategy for analysing lymphocytes collected from mouse intestinal tissue following infection with *C. rodentium.* A) Representative contour plots to illustrate dead cell exclusion by Live/Dead viability staining, exclusion of doublets, and gating based on the expression of cell surface markers CD3, B220, NK1.1 and γδ T to distinguish between B cells (defined as $B220^{+}$ $CD3^{-}$), NK cells ($NK1.1^{+}$ $CD3^{-}$), NKT cells ($NK1.1^{+}$ $CD3^{+}$) and T cells ($NK1.1^{-}$ $CD3^{+}$). T cell were further gated on γδ T cell receptor expression to assess γδ-T cell frequency compared to the fluorescence minus one (FMO) negative control (shown in grey). **B)** The gating strategy for $CD4^{+}$ $CD8^{-}$ and $CD4^{-}$ $CD8^{+}$ cells. Representative contour plots illustrate dead cell exclusion by Live/Dead viability staining, lymphocyte gating based on FSC-A, singlet gating, and gating based on the expression of cell surface markers CD4 and CD8. Numbers represent percentages of total cells within the parent gate.

The relative frequencies of macrophages (defined as $Ly6G^{-}$ $CD11b^{+}$ $F4/80^{+}$), dendritic cells ($CD11c^{+}$ $CD11b^{-}$) and neutrophils ($Ly6G^{+}$ $CD11b^{+}$) are known to become elevated during *C. rodentium* infection, starting to increase within days of infection and remaining high until day 14 or 21 post-infection (Crepin *et al.*, 2015; Manta *et al.*, 2013; Ale *et al.*, 2017). In this experiment, we saw a significant enrichment of the dendritic cell population in infected colons at day 14 and day 21 post-infection (**Figure 4.8A**). Similarly, an increase was observed for macrophages and neutrophils, reaching statistical significance on day 14 and day 21 post-infection (**Figures 4.8B-C**).

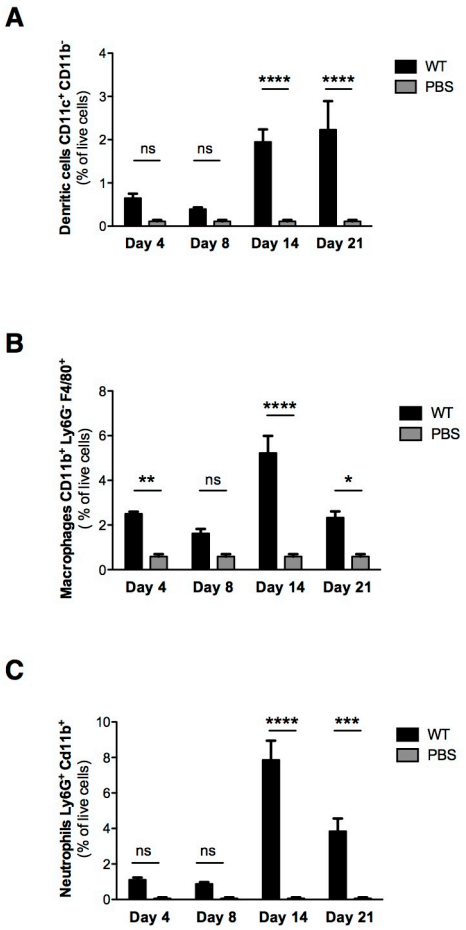


Figure 4.8 | Changes in myeloid cell populations in the mouse intestine during *C. rodentium* infection. Colons of mock-infected (PBS) or WT *C. rodentium*-infected mice (n=4-5 per timepoint) were homogenised and immunophenotyped by flow cytometry at defined days post-infection. **A)** The relative frequency of neutrophils, defined as $Ly6G^+$ $CD11b^+$ cells. **B)** The relative frequency of macrophages, defined as $Ly6G^-$ $CD11b^+$ $F4/80^+$. **C**) The relative frequency of dendritic cells, defined as $CD11c^+$ $CD11b^-$. Error bars represent ±SEM; statistical significance assessed by two-way ANOVA with Bonferroni's post-test; ns = non-significant, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

Importantly, a pronounced Th1-biased T cell response is required for clearing *C. rodentium* from the murine gut and antigen-specific $CD4^+$ T cells are pivotal for sterilising immunity. Comparing the frequency of T cells in gut tissue from *C. rodentium*-infected and uninfected mice confirmed an overall increase in T cells ($NK1.1^-$ $CD3^+$) from day 8 post-infection, increasing further until day 21 (**Figure 4.9D**). Looking at $CD4^+$ and $CD8^+$ T cells individually supported an increase in their relative frequencies (significant by day 14 post-infection; **Figure 4.9F-G**). B cells ($B220^+$ $CD3^-$) were enriched as early as day 4 post-infection (slightly earlier than expected; **Figure 4.9A**) and NK cells ($NK1.1^+$ $CD3^-$) on day 8, increasing further during the course of the infection (**Figure 4.9B**). An increase in NKT cells ($NK1.1^+$ $CD3^+$) was also observed on day 21 post-infection (**Figure 4.9C**). In recent years, it has become clear that γδ-T cells contribute to immunity against enteric pathogens, but there was no difference in their relative frequency in this experiment (**Figure 4.9E**).

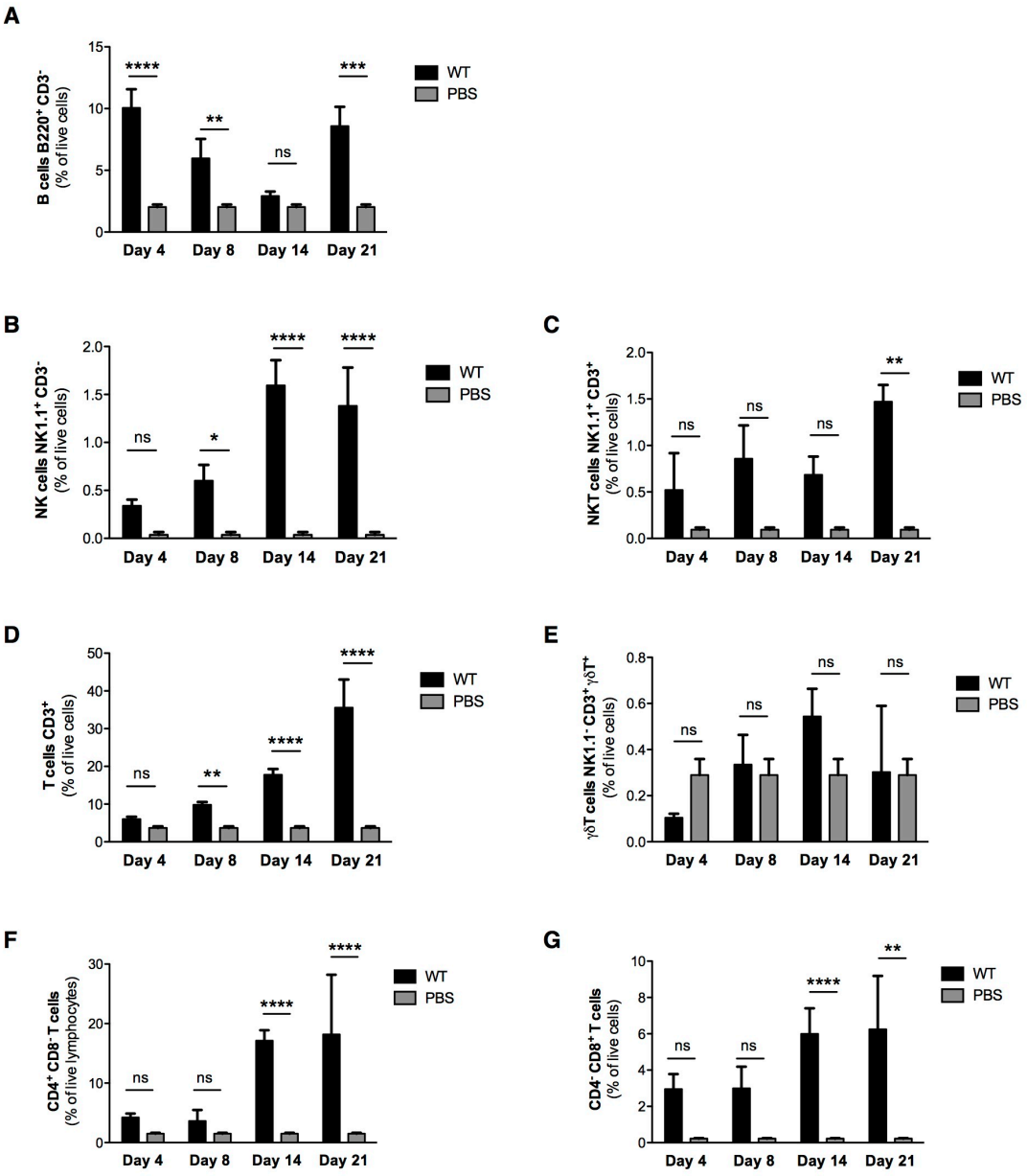


Figure 4.9 | Changes in immune cell populations during *C. rodentium* infection *in vivo*. Colons of mock-infected (PBS) or WT *C. rodentium*-infected mice (n=4-5 per timepoint) were homogenised and immunophenotyped by flow cytometry at defined days post-infection. The expression of cell surface markers CD11c, CD11b, Ly6G and F4/80 was used to distinguish between neutrophils ($Ly6G^+$ $CD11b^+$) and macrophages ($Ly6G^-$ $CD11b^+$ $F4/80^+$). The cell surface markers CD3, B220, NK1.1 and γδ T were used to distinguish between B cells ($B220^+$ $CD3^-$), NK cells ($NK1.1^+$ $CD3^-$), NKT cells ($NK1.1^+$ $CD3^+$), T cells ($NK1.1^-$ $CD3^+$) and γδ T cells ($NK1.1^-$ $CD3^+$ $\gamma\delta^+$). CD4 and CD8 were used to distinguish between $CD4^+$ cells ($CD4^+$ $CD8^-$) and $CD8^+$ cells ($CD4^-$ $CD8^+$). **A)** The relative frequency of B cells. **B)** The relative frequency of NK cells. **C)** The relative frequency of NKT cells. **D)** The relative frequency of T cells. **E)** The relative frequency of γδ T cells. **F)** The relative frequency of $CD4^+$ $CD8^-$ cells. **G)** The relative frequency of $CD4^-$ $CD8^+$ cells. Error bars represent ±SEM; statistical significance tested by two-way ANOVA with Bonferroni's post-test; ns = non-significant, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$.

4.5. The role of NleG1 in C. rodentium colonisation

Having described the WT *C. rodentium* infection *in vivo* both in terms of colonisation dynamics and increase in relative frequencies of specific immune cells, the focus was shifted to the contribution of NleG proteins to the infection. The contribution of NleG1 to was considered first. The *in vivo* impact of knocking out *nleG1* was investigated by orally inoculating C57BL/6 mice with WT and Δ*nleG1 C. rodentium* and enumerating the viable bacteria shed with the stools. The Δ*nleG1* strain displayed similar colonisation dynamics to wildtype *C. rodentium* until the peak of infection (day 8 post-infection, see **Figure 4.10**). A small number of mice were used for this early experiment, but the mice were nevertheless consistently colonised to expected WT-levels.

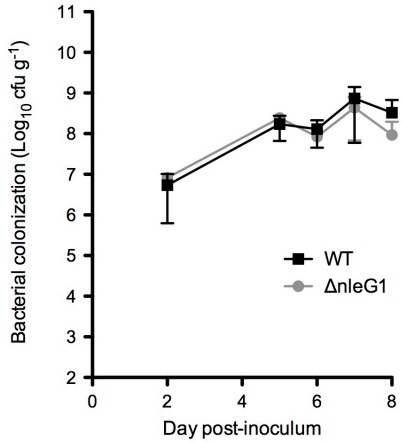


Figure 4.10| Mice infected with the *C. rodentium nleG1* deletion mutant show similar levels of bacterial colonisation as wild-type (WT) infected mice. C57BL/6 mice (n=5 for WT, n=2 for Δ*nleG1*) were infected with *C. rodentium* by oral gavage. The colonisation dynamics up to day 8 post-infection were identical between the WT and Δ*nleG1* strains. Error bars represent ±SEM.

4.6. The role of NleG1 in immune cell recruitment in response to C. rodentium infection

The identical colonisation between Δ*nleG1* and WT *C. rodentium* was a prerequisite for being able to compare immune cell recruitment between Δ*nleG1* and WT *C. rodentium*-infected tissues. To assess the impact of knocking out *nleG1* on early immune cell recruitment, immune cell populations in the colon were analysed by flow cytometric immunophenotyping using the gating strategies established in **Section 4.4**. Even though changes in certain immune cell populations had been subtle on day 8 post-infection (**Figure 4.9**), we hypothesised that if NleG1 were involved in down-regulating specific immune cell recruitment during infection, the Δ*nleG1 C. rodentium* strain should elicit more pronounced immune cell recruitment compared to WT from this time-point onwards. However, there were no obvious differences between the *nleG1*

knockout and WT group for the myeloid cell populations (**Figure 4.11A-B**), indicating that NleG1 does not target their recruitment to the site of infection during the early immune cell response against *C. rodentium*. Similarly, comparing the relative frequencies of T cells revealed no differences for overall T cell frequency (NK1.1$^-$ CD3$^+$), CD4$^+$ lymphocytes (CD4$^+$ CD8$^-$) or CD8$^+$ T lymphocytes (CD4$^+$ CD8$^-$) (**Figure 4.11C-D** and **Figure 4.11G**). In line with this, there were no differences in the recruitment of NK cells (NK1.1$^+$ CD3$^-$), B cells (B220$^+$ CD3$^-$) or γδ-T cells (NK1.1$^-$ CD3$^+$ γδ$^+$) between the WT and Δ*nleG1 C. rodentium*-infected tissues.

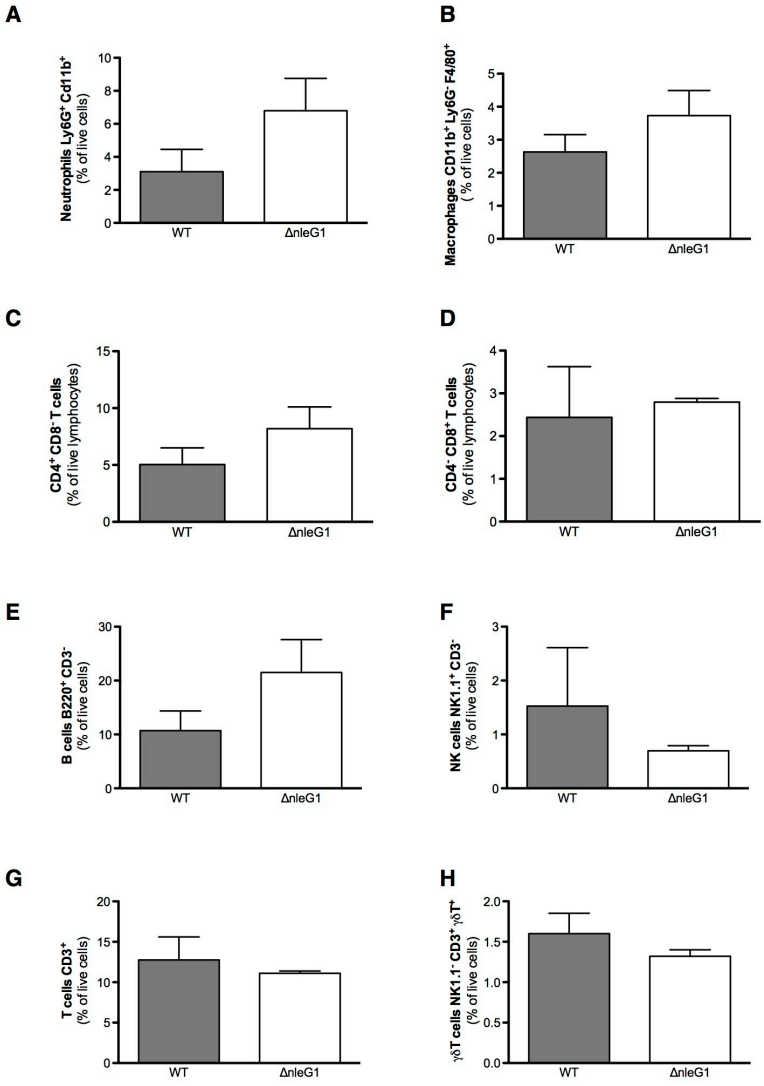


Figure 4.11 | *In vivo* infection with the *C. rodentium nleG1* knockout mutant does not significantly alter the recruitment of myeloid cells or lymphocytes in the mouse intestine. Colons of *C. rodentium*-infected mice were homogenised and immunophenotyped by flow cytometry. The expression of cell surface markers CD11c, CD11b, Ly6G and F4/80 was used to distinguish between neutrophils (Ly6G$^+$ CD11b$^+$) and macrophages (Ly6G$^-$ CD11b$^+$ F4/80$^+$). The cell surface markers CD3, B220, NK1.1 and γδ T were used to distinguish between B cells (B220$^+$ CD3$^-$), NK cells (NK1.1$^+$ CD3$^-$), NKT cells (NK1.1$^+$ CD3$^+$), T cells (NK1.1$^-$ CD3$^+$) and γδ T cells (NK1.1$^-$ CD3$^+$ γδ$^+$). CD4 and CD8 were used to distinguish between CD4$^+$ cells (CD4$^+$ CD8$^-$) and CD8$^+$ cells (CD4$^-$ CD8$^+$). **A)** The relative frequency of neutrophils. **B)** The relative frequency of macrophages. **C)** The relative frequency of CD4$^+$ CD8$^-$ cells. **D)** The relative frequency of CD4$^-$ CD8$^+$ cells. **E)** The relative frequency of B cells. **F)** The relative frequency of NK cells. **G)** The relative frequency of T cells. **H)** The relative frequency of γδ T cells. Error bars represent ±SEM.

4.7. The role of NleG1 in Lcn2 release during C. rodentium infection

The next *in vivo* experiment was a repeat of the Δ*nleG1 C. rodentium* infection, but also encompassed the NleG1$_{P163K}$ genomic mutant strain that expresses E3 ligase-dead NleG1 *in lieu* of WT NleG1 (Δ*nleG1* + NleG1$_{P163K}$). It was included as an additional control, so that if a phenotype were found for the Δ*nleG1* mutant, it should be phenocopied by the NleG1$_{P163K}$ mutant if the phenotype resulted from the lack of NleG1 E3 ligase activity. Prior to the murine infection, the growth of the strains was monitored by growth curves in M9 minimal media, which confirmed that the mutant strains had no growth defects *in vitro* (**Figure 4.12**). The ability of the NleG1$_{P163K}$ mutant to infect cultured cells and trigger the formation of actin pedestals was also verified prior to *in vivo* experiments (data not shown).

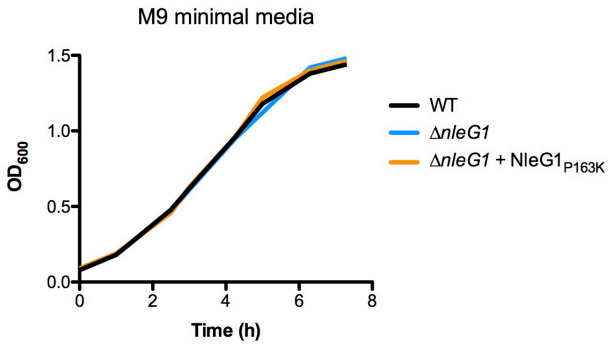


Figure 4.12 | *C. rodentium* Δ*nleG1* and Δ*nleG1* chromosomally complemented with NleG1$_{P163K}$ (catalytically inactive) have no growth defect in liquid minimal media. The growth of WT, and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* was assessed by measuring the OD_{600} at regular intervals during growth in M9 minimal media. All strains tested produced identical growth curves.

C57BL/6 mice were infected with WT, Δ*nleG1* and NleG1$_{P163K}$ *C. rodentium*. As before, the Δ*nleG1* mutant behaved like the WT strain in terms of bacterial colonisation as enumerated by plating serial dilutions of homogenised faeces (**Figure 4.13A**). This was also true for NleG1$_{P163K}$ *C. rodentium.* Next, we tested if the lack of NleG1 or NleG1 catalytic activity had an impact on colonic crypt hyperplasia – the hallmark pathological feature of *C. rodentium* infection in mice. Hematoxylin and eosin (H&E)-staining of fixed colonic section was performed to see if there were any changes in colonic crypt lengths and epithelial layer architecture (as described in Crepin *et al.*, 2016). Measurements of colonic crypt lengths revealed that WT, Δ*nleG1* and NleG1$_{P163K}$ *C. rodentium* infections resulted in significant elongation of colonic crypt compared to the PBS mock-infected control (**Figure 4.13B**; representative images in **Figure 4.13C**). The colonic crypt hyperplasia caused by Δ*nleG1 C. rodentium* was indistinguishable from WT-levels. Intriguingly, NleG1$_{P163K}$ *C. rodentium* gave a slight but statistically significant increase in colonic crypt hyperplasia compared to WT and Δ*nleG1* strains.

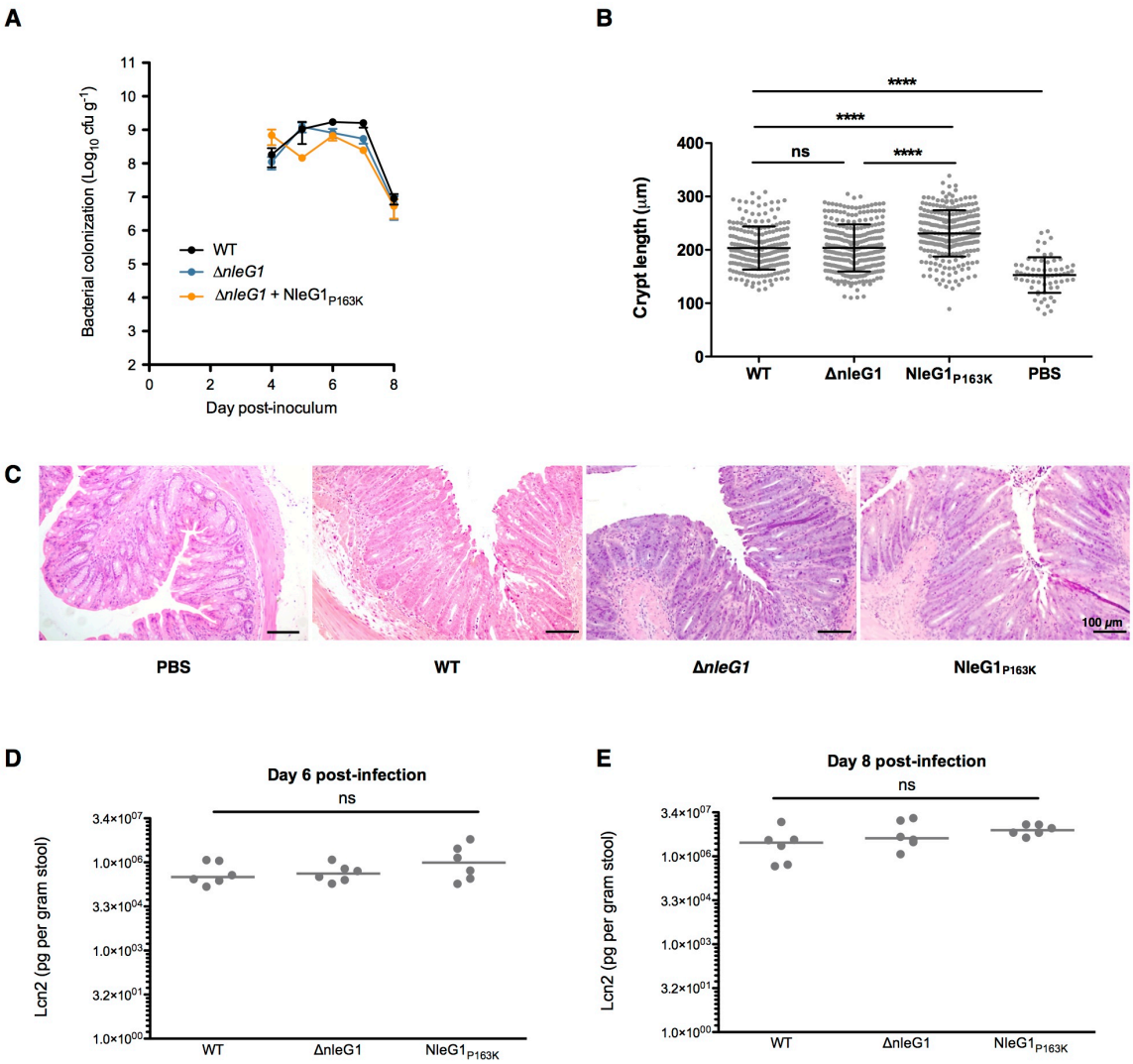


Figure 4.13 | *C. rodentium* lacking NleG1 or NleG1 catalytic activity colonises the murine gut to WT-levels and does exhibit altered stool concentrations of Lcn2, a biomarker of intestinal inflammation. Mice (n=6 per treatment group, n=4 for PBS mock-infected group) were infected with WT *C. rodentium*, a *ΔnleG1* mutant or a NleG1$_{P163K}$ chromosomal mutant (*ΔnleG1* chromosomally complemented with catalytically inactive NleG1). **A)** All *C. rodentium* infection groups displayed similar colonisation dynamics, measured by two-way ANOVA and Bonferroni's post-test. **B)** Mice were sacrificed at day 8 post-infection and the distal colon analysed for colonic crypt hyperplasia by H&E staining. Measurements of crypt lengths revealed no difference between WT *C. rodentium* and the *ΔnleG1* mutant. A significant increase in hyperplasia was seen for the NleG1$_{P163K}$ mutant. All mutants and WT *C. rodentium* had increased colonic crypts lengths compared to the PBS mock-infected control. Statistical differences determined by one-way ANOVA and Tukey's post test. **C)** Representative images from B. Scale bar 100 µm. **D-E)** The concentration of Lcn2 in stool supernatant was measured by ELISA for day 6 and 8 post-infection, but showed no significant differences between the groups (as measured by one-way ANOVA and Tukey's post test). Line represents median Lcn2 concentration. Error bars represent ±SEM. ns = non-significant, ****p ≤ 0.0001.

Instead of performing immunophenotyping of colonic tissue in this experiment, the inflammatory state of the intestine was assessed non-invasively. Recent reports had highlighted ELISA-based quantification of Lcn2 levels in faecal matter as a rapid, stable and cost-effective way of assessing low- as well as high-level gut inflammation in a murine model (Chassaing *et al.*, 2012; Abella *et al.*, 2015). Lcn2 is secreted by neutrophils and host epithelial cells (including

enterocytes) to bind the bacterial siderophore enterobactin (Ent), disrupting bacterial acquisition of ferric Ent and thereby compromising its survival in the gut environment. Using Lcn2 as a biomarker of inflammation could reveal subtle differences in intestinal inflammation that could be informative in the context of *C. rodentium* infection both with regards to this project and beyond the study of NleG proteins. Measuring the Lcn2 concentration in the stools of infected mice by ELISA showed that Lcn2 concentration increased as the infection progressed (between day 6 and day 8; **Figure 4.13D-E**). However, there were no significant differences between the treatments groups. This was in accordance with the previous immunophenotyping experiment, suggesting that the lack of NleG1 – or the lack of *only* NleG1 and not the other NleG proteins – does not have an overall impact on the inflammatory state of the intestine at the peak of *C. rodentium* infection.

4.8. A TMT 10plex proteomics approach to study the impact of NleG1 on the host proteome during in vivo infection

Phenotypes conferred by individual T3SS effectors *in vivo* can be subtle and difficult to ascertain. One way to start inferring functions of effectors is to compare proteomes of cells infected with a WT and mutant bacterial strain. This has been done successfully for bacterial E3 ligases in the literature *in vitro*. Moreover, advances in isobaric labelling techniques and mass spectrometry analyses has allowed such analyses to be performed on tissues from *in vitro* infection samples (Berger *et al.*, 2009a; Pollard *et al.*, 2018). In collaboration with Dr Jyoti Choudhary at the Institute of Cancer Research, London, we combined our murine infections with high sensitivity TMT 10plex mass spectrometry analysis. The goal was to compare proteomes of enterocytes infected *in vivo* with WT, Δ*nleG1* or NleG1$_{P163K}$ *C. rodentium* to gain insight into which cellular pathways are targeted by NleG1. Such comparisons seemed particularly fitting to the study of a bacterial E3 ubiquitin ligase, as the ubiquitination of host proteins often involves a change in the abundance of the target protein or in the abundance of proteins further downstream in the targeted cellular pathway.

An overview of the strategy for the TMT 10plex approach is shown in **Figure 4.14B**. Mice were infected with WT, Δ*nleG1* and NleG1$_{P163K}$ *C. rodentium* (see **Figure 4.14A** for bacteria shed with faeces) alongside an uninfected control and a further six *C. rodentium* mutant infections from colleagues in the lab. A total of 10 different infection groups were thereby generated, suitable for the 10plex workflow. Enterocytes were isolated from the collected tissues at day 8 post-infection. Isobaric peptide labelling (tandem mass tag, TMT) and MS3 quantification was used,

which has previously been shown to reproducibly detect changes in protein abundance between infected and mock-infected enterocytes (Berger *et al.*, 2017).

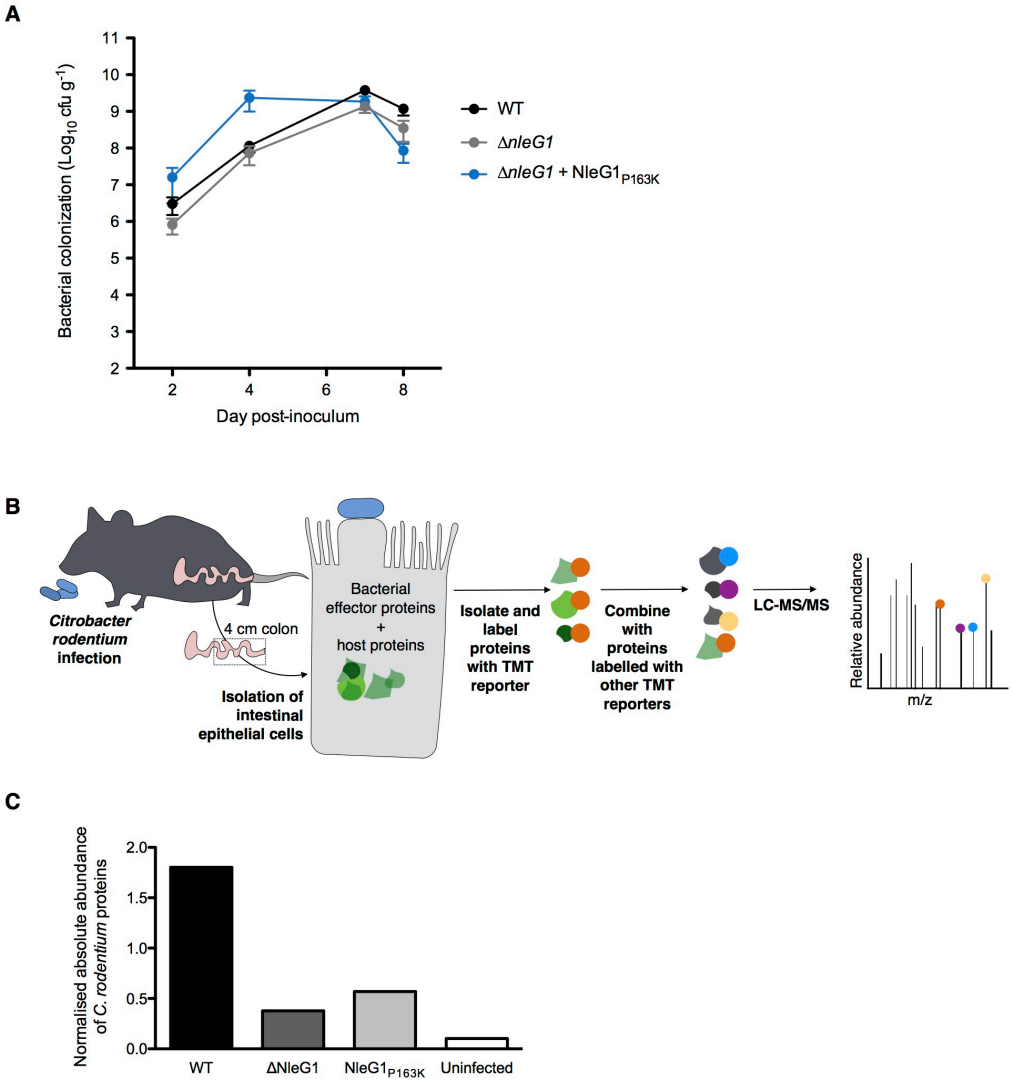


Figure 4.14 | A mass spectrometry-based approach to compare proteomes of enterocytes from mice infected with WT, ΔnleG1 or NleG1$_{P163K}$-expressing *C. rodentium*. A) Mice (n=5 per treatment group) were infected with WT *C. rodentium*, a *ΔnleG1* mutant or a NleG1$_{P163K}$ chromosomal mutant (*ΔnleG1* chromosomally complemented with catalytically inactive NleG1). All groups displayed similar colonisation dynamics and levels of colonisation, with no significant difference as measured by two-way ANOVA with Bonferroni's post-test. **B)** Schematic representation of the experimental approach using TMT10plex. Mice were infected with different strains of *C. rodentium* alongside an uninfected control (for a total of 10 treatment groups). On day 8 post-infection, mice were sacrificed and 4 cm of the distal colon was collected. Enterocytes were isolated from each colon and grouped based on the treatment condition. Cells were lysed and proteins labelled with isobaric tandem-mass-tag reporter tags (of distinct 13C and 15N isotope composition). The 10 resulting samples were mixed before sample fractionation and clean up, followed by analysis by high resolution LC-MS/MS. Data were then analysed to identify peptides and quantify reporter ion relative abundance. **C)** Comparison of the normalised absolute abundance of *C. rodentium* proteins in the samples relevant to this project revealed very low levels of *C. rodentium* proteins in the *ΔnleG1* mutant group and NleG1$_{P163K}$ chromosomal mutant group compared to the WT group.

Approximately 7400 mouse proteins were identified and quantified by shotgun proteomics analysis of labelled enterocyte proteins (with a FDR-corrected *P* value <0.01). Almost 3400 proteins abundances were changed 1.5-fold up or down in WT-infected enterocytes compared to uninfected enterocytes, highlighting the significant altering of a host proteome following enteric infection and the complexity of the infectious process. Prior to analysis of changes in protein abundances in the NleG1 groups and further gene ontology (GO) term analyses, the abundance of *C. rodentium* proteins in each sample was assessed. This revealed that despite only a small statistically insignificant difference in colonisation levels between WT, Δ*nleG1* and $NleG1_{P163K}$ *C. rodentium* at day 8 post-infection (**Figure 4.14A**), there was a pronounced difference in the normalised absolute abundance of *C. rodentium proteins* in the samples (**Figure 4.14C**). The abundance of *C. rodentium* proteins in the Δ*nleG1* and $NleG1_{P163K}$ samples were closer the uninfected control than to the WT sample. Therefore, further analyses of these samples were abandoned as the level of infection must be comparable between the samples. Due to the time constraints of this project, the experiment could regrettably not be repeated.

4.9. The role of NleG8 in C. rodentium colonisation

In parallel to studying NleG1 *in vivo*, animal experiments were conducted with the Δ*nleG8 C. rodentium* strain to assess the contribution of NleG8 to the hallmark pathologies of *C. rodentium*. Mice were infected with WT and Δ*nleG8 C. rodentium* by oral gavage and the viable bacteria shed in the stool enumerated as previously described. A Δ*nleG8* strain transformed with a bacterial expression vector pACYC encoding WT NleG8 was also included, as this would rescue any phenotypic changes in colonisation caused by the lack of NleG8. The Δ*nleG8* strain displayed similar colonisation levels as wildtype *C. rodentium* until day 6 post-infection, with a slight decrease in colonisation at day 10 that did not reach statistical significance (**Figure 4.15**). Importantly, expressing NleG8 from pACYC did not increase the *C. rodentium* CFU at day 10 post-infection, suggesting that the loss of NleG8 does not significantly impair the ability of *C. rodentium* to colonise the murine gut up to day 10 post-infection.

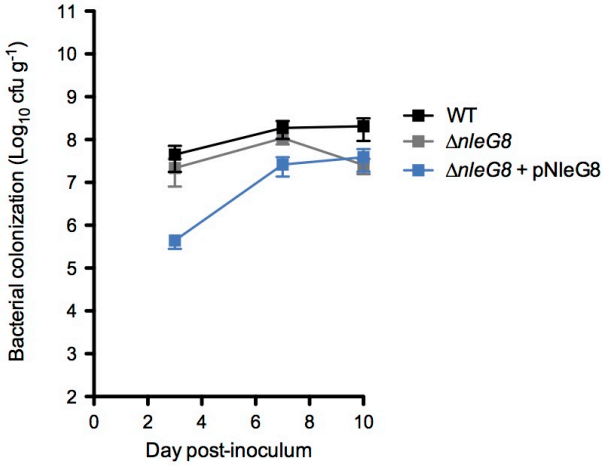


Figure 4.15 | Mice infected with the *C. rodentium nleG8* deletion mutant show similar levels of bacterial colonisation as WT infected mice. C57BL/6 mice (n=5 for all groups) were infected with *C. rodentium* ICC169 (WT), Δ*nleG8* and Δ*nleG8* expressing NleG8 from the pACYC plasmid (Δ*nleG8* + pNleG8) by oral gavage. There were no statistically significant differences between the number of viable bacteria between the groups as measured by two-way ANOVA with Bonferroni's post-test. Error bars represent ±SEM.

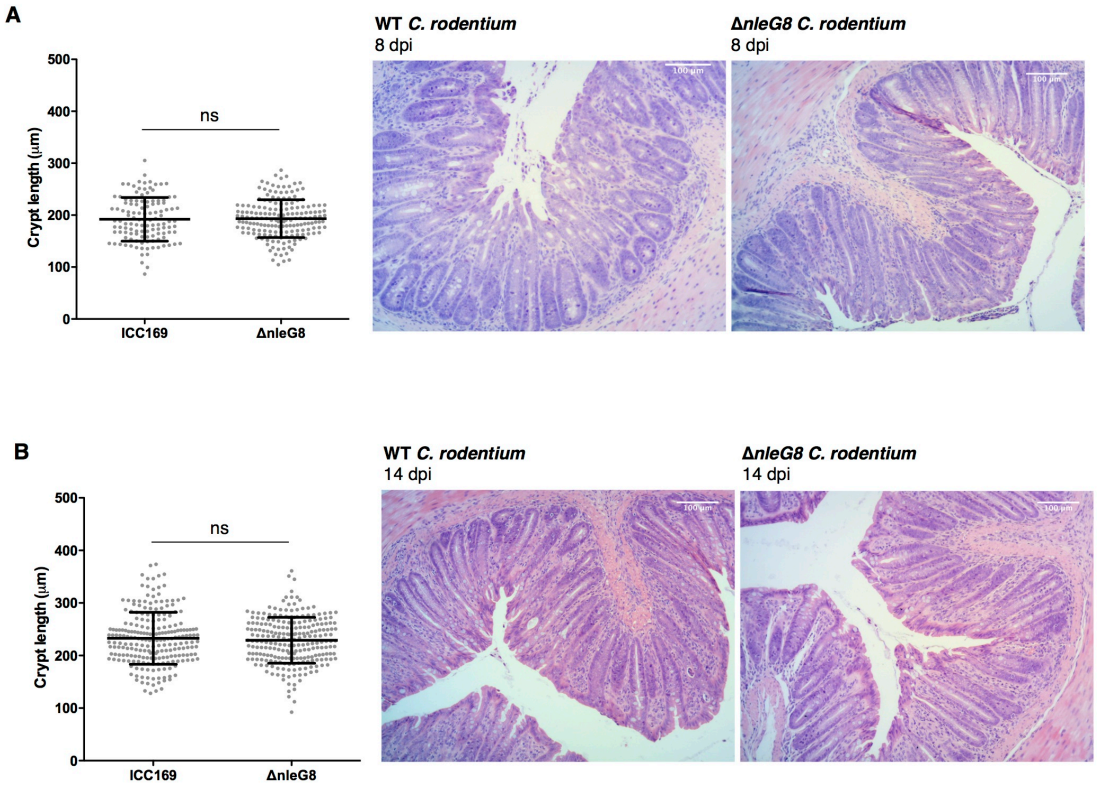


Figure 4.16 | Mice infected with the *C. rodentium nleG8* deletion mutant develop WT-levels of colonic crypt hyperplasia. C57BL/6 mice (n=5 for all groups) were infected with WT and Δ*nleG8 C. rodentium* ICC169 by oral gavage. Mice were sacrificed at the peak of infection (8 days post-infection; dpi) or at the peak of colonic crypt hyperplasia (14 dpi). The distal colon was analysed for colonic crypt hyperplasia by H&E staining. **A)** Measurements of crypt lengths revealed no difference between the wild-type and Δ*nleG8 C. rodentium* at 8 dpi; representative images on the right. **B)** Measurements of crypt lengths revealed no difference between the wild-type and Δ*nleG8 C. rodentium* at 14 dpi; representative images on the right. Error bars represent ±SEM, ns = non-significant by Students t-test. Scale bar for representative measurements is 200 µm.

4.10. C. rodentium lacking nleG8 induces colonic crypt hyperplasia

Next, the ability of the Δ*nleG8 C. rodentium* mutant to induce colonic crypt hyperplasia was tested. The most distal 0.5 cm segment of the colon was collected and fixed at either day 8 or day 14 post-infection. Measurements of H&E stained crypts showed no difference in average colonic crypt length between WT and Δ*nleG8 C. rodentium*-infected colonic tissue on day 8 post-infection (**Figure 4.16A**) or on day 14 post-infection (**Figure 4.16B**). As Δ*nleG8* induced colonic crypt hyperplasia to WT-levels, we concluded that NleG8 does not contribute to triggering or regulating the induction of colonic crypt hyperplasia or altering the crypt architecture.

4.11. The effect of nleG8 deletion on myeloid cell recruitment

The ability of the Δ*nleG8 C. rodentium* mutant and the complemented mutant to trigger immune cell infiltration was then examined by a flow cytometric approach as for NleG1 in **Section 4.6**. The same staining panels and gating strategies as before were employed for colonic tissues collected and homogenised on day 8 post-infection. There were no significant changes in relative frequencies of immune cells between the *C. rodentium* strains considered, as shown in **Figure 4.17** for neutrophils, macrophages and dendritic cells.

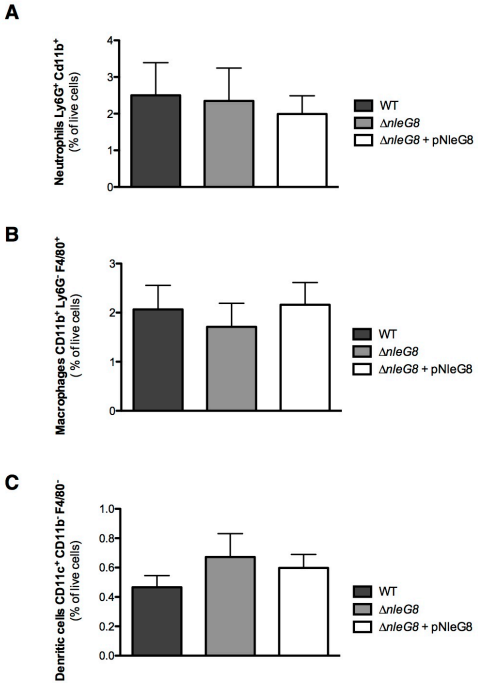


Figure 4.17 | Intestines of mice infected with the *C. rodentium nleG8* deletion mutant and plasmid-complemented mutant have the same levels of myeloid cell recruitment as WT-infected mice. C57BL/6 mice (n=5 for all groups) were infected with *C. rodentium* ICC169 (WT), Δ*nleG8* and Δ*nleG8* expressing NleG8 from the pACYC plasmid (Δ*nleG8* + pNleG8) by oral gavage. On day 8 post-infection, their intestines were analysed for immune cells by flow cytometry. **A)** The relative frequency of neutrophils. **B)** The relative frequency of macrophages. **C**) The relative frequency of dendritic cells. There were no significant differences between the groups for any cell-type, as measured by ANOVA with Tukey's post test. Error bars represent ±SEM.

4.12. *The impact of NleG8 overexpression on C. rodentium colonisation*

The previous chapter in this thesis (**Chapter 3**) established that NleG8 is the only NleG homologue in *C. rodentium* that confers cytotoxicity in petite yeast dependent on its E3 ligase activity. Knowing that expression of NleG8 can cause cell death in eukaryotic cells, it was tempting to speculate that overexpression of NleG8 in *C. rodentium* might somehow affect the ability of the pathogen to colonise the mouse gut or alter the colonisation dynamics. Therefore, WT *C. rodentium* transformed with empty pACYC, pACYC-NleG8 or pACYC-NleG8$_{P177K}$ were compared in an *in vivo* mouse infection experiment. This showed that the colonisation efficiency and dynamics were identical between the strains (**Figure 4.18**), regardless of the overexpression of NleG8. Further experiments therefore focused on single, double and triple *nleG* deletion strains.

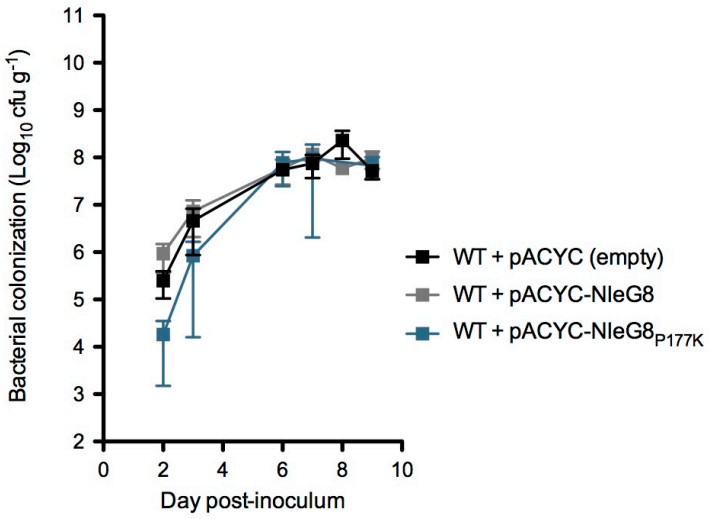


Figure 4.18 | Expressing NleG8 or catalytically inactive NleG8$_{P177K}$ from the pACYC plasmid does not impact the ability of *C. rodentium* to colonise the murine gut. C57BL/6 mice (n=5 per group) were infected with *C. rodentium* ICC169 (WT) transformed with pACYC (empty), pACYC-NleG8 or pACYC-NleG8$_{P177K}$. There were no significant differences between the number of viable bacteria between the groups as measured by two-way ANOVA with Bonferroni's post-test. Error bars represent ±SEM. Plating of serial dilutions was performed on Nal plates and Tet plates in parallel to ensure plasmid carriage throughout the experiment.

4.13. *ΔnleG7 ΔnleG8 C. rodentium stimulates less Lcn2 release than WT C. rodentium*

Next, an *in vivo* experiment to compare the three single *nleG* deletion mutants (Δ*nleG1*, Δ*nleG7*, and Δ*nleG8*) and the Δ*nleG7* Δ*nleG8* double *nleG* deletion mutant to WT *C. rodentium* was performed. No significant differences in bacterial colonisation was observed between the strains until day 8-post-infection, when the mice were sacrificed (**Figure 4.19A**). At the end of the experiment, the colons of each mouse were collected, sectioned and analysed by H&E staining.

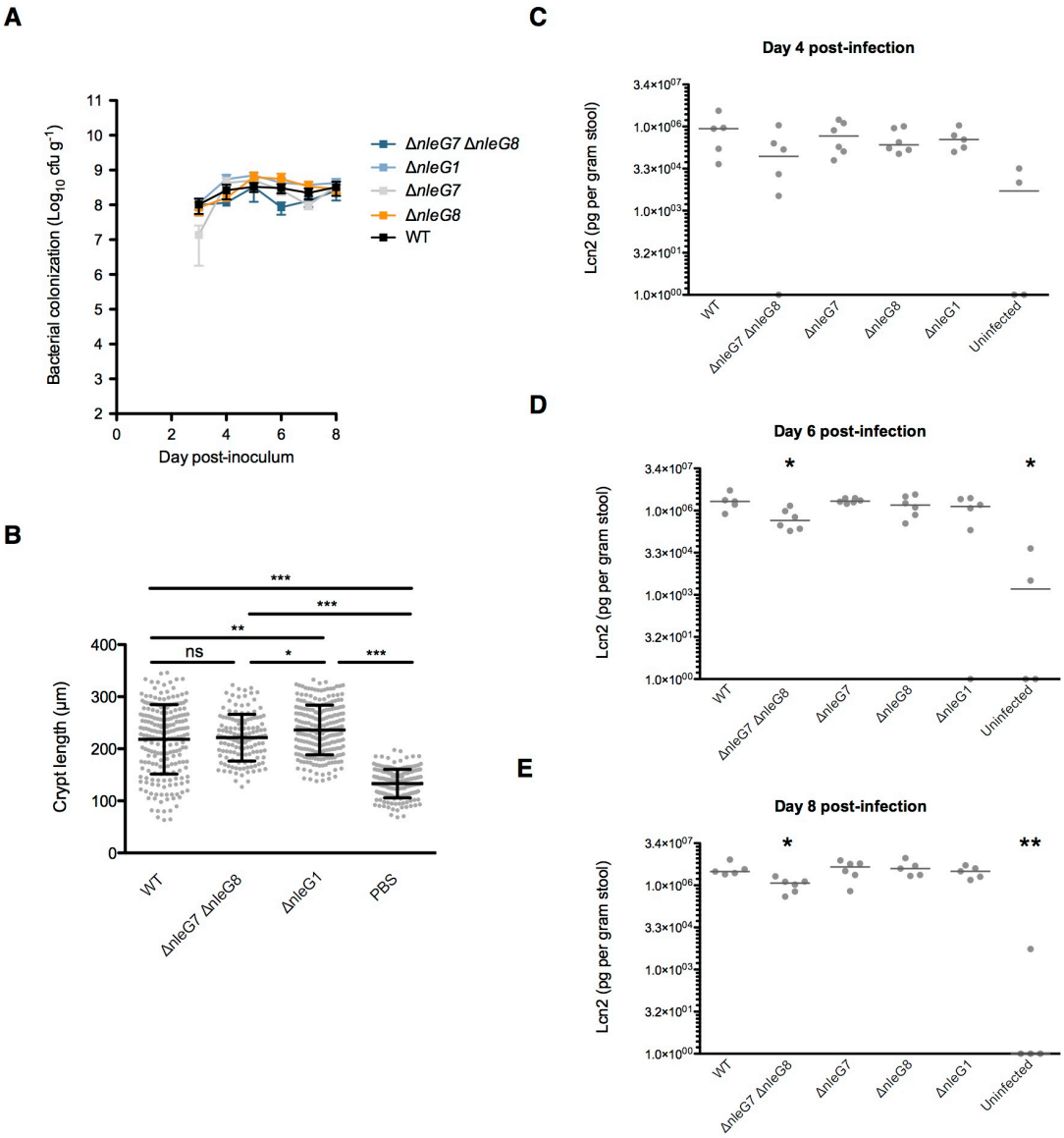


Figure 4.19 | A double Δ*nleG7* Δ*nleG8* *C. rodentium* displays a decreased level of Lcn2, a biomarker of intestinal inflammation, in stools upon *in vivo* infection of C57BL/6 mice despite no attenuation in bacterial colonisation or colonic crypt hyperplasia. Mice (n=5-6 per treatment group, n=4 for PBS mock-infected group) were infected with WT *C. rodentium*, a *ΔnleG1* mutant, a *ΔnleG7* mutant, a *ΔnleG8* mutant or a double *ΔnleG7 ΔnleG8* mutant alongside a PBS mock-infection control. **A)** All *C. rodentium* infection groups displayed similar colonisation dynamics and levels of colonisation until the peak of infection. **B)** Mice were sacrificed at day 8 post-infection and the distal colon analysed for colonic crypt hyperplasia by H&E staining. Measurements of crypt lengths revealed no difference between WT *C. rodentium* and the double *ΔnleG7 ΔnleG8* mutant. A small increase in hyperplasia was seen for the *ΔnleG1* mutant. All mutants and WT *C. rodentium* had significantly increased colonic crypts lengths compared to the PBS mock-infected control. **C-E)** The concentration of Lcn2 in the stool supernatant was measured by ELISA for day 4, 6 and 8 post-infection. There was a decrease in the Lcn2 concentration in stools from mice infected with the double *ΔnleG7 ΔnleG8* mutant compared to WT *C. rodentium* on day 6 and day 8 post-infection. Error bars represent ±SEM; *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

Measurements of crypt lengths (**Figure 4.19B**) revealed no difference between WT *C. rodentium* and the *ΔnleG7 ΔnleG8* mutant. All mutants and WT *C. rodentium* had significantly increased colonic crypts lengths compared to the PBS mock-infected control. A small increase in hyperplasia was seen for the *ΔnleG1* mutant, inconsistent with previous crypt measurements of *ΔnleG1 C. rodentium* infected tissues, but consistent with the observation that the $NleG1_{P163K}$ mutant displayed slightly enhanced colonic crypt hyperplasia. These subtle changes in colonic crypt hyperplasia for NleG1 were not investigated further but strengthened motivations to study a full *nleG* knockout strain *in vivo*. The concentration of Lcn2 in the stools of infected mice was monitored throughout the experiment. Importantly, there was a reduction in Lcn2 levels for the *ΔnleG7 ΔnleG8* mutant compared to WT *C. rodentium* that reached statistical significance on day 6 and day 8 post-infection (**Figure 4.19C-E**). The single deletion mutants, on the other hand, showed Lcn2 concentrations that were not different from WT-levels throughout the duration of the experiment. This could mean that the NleG proteins somehow function in concert to contribute to the inflammatory response in the murine gut upon infection, and prompting the hypothesis that a triple *nleG* knockout strain might show even lower levels of Lcn2 in the faeces of infected mice.

4.14. The effect of C. rodentium infection on endogenous ubiquitin

The pool of evidence supporting manipulation of ubiquitin-dependent host processes by enteric pathogens is expanding rapidly. *In vitro*, *C. rodentium* infection however does not alter the *overall* distribution of ubiquitin, as shown in **Figure 4.20A** where cultured cells were infected with ICC169 and analysed by fluorescence microscopy. The overall quantity of ubiquitin in cell infected with various *nleG* mutants and WT *C. rodentium* was also assessed by lysing infected HeLa cells (1 or 3 hour infections) and analysing lysates by SDS-PAGE and western blot using an anti-ubiquitin antibody (**Figure 4.20B-C**). However, this gave inconsistent ubiquitin smears of varying intensities, and further efforts to study ubiquitin levels *in vitro* in this manner were not pursued.

In terms of ubiquitination during *in vivo* infection, there have been reports of ubiquitin-rich aggregates forming in colonic cells in response to *C. rodentium* infections (Kaniuk *et al.*, 2011) but no report of overall changes in ubiquitin level in infected cells. To test if a change in ubiquitination in infected enterocytes compared to uninfected enterocytes was detectable, and to whether these results depended on the NleG proteins, enterocytes were collected from the mice in the *in vivo* experiment already described in **Section 4.8**. Enterocytes were lysed and the

enterocyte protein analysed by SDS-PAGE and western blot using an anti-ubiquitin antibody. Enterocytes from two mice per group were analysed. However, there was significant variation between the two mice in both the WT and uninfected groups (**Figure 4.21**) and we concluded that this was not a robust way study ubiquitination in enterocytes following *C. rodentium* infection.

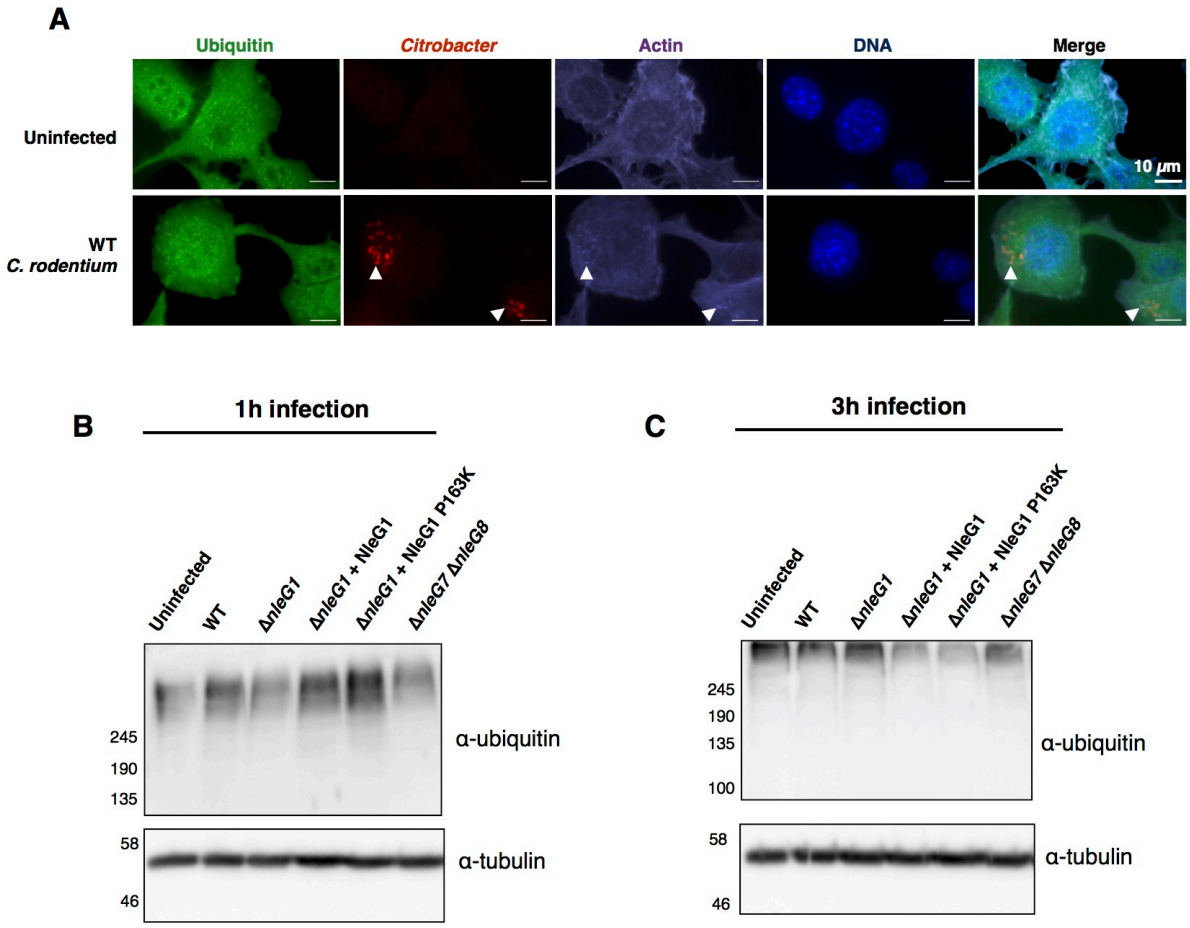


Figure 4.20 | Ubiquitin distribution and total ubiquitin in cultured cells infected with *C. rodentium in vitro*. A) Swiss cells were infected with *C. rodentium* for 5 hours and analysed by immunofluorescent staining. Actin shown in purple, *C. rodentium* in red, ubiquitin in green and DNA in blue. Actin polymerisation was observed under bacteria and there was no re-distribution of endogenous ubiquitin in infected compared to uninfected cells. Scale bar 10 µm. **B)** HeLa cells were infected with *C. rodentium* strains: WT, Δ*nleG1*, Δ*nleG1* chromosomally complemented with WT NleG1, Δ*nleG1* chromosomally complemented with the catalytically inactive $NleG1_{P163K}$ and Δ*nleG7* Δ*nleG8*. Cells were lysed and collected after 1 hour infection for analysis by SDS-PAGE and western blot. Total ubiquitin was visualised using an anti-ubiquitin antibody. Anti-tubulin was used as a loading control. **C)** Infection carried out as in (B) but continued for 3 hours prior to lysis. For both 1- and 3-hour infection, there was inconsistent variation between the tested conditions.

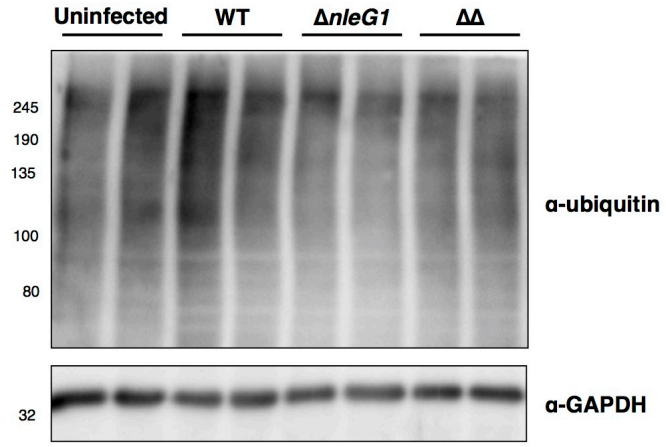


Figure 4.21 | Total ubiquitin smear in enterocytes infected with *C. rodentium* WT and Δ*nleG1* and Δ*nleG7* Δ*nleG8* strains. Mice (n=5-6 per treatment group, n=4 for PBS mock-infected group) were infected with WT *C. rodentium*, a *ΔnleG1* mutant, or a double *ΔnleG7 ΔnleG8* mutant alongside a PBS mock-infection control. Enterocytes were isolated from the distal colon of each mice at day 8 post-infection. Enterocytes were lysed and the enterocyte protein analysed by SDS-PAGE and western blot. Total endogenous ubiquitin was assessed using an anti-ubiquitin antibody. Anti-GAPDH was used as a loading control. 10 µg protein was loaded per lane and two mice were analysed per infection group.

4.15. In vivo phenotypes of a triple nleG deletion mutant C. rodentium

To fully appreciate and study the contribution of NleG proteins to *C. rodentium* pathogenesis *in vivo*, it was necessary to perform mouse infections with the *nleG* null strain. Prior to infecting mice, the mutant's ability to grow in liquid media (DMEM, M9 minimal media and LB) was demonstrated by growth assays (**Figure 4.22**).

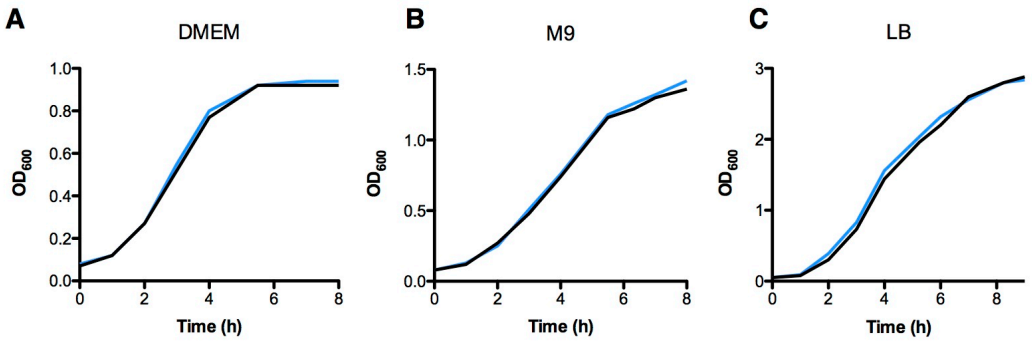


Figure 4.22 | A full nleG knockout C. rodentium strain displays no growth defect in liquid media *in vitro*. The growth of WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* was assessed by measuring the OD_{600} at regular intervals in **A)** Dulbecco's Modified Eagle's Medium (DMEM), **B)** M9 minimal media and **C)** LB. The growth of WT (in black) and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* (in blue) were identical in all growth media tested.

4.15.1. Reduced colonisation of a triple nleG deletion mutant and lower concentrations of Lcn2 in stools

Knowing that knocking out all NleG proteins bore no obvious fitness cost to *C. rodentium* in terms of replication and having previously established that the mutant could form pedestals upon infection of cultured cells *in vitro*, the mutant was deemed suitable for *in vivo* analyses. Mice were orally gavaged with WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* and bacterial shedding was monitored until day 8 post-infection. This revealed a slight approximately 10-fold reduction in colonisation on days 6, 7 and 8- post-infection (**Figure 4.23A**). However, the difference only reached statistical significance on day 7 post-infection. There was no difference in colonic crypt hyperplasia between WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium*, as established by measuring by the length of the colonic crypt of H&E stained sections (**Figure 4.23B**).

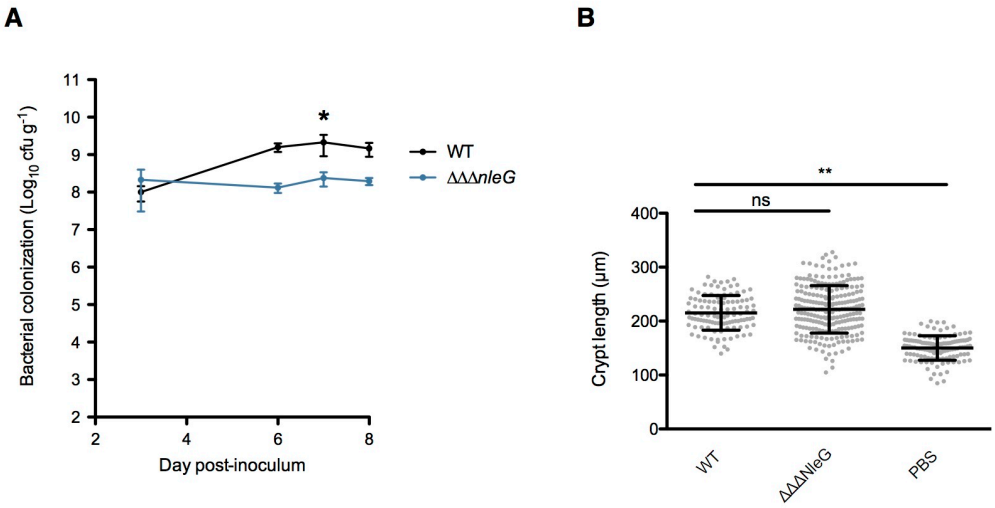


Figure 4.23 | *In vivo* colonisation of a full *nleG* knockout *C. rodentium* strain. C57BL/6 mice (n=5 for all groups) were infected with ICC169 (WT) or Δ*nleG1* Δ*nleG7* Δ*nleG8 (*ΔΔΔ*nleG) C. rodentium*. **A)** An average 10-fold decrease in colonisation for the *nleG* knockout reached statistical significance only on day 7 post-infection as measured by ANOVA with Bonferroni's post-test. **B)** There was no difference between Δ*nleG1* Δ*nleG7* Δ*nleG8* and WT *C. rodentium* in terms of colonic crypt hyperplasia. Error bars represent ±SEM; ns = non-significant, *p ≤ 0.05, **p ≤ 0.01.

The *in vivo* infection with the triple *nleG* knockout was repeated to test if the mild colonisation attenuation was a reproducible phenotype. As before, a slight decrease in colonisation was noted throughout the infection (**Figure 4.24**). On day 6 post-infection the difference reached statistical significance.

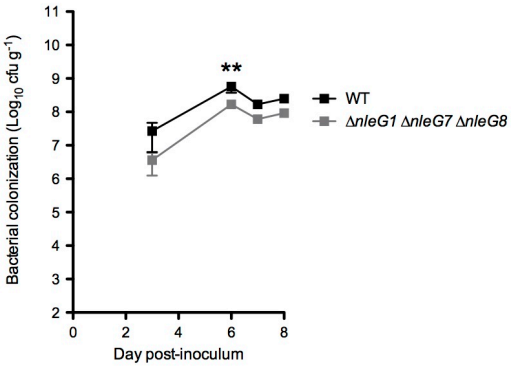


Figure 4.24 | *In vivo* colonisation of a full *nleG* knockout *C. rodentium*. C57BL/6 mice (n=4-5 for all groups) were infected with ICC169 (WT) or Δ*nleG1* Δ*nleG7* Δ*nleG8 (*ΔΔΔ*nleG) C. rodentium*. **A)** A decrease in colonisation for the ΔΔΔ*nleG* strain (grey) compared to WT (black) was significant on day 6 post-infection as measured by two-way ANOVA with Bonferroni's post-test. Error bars are ±SEM. **p ≤ 0.01.

Measurements of total crypt lengths from histological tissue sections showed a slight decrease in crypt hyperplasia for the triple mutant compared to WT *C. rodentium* (**Figure 4.25A-B**). However, further measurements of proliferating cells within the crypt (by staining for the proliferation marker Ki67 and measuring the percentage of the crypt that contained $Ki67^+$ cells) showed no significant difference between the mutant and the WT (**Figure 4.25C-D**). As the difference in crypt length was not a reproducible phenotype and the Ki67 staining showed no difference between the WT and the mutant, this was not deemed an interesting phenotype to follow up on.

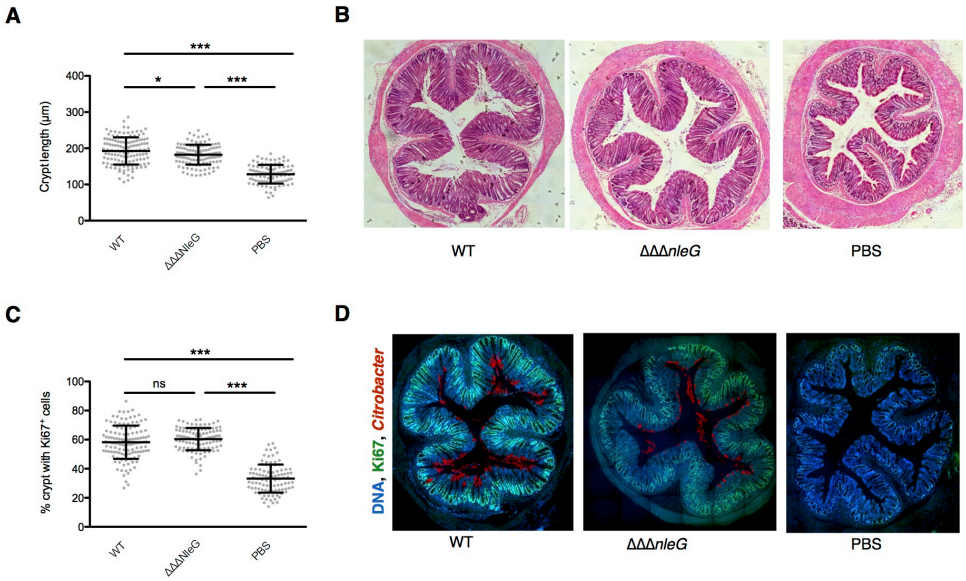


Figure 4.25 | Comparison of colonic crypts from mice infected with WT and Δ*nleG1* Δ*nleG7* Δ*nleG8 C. rodentium* at day 8 post-infection. C57BL/6 mice were infected with ICC169 (WT, n=4) or Δ*nleG1* Δ*nleG7* Δ*nleG8 (*ΔΔΔ*nleG,* n=5) *C. rodentium* alongside a PBS mock-infected control (n=5). Tissues were harvested in day 8 post-infection, fixed, sectioned and analysed by staining and microscopy. **A)** Measurements of total crypt lenght from H&E-stained sections revealed a slight reduction in colonic crypt hyperplasia for the Δ*nleG1* Δ*nleG7* Δ*nleG8* mutant. **B)** Representative images of H&E analysis. **C)** Percentage of total crypt length with Ki67-positive cells, showing no difference between the Δ*nleG1* Δ*nleG7* Δ*nleG8* mutant and WT. **D)** Representative images of immunofluorescence analysis of colonic crypts, with DNA in blue, Ki67 in green and *C. rodentium* in red. Error bars represent ±SEM; groups compared with an ANOVA with Tukey's post-test; ns = non-significant, *p ≤ 0.05, ***p ≤ 0.001.

However, as the double (Δ*nleG7* Δ*nleG8*) deletion mutant had shown a decrease in faecal Lcn2 levels, an Lcn2 ELISA was also carried out for this *in vivo* experiment to compare faecal Lcn2 levels in WT- and triple *nleG* mutant-infected mice Strikingly, the triple *nleG* knockout had an obvious and statistically significant reduction in Lcn2 concentration in the stool compared to WT (**Figure 4.26**). This suggested that the triple *nleG* knockout strain is less able to induce inflammation in the mouse gut, hinting at a role for NleG proteins in the induction or regulation of Lcn2-related responses during *C. rodentium* infection *in vivo*.

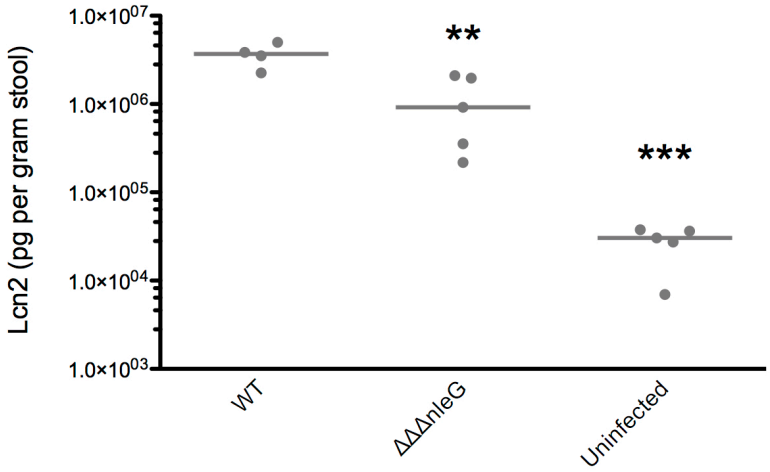


Figure 4.26 | *In vivo* infection using triple *nleG* deleted *C. rodentium* induces lower concentrations of Lcn2 in the stool. C57BL/6 mice (n=4-5 for all groups) were infected with ICC169 (WT) or Δ*nleG1* Δ*nleG7* Δ*nleG8* (ΔΔΔ*nleG*) *C. rodentium*. The concentration of Lcn2 in homogenised stool supernatant was measured by ELISA for day 8 post-infection. There was a decrease in the Lcn2 concentration in stools from mice infected with the Δ*nleG1 ΔnleG7 ΔnleG8* mutant compared to WT *C. rodentium* as measured by one-way ANOVA with Tukey's post-test. Error bars represent ±SEM. **p ≤ 0.01, ***p ≤ 0.001.

To verify that the slight decrease in colonisation in the triple *nleG* deletion mutant was attributable to the lack of a full repertoire of NleG proteins during infection, genes encoding NleG1 or the E3 ligase-dead NleG1$_{P163K}$ point mutant were introduced back onto the genome into the native site. This generated two strains: ΔΔΔ*nleG* + NleG1 that should phenocopy the double Δ*nleG7* Δ*nleG8* deletion mutant and ΔΔΔ*nleG* + NleG1$_{P163K}$ that should phenocopy the triple *nleG* deletion mutant. However, when these strains were tested *in vivo*, no statistical differences were observed between any of the strains (**Figure 4.27**). Therefore, this experiment on its own could neither confirm nor deny that the previously observed differences were due to the lack of NleG proteins. To establish this, the complementation experiment should be repeated and it should also encompass Lcn2 ELISA experiments to verify the contribution of NleG proteins to the Lcn2 phenotype.

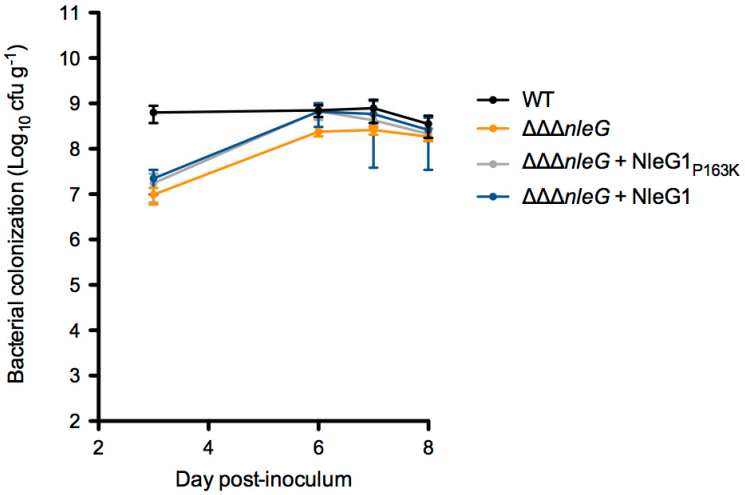


Figure 4.27 | Bacterial colonisation of the mouse intestine by full *nleG* knockout *C. rodentium* and the same strain chromosomally complemented with either WT of catalytically inactive NleG1. C57BL/6 mice (n=5 for all groups) were infected with ICC169 (WT) Δ*nleG1* Δ*nleG7* Δ*nleG8* (ΔΔΔ*nleG*) *C. rodentium* or ΔΔΔ*nleG* complemented with NleG1 or NleG1P163K (on the genome, in the native site). There were no statistically significant differences between the number of viable bacteria between the strains (measured by two-way ANOVA with Bonferroni's post-test. Error bars represent ±SEM.

4.16. Discussion

The *C. rodentium* model has been pivotal in determining the *in vivo* importance of T3SS effector proteins during A/E pathogenesis. Here, we show that single *nleG* deletion mutant *C. rodentium* strains and a Δ*nleG7 ΔnleG8* double knockout strain displayed identical gut colonisation dynamics to the wild-type strain in this model. A *C. rodentium* triple NleG deletion mutant on the other hand displayed a minor colonisation attenuation, consistently giving an approximately 10-fold reduction in bacterial load during the peak of infection. Nevertheless, this only reached statistical significance in two out of three biological repeats and should therefore be interpreted with caution. A full repertoire of NleG proteins is clearly not essential for bacterial colonisation of the murine gut. Loss of essential effector proteins, such as EspZ, render *C. rodentium* completely unable to colonise the intestinal mucosa to its full potential, which was not the case for NleGs (Deng *et al.*, 2004). As knocking out *nleG* genes individually or *nleG7*/*8* together resulted in no reduction in colonisation, the subtle difference in colonisation for the triple deletion mutant may, interestingly, be reflective of NleG proteins contributing to virulence with some degree of functional redundancy. To establish if this hypothesis holds true, further complementation experiments should be performed, as was started at the end of this chapter.

Assuming a degree of functional redundancy *in vivo*, true phenotypes attributable to the NleG proteins would only be picked up by experiments using the triple knockout strain. Therefore, it would be informative to repeat the immunophenotyping experiments performed in this project (showing no change in immune cell recruitment for either the *ΔnleG1* or *ΔnleG8* mutant) but using the triple deletion mutant. Other bacterial E3 ligase effector proteins have in recent years been described to target intracellular pathways that concern immune signalling and ultimately affect the release of inflammatory immunomodulators (Zhang *et al.*, 2006; Fiskin *et al.*, 2016; Maculins *et al.*, 2016b; Wickham *et al.*, 2007). Whether NleG proteins evolved to hold similar functions remains to be determined, but their conservation in enteric pathogenic bacteria suggest that they play a role in virulence.

The TMT 10plex MS-based strategy comparing proteomes of enterocytes infected with WT, *ΔnleG1* and $NleG1_{P163K}$ *C. rodentium*, respectively, unfortunately yielded no useful data for bioinformatic analysis as the abundance of *C. rodentium* peptides was inconsistent across the groups. The ability to pair *in vivo* experiments of high relevance to human infection with state-of-the-art high throughput proteomics, nevertheless, holds great potential and is enabling the identification of very subtle changes in the host caused by individual effector proteins (Berger *et al.*, 2017; Pollard *et al.*, 2018). It may still be used to answer the question of which pathway(s) the NleG proteins are targeting in the host, ideally using the triple knockout mutant (and perhaps a triple ligase-dead NleG mutant) with robust complementation controls and adequate colonisation. Further, the workflow could be coupled to enrichment processes that would, for instance, enrich for polyubiquitinated proteins. This presents certain challenges for data analysis as it remains unknown whether NleG-dependent ubiquitination triggers proteasomal degradation of host proteins, changes target protein half-life (shorter or longer), or in fact has no effect on target protein turnover. While an *in vivo* proteomics approach might therefore not reveal one true target of NleG proteins – a fundamental reason that *in vitro* approaches are better suited for finding binding partners of E3 ligase proteins – this technique could still highlight cellular responses and processes associated with *C. rodentium* infection dependent on the NleGs, giving huge insight into NleG function. We briefly considered the overall ubiquitination state of infected cells for any changes detectable by immunofluoresence (*in vitro* experiments) or by western blot (*in vivo* experiments). Due to the inherent complexity and scale of cellular ubiquitinomes we concluded that more sensitive approaches, such as MS-based methodologies, are required to find what is likely to be subtle changes between WT and *ΔnleG* samples.

A key finding in this chapter concerns the host Lcn2 response, which is a way for the host to limit uptake of iron by bacteria in the gut. Gram negative bacteria secrete the siderophore Ent to bind iron, and Lcn2 secreted by neutrophils and host mucosal cells bind ferric and aferric Ent to inhibit its re-uptake by bacteria. Thus, bacterial iron acquisition is compromised. The importance of this is illustrated by Lcn2-deficient mice rapidly succumbing to infections by *E. coli* or *Klebsiella pneumoniae* strains that use Ent for iron uptake (Berger *et al.*, 2006; Cramer *et al.*, 2017). Additionally, the combined release of Ent and Lcn2 in the gut has been shown to enhance IL-8 secretion (Holden *et al.*, 2014). To evade Lcn2-mediated effects, some bacteria secrete siderophores that cannot be bound by Lcn2 (for example glycosylated Ent). Recently, the Lcn2 response was shown to be a promising non-invasive biomarker for intestinal inflammation in laboratory mouse models, and it was therefore tested here. Single *nleG* deletion mutants exhibited WT-concentrations of Lcn2 in their stools. The double Δ*nleG7 ΔnleG8* strain had a decrease in Lcn2 concentration that could not be explained by a decrease in colonisation as the mutant colonises the mouse gut to WT-levels. The full *nleG* deletion mutant showed an even more pronounced decrease in Lcn2 levels. This suggests that NleG proteins contribute to a host inflammatory response in the gut, motivating future efforts to complement the phenotype by expressing NleG proteins ectopically and by inserting the *nleG* genes back into their original chromosomal locations for regulations under the native promoter and performing more Lcn2 ELISAs. Ideally, this should be performed in parallel to immunophenotyping experiments, to clarify the potential contributions of NleG proteins to host inflammation. If shown to be true, this would be the first *in vivo* phenotype attributable to the NleG proteins.

Chapter 5:
Identifying interactors of C. rodentium NleG proteins using mass spectrometry

5. Identifying interactors of C. rodentium NleG proteins using mass spectrometry

5.1. Introduction

Identifying host interacting partners or substrates of the NleG proteins could lead us to understand their specific function during *C. rodentium* infection. However, as discussed in **Chapter 3**, finding biologically relevant targets of E3 ligases is notoriously challenging and remains an inherent hurdle in E3 ligase research. E3 ligases generally target more than one substrate, and one substrate may conversely be targeted by more than one E3 ligase (O'Connor & Huibregtse, 2017). Further, interactions between E3 ligases and their substrates are generally of low affinity and highly transient in nature. If the outcome of the interaction is proteasomal degradation of the substrate, the substrate becomes difficult to detect. Consequently, only a handful of the hundreds of known E3 ligases have been assigned high-confidence cellular targets.

In recent years, mass spectrometry-based methods have emerged as the most promising technologies for identifying E3 ligase substrates. Affinity purification coupled with mass spectrometry-based proteomics is widely used to identify and characterise novel protein-protein interactions. Affinity-tagged E3 ligases expressed in cells *in vitro* can co-immunoprecipitate with their substrates or binding partners, permitting their identification by mass spectrometry.

The aim of this part of the NleG project was to employ affinity purification mass spectrometry to identify host interactors of the *C. rodentium* NleG proteins and to characterise putative interactions *in vitro* and *in vivo*. The mass spectrometry in this chapter was performed in collaboration with the Ivan Đikić laboratory in Germany (Institute of Biochemistry II, Medical Faculty, Goethe University Frankfurt). Dr Timurs Maculins (together with their student Gabor Varga) kindly generated a number of transgenic cell lines with inducible expression of NleG proteins. I joined Tim and Gabor for four weeks in Frankfurt to perform pull-downs and preparation of samples for mass spectrometry analysis. The mass spectrometry itself was kindly performed by Dr Florian Bonn, who also did the initial analysis of label free quantification (LFQ) intensities in MaxQuant. I analysed the mass spectrometry data further using Perseus and performed the follow up experiments in the Frankel Lab at Imperial College London.

5.2. A mass spectrometry-based strategy to identify NleG interactors

A crucial challenge with affinity purification coupled with mass spectrometry-based proteomics is distinguishing between biologically important interactions and interfering background noise. Weak or transient binding partners may easily be overlooked. In the context of E3 ligase research, assays can be performed with enzymatically inactive E3 ligase mutants to circumvent this problem. Inactive E3 ligases bind substrates without ubiquitinating them, effectively trapping the substrates in complex with the E3 ligase. As the interaction is thereby rendered less transient and substrates are not targeted for proteasomal degradation, interactions become considerably more detectable by mass spectrometry. With this in mind, an experimental workflow was designed for identification of host interactors of the NleGs. The workflow is shown in **Figure 5.1**, which comprises a schematic of the Retro-X™ Tet-ON® Advanced Inducible Expression System used to express NleG proteins (**Figure 5.1A**) and an overview of key experimental steps (**Figure 5.1B**).

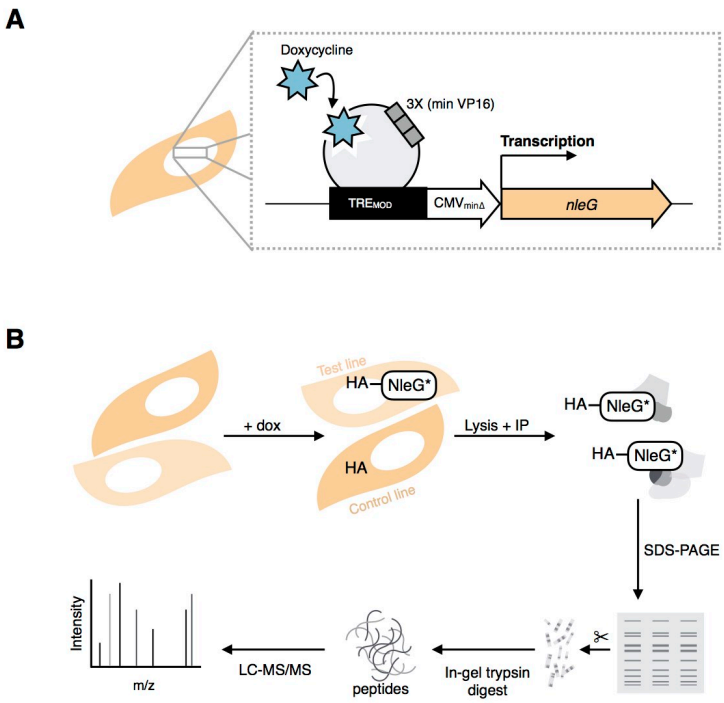


Figure 5.1| Schematic overview of the approach used in this study to identify interacting partners of NleGs. A) The Retro-X™ Tet-ON® Advanced Inducible Expression System was used for regulated *nleG* gene induction in cultured cells. It uses a doxycycline-controlled transactivator (rtTA-Advanced) that is a "reverse" Tet repressor, as it contains five amino acid changes that reverse the protein's original response to doxycycline. Its DNA-binding domain is joined to three minimal transcription activator domains. These activate transcription when bound to the tetracycline response element (TRE) upstream of the minimal immediate promoter of cytomegalovirus ($CMV_{min\Delta}$). This promoter is silent in the absence of doxycycline. **B)** In this study, doxycycline-inducible cell lines expressing HA-tagged enzymatically inactive NleGs ($NleG1_{P163K}$ and $NleG8_{P177K}$; test lines) were generated in parallel to cells expressing the HA-tag on its own (control line). Following transgene expression, NleG-containing complexes were immunoprecipitated using anti HA-beads. Proteins were separated by SDS-PAGE and peptides extracted by in-gel trypsin digest. Peptides were then identified by LC-coupled tandem mass spectrometry (LC-MS/MS) and analysed for protein enrichment in NleG samples compared to control samples.

Briefly, stably transduced mammalian cells (HCT116 and immortalised bone marrow macrophages, iBMMs) with inducible expression of HA-tagged E3 ligase dead NleG1 or NleG8 were generated. Control lines were also generated to express only the HA tag. Using anti HA-beads, $NleG_{P163K}$ or $NleG8_{P177K}$-complexes were immunoprecipitated and separated by SDS-PAGE. Peptides were then extracted by in-gel tryptic digest and analysed in technical triplicate by liquid chromatography-coupled MS/MS. By comparing peptides pulled down with HA-tagged ligase dead NleG to those pulled down with the HA-tag on its own, we could determine which proteins were specifically and significantly enriched in the NleG sample. The experiments are explained in more detail in the following few sections.

5.3. Generation of stably transduced cell lines

Dr Timurs Maculins and Gabor Varga generated cell lines to allow expression of tagged NleG proteins in an inducible manner in cultured cells. The Retro-X™ Tet-ON® Advanced Inducible Expression System was used, where a stably transduced cell line is generated with a basal "no transgene expression" state that is maintained in the absence of doxycycline in the cell culture medium. Induction is achieved by adding doxycycline to the cell culture medium at low concentrations (sub-lethal to cultured mammalian cells). Further, the expression levels can be fine-tuned by titrating the concentration of doxycyline. When making the cells lines, tag size and positioning became important considerations. We opted for a small N-terminal HA tag, separated from the NleG by a short linker region in an attempt to prevent any steric hindrance of biologically relevant protein-protein interactions. Importantly, there are high-quality enrichment reagents available for HA-tagged proteins, facilitating the pulldown of NleG-containing complexes.

Cell lines expressing HA, HA-NleG1, HA-$NleG1_{P163K}$, HA-NleG8 and HA-$NleG8_{P177K}$ were generated in two different parental backgrounds: HCT116 cells and immortalised bone-marrow derived macrophages (iBMMs). Cells expressing WT NleG1 and NleG8 were generated for follow up experiments after the initial mass spectrometry screen for interactors. HCT116 cells are a human colorectal carcinoma line with epithelial cell physiology, which is relevant to study NleG proteins as *C. rodentium* infects intestinal epithelial cells. The iBMM line was a murine cell line, and although *C. rodentium* does not primarily infect macrophages in the mouse gut, the *C. rodentium* NleG proteins might recognise binding partners better in a mouse cell line compared to a human-derived cell line. Ultimately, performing the experiments in two distinct cell lines and assessing

the overlap in enriched proteins between the two will strengthen the data (O'Connor & Huibregtse, 2017).

The expression of NleG proteins was verified by analysing cell lysates by SDS-PAGE and western blot using an anti-HA antibody. This showed robust expression of HA-NleG1 and HA-NleG1$_{P163K}$ (**Figure 5.2**). The expression was lower in the iBMM cells compared to HCT116 cells, but detectable by western blot in both cases. Expression of NleG8 proteins could not be detected by western blot (data not shown). However, owing to the high sensitivity of the mass spectrometry analysis that the cells were made for, we still decided to include the NleG8$_{P177K}$-expressing cells in the first biological repeat of the experiment.

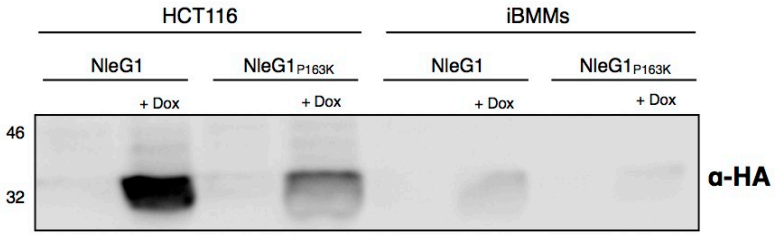


Figure 5.2| Doxycycline-inducible expression of the NleGs in cultured cells. Expression of NleG1 and NleG1$_{P163K}$ in the inducible cell lines following 24 hours in media containing 1 µM doxycycline (Dox) was tested by analysing cell lysates by SDS-PAGE and western blot using an anti-HA antibody. NleG1 and NleG1$_{P163K}$ expressed well in HCT116s, with lower expression iBMMs.

5.4. Immunoprecipitation of NleG-containing complexes

The next step was to express the enzymatically inactive HA-NleGs in a suitably large format for purifying them using antibodies against the epitope tag, with the aim of co-purifying their interacting partners. Cells were grown for in culture media containing 1 µM doxycyline in 10 cm diameter dishes. 1 × 10^6 HCT116s or 1.5 × 10^6 iBMMs were seeded per dish, as iBMMs are smaller and showed lower transgene expression. HA-expressing cells were included as control lines. Cell lysates were then subjected to immunoprecipitation (IP) using anti-HA beads. A western blot was performed to confirm that the IP had worked (**Figure 5.3A**). A band of approximately 30 kDa corresponding to HA-NleG1 was present in input samples and enriched in IP samples (confirming success of the IP). It was present in the unbound sample, indicating saturation of the anti-HA beads. Intriguingly, bands of higher molecular weight were also observed, corresponding approximately to the molecular weights of mono- and di-ubiquitnated NleG1. NleG$_{P163K}$ might be ubiquitinated by host E3 ligases in the cell, or alternatively the HA antibody cross-reacts with proteins co-immunoprecipitated with NleG$_{P163K}$. Bands corresponding to the heavy and light chains of IgG (used to make the anti-HA beads) were also present in IP samples, with molecular weights of approximately 50 kDa and 25 kDa. These were the only bands

detectable for NleG8$_{P177K}$ by anti-HA western blot, as expected from the low expression of NleG8$_{P177K}$ discussed above. Ponceau staining of the membrane (top panel) was used as a loading control.

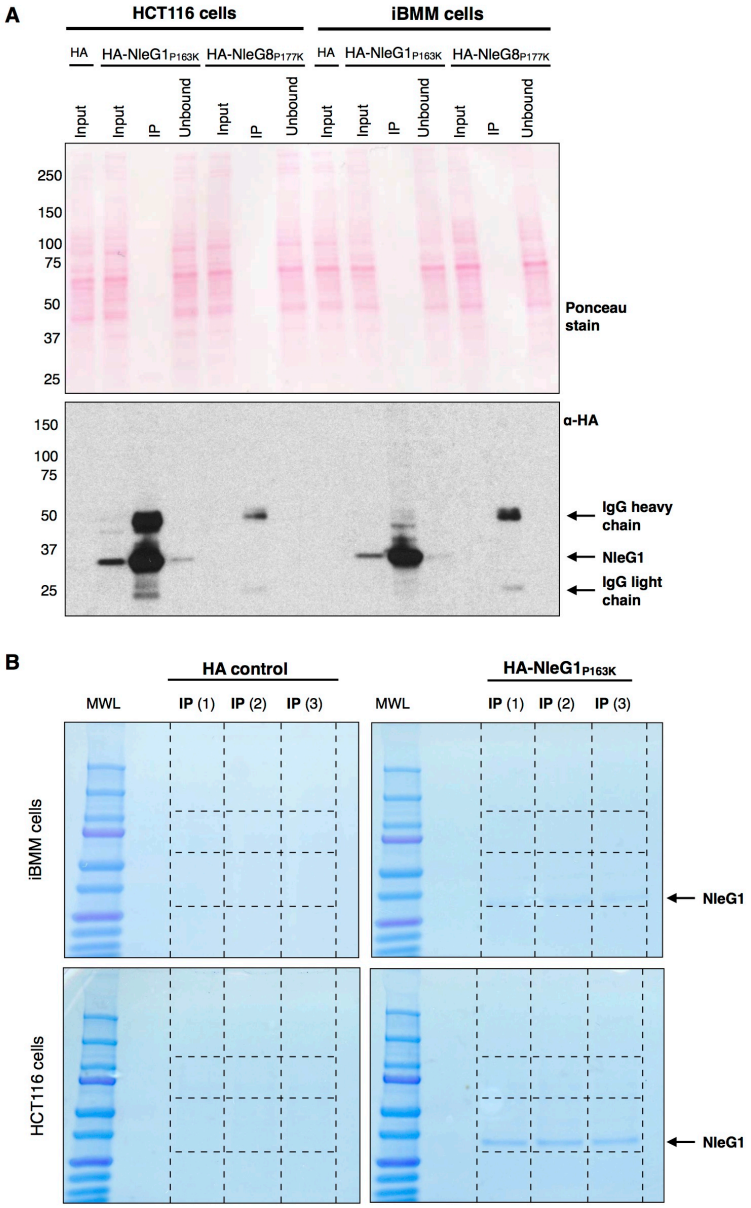


Figure 5.3| Immunoprecipitation (IP) of HA-tagged NleG proteins from cell lysates for mass spectrometry analysis. A) HCT116 and iBMM cell lines generated in this study were grown in 1 µM doxycycline for 24 hours to induce expression of HA-NleG1$_{P163K}$, HA-NleG8$_{P177K}$ or a HA epitope tag control. An IP with anti-HA beads was performed using the cell lysates and analysed by SDS-PAGE and immunoblotting with an anti-HA antibody. The western blot membrane was also stained with Ponceau as a loading control. **B)** The IP samples generated in (A) were separated into technical triplicates and run in parallel lanes on 4-20% gradient SDS-PAGE gels. Gels were stained with Instant*Blue* and a band was easily observed for NleG1 (black arrow; gels for NleG8 not shown). Each lane (ie technical triplicate, as indicated in brackets in the figure) was divided into four separate gel fragments based on molecular weight, as illustrated by the dotted lines. Each resulting fragment was cut into 1mm^3 pieces and collected in a 96-well plate (one well per original gel fragment). Peptides could then be extracted from the gel pieces by in-gel tryptic digested and prepared for identification by liquid chromatography-coupled tandem mass spectrometry.

Knowing that the IP has successfully pulled down HA-NleG1, IP samples were separated in technical triplicate by SDS-PAGE using 4-20% gradient SDS-PAGE gels and processed for MS. The gels were stained with Instant*Blue* and the gels for HA-NleG1 IPs and their corresponding HA controls are shown in **Figure 5.3B**. As indicated by black arrows, a clear band for HA-NleG1$_{P163K}$ was seen in both cell background. Each gel lane was carefully divided into four separate gel fragments based on molecular weight, as indicated by dotted lines in the figure. Next, each resulting fragment was cut into 1 mm^3 pieces and collected in a 96-well plate. One well was used per original gel fragment, and the pieces were then processed in a 96-well format. Peptides were extracted from the pieces by in-gel tryptic digest and prepared for identification by LC-coupled tandem MS.

5.5. Mass spectrometry analysis identifies putative interactors of NleG1 and NleG8

Dr Florian Bonn loaded the prepared samples onto an Orbitrap mass spectrometer for liquid chromatography-coupled tandem mass spectrometry. He kindly collected the data and compiled the resulting LFQs in MaxQuant so that I could analyse them further by comparing LFQ intensities using Perseus.

NleG1 or NleG8 were identified and highly enriched in the appropriate IP samples (despite the undetectable expression of NleG8 by western blot), meaning that the data was suitable for further analysis. Affinity purification commonly gives hundreds of nonspecific proteins binding to the support matrix, necessitating a quantitative filter to discriminate *bona fide* interacting partners from background binding proteins and contaminants (Smits & Vermeulen, 2016). Therefore, the NleG samples were compared to the HA control samples and putative interactors were defined as having a significant minimum 1.5-fold enrichment in NleG samples. Background proteins on the other hand should have equal presence in both samples. Further, proteins that were not detected in at least 2 of the 3 technical repeats for the NleG IP samples were excluded from the analysis, as were peptides that were not identified by at least 1 unique peptide sequence. The false-discovery rate (FDR) was set to 0.05 and FDR-corrected paired t-tests were used to identify significantly enriched proteins. This generated a list of significantly enriched proteins and thus putative interactors of NleG1 and NleG8 in both HCT116 cells and iBMMs (full lists in **Tables 5.1, 5.2, 5.3** and **5.4**). All datasets were also visualised using volcano plots (**Figures 5.4** and **5.5**), where the ten most highly enriched hits were highlighted as black circles. The following pages will highlight the interesting putative interactors identified in the datasets, with a

focus on hits that are conserved between the cell lines. A second biological repeat was performed for NleG1 in HCT116 cells, the result of which is shown in **Table 5.5** and **Figure 5.7**. A comparison and overlap analysis between the enriched hits in the different samples is provided in **Figure 5.8** towards the end of this section. Attention is drawn to proteins that represent likely or interesting targets during manipulation of host cell responses by enteric pathogens.

5.5.1. Novel putative interactors of NleG1 include UBE2O and C1QBP

When doing an affinity purification mass spectrometry screen for interactors, it is extremely helpful to have a known substrate for the bait protein to evaluate the success of the experiment and determine whether set thresholds are suitable to best analyse the data (O'Connor & Huibregtse, 2017). As this was not applicable here, comparing the list of enriched proteins in two different cell lines and assessing their overlap was used to strengthen the validity of hits. We hypothesised that true interactors of NleG1 would be significantly enriched in both cell lines. 79 proteins were identified as putative interactors for NleG1 in HCT116 cells (listed in **Table 5.1** and shown graphically in **Figure 5.4B**). In iBMMs cells, 31 proteins were significantly enriched for NleG1 (listed in **Table 5.2** and shown in **Figure 5.4A**). Eight proteins were significantly enriched in both.

Complement component 1 Q subcomponent-binding protein (C1QBP) was the most enriched host protein in both cell lines (almost 120-fold in iBMMs and 1600-fold in HCT116s). Also known as gC1qr, p32, HAPB1 and SF2p32, it is a ubiquitously expressed multi-ligand protein involved in a range of biological processes. These include inflammation, infection, cell signalling and chemotaxis (Tye *et al.*, 2001; McGee *et al.*, 2011; Itahana & Zhang, 2008; Shi *et al.*, 2017). C1QBP was originally isolated from the plasma membrane as a receptor for the C1q a component of the complement system (Eggleton *et al.*, 1994; Lim *et al.*, 1996). It has since been reported to localise to a number of cellular compartment but is widely accepted to be predominantly mitochondrial. It is an inhibitor of mitochondrial-driven cell death in the mitochondrial matrix (McGee & Baines, 2011), reliant on its binding to cyclophilin D, a regulator of the mitochondrial permeability transition pore. A cytoprotective role for C1QBP is also supported by its expression being up-regulated in HeLa cells in response to cisplatin stimulation, which induces apoptosis (Kamal & Datta, 2006). Further, C1QBP is up-regulated in many cancers (Rubinstein *et al.*, 2004; Dembitzer *et al.*, 2012; Scully *et al.*, 2015) including colorectal cancers (Kim *et al.*, 2017). C1QBP overexpression is particularly prominent in metastatic cancers (Ghosh *et al.*, 2004) and its expression in tumours is correlated with unfavourable prognosis in breast cancer patients (Chen

et al., 2009; Wang *et al.*, 2015) and in endometrial cancer patients (Zhao *et al.*, 2015). It has also been shown to interact with PRKZ3, an important modulator of tumour cell chemotaxis (Zhang *et al.*, 2013a). Other reported interacting proteins of C1QBP include a number of viral proteins, suggesting a role in virus-host interactions (Tye *et al.*, 2001). Multiple siRNA studies have shown that C1QBP suppresses antiviral immune responses (Xu *et al.*, 2009; Wang *et al.*, 2014; Liu *et al.*, 2015b; Sasaki *et al.*, 2017), by associating with MAVS to suppress RIG-1 and MDA-5 dependent interferon responses (Xu *et al.*, 2009). Little is known about the importance of C1QBP during responses to bacterial infections. However, if validated, C1QBP could be a very interesting binding partner of NleG1.

Another protein that was highly enriched in the NleG sample compared to the control IP sample was UBE2O. UBE2O is an atypical E2 enzyme, as it possesses both E2 and E3 activities. It is widely conserved, expressed and has been implicated in a range of cellular pathways *in vitro*. Like C1QBP, it is known to be amplified in many cancers, such as breast, bladder, liver and lung carcinomas. The *Ube2O* gene was knocked out in tumour-spontaneous mouse lines which resulted in a delay in tumour initiation and growth as well as a reduction in rates of metastasis (Vila *et al.*, 2017). It was shown to bind and ubiquitinate a catalytic subunint of AMP-activated protein kinase (AMPKα2), which is subsequently degraded. AMPKα2 has a number of tumour suppressive functions, such as inactivating the mTOR signalling pathway (Gwinn *et al.*, 2008), downregulating the glycolytic pathway to give an anti-Warburg effect (Faubert *et al.*, 2013), contributing to cell cycle arrest and inhibiting epithelial mesenchymal transition (Chou *et al.*, 2014). Further, it was shown to negatively regulate TRAF6-dependent activation of NFκB by binding TRAF6, inhibiting its K63-poly-ubiquitination which disrupts its binding to MyD88 and prevents NFκB activation (Zhang *et al.*, 2013b). Recently, UBE2O was described as an inducible quality control factor that recognises juxtaposed basic and hydrophobic patches on misfolded, unassembled and mislocalised proteins, ubiquitinating them and targeting them for degradation (Yanagitani, Juszkiewicz & Hegde, 2017). An analogous quality control ligase San1 exists in fungi (Rosenbaum *et al.*, 2011). In line with this function, UBE2O was also shown to drive wide-scale proteome remodelling in reticulocytes, including the elimination of intact ribosomes (Nguyen *et al.*, 2017). Importantly, proteomes of *C. rodentium*-infected intestinal epithelial cells extracted from mouse colons at the peak of infection showed a reduced abundance of UBE2O (unpublished data, Frankel lab). This suggests that UBE2O could be specifically targeted for degradation by *C. rodentium*, perhaps by NleG1.

Other proteins enriched in NleG samples in both cell lines were Mac-2BP (discussed later in **Section 5.5.2**), TRIM21, Pentatricopeptide repeat domain containing protein 2 (Ptcd3), Zinc finger CCCH-type antiviral protein 1 (ZAP1) and RNA-binding protein 39 (Rbm39, also known as Co-activator of Activating Protein-1 and Estrogen Receptors, CAPER; discussed further in **Section 5.5.2**). Ptcd3 is an RNA-binding protein thought to facilitate mitochondrial gene expression (Lightowlers & Chrzanowska-Lightowlers, 2013). ZAP1 also binds RNA, but specifically recognises ZAP-responsive elements within viral mRNA. It is an antiviral protein that recruits RNA degradation machineries to viral mRNA, thus inhibiting viral replication in host cells (Kerns, Emerman & Malik, 2008; Zhu *et al.*, 2011; Wang *et al.*, 2012). One of its isoforms also induces RIG-1 signalling, resulting in expression of type I interferons and interferon-stimulated genes (ISGs; Hayakawa *et al.*, 2011). TRIM21 belongs to the TRIM family of E3 ligase proteins. Other TRIM-family members were previously reported to be targeted by bacterial E3 ligase proteins (Kamanova *et al.*, 2016; Fiskin *et al.*, 2017), making it an intriguing protein to find enriched in the NleG1 sample. TRIM21 acts as a receptor for intracellular Fc receptors of immunoglobulins (predominantly IgG, but also IgA and IgM; Rhodes & Isenberg, 2017). TRIM21 mediates antibody-dependent intracellular neutralisation (ADIN) and intracellular antibody-mediated degradation (IAMD) (Mcewan *et al.*, 2011). By targeting intracellular viruses for degradation, TRIM21 also permits the recognition of viral nucleic acids by sensors such as RIG-1, which trigger inflammatory responses (Watkinson *et al.*, 2015). TRIM21 has been linked to autophagy and its overexpression results in cellular structures that resemble autophagosomes (Rhodes *et al.*, 2002). Further, TRIM21 co-localises with LC3 and *S. enterica* in infected cells (Rakebrandt *et al.*, 2014), implying a role in restriction of enteric pathogens.

Interestingly, the protein calcium- and zinc-binding protein S100-A9 (belonging to the S100 family of antimicrobial peptides) was significantly enriched for NleG1 in HCT116s but not in iBMMs. The exact opposite was true for S100-A8, which commonly forms a complex with S100-A9 called calprotectin. Calprotectin can induce neutrophil adhesion, chemotaxis and degranulation (Ryckman *et al.*, 2003; Simard, Girard & Tessier, 2010). Indeed, S100-A8 and S100-A9 hold roles in regulating acute inflammatory processes and immune responses (Miyasaki *et al.*, 1993; Koike *et al.*, 2012). There are also reports of S100-A8 being involved in apoptosis (Nakatani *et al.*, 2005) and autophagy (Ghavami *et al.*, 2010). Their roles immunity make S100 proteins interesting candidates for being targeted by the NleG proteins during *C. rodentium* infection.

Some of the proteins significantly enriched here were recognised to be common contaminants frequently detected in affinity purification mass spectrometry datasets. These included tubulins,

heat shock proteins, histone proteins, ribonuclear proteins and ribosomal proteins (Mellacheruvu *et al.*, 2013).

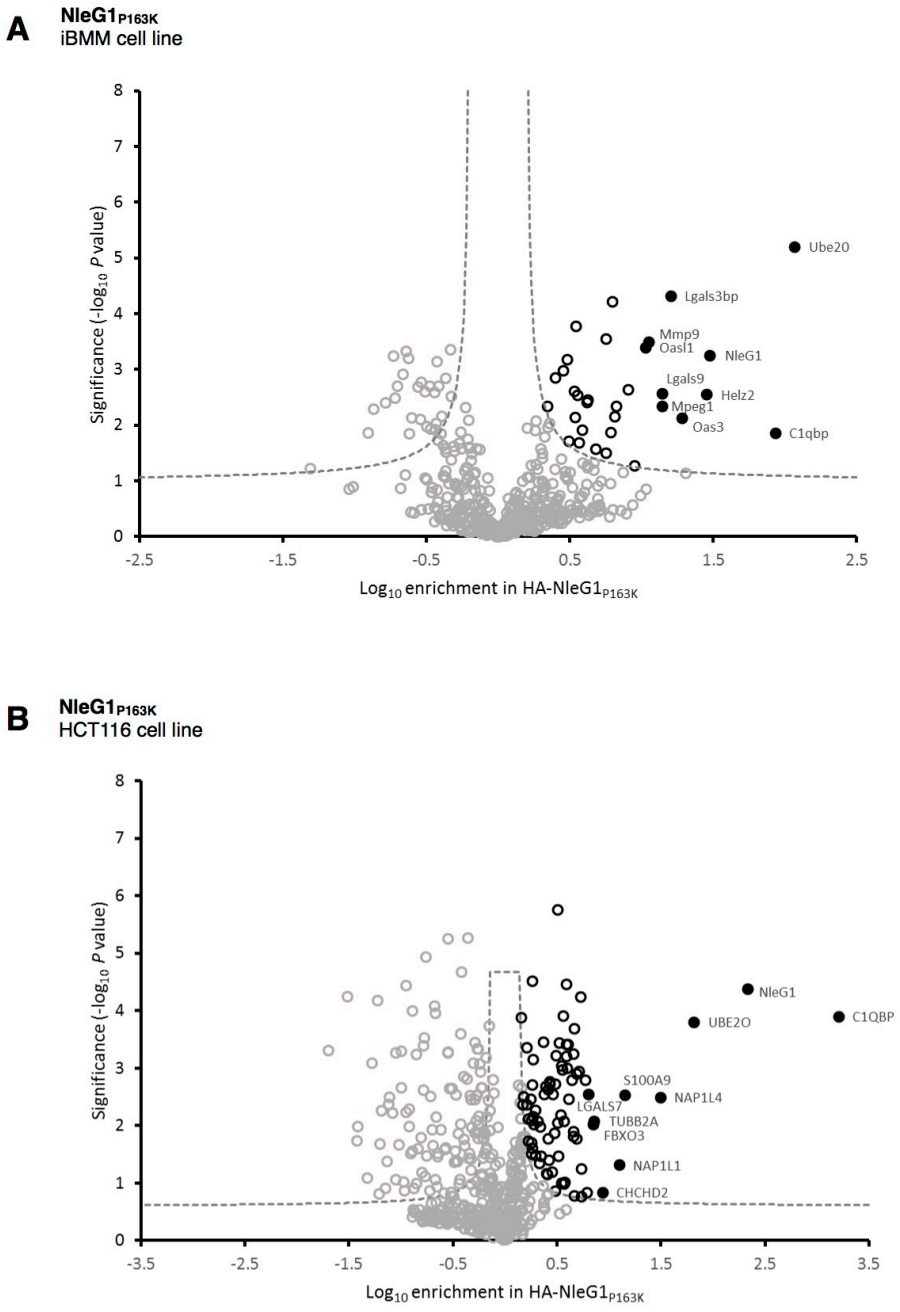


Figure 5.4 | Tandem mass spectrometry analysis of immunoprecipitated NleG1$_{P163K}$–containing complexes from mammalian cell lysates identifies putative interacting partners of NleG1. Volcano plots show the distribution of quantified proteins according to -log$_{10}$ p-value and log$_{10}$ fold enrichment, where enrichment is the ratio of label free quantification (LFQ) intensities for proteins in HA-NleG1$_{P163K}$ IP samples to the corresponding LFQ intensities in control IP samples. Technical triplicates were compared by paired t-tests and the dotted line indicates the S0 threshold for significance. Proteins marked with a black outline are significantly enriched more than 1.5-fold in the HA pulldowns from HA-NleG1$_{P163K}$ expressing cells. The ten most highly enriched proteins are indicated by a filled black circle; labels indicate their respective gene names. **A)** 31 proteins were significantly enriched in the NleG1$_{P163K}$ pull-down sample from iBMMs, a murine cell line. **B**) 79 proteins were significantly enriched in the NleG1$_{P163K}$ sample from HCT116 cells, a human-derived cell line.

Table 5.1| Putative interactors of NleG1 in HCT116 cells as identified by AP-MS/MS. *Repeat 1.*

Significant	-log$_{10}$(p-value)	Log$_{10}$ difference	Protein name	Gene name
+	3.90211	3.205339909	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP
+	4.377181	2.33015871	NleG1	NleG1
+	3.807457	1.812184811	E2/E3 hybrid ubiquitin-protein ligase UBE2O	UBE2O
+	2.487969	1.49616766	Nucleosome assembly protein 1-like 4	NAP1L4
+	2.534586	1.148663362	Protein S100-A9	S100A9
+	1.316092	1.099115213	Nucleosome assembly protein 1-like 1	NAP1L1
+	0.84596	0.935656706	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2, mitochondrial	CHCHD2; CHCHD2P9
+	2.081467	0.857372602	Tubulin beta-2A/B chain	TUBB2A; TUBB2B
+	2.028775	0.844871203	F-box only protein 3	FBXO3
+	2.549307	0.804490725	Galectin-7	LGALS7
+	0.834764	0.786196391	Thioredoxin	TXN
+	2.792964	0.771146456	Cornifin-B; Cornifin-A	SPRR1B; SPRR1A
+	1.258842	0.734511058	Serpin B12	SERPINB12
+	0.775221	0.734103362	Cystatin-S; Cystatin-SA	CST4; CST2
+	4.249469	0.722990672	Putative RNA-binding protein Luc7-like 2	LUC7L2; C7orf55-LUC7L2
+	2.951346	0.71276172	39S ribosomal protein L38, mitochondrial	MRPL38
+	1.780694	0.689857006	Lysozyme C	LYZ
+	2.911328	0.684075514	Protein arginine N-methyltransferase 5	PRMT5
+	3.697245	0.662002087	Plakophilin-1	PKP1
+	0.782463	0.661101023	Ig alpha-1/2 chain C region	IGHA1; IGHA2
+	1.812861	0.654348691	Protein-glutamine gamma-glutamyltransferase E	TGM3
+	3.246791	0.654321353	Zinc-alpha-2-glycoprotein	AZGP1
+	1.897631	0.652485847	28S ribosomal protein S30, mitochondrial	MRPS30
+	2.795519	0.638919195	Catalase	CAT
+	2.464935	0.608421326	Caspase-14	CASP14
+	3.411658	0.603225231	Keratinocyte proline-rich protein	KPRP
+	3.003852	0.594897111	Protein POF1B	POF1B
+	4.464498	0.588870366	39S ribosomal protein L15, mitochondrial	MRPL15
+	3.212212	0.587715785	Desmocollin-1	DSC1
+	3.414433	0.585547924	Desmoplakin	DSP
+	1.024729	0.570189317	Corneodesmosin	CDSN

+	0.992981	0.566965103	F-box only protein 50	NCCRP1
+	2.080306	0.563371181	Protein-glutamine gamma-glutamyltransferase K	TGM1
+	3.919716	0.558112144	Desmoglein-1	DSG1
+	2.98187	0.550921917	39S ribosomal protein L44, mitochondrial	MRPL44
+	1.002525	0.540954749	Protein S100-A8; Protein S100-A8, N-terminally processed	S100A8
+	3.04375	0.539749781	Ensconsin	MAP7
+	2.193669	0.535148303	Tubulin beta chain	TUBB
+	3.445207	0.521766822	Transcription intermediary factor 1-beta	TRIM28
+	1.470516	0.510665735	Squamous cell carcinoma antigen recognized by T-cells 3	SART3
+	2.046383	0.507249991	39S ribosomal protein L4, mitochondrial	MRPL4
+	3.22688	0.491475264	Serpin B3	SERPINB3
+	0.865918	0.484252294	39S ribosomal protein L12, mitochondrial	MRPL12
+	2.721035	0.479595979	Pentatricopeptide repeat domain-containing protein 3, mitochondrial	PTCD3
+	1.873695	0.470035553	Keratin, type II cytoskeletal 2 epidermal	KRT2
+	2.553007	0.459325314	39S ribosomal protein L39, mitochondrial	MRPL39
+	1.191504	0.454188665	39S ribosomal protein L49, mitochondrial	MRPL49
+	2.725356	0.42631038	E3 ubiquitin-protein ligase TRIM21	TRIM21
+	2.773581	0.423904737	Skin-specific protein 32	XP32
+	1.405587	0.422966639	Protein LTV1 homolog	LTV1
+	2.627509	0.410881996	Zinc finger CCCH-type antiviral protein 1	ZC3HAV1
+	1.778718	0.409504096	39S ribosomal protein L46, mitochondrial	MRPL46
+	1.15332	0.40498654	Ig gamma-1 chain C region	IGHG1
+	1.168327	0.393713474	Growth factor receptor-bound protein 2	GRB2
+	2.687922	0.385924657	Protein LYRIC	MTDH
+	2.554081	0.378019969	Heat shock cognate 71 kDa protein	HSPA8
+	3.459994	0.368162632	Alpha-enolase	ENO1
+	1.479175	0.340453148	Cystatin-A;Cystatin-A, N-terminally processed	CSTA
+	1.976993	0.33897241	Stress-70 protein, mitochondrial	HSPA9
+	1.346363	0.330154101	Dermcidin	DCD
+	2.08106	0.313865503	RNA-binding protein 39	RBM39

+	1.489447	0.291922887	L-lactate dehydrogenase B chain	LDHB
+	2.273888	0.288650672	La-related protein 4B	LARP4B
+	2.021428	0.27162679	U1 small nuclear ribonucleoprotein 70 kDa	SNRNP70
+	3.157109	0.267449697	Galectin-3-binding protein	LGALS3BP
+	2.156848	0.264835199	78 kDa glucose-regulated protein	HSPA5
+	1.607682	0.262638887	Cold shock domain-containing protein E1	CSDE1
+	4.515513	0.260165215	RNA-binding protein FUS	FUS
+	2.710174	0.258035342	Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3
+	1.707177	0.25210921	Proliferating cell nuclear antigen	PCNA
+	2.098574	0.25114584	Protein argonaute-2	AGO2
+	1.517117	0.249058088	Junction plakoglobin	JUP
+	2.462094	0.247227987	Keratin, type I cytoskeletal 18	KRT18
+	2.116822	0.221522172	Heterochromatin protein 1-binding protein 3	HP1BP3
+	1.731922	0.220609665	Protein argonaute-1	AGO1
+	3.36144	0.209361553	ATP synthase subunit alpha, mitochondrial	ATP5A1
+	2.368242	0.203916868	Protein TFG	TFG
+	2.512698	0.173561096	Endoplasmin	HSP90B1
+	2.362995	0.172349771	La-related protein 4	LARP4
+	3.892087	0.154362202	Kelch-like protein 12	KLHL12

+ indicates a significant difference compared to the HA control, as determined by a paired T-test with the $-\log_{10}$ (p-value) provided.

Difference is the $\log_{10}$-transformed fold change in abundance of proteins between the HA control and the NleG test sample.

Table 5.2| Putative interactors of NleG1 in iBMM cells as identified by AP-MS/MS.

Significant	-log_{10}(p-value)	Log_{10} difference	Protein name	Gene name
+	5.197345146	2.064998309	E2/E3 hybrid ubiquitin-protein ligase UBE2O	Ube2o
+	1.853839868	1.933333397	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1qbp
+	3.249241923	1.476490339	NleG1	NleG1
+	2.549846766	1.453422705	Helicase with zinc finger domain 2	Helz2
+	2.129650538	1.28278923	2-5-oligoadenylate synthase 3	Oas3
+	4.321327619	1.20333306	Galectin-3-binding protein	Lgals3bp
+	2.568166648	1.144025803	Galectin-9	Lgals9
+	2.349167729	1.142185211	Macrophage-expressed gene 1 protein	Mpeg1
+	3.490241641	1.046666622	Matrix metalloproteinase-9	Mmp9
+	3.40132854	1.026666641	2-5-oligoadenylate synthase-like protein 1	Oasl1
+	1.272706283	0.951799075	Putative helicase MOV-10	Mov10
+	2.640532206	0.904925187	Interferon-induced protein with tetratricopeptide repeats 1	Ifit1
+	2.340419555	0.823496342	39S ribosomal protein L37, mitochondrial	Mrpl37
+	2.158375097	0.809846719	Probable ATP-dependent RNA helicase DHX58	Dhx58
+	4.221772293	0.793179353	39S ribosomal protein L3, mitochondrial	Mrpl3
+	1.872061078	0.783994198	28S ribosomal protein S30, mitochondrial	Mrps30
+	1.503151295	0.750284195	Cis-aconitate decarboxylase	Irg1
+	3.553954533	0.748476664	Spermatogenesis-associated protein 5	Spata5
+	1.582147375	0.680061817	Poly [ADP-ribose] polymerase 12	Parp12
+	2.457721745	0.623957157	NFX1-type zinc finger-containing protein 1	Znfx1
+	2.413983481	0.616666635	40S ribosomal protein S6	Rps6
+	1.923087347	0.585172494	Pentatricopeptide repeat domain-containing protein 3, mitochondrial	Ptcd3
+	1.69507491	0.564584255	Zinc finger CCCH-type antiviral protein 1	Zc3hav1
+	2.536719897	0.552420775	2-5-oligoadenylate	Oasl2

			synthase-like protein 2	
+	3.782940409	0.539999962	Complement C1q subcomponent subunit B	C1qb
+	2.145668171	0.536666552	40S ribosomal protein S2	Rps2; Gm5786; Gm18025; Gm6576; Gm8225
+	2.608214806	0.527822495	Pyrin and HIN domain-containing protein 1	Pyhin1
+	1.723282552	0.490472635	ATP synthase subunit gamma, mitochondrial	Atp5c1
+	3.181031541	0.478123188	RNA-binding protein 39	Rbm39
+	2.978573416	0.453333378	60S ribosomal protein L7a	Rpl7a
+	2.856424633	0.399999936	E3 ubiquitin-protein ligase TRIM21	Trim21
+	2.340622474	0.343333403	Ubiquitin-40S ribosomal protein; Ubiquitin	Rps27a

+ indicates a significant difference compared to the HA control, as determined by a paired T-test with the $-log_{10}$ (p-value) provided.
Difference is the log_{10}-transformed fold change in abundance of proteins between the HA control and the NleG test sample.

5.5.2. Identifying putative interactors of NleG8

Next, the mass spectrometry datasets for NleG8 were analysed for proteins significantly enriched for in NleG8 IP samples compared to HA-control IP samples. This generated a list of 19 proteins significantly enriched in the NleG8-expressing iBMMs (**Table 5.4**), compared to 72 proteins in HCT116 cells (**Table 5.3**). Volcano plots for these data sets are shown in **Figure 5.5**. Surprisingly, there was a very small overlap between enriched proteins between the cell lines. These were the two proteins Ighg1, a component of the heavy chain of immunoglobulins (likely present in the sample from the IgG used to make anti-HA beads) and Mac-2BP. Mac2-BP is also known as Galectin-3 binding protein. Galectin-3, expressed in mucosal epithelial cells, binds oligosaccharides containing β-galactoside and modulate cell-cell and cell-matrix interactions. Galectins also have pro-inflammatory functions such as promoting immune cell recruitment, activation and proliferation as well as stimulating NFκB activation in colonic lamia propria fibroblasts (Lippert *et al.*, 2007). However, Gal-3 was shown to play only a minor role in the immune response to *C. rodentium* (Curciarello *et al.*, 2014).

Importantly, Mac2-BP was the only protein identified across all test samples in the mass spectrometry screen for NleG interactors, i.e. for both NleG1 and NleG8 in both HCT116 and iBMM cells. Galectin-7 was also found for NleG1 in HCT116 cells, while Galectin-9 was found for NleG1 and NleG8 in iBMMs. On the one hand, this finding could be interesting if one assumes

that NleG proteins target the same substrates or complexes *in vivo*. On the other hand, we have previously hypothesized that the variable N-termini of NleG proteins reflect their ability to recognize different host partners. Following this hypothesis, interactors found for both NleG1 and NleG8 are unlikely to be true interactors. Moreover, many of the top 10 most enriched proteins for NleG8 in HCT116 cells and iBMMs were equally enriched in the equivalent NleG1-expressing line. For example, the antimicrobial peptides Oas3 and Mmp9 were identified for NleG1 in iBMMs and for NleG8 in iBMMs. Considering the low expression of NleG8, which may compromise how robustly true interactors can be pulled down, it is tempting to assume that these cell-line specific interactors reflect background binding proteins rather than biologically relevant interactors. Importantly, none of the interactors identified for NleG8 were as enriched as C1QBP and UBE2O were for NleG1 (in terms of fold change), and C1QBP and UBE2O were not recognized as enriched in the NleG8 data sets. This finding suggested that the C1QBP and UBE2O putative interactions with NleG1 may be specific to NleG1, making them interesting targets to follow up on. Further efforts therefore focused on validating putative interactors of NleG1.

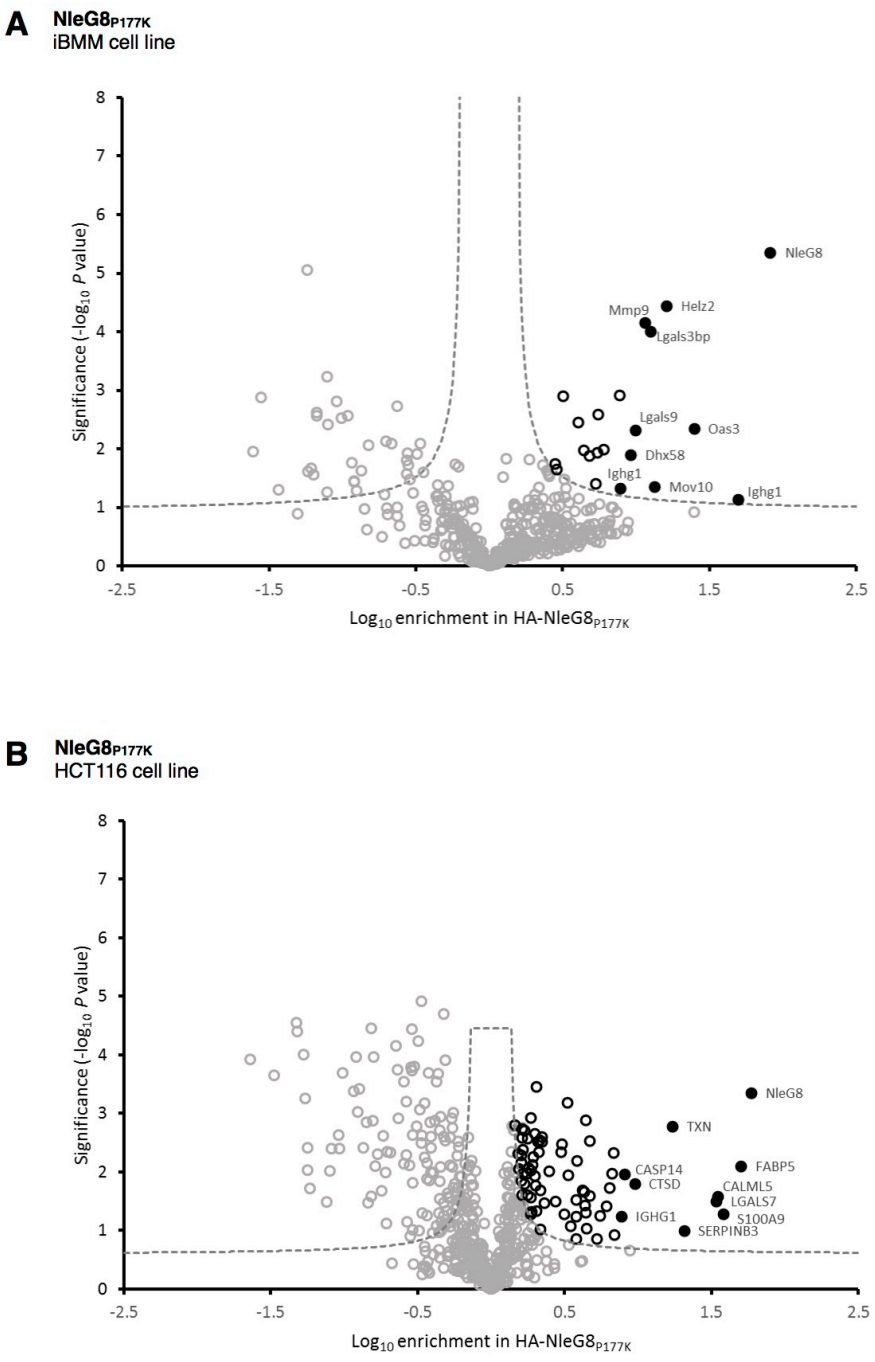


Figure 5.5 | Tandem mass spectrometry analysis of immunoprecipitated NleG8$_{P177K}$–containing complexes from mammalian cell lysates identifies putative interacting partners of NleG8. Volcano plots show the distribution of quantified proteins according to -$\log_{10}$ p-value and $\log_{10}$ fold enrichment, where enrichment is the ratio of label free quantification (LFQ) intensities for proteins in HA-NleG8$_{P177K}$ IP samples to the corresponding LFQ intensities in control IP samples. Technical triplicates were compared by paired t-tests and the dotted line indicates the S0 threshold for significance. Proteins marked with a black outline are significantly enriched more than 1.5-fold in the HA pulldowns from HA-NleG8$_{P177K}$ expressing cells. The ten most enriched proteins are indicated by a filled black circle; labels indicate their respective gene names. **A)** 19 proteins were significantly enriched in the NleG8$_{P177K}$ pull-down sample in iBMMs. **B**) 72 proteins were significantly enriched in the NleG8$_{P163K}$ sample in HCT116 cells.

Table 5.3| Putative interactors of NleG8 in HCT116 cells as identified by AP-MS/MS.

Significant	-log$_{10}$(p-value)	Log$_{10}$ difference	Protein name	Gene name
+	3.354623922	1.768909136	NleG8	NleG8
+	2.096254541	1.701605161	Fatty acid-binding protein, epidermal	FABP5
+	1.288400903	1.580979029	Protein S100-A9	S100A9
+	1.589012174	1.545621236	Calmodulin-like protein 5	CALML5
+	1.495819974	1.534620603	Galectin-7	LGALS7
+	1.000161146	1.31488657	Serpin B3	SERPINB3
+	2.778909396	1.232647896	Thioredoxin	TXN
+	1.797617049	0.976823171	Cathepsin D	CTSD
+	1.963139313	0.908246199	Caspase-14	CASP14
+	1.242837895	0.888128757	Ig gamma-1 chain C region	IGHG1
+	0.933988582	0.836250941	Protein S100-A8; N-terminally processed	S100A8
+	2.329349379	0.834765116	Zinc-alpha-2-glycoprotein	AZGP1
+	1.976990229	0.820710659	Dermcidin	DCD
+	1.737716163	0.80576245	Clathrin light chain A	CLTA
+	1.425400998	0.782426834	Plakophilin-1	PKP1
+	1.260787397	0.741978645	Protein POF1B	POF1B
+	0.866660856	0.718961398	Protein S100-A7; Protein S100-A7A	S100A7; S100A7A
+	2.540282407	0.671478907	Lysine--tRNA ligase	KARS
+	1.592628574	0.669306596	Cystatin-A; N-terminally processed	CSTA
+	1.042111942	0.648525397	Serpin B12	SERPINB12
+	1.306750216	0.645599206	Desmocollin-1	DSC1
+	2.894245136	0.642815272	Protein arginine N-methyltransferase 5; N-terminally processed	PRMT5
+	1.43496052	0.638791879	Protein S100-A6	S100A6
+	1.692044978	0.62052997	Lysozyme C	LYZ
+	2.195624351	0.584262848	Proteasome subunit beta type-6	PSMB6
+	1.525997773	0.579164346	Desmoplakin	DSP
+	1.247547412	0.578226884	Ig alpha-1 chain C region	IGHA1; IGHA2
+	0.85837103	0.577341874	Cornifin-B; Cornifin-A	SPRR1B; SPRR1A
+	1.086260776	0.541355451	Lactotransferrin	LTF
+	1.948653593	0.524381955	Prolactin-inducible protein	PIP
+	3.183908554	0.516386986	Putative RNA-binding protein Luc7-like 2	LUC7L2; C7orf55-LUC7L2
+	1.281583118	0.497286002	Protein-glutamine gamma-glutamyltransferase E	TGM3
+	2.485631648	0.478013833	Ensconsin	MAP7
+	2.345742774	0.474473953	Alpha-enolase	ENO1
+	1.505669955	0.434657415	Catalase	CAT

+	2.0238978	0.395350297	Protein LTV1 homolog	LTV1
+	1.470355053	0.358735085	Prohibitin	PHB
+	2.51926766	0.344656626	40S ribosomal protein S21	RPS21
+	2.598689402	0.343811989	Heat shock cognate 71 kDa protein	HSPA8
+	1.025252125	0.332054138	Arginase-1	ARG1
+	1.685105339	0.331871033	60S acidic ribosomal protein P2	RPLP2
+	2.351157499	0.320934931	Protein LYRIC	MTDH
+	2.516268933	0.315425555	60 kDa heat shock protein, mitochondrial	HSPD1
+	3.458208547	0.30803299	Tubulin beta-4A/B chain	TUBB4B; TUBB4A
+	1.336026094	0.305554549	39S ribosomal protein L11, mitochondrial	MRPL11
+	1.766792475	0.301501115	Nucleolysin TIAR	TIAL1
+	2.654543943	0.29834191	39S ribosomal protein L43, mitochondrial	MRPL43
+	2.265432779	0.282560825	Stress-70 protein, mitochondrial	HSPA9
+	2.132470652	0.278467178	RNA-binding protein FUS	FUS
+	1.317725919	0.272942861	Peroxiredoxin-1	PRDX1
+	1.281934743	0.270769914	DNA-directed RNA polymerases I, II, and III subunit RPABC1	POLR2E
+	1.575459326	0.269154549	Keratin, type I cytoskeletal 18	KRT18
+	2.926996399	0.268288453	DnaJ homolog subfamily C member 10	DNAJC10
+	2.033117205	0.267960866	U1 small nuclear ribonucleoprotein 70 kDa	SNRNP70
+	1.303621936	0.25704368	60S ribosomal protein L28	RPL28
+	2.07009778	0.251319091	Protein-L-isoaspartate O-methyltransferase	PCMT1
+	2.573457426	0.247499784	Protein argonaute-2	AGO2
+	1.615734165	0.24464639	Eukaryotic translation initiation factor 3 subunit A	EIF3A
+	1.982818511	0.239500046	40S ribosomal protein S2	RPS2
+	2.718852686	0.227841854	Ribosomal RNA processing protein 1 homolog B	RRP1B
+	1.797763522	0.226785342	60S ribosomal protein L13a	RPL13A; RPL13a
+	2.389092761	0.216133595	La-related protein 4B	LARP4B
+	2.747557381	0.213949362	28S ribosomal protein S28, mitochondrial	MRPS28
+	1.981786216	0.213363965	28S ribosomal protein S6, mitochondrial	MRPS6
+	1.606220403	0.211865743	40S ribosomal protein S15a	RPS15A
+	2.586859971	0.210893472	40S ribosomal protein S19	RPS19

+	2.738230187	0.210881392	La-related protein 4	LARP4
+	2.284211872	0.205803076	Actin-like protein 6A	ACTL6A
+	2.16622716	0.204910596	Galectin-3-binding protein	LGALS3BP
+	1.859169148	0.203128815	Spermatogenesis-associated serine-rich protein 2	SPATS2
+	2.065505483	0.190209548	Probable ATP-dependent RNA helicase DDX17	DDX17
+	2.316188321	0.181536039	Protein TFG	TFG
+	2.801654179	0.161305586	SRSF protein kinase 1	SRPK1

+ indicates a significant difference compared to the HA control, as determined by a paired T-test with the $-\log_{10}$ (p-value) provided.

Difference is the $\log_{10}$-transformed fold change in abundance of proteins between the HA control and the NleG test sample.

Table 5.4| Putative interactors of NleG8 in iBMM cells as identified by AP-MS/MS.

Significant	-log_{10}(p-value)	Log_{10} difference	Protein name	Gene name
+	5.353031077	1.912561893	NleG8	NleG8
+	1.134257119	1.69443051	Ig gamma-1 chain C region	Ighg1
+	2.351882715	1.394205411	2-5-oligoadenylate synthase 3	Oas3
+	4.437923191	1.203560988	Helicase with zinc finger domain 2	Helz2
+	1.3573635	1.124204159	Putative helicase MOV-10	Mov10
+	4.009114986	1.096828302	Galectin-3-binding protein	Lgals3bp
+	4.151412099	1.056671937	Matrix metalloproteinase-9	Mmp9
+	2.325013467	0.995346546	Galectin; Galectin-9	Lgals9
+	1.897780286	0.964145819	Probable ATP-dependent RNA helicase DHX58	Dhx58
+	1.331732371	0.889843623	Ig gamma-1 chain C region	Ighg1
+	2.918985438	0.887506326	2-5-oligoadenylate synthase-like protein 1	Oasl1
+	1.99538535	0.774681409	ATP-dependent RNA helicase DDX3Y	Ddx3y
+	2.596586906	0.737702211	Squamous cell carcinoma antigen recognized by T-cells 3	Sart3
+	1.936239154	0.731386979	Interferon-induced protein with tetratricopeptide repeats 1	Ifit1
+	1.407244227	0.721748352	Small nuclear ribonucleoprotein G	Snrpg
+	1.881387275	0.680174828	Ig kappa chain C region	Igkc
+	1.982383898	0.641870181	Interferon-induced protein with tetratricopeptide repeats 3	Ifit3
+	2.462127264	0.601484299	Poly [ADP-ribose] polymerase 12	Parp12
+	2.907434887	0.49883906	Protein FAM98A	Fam98a
+	1.658521704	0.459768613	E3 ubiquitin-protein ligase TRIM21	Trim21
+	1.750624297	0.445387363	Small nuclear ribonucleoprotein F	Snrpf

+ indicates a significant difference compared to the HA control, as determined by a paired T-test with the –log_{10} (p-value) provided.

Difference is the log_{10}-transformed fold change in abundance of proteins between the HA control and the NleG test sample.

5.5.3. UBE2O, C1QBP and CAPER and reproducibly pulled down with NleG1 but not NleG8 in mammalian cells

To start validating the interactors of NleG1 found in the initial mass spectrometry screen, a second biological repeat of the affinity purification mass spectrometry experiment was performed in HCT116 cells. This cell line was chosen for the repeat as it is more physiologically relevant and had given a higher number of putative interactors in the first experiment. The second repeat identified 47 host interactors as significantly enriched for NleG1 (see **Figure 5.6** and **Table 5.5** for details). Importantly, Mac-2BP was not identified in this repeat, supporting the decision to focus on proteins exclusively enriched for NleG1 and not focus on proteins enriched for both NleG1 and NleG8. Certain proteins, such as Nucleosome assembly protein 1-like 4 (NAP1L4) were found enriched in first repeat (over 30-fold) and second repeat (26-fold), although it was not enriched in the iBMM cell line in the first repeat.

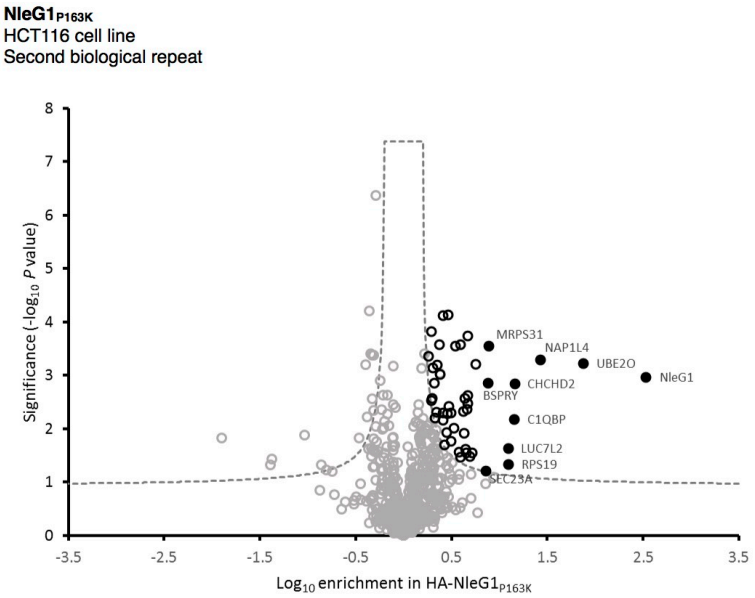
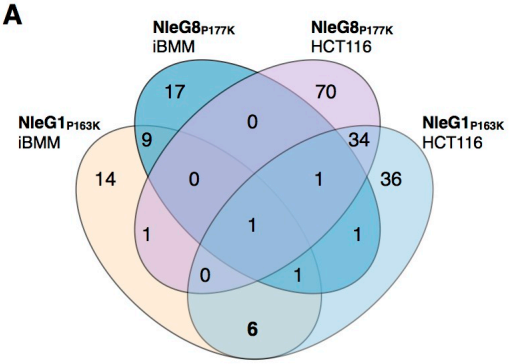


Figure 5.6 | A number of proteins are reproducibly enriched in HA-immunoprecipitation samples from lysates of HCT116 cells expressing HA-NleG1$_{P163K}$. A biological repeat of the NleG1$_{P163K}$ pulldown in HCT116 lysates. Volcano plots showing the distribution of quantified proteins according to -log$_{10}$ p-value and log$_{10}$-fold enrichment, where enrichment is the ratio of label free quantification (LFQ) intensities for proteins in the HA-NleG1$_{P163K}$ IP samples to the corresponding LFQ intensities in the control IP samples. Technical triplicates were compared by paired t-tests; the dotted line indicates the S0 threshold for significance. Proteins marked with a black outline are significantly enriched more than 1.5-fold in the HA pulldowns from HA-NleG1$_{P163K}$ expressing cells. The ten most highly enriched proteins are indicated by a filled black circle; labels indicate their gene names. 47 proteins were significantly enriched in the NleG1$_{P163K}$ sample from HCT116 lysates, 9 of which were also enriched in the first repeat of this experiment.

Venn-diagrams of number of overlapping hits between the different cell lines and repeats and summary tables of proteins identified in NleG1 samples but not NleG8 samples are provided in **Figure 5.7**. C1QBP and UBE2O remained among the top 10 most-enriched proteins in the second repeat of the experiment, and together with CAPER were the only 3 hits that were identified for

NleG1 in HCT116 cells (both repeats) and iBMMs but were not identified as hits for either NleG8 sample. CAPER is an RNA binding protein that functions as a pre-mRNA splicing factor and transcriptional co-activator (Loerch *et al.*, 2014). Validating CAPER as an interactor was not deemed a priority, for three reasons. Firstly, ribonuclear proteins and transcription factors are common contaminants in affinity purification mass spectrometry. Secondly, CAPER is not known to contribute to inflammatory processes or pathways that represent obvious targets for manipulation by *C. rodentium* effector proteins. Thirdly, CAPER localises to nuclear speckles, while NleG1 localisation in cultured mammalian cells is predominantly cytosolic and diffuse. Importantly, however, prioritizing identified protein-protein interactions for further analyses is not trivial and represents a fundamental challenge in mass spectrometry follow-up studies. Based on the data generated in the mass spectrometry experiments, we chose to commence validation experiments focusing on the two proteins C1QBP and UBE2O, which both display cell-wide localisation akin NleG1.

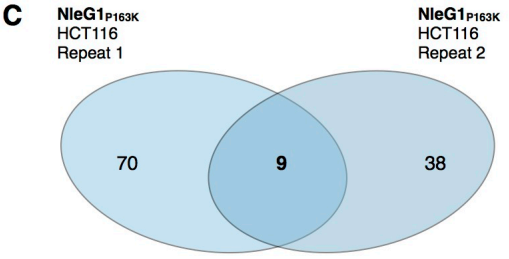
A



B

Enriched in both NleG1$_{P163K}$ samples *iBBM and HCT116 cells*	Uniquely enriched in NleG1$_{P163K}$ samples *Not enriched in NleG8$_{P177K}$ samples*
UBE2O	+
C1QBP	+
MRPS30	+
PTCD3	+
ZAP	+
CAPER	+
TRIM21	
Mac-2BP	

C



D

Enriched in both NleG1$_{P163K}$ biological repeats *HCT116 cells only*	Uniquely enriched in NleG1$_{P163K}$ samples (HCT116/iBMMs) in repeat 1 *Not enriched in NleG8$_{P177K}$ samples*
UBE2O	+
C1QBP	+
CAPER	+
NAP1L4	
LUC7L2	
SNRNP70	
TUBB	
ATP5A1	
CHCHD2	

Figure 5.7| Overview of the proteins identified as putative interacting partners of NleG1 and NleG8 by affinity purification-coupled mass spectrometry experiments in this study. A) Venn-diagram showing the overlap of significantly enriched proteins between the four test samples compared to HA-controls in this study. One protein, Mac-2BP, was enriched across all NleG IP samples. **B)** List of the 8 proteins enriched for NleG1$_{P163K}$ in both cell backgrounds. Six of these were not enriched in either NleG8$_{P177K}$ sample (indicated by +). **C)** Venn diagram showing the overlap in specifically enriched proteins in the two biological repeats of the HA-IP and mass spectrometry analysis of lysates from HCT116 expressing HA-NleG1$_{P163K}$. **D)** The 9 proteins enriched in NleG1$_{P163K}$ HCT116 lysates in both repeats. + indicates proteins not enriched in either NleG8$_{P177K}$ sample in the first repeat, therefore thought to be specifically enriched for NleG1$_{P163K}$.

Table 5.5| Putative interactors of NleG1 in HCT116 cells as identified by AP-MS/MS. *Repeat 2.*

Significant	-log$_{10}$(p-value)	Log$_{10}$ difference	Protein name	Gene name
+	2.969252	2.529656887	NleG1	NleG1
+	3.235463	1.875426133	E2/E3 hybrid ubiquitin-protein ligase UBE2O	UBE2O
+	3.304311	1.427687486	Nucleosome assembly protein 1-like 4	NAP1L4
+	2.853226	1.160581271	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2; mitochondrial	CHCHD2; CHCHD2P9
+	2.184581	1.154973507	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP
+	1.345715	1.093058427	40S ribosomal protein S19	RPS19
+	1.640254	1.089846134	Putative RNA-binding protein Luc7-like 2	LUC7L2
+	3.550594	0.890048345	28S ribosomal protein S31, mitochondrial	MRPS31
+	2.857496	0.882370949	B box and SPRY domain-containing protein	BSPRY
+	1.212557	0.855055968	Protein transport protein Sec23A	SEC23A
+	3.221287	0.749967893	Serine/arginine repetitive matrix protein 1	SRRM1
+	1.556838	0.713104089	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	SLC9A3R1
+	1.49098	0.691506068	Uncharacterised protein C11orf98	C11orf98
+	2.634585	0.670532544	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	PPP2R1A
+	2.483195	0.665207386	28S ribosomal protein S5, mitochondrial	MRPS5
+	3.75431	0.665068626	39S ribosomal protein L1, mitochondrial	MRPL1
+	2.370336	0.657614708	39S ribosomal protein L43, mitochondrial	MRPL43
+	1.553899	0.653517564	Fatty acid synthase	FASN
+	1.628227	0.642223994	T-complex protein 1 subunit delta	CCT4
+	2.574456	0.63827308	Cullin-1	CUL1
+	1.92841	0.626882076	T-complex protein 1 subunit gamma	CCT3
+	2.328333	0.624972661	Luc7-like protein 3	LUC7L3
+	3.590575	0.591479619	RNA-binding protein 39	RBM39
+	1.475297	0.591413021	DNA-directed RNA polymerase II subunit RPB2	POLR2B

+	1.564652	0.574618816	26S proteasome non-ATPase regulatory subunit 1	PSMD1
+	3.552424	0.536095937	Protein transport protein Sec61 subunit alpha isoform 1; Protein transport protein Sec61 subunit alpha isoform 2	SEC61A1; SEC61A2
+	2.018262	0.523085594	U1 small nuclear ribonucleoprotein 70 kDa	SNRNP70
+	2.305968	0.489367008	Proteasome subunit beta type-5	PSMB5
+	1.771754	0.488970439	ADP/ATP translocase 2; ADP/ATP translocase 2, N-terminally processed; ADP/ATP translocase 1	SLC25A5; SLC25A4
+	2.426531	0.469517072	60S ribosomal protein L19	RPL19
+	4.142734	0.46167024	39S ribosomal protein L22, mitochondrial	MRPL22
+	2.297055	0.45552063	40S ribosomal protein S3	RPS3
+	1.936844	0.448682308	Peptidyl-tRNA hydrolase ICT1, mitochondrial	ICT1
+	1.710899	0.42594258	60S ribosomal protein L24	RPL24
+	2.164064	0.40877978	Tubulin beta chain	TUBB
+	4.133592	0.407900174	60S ribosomal protein L13	RPL13
+	2.308229	0.407875697	Glutamine-dependent carbamoyl-phosphate synthase; Aspartate carbamoyltransferase	CAD
+	3.032038	0.375442505	Ribosomal L1 domain-containing protein 1	RSL1D1
+	3.578879	0.368048986	ATP synthase subunit alpha, mitochondrial	ATP5A1
+	3.204279	0.348578135	Elongation factor Tu, mitochondrial	TUFM
+	2.320855	0.339098612	Tubulin beta-4A/B chain	TUBB4A/B
+	2.203742	0.323760986	Drebrin	DBN1
+	2.867212	0.31496493	Proteasome subunit beta type-6	PSMB6
+	3.143014	0.305315971	RNA-binding protein 12B	RBM12B
+	2.574109	0.292956988	Pre-rRNA-processing protein TSR1 homolog	TSR1
+	2.540086	0.290406704	Bifunctional glutamate/proline--tRNA ligase	EPRS
+	3.828636	0.286535899	60S ribosomal protein L6	RPL6
+	3.372375	0.257551988	LIM domain and actin-binding protein 1	LIMA1

+ indicates a significant difference compared to the HA control, as determined by a paired T-test with the $-\log_{10}$ (p-value) provided.

Difference is the $\log_{10}$-transformed fold change in abundance of proteins between the HA control and the NleG test sample.

5.6. Validating NleG1 interactions with C1QBP and UBE2O by repeated immunoprecipitations and western blot analysis

The sensitivity of the mass spectrometry is its core strength as a technique. However, it also calls for validation experiments to establish the biological relevance of putative protein-protein interactions. The following half of this chapter describes efforts to validate the mass spectrometry data, with a range of cell-free, cell-based and functional assays. A starting point for validation experiments was to repeat the IP performed for mass spectrometry analysis, and analysing IP samples for putative interactors by western blot. Owing to the difference in NleG expression between the iBMMs and HCT116 cells, we opted for performing all validation experiments in the more physiologically relevant HCT116 cell lines, with more robust transgene expression and two associated biological repeats of the mass spectrometry data. Similarly, NleG8 was not used as a specificity control due to the aforementioned low expression of NleG8.

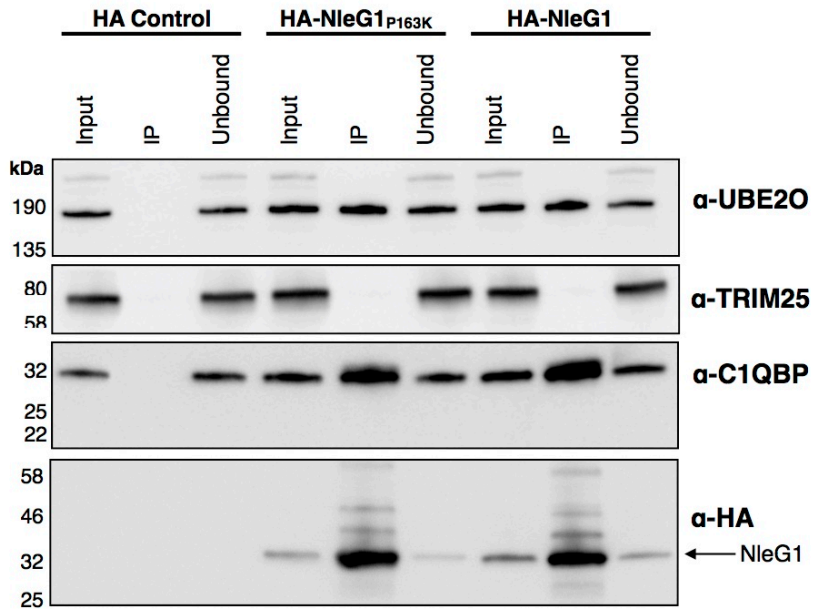


Figure 5.8| Endogenous UBE2O and C1QBP co-immunoprecipitate with wildtype and catalytically inactive HA-NleG1 in HCT116 cells. Lysates from cells expressing HA, HA-NleG1 or HA-NleG1$_{P163K}$ were subjected to anti-HA IP followed by SDS-PAGE and immunoblotting. Probing with an anti-HA antibody showed that the IP was successful in pulling down HA-NleG1 proteins (indicated by arrow; bottom panel). Immunoblotting using anti-UBE2O and anti-C1QBP antibodies gave strong bands in the IP lanes for NleG1 and HA-NleG1$_{P163K}$, but no band in the HA control IP lane, indicative of an interaction. An anti-TRIM25 antibody was used as a negative control, see text for more detail.

Analysing IP samples by SDS-PAGE and immunoblotting revealed that both UBE2O and C1QBP robustly and reproducibly co-immunprecipitated with NleG1 (**Figure 5.8,** as indicated by bands in lanes 5 and 8, but not in the control IP in lane 2). This held true for both the catalytic dead mutant NleG1$_{P163K}$ (used as a bait in the mass spectrometry screen) and wildtype NleG. TRIM25,

which was unchanged in the MS data, was probed for as a negative control. As expected, TRIM25 did not co-immunoprecupitate with NleG1, reassuring that the stringency of the mass spectrometry analysis had not been excessive.

A HA blot was performed to confirm that the IP has successfully pulled down HA-NleG from the cell lysate (**Figure 5.8, bottom panel**). This showed pulldown of NleG1, but also revealed the multiple banding pattern previously observed. To test if the banding pattern corresponded to mono-, di-, and poly-ubiqutinated HA-NleG1, we blotted for ubiquitin. Surprisingly, the banding was not due to ubiquitination of NleG1, as shown in **Figure 5.9**, as there were no bands in the IP lanes.

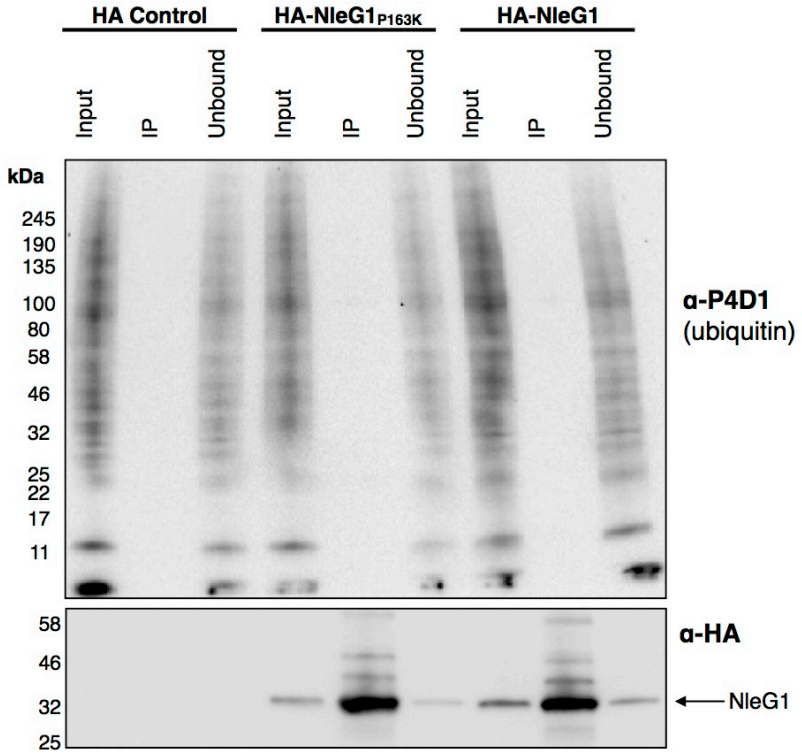


Figure 5.9| NleG1 and NleG1$_{P163K}$ expressed in HCT116 cells is not ubiquitinated. Lysates from cells transduced with HA, HA-NleG1 or HA-NleG1$_{P163K}$ were subjected to anti-HA IP followed by SDS-PAGE and immunoblotting. Probing with an anti-HA antibody showed banding of NleG. Probing with an anti-ubiquitin antibody (clone P4D1) that recognises mono- as well as di- and poly-ubiquitination revealed that the multiple bands were not due to ubiquitination of the NleG as there were no bands of the corresponding molecular weights in lanes 5 and 8. Further, it showed that the pull down did not enrich for ubiquitinated proteins to a level detectable by western blot.

Having established that the bands were not ubiquitinated NleG, we tested if they were NleG covalently modified by other small ubiquitin-like modifiers. Anti-NEDD8, anti-SUMO1 and anti-SUMO2/3 antibodies were used to test if NleG was modified by NEDD8 (a 9 kDa modifier) or a SUMO protein (a 12 kDa modifier) (**Figure 5.10**). This revealed no bands for NEDD8 and SUMO2/3. Intriguingly, however, one band was observed using anti-SUMO1 for the IP with HA-

NleG1$_{P163K}$. Since its approximate molecular weight was slightly lower than the size that HA-NleG1 runs at (and over 10 kDa lower than the expected size of sumolylated HA-NleG1 of 44 kDa), it was unlikely that this band was sumoylated NleG1. Instead, it was more likely that a host protein pulled down with NleG1$_{P163K}$ is sumoylated in HCT116 cells. Although this was interesting, the time constraints of the project did not allow for further analysis of this unknown sumoylated protein, and priority was given to validating the interactors of NleG1.

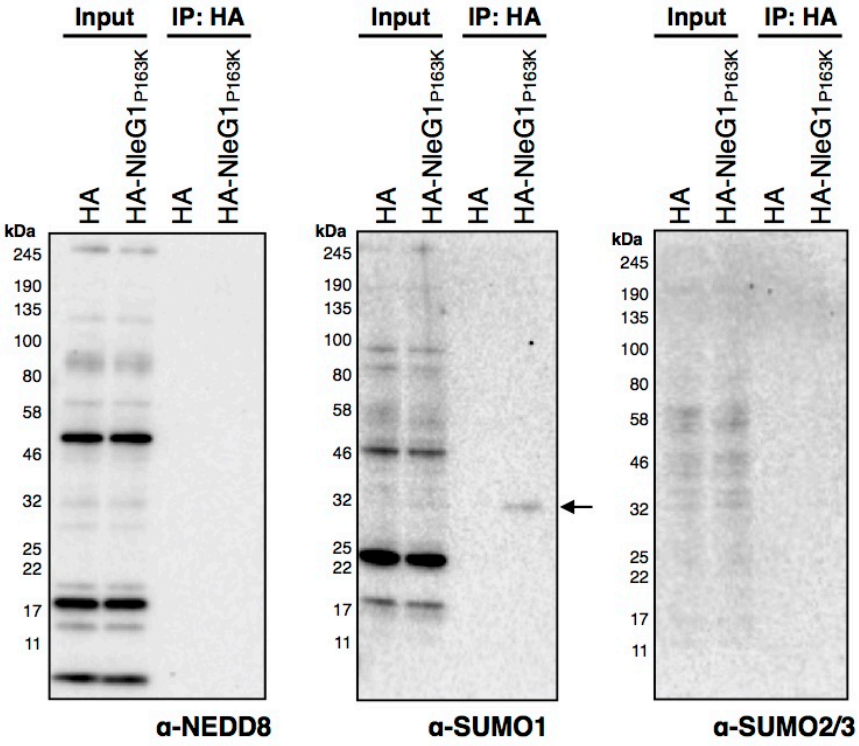


Figure 5.10| Banding of NleG1 and NleG1$_{P163K}$ from IPs using HA-beads is not due to neddylation or sumoylation of the NleG. Lysates from cells transduced with HA, HA-NleG1 or HA-NleG1$_{P163K}$ were subjected to anti-HA IP followed by SDS-PAGE and immunoblotting. Probing with anti-NEDD8, anti-SUMO1 or anti-SUMO2/3 antibodies showed that the multiple banding seen for NleG in IP lanes was not due to neddylation or sumoylation. However, one band was seen in the anti-SUMO1 blot (middle panel, see arrow), indicating that NleG1$_{P163K}$ can be mono-sumoylated in HCT116 cells, or that a protein pulled down in the IP of roughly the same molecular weight as NleG1 is sumoylated.

5.7. The interaction of NleG1 and NleG8 with C1QBP in HeLa cell lysates

The low expression of NleG8 in the transgenic cell lines ruled out direct comparisons between the NleG1- and NleG8-expressing cells. Another way to determine whether the putative C1QBP-NleG1 interaction was specific to NleG1 was to pull down C1QBP from HeLa lysates using immobilised purified recombinant His-NleG1 and His-NleG8. His-tags were used in lieu of HA-tags, as true interactions should be detectable regardless of the tag used (provided that the tag itself does not block the interaction due to its size or positioning). WT NleG1 and NleG8 (not the E3 ligase null mutants) were used as NleG8$_{P177K}$ is insoluble (**Chapter 3, Section 3.1.4**), but the

interaction between NleG1 and C1QBP had been observed for WT HA-NleG1 in HCT116 cells. Nevertheless, no interaction was observed by this methodology (**Figure 5.11**). Endogenous C1QBP did not co-immumoprecipitate with His-NleG1 or His-NleG8, demonstrated by the lack of band for C1QBP following elution of His-NleGs from the TALON resin.

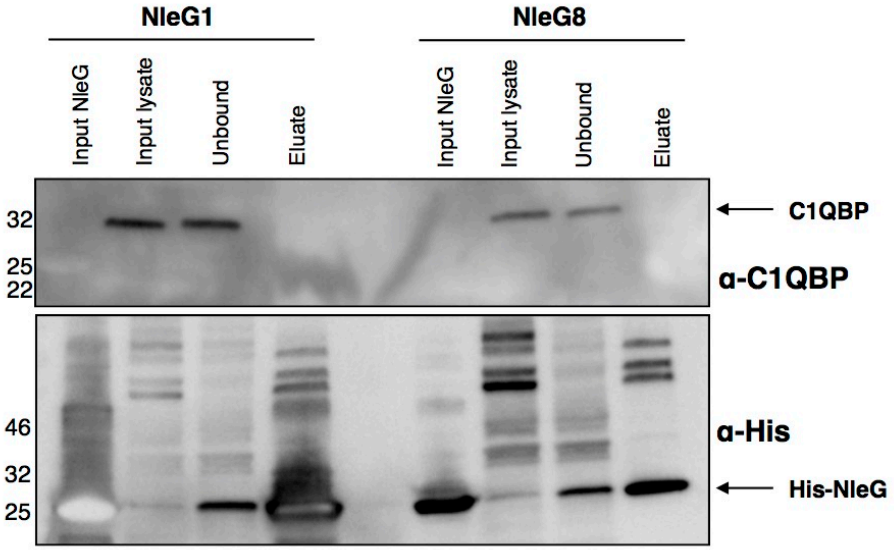


Figure 5.11| Recombinant His-NleG1 and His-NleG8 do not interact with C1QBP from HeLa cell lysates. Purified recombinant His-NleG1 and His-NleG8 were applied to TALON His-resin subsequently incubated with HeLa cell lysate, allowing bound NleGs to complex with proteins within HeLa lysate. Background binders were removed by washing; NleGs and interacting partners were eluted from using 100 mM imidazole. Input NleG, input lysate, unbound and eluate samples were separated by SDS-PAGE and analysed by immunoblotting. Western blot using an anti-His antibody showed that the NleGs were enriched in the eluate (bottom panel; indicating that the IP itself had worked). However, blotting with an anti-C1QBP antibody (top panel) showed that the C1QBP and NleG did not co-immunoprecipitate under these experimental conditions.

5.8. Reciprocal IP for the NleG1-UBE2O interaction

We used a cell-based system to perform validation experiments for UBE2O. A reciprocal IP to the original IP was performed, where the aim is to pull down the original bait protein using the mass spectrometry-identified prey protein. A FLAG-UBE2O construct (kindly gifted by Dr Ramanujan Hegde, MRC Laboratory of Molecular Biology, Cambridge) was transfected into HCT116 cells in parallel to FLAG-GFP transfection. The HCT116 cells were induced to express HA-NleG1 or HA-NleG1$_{P163K}$. Lysates were then subjected to anti-FLAG IP followed by analysis by SDS-PAGE and western blot. This showed that FLAG-tagged UBE2O was successfully pulled down (**Figure 5.12**, top two panels, see lanes 6 and 8). This was also observed for FLAG-GFP (anti-GFP blot, lanes 5 and 7) where relevant. Nevertheless, the anti-HA blot (bottom panel) revealed no pull down of HA-tagged proteins. It was therefore concluded that NleG1 (WT and ligase dead) was not able to interact with UBE2O under these conditions. This started to raise concerns about the validity and

biological importance of the NleG1-UBE2O interaction identified in the mass spectrometry screen.

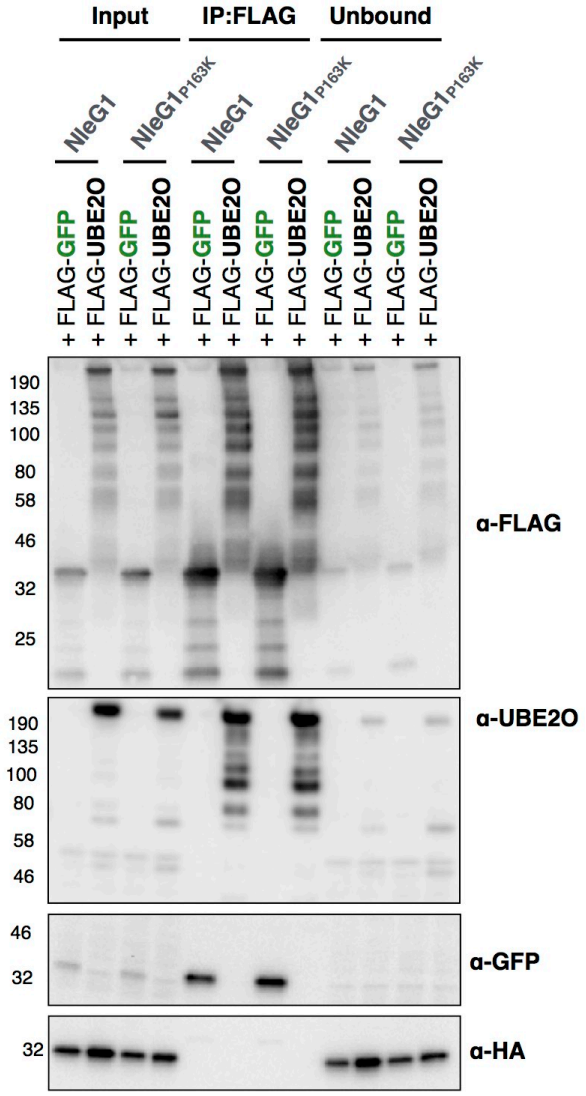


Figure 5.12| *In vitro* FLAG pulldown shows no interaction between NleG1 and UBE2O when the proteins are co-expressed in HCT116 cells. Stably transduced cells expressing HA-NleG1 or HA-NleG1$_{P163K}$ were transiently transfected with a FLAG-UBE2O or FLAG-GFP construct for 24 hours. Cell lysates were subjected to anti-FLAG immunoprecipitation followed by SDS-PAGE and immunoblotting. Western blot with anti-FLAG and anti-UBE2O confirmed pulldown of FLAG-UBE20 (top two panels; lanes 6 and 8); western blot with the anti-GFP antibody confirmed pulldown of FLAG-GFP (lanes 5 and 7). However, western blot with an anti-HA antibody (bottom panel) detected no HA-tagged proteins in the pulldown when FLAG-UBE2O and HA-NleG1 (wildtype or ligase dead) were expressed in the same cell (lanes 6 and 8), indicating that NleG1 does not interact with UBE2O under these conditions.

5.9. Studying the interaction between C1QBP and NleG7 or NleG8 by direct yeast two-hybrid

As discussed in **Chapter 3**, yeast direct two-hybrid methods can be valuable when testing putative interactions. The assay was therefore used to test interactions with NleG1 here. Only the interaction with C1QBP was tested, owing to UBE2O being difficult to clone (due to its large size). However, no interaction was observed between NleG1 and C1QBP (full length or the mature form), evidenced by lack of growth on QDO media of yeast co-transformed with pGADT7 and pGBKT7 constructs (**Figure 5.13**). As highlighted in the **Chapter 3**, it remains uncertain whether the expression of NleG1 is sufficiently high in yeast AH109. However, NleG7 and NleG8 gave positive interactions with ubiquitin previously, meaning that these proteins are expressed in the yeast. The lack of growth on QDO following yeast co-transformation of the prey pGADT7-C1QBP and bait pGBKT7-NleG7/8 constructs was therefore more informative, telling us that C1QBP was unable to interact with NleG7 or NleG8 in this yeast system.

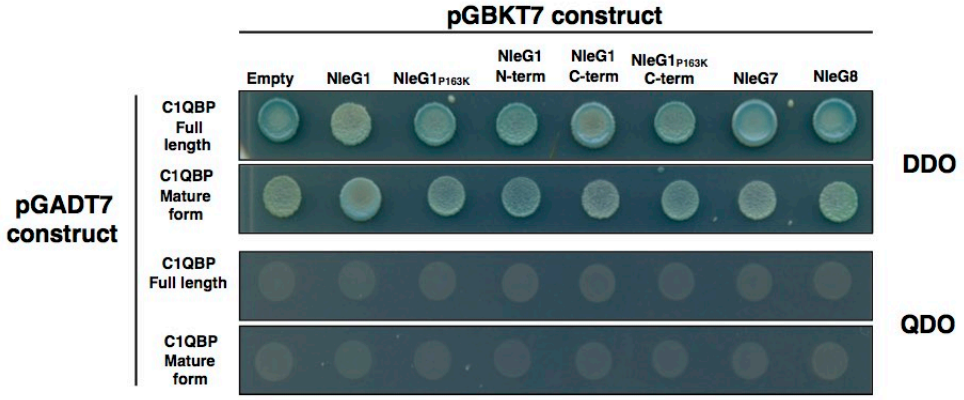


Figure 5.13| Direct yeast two-hybrid assay did not confirm an interaction between C1QBP and *C. rodentium* NleG proteins. Yeast AH109 were co-transformed with a pGADT7-C1QBP construct and a pGBKT7-NleG construct (or empty vector). Co-transformants were plated on DDO and QDO plates. C1QBP did not interact with *C. rodentium* NleG1 (full length, C-terminal or N-terminal), NleG7 or NleG8, as indicated by the lack of growth on QDO. This was also true for enzymatically inactive NleG1$_{P163K}$ (full length or C-terminal). Images were taken after 3 days of growth at 30°C.

As this methodology failed to validate the interaction between NleG1 and C1QBP, and validations for UBE2O had also been fruitless, it was deemed more interesting to turn attention to whether NleG1 can ubiquitinate these proteins in cell-free or cell-based systems.

5.10. NleG1 family members can ubiquitinate C1QBP in vitro

In the context of ubiquitin ligases and their partners, it is useful to determine whether the outcome of an interaction results in the partner being ubiquitinatied. A partner can also be part

of a larger complex where other components of the complex are the true substrates for the E3 ligase. Further, it is useful to ascertain the outcome of ubiquitination. In many cases, the result is proteasomal degradation of the partner. Owing to the complexity of the ubiquitin code, ubiquitintion may also result in changes in activation status, subcellular localisation, protein stability or ability to bind proteins in downstream pathways. Linkage-specific ubiquitination in particular can give important clues regarding the outcome of an interaction where a substrate is ubiquitinated.

In this section, cell-free ubiquitination assays were performed with C1QBP. This experiment called for recombinant C1QBP, which was kindly purified by Dr Olga Kotik. Purification was attempted with both full-length and mature C1QBP. As shown in **Figure 5.14A**, full length C1QBP was insoluble when expressed in *E. coli*, while mature C1QBP [74-282] was soluble. A His-tag was used to facilitate purification, but this could be cleaved off using a TEV protease (**Figure 5.14B**) when necessary.

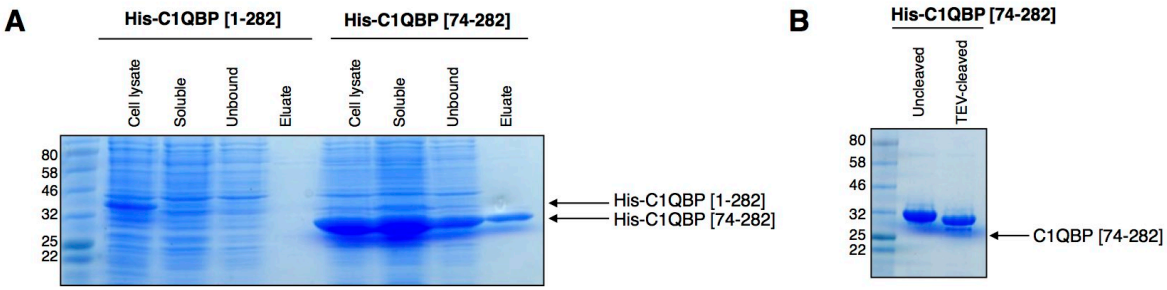


Figure 5.14| Preparation of recombinant C1QBP for *in vitro* assays. A) Full length C1QBP or mature C1QBP (amino acid residues 74 to 282) with a TEV-cleavable N-terminal His-tag was expressed in *E. coli* B21 Star. Cleared cell lysates were applied to TALON His-tag purification resin for protein purification. Whole cell lysates, soluble fractions, unbound fractions and eluates were separated by SDS-PAGE and stained with Coomassie. Only the mature form was soluble. **B**) The His-tag was cleaved from His-C1QBP$_{74-282}$ using TEV protease and the protein re-applied to the TALON His-tag purification resin to remove the His-tag. Cleaved and uncleaved protein compared by SDS-PAGE and Coomassie. Relevant proteins are indicated by arrows. *Protein expression and purification was performed by Dr. Olga Kotik.*

Purified C1QBP was then used in *in vitro* ubiquitination assays, performed as previously described (see **Chapter 3**) but with C1QBP added as a substrate. The ubiquitination reactions contained all the components necessary for NleG to exert its E3 ligase function (either auto-ubiquitination or ubiquitinating a substrate), and control reactions lacked one or more of the components (as indicated in the figure). NleG1 and NleG8 were both tested, along with NleG1$_{P163K}$. Reactions were analysed by SDS-PAGE and western blot. As shown in **Figure 5.15**, NleG1 and NleG8 were able to ubiquitinate C1QBP under the assay conditions. Ubiquitin smears

indicated that ubiquitination had occurred, while C1QBP smears indicated that the observed ubiquitin smears were not simply due to autoubiquitination of the E3 ligase. This tells us that in this system, C1QBP and NleG proteins form an interaction sufficient for a ubiquitin ligation reaction to occur.

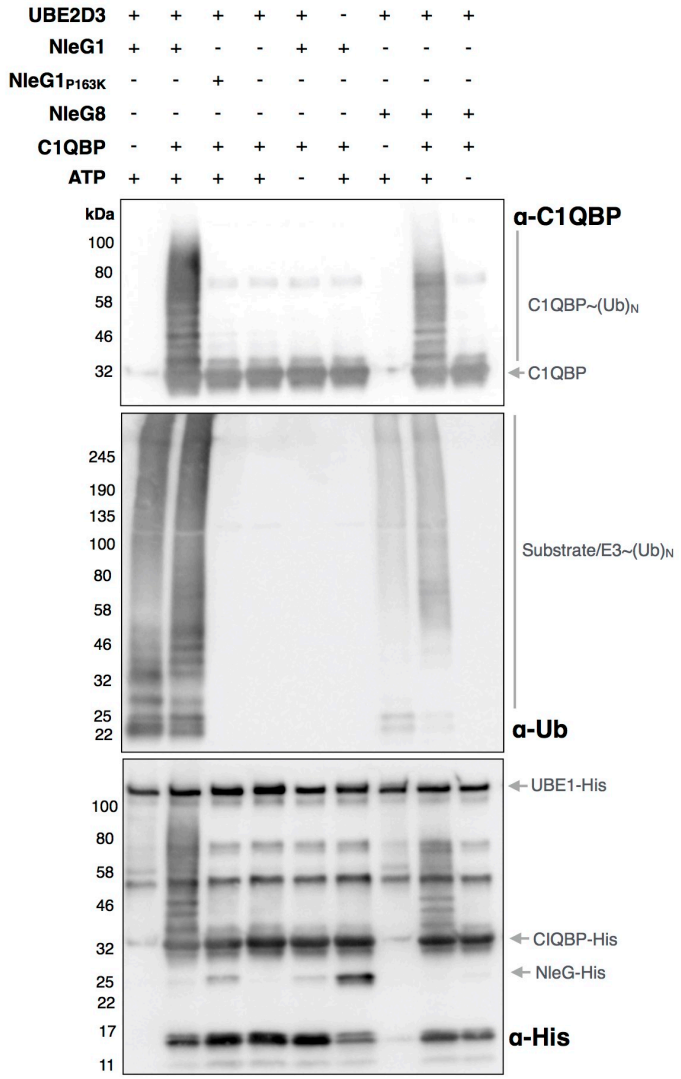


Figure 5.15| *C. rodentium* NleG1 and NleG8 can support ubiquitination of C1QBP *in vitro*. Ubiquitination assays were performed using His-NleG1 [10-194], His-NleG1$_{P163K}$ [10-194], His-NleG8 [11-210] as E3 ligases and with His-C1QBP [74-282] as a substrate. Reactions were analysed by SDS-PAGE and western blot using an anti-ubiquitin (FK2) antibody to detected ubiquitin conjugation. An anti-C1QBP was used to detect C1QBP and anti-His to detect His-tagged proteins. All reactions contained ubiquitin and UBE1-His. UBE2D3, E3 ligases, C1QBP and ATP were added as indicated.

The ubiquitination assay reactions were also analysed for linkage-specific ubiquitination. Blotting with linkage-specific antibodies revealed that NleG1 can support K48- and K63-linked ubiqtuitination *in vitro*, when paired with UBE2D3 as the E2 ligase (**Figure 5.16**). The ubiquitin linkage typically depends on the E2 for RING-type E3 ligases, but E3 ligases may still display a

preference for linkage type. Some E3s are inherently incapable of supporting certain linkage types (sometimes due to steric hindrance). Knowing that NleG1 can support K48- and K63-linked ubiquitination does not tell us much about NleG function, but is nonetheless an additional bit of knowledge regarding NleG biology.

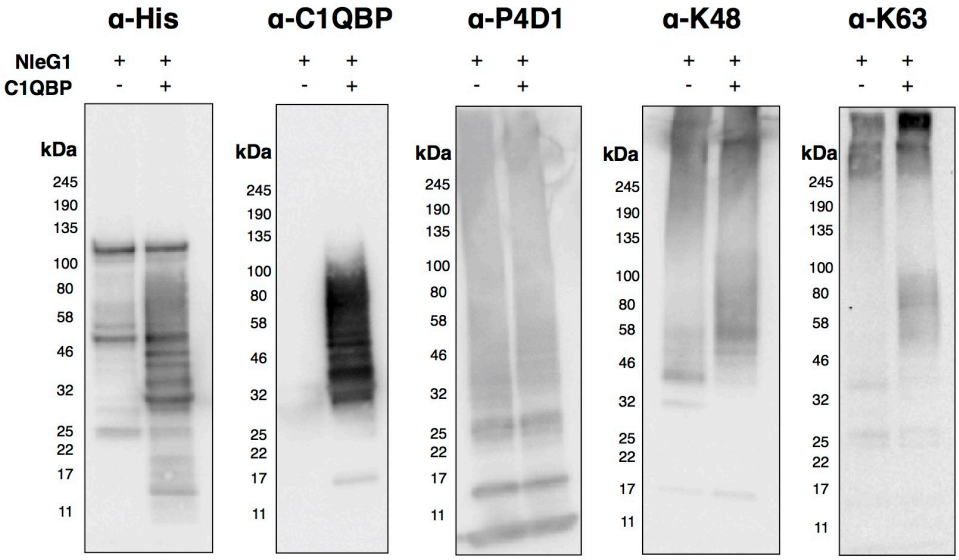


Figure 5.16| *C. rodentium* NleG1 can support K48- and K63-linked ubiquitination of itself and of C1QBP *in vitro*. Ubiquitination assays were performed with His-NleG1 [10-194] as an E3 ligase and His-C1QBP [74-282] as a putative substrate. Reactions were analysed by SDS-PAGE and western blot using an anti-ubiquitin (P4D1) antibody to detected mono-ubiquitin and ubiquitin conjugation alongside antibodies specific to K48- and K63-linked ubiquitination. An anti-C1QBP was used to detect C1QBP and anti-His to detect His-tagged proteins. All reactions contained ubiquitin, UBE1-His, UBE2D3, NleG1 and ATP. C1QBP was added as indicated.

Having demonstrated ubiquitination of C1QBP by both NleG1 and NleG8, we investigated whether the ubiquitination of C1QBP in this assay was NleG-specific. StoD and NleL were therefore included as additional controls. NleL is also a bacterial E3 ligase, but it has a HECT-type rather than RING-type domain. This showed that the NleG family members NleG1, NleG8 and StoD all ubiquitinated C1QBP (**Figure 5.17**). NleL, however, did not (full length or catalytic domain on its own). Regrettably, no purified recombinant non-NleG family member RING-type bacterial E3 ligase was available to use, and it was therefore not possible to ascertain whether ubiquitination of C1QBP was due to the presence of a RING-type ligase or an NleG family member. Importantly, observing ubiquitination of C1QBP in this cell-free assay does not mean that this occurs in cells *in vivo* and during *C. rodentium* infection. The focus was therefore switched back to cell-based systems.

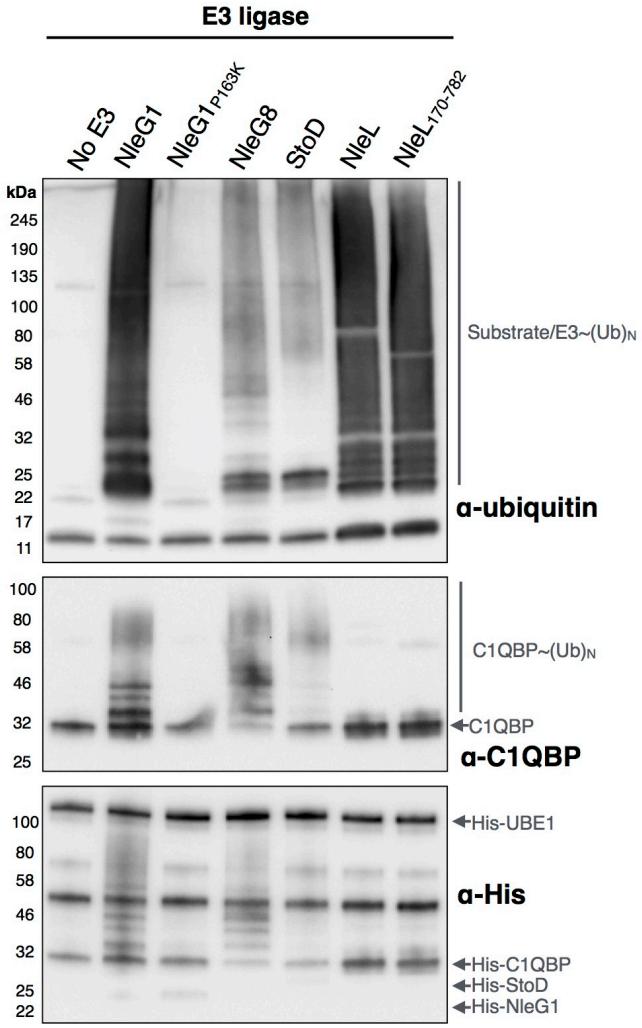


Figure 5.17| The NleG family members NleG1, NleG8 and StoD but not the HECT-type bacterial E3 ligase NleL can ubiquitinate C1QBP *in vitro*. Ubiquitination assays were performed to test the ability of the E3 ligases His-NleG1 [10-194], His-StoD, His-NleG8 [11-210], NleL [1-782] and NleL [170-782] to ubiquitinate His-C1QBP [74-282]. E3 ligase-null NleG1 ($NleG1_{P163K}$) was included as a negative control. All reactions contained an E3 ligase, His-C1QBP, ubiquitin, UBE1-His, UBE2D3 and ATP. Reactions were analysed by SDS-PAGE and western blot using an anti-ubiquitin antibody to detected ubiquitin conjugation alongside antibodies against C1QBP and His-tagged proteins. While autoubiquitination was observed for all above E3 ligases (indicated by a ubiquitin smear) bar the negative control, only the NleG family proteins could ubiquitinate C1QBP (indicated by a C1QBP banding and smear).

5.11. Ubiquitination of C1QBP and UBE2O by NleG1 in cultured HCT116 cells

A way to assess the ubiquitination state of a protein in the cell is by subjecting the lysate to an enrichment of ubiquitin using commercially available Tandem Ubiquitin-Binding Entities (TUBEs) (Hjerpe *et al.*, 2009). TUBEs carry an array of domains with high avidity and affinity for polyubiquitin chains. They are a specialised tool for enriching for ubiquitinated substrates by effectively performing an IP. TUBE-IPs were originally used to describe cell ubiquitinomes, but, are now commonly used to validate substrates of orphan E3 ligases. TUBE-IP (or Ubi-IPs) were

performed here on lysates from the HA-NleG1- and HA-NleG1$_{P163}$-expressing HCT116 cells alongside the HA control. Prior to cell lysis, cells were maintained with both a proteasomal inhibitor (20 µM MG132) and a de-ubiquitinase inhibitor (5 µM PR-619) to prevent proteasomal degradation of ubiquitinated proteins and to maintain ubiquitin chain.

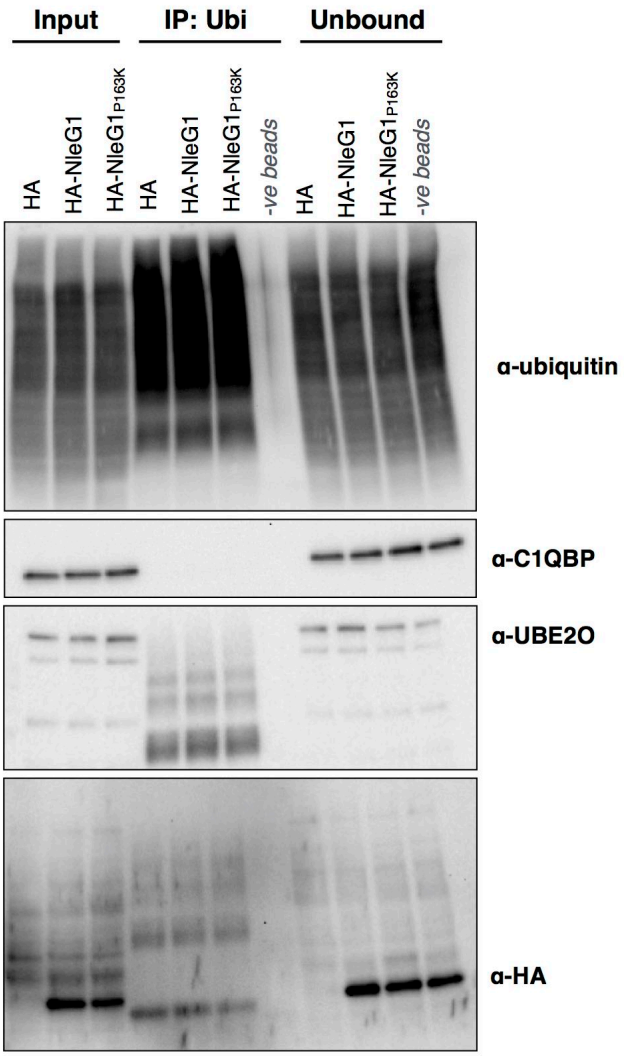


Figure 5.18 | NleG1 is not ubiquitinated and does not ubiquitinate C1QBP or UBE2O in HCT116 cells. Lysates from cells expressing HA, HA-NleG1 or HA-NleG1$_{P163K}$ were subjected to TUBE-IP (immunoprecipitation of ubiquitinated proteins) and subsequent analysis by SDS-PAGE and immunoblotting. Prior to lysis, cells were maintained in 20 µM MG132 (to inhibit proteasomal degradation) and 5 µM PR619 (to inhibit de-ubiquitinating proteins). Transgenes were expressed for 24 hours prior to lysis. Probing with an anti-ubiquitin antibody confirmed enrichment of ubiquitinated proteins in TUBE-IP samples, absent in the TUBE-negative bead control (top panel). No bands were present for UBE2O or C1QBP in the NleG1 TUBE-IP lane (lane 5) that were absent in control IPs (HA and TUBE-negative beads; lanes 4 and 7) and for E3 ligase dead NleG1 (NleG1$_{P163K}$, lane 6), indicating that they are not ubiquitinated by NleG1 under these conditions. Western blot for HA showed no enrichment for HA-NleGs following TUBE-IP, indicating that they were not ubiquitinated.

Samples were separated by SDS-PAGE and analysed by western blot (**Figure 5.18**). Using an anti-ubiquitin antibody, enrichment of ubiquitinated proteins was clearly seen for TUBE-IP samples (top panel). As expected, the unbound samples displayed a less prominent ubiquitin smear, confirming that ubiquitin proteins had been pulled down from the lysates to the full capacity of the TUBE-beads. For the TUBE-negative control beads, no ubiquitin enrichment was seen and the corresponding unbound fraction had more ubiquitinated proteins compared to unbound fractions following TUBE-IP. This confirmed that the assay itself had worked. However, the assay revealed that NleG1 does not ubiquitinate C1QBP or UBE2O in HCT116 cells. Ubiquitination by NleG1 would have entailed bands for the ubiquitinated proteins in the NleG1 TUBE-IP lane (lane 5) and no bands in control samples: HA- control in lane 4, TUBE-negative bead control in lane 7, and E3 ligase-dead NleG1 ($NleG1_{P163K}$) in lane 6. Moreover, bands were not observed in the HA-blot either, supporting the previous observation that NleG proteins are not ubiquitinated in HCT116 cells, even though E3 ligases often self-ubiquitinate *in vivo*.

5.12. Ubiquitination of UBE2O by NleG1 during co-expression in cultured HCT116 cells

In vitro ubiquitin assays in this chapter had showed that the anti-C1QBP antibody could detect ubiquitinated C1QBP. This had not been confirmed for the anti-UBE2O antibody. To avoid missing ubiquitination of UBE2O in cultured cells due to an inability of the anti-UBE2O antibody to recognise ubiquitinated UBE2O, the above TUBE-experiment was repeat following transient transfection of FLAG-UBE2O into cells expressing HA, HA-NleG1 or HA-$NleG1_{P163K}$. Ubiquitinated UBE2O, should it be present, would then be detected using an anti-FLAG antibody. As before, the assay worked (**Figure 5.19**, shown by the ubiquitin blot). This time, faint bands were seen for UBE2O by anti-FLAG western blot. However, any ubiquitination here did not depend on the presence or enzymatic activity of NleG1, as the faint band was present in lanes 4, 5 and 6.

Surprisingly, probing with an anti-HA antibody to control that the HA-tagged genes were expressed revealed distinct banding in cell lysate of HA-$NleG1_{P163K}$- expressing HCT116s (lane 3). The band was seen best upon longer exposure of the blot (bottom panel). The banding differed from the bands seen in lysates from cells expressing HA-NleG1, by having a more pronounced higher band with an approximately molecular weight of around 15 kDa higher than HA-NleG1. This band was seen also in the TUBE-IP lane (lane 6) and in the unbound fraction (lane 8). Considering the recent report that a function of UBE2O is to recognise aberrantly folded proteins

in the cytosol in order to ubiquitinate them and target them for degradation, this raised the concern that perhaps NleG1$_{P163K}$ was being ubiquitinated by UBE2O, not *vice versa*. This would explain why UBE2O was enriched in the original mass spectrometry screen, and perhaps the lack of detectable ubiquitinated NleG1$_{P163K}$ was due to its proteasomal degradation.

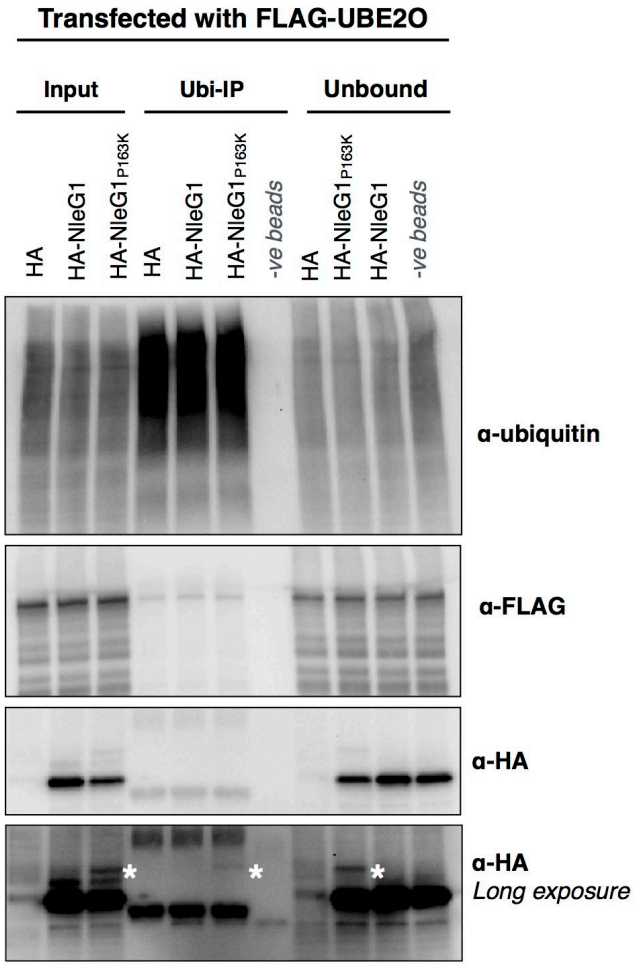


Figure 5.19| Immunoprecipitation of ubiquitinated proteins to assess whether UBE2O transiently expressed in NleG1-expressing HCT116 cells ubiquitinates NleG1. Stably transduced cells expressing HA, HA-NleG1 or HA-NleG1$_{P163K}$ were transiently transfected with a FLAG-UBE2O construct for 24 hours. Transgenes were expressed for 24 hours prior to lysis. Cell lysates were subjected to TUBE immunoprecipitation followed by SDS-PAGE and immunoblotting. Probing with an anti-ubiquitin antibody confirmed enrichment of ubiquitinated proteins in TUBE-IP (labelled Ubi-IP) samples, absent in the TUBE-negative bead control (top panel). Probing with an anti-FLAG antibody revealed a faint band for FLAG-UBE2O in ubiquitination-enriched (Ubi-IP) samples (lanes 4-6). Probing with an anti-HA antibody showed distinct banding in lane 3 (see white asterisk) in lysates from cells co-expressing NleG1$_{P163K}$ and UBE2O. The band was present following TUBE-IP (lane 6, white asterisk), but was also seen in the unbound fraction (lane 9). There was no other obvious band for ubiquitinated NleG1 in the Ubi-IP lanes, apart from background IgG heavy/light chain.

To test this hypothesis, cells expressing NleG1$_{P163K}$ or HA-NleG1 were transfected with a FLAG-UBE2O construct (WT or ligase dead, UBE2O$_{C1040S}$) in the presence or absence of MG-132. Protein abundances in cell lysates were then evaluated by western blot (**Figure 5.20**). Anti-UBE2O and anti-FLAG blots confirmed that the transfections had worked. However, the

abundance of HA-NleG proteins was not increased in the presence of MG-132. This means that HA-NleG1 and HA-NleG1$_{P163K}$ must not be degraded by the proteasome in the absence of a proteasomal inhibitor. Further, there were no differences between NleG levels when transfected with FLAG-UBE2O and FLAG-UBE2O$_{C1040S}$.

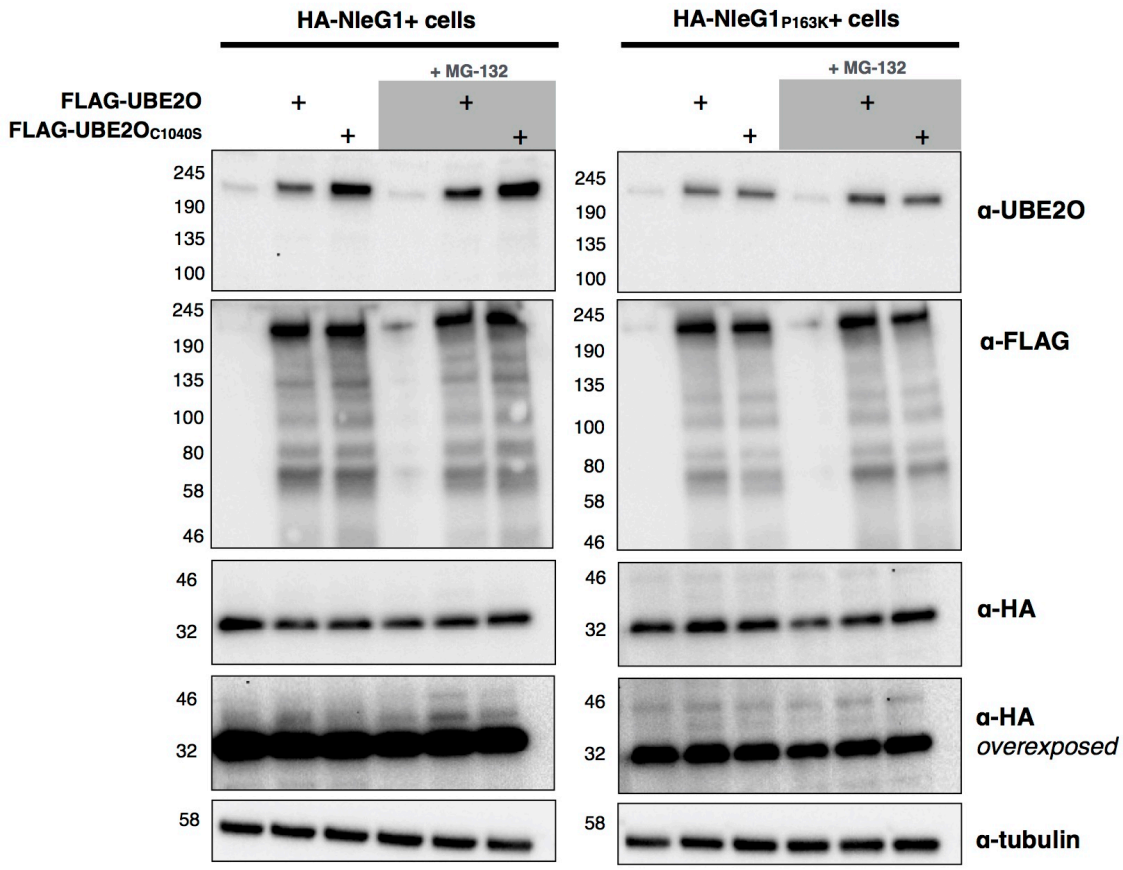


Figure 5.20| Co-expression of UBE2O (WT or E3 ligase dead) and NleG1 (WT or E3 ligase dead) in HCT116 cells in the presence and absence of a proteasomal inhibitor. Stably transduced HCT116 cells expressing HA-NleG1 (left side) or HA-NleG1$_{P163K}$ (right side) were transiently transfected with FLAG-UBE2O or E3 ligase dead FLAG-UBE2O$_{C1040S}$ constructs as indicated for 24 hours. This was performed with and without the addition of 10 µM MG-132 for 4 hours. NleGs were expressed for 24 hours prior to transfection. Cell lysates were analysed by SDS-PAGE and immunoblotting. Blotting with an anti-UBE2O antibody confirmed enhanced expression of UBE2O in transfected cells compared to untransfected cells, where a fainter band for endogenous UBE2O was seen (top panels). The presence of MG-132 gave a very slight increase in the level of UBE2O (WT and C1040S) in both transgenic cell lines. These slight differences could not be seen when probing with the anti-FLAG antibody, which suggested no change in FLAG-UBE2O levels between transfected samples. Probing with the anti-HA antibody revealed no reproducible changes in levels of HA-NleG1 between the samples. When overexposed, the blots showed some of the higher HA banding previously seen, but this was not dependent on transfection of UBE2O with E3 ligase activity. Anti-tubulin blots were used as loading controls (bottom panels). Representative of two independent repeats.

The HA banding pattern for all NleG1$_{P163K}$-containing samples were similar, with the slightly higher band seen in all lanes, regardless of the overexpression of WT or ligase-dead UBE2O. This suggests that UBE2O does not ubiquitinate NleG1$_{P163K}$. A band of approximately 40 kDa was seen

for NleG1 when cells were transfected with WT UBE2O in the presence of MG132. However, further investigating this was deemed unnecessary, with the accumulating evidence that UBE2O might not be a biologically relevant interacting partner or substrate of NleG1. Last efforts to determine if there was any relevance to the putative NleG1-UBE2O interaction identified in this chapter were therefore performed in the context of infection.

5.13. Evaluating abundance of C1QBP and UBE2O in the presence and absence of NleG proteins

Even if NleG1 and UBE2O do not bind directly, as the validation data in this chapter suggests, NleG1 could target a complex that UBE2O is also a part of. Similarly, even if NleG1 does not ubiquitinate UBE2O, it may still change the stability of a larger complex that has downstream consequences on UBE2O levels in the cell. As mentioned above, proteomics data from the Frankel lab have shown a reduction of UBE2O abundance in intestinal epithelial cells during *C. rodentium in vivo* infection. For C1QBP, no such changes had been seen. We wanted to test if the change in UBE2O level in response to *C. rodentium* infection could be attributable to the NleG proteins, which could indicate that NleG1 may target a biological pathway that UBE2O is involved in.

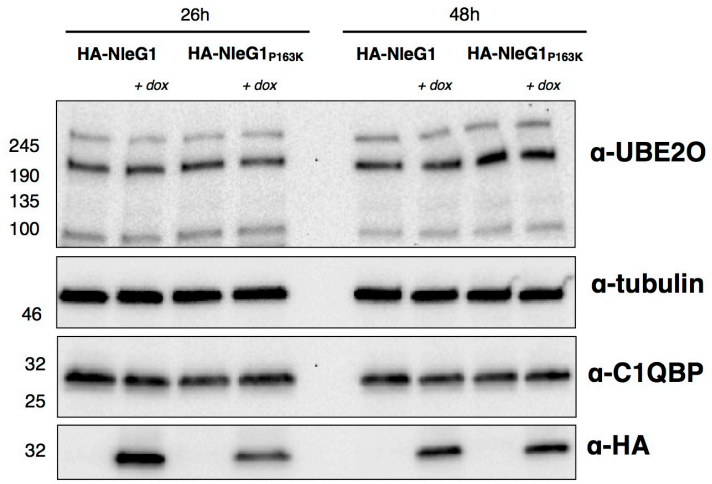


Figure 5.21| Expressing NleG1 and NleG1$_{P163K}$ in HCT116 cells does not alter C1QBP and UBE2O protein levels. The stably transduced HCT116 cell lines used in this study were grown in cell culture media supplemented with 1 µM doxycycline (Dox) for 26 or 48 hours to induce the expression of HA-NleG1 and HA-NleG1$_{P163K}$. At both time points, cell lysates were analysed for changes in endogenous UBE2O and C1QBP levels (alongside a tubulin control) by SDS-PAGE (13.5 µg protein per well) and western blotting using anti-UBE2O, anti-C1QBP and anti-tubulin antibodies. Expression of the transgene was also verified by western blot using an anti-HA antibody.

First, however, it was ensured that the inducible expression of NleG1 in HCT116 cells had no impact on UBE2O abundance. This was shown by western blot (**Figure 5.21**) after 26- and 48-

hour doxycycline treatment. There was no change in UBE2O or C1QBP levels, as expected from the input controls for these conditions already shown in this chapter. This result was replicated by immunofluorescence microscopy of HCT116 cells, that showed cell-wide diffuse UBE2O staining with and without the induced expression of NleG1-HA in the cell (**Figure 5.22**).

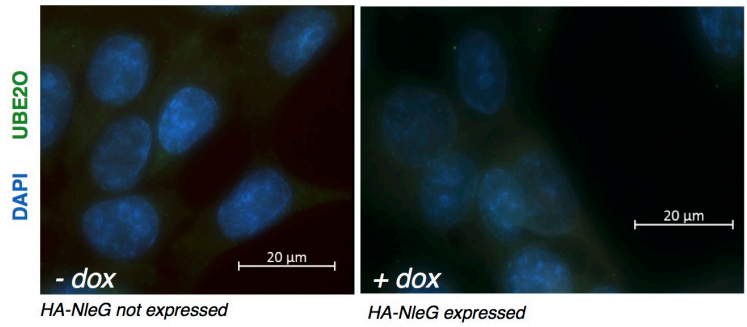


Figure 5.22| Endogenous UBE2O is detectable and cytoplasmic in HCT116 cells with and without the induced expression of NleG1. HCT116 cells with doxycycline-inducible expression of HA-NleG1 were seeded onto glass coverslips. HA-NleG1 expression was induced for 24 hours and cells were then methanol-fixed and analysed by immunofluorescence microscopy, with DNA shown in blue and endogenous UBE2O in green.

Next, HeLa cells were infected for 0.5 hours or 3 hours with WT or a full *nleG* knockout *C. rodentium* strain. Lysates were analysed for UBE2O and C1QBP levels as above, which showed no change in either protein abundance (**Figure 5.23**). Unexpectedly, there was no decrease in UBE2O levels in response to infection, which was anticipated based on the *in vivo* proteomics.

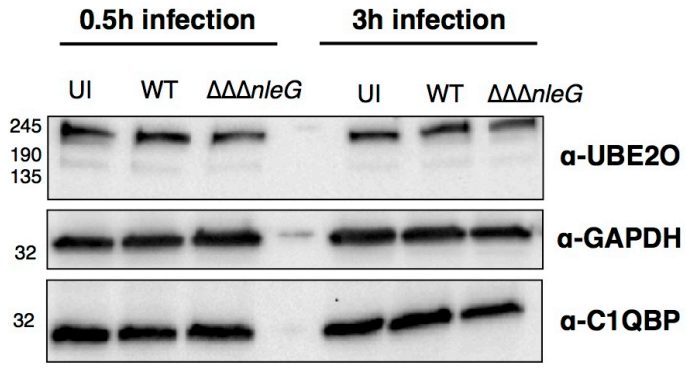


Figure 5.23| Endogenous levels of UBE2O and C1QBP in HeLa cells remains unchanged during *C. rodentium* infection, regardless of presence of NleG proteins. HeLa cells were infected with WT and triple *nleG* knockout (Δ*Δ*Δ*nleG*) *C. rodentium* for 0.5 or 3 hours. Cell lysates were evaluated for levels of UBE2O and C1QBP (alongside a GAPDH control) by SDS-PAGE and western blot analysis with anti-UBE2O, anti-GAPDH and anti-C1QBP antibodies.

Moreover, the 3 hour Hela cells infection using WT and full *nleG* knockout *C. rodentium* was also performed on glass coverslips for analysis by immunofluorescent microscopy. As shown in **Figure 5.24** there was a slight decrease in UBE2O staining intensity for the infected cells, but this did

not depend on the presence of intact *nleG* genes in *C. rodentium*. As the *in vitro* infections were somewhat inconclusive with regard to UBE2O reduction during *C. rodentium* infection, murine intestinal epithelial cells isolated form the peak of infection were instead analysed for UBE2O and C1QBP levels.

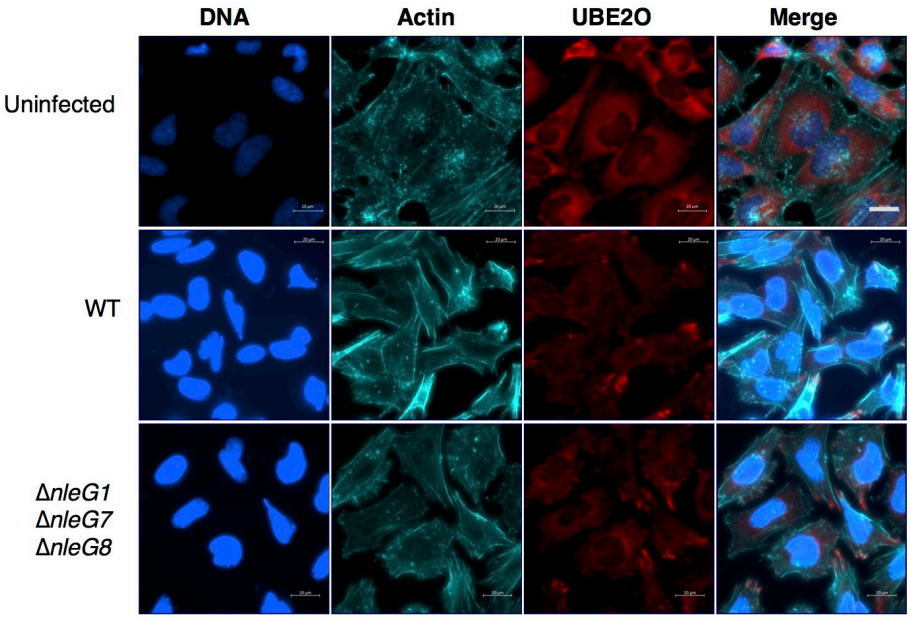


Figure 5.24| Assessing the level and distribution of endogenous UBE2O during *C. rodentium* infection. HeLa cells were infected with WT and triple *nleG* knockout (ΔΔΔ*nleG*) *C. rodentium* for 3 hours on glass coverslips. Cells were fixed with PFA and analysed by fluorescence microscopy, with DNA in blue, actin in teal and UBE2O in red. Scale bar 20 µm.

Two biological repeat experiments of the *in vivo* infections were analysed in parallel. The colonisation up to day 8 (as measured by shedding of viable bacteria in the stool) is shown in **Figure 5.25A** and **C**. For more details on the *in vivo* data, refer back to **Chapter 4**. UBE2O and C1QBP levels were analysed by SDS-PAGE and western blot performed on intestinal epithelial cell lysate. Mice had been infected with either WT or a full *nleG* knockout *C. rodentium* strain for 8 days on the harvest day. In the first repeat (**Figure 5.25B**), only two WT enterocyte protein samples were available, but they showed a slight decrease in UBE2O levels compared to the uninfected control. This was true both when the two samples were analysed individually and when they were grouped. The *nleG* knockout group had one mouse with a strong band for UBE2O, while the rest of the mice in the group exhibited UBE2O levels lower than the uninfected control mice. In the second *in vivo C. rodentium* infection, WT levels of UBE2O were markedly lower than the levels in uninfected mice. Strikingly, the *nleG* knockout group displayed UBE2O levels that were lower than for uninfected mice but not quite as low as WT-infected mice, giving an intermediate phenotype (**Figure 5.25D**). To determine whether this is a robust phenotype attributable to the NleGs, complementation experiments must be performed. It may be that the

change in UBE2O level reflects the very slight difference in colonisation level between WT and *nleG* null *C. rodentium,* as discussed in **Chapter 4.** No changes were seen in C1QBP levels between the groups.

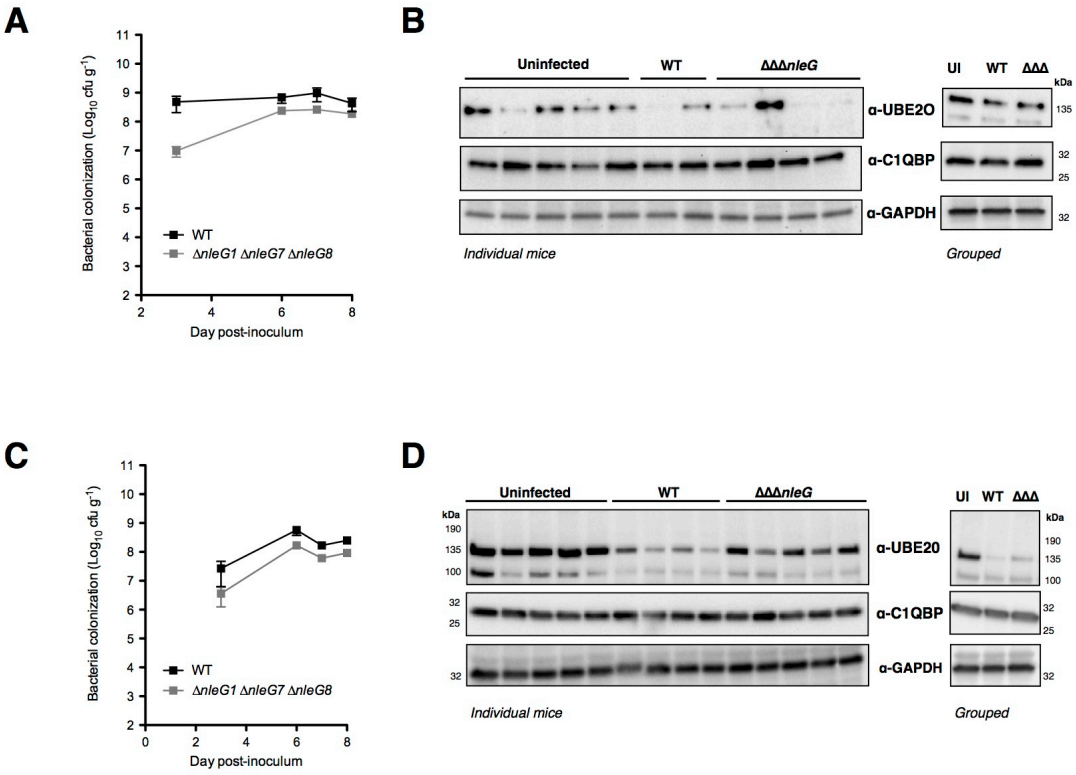


Figure 5.25| UBE2O but not C1QBP protein abundance in intestinal epithelial cells (IECs) is reduced during *C. rodentium* infection. A) C57BL/6 mice (n=5) were infected with ICC169 (WT) or Δ*nleG1* Δ*nleG7* Δ*nleG8 (*ΔΔΔ*nleG) C. rodentium* and colonisation was monitored by enumerating viable bacteria shed in the stool. *This data is also presented in Figure 4.27A.* **B)** IECs were isolated from the colon of mice sacrificed at day 8 post-infection. IEC cell lysates were evaluated for levels of UBE2O and C1QBP (alongside a GAPDH control) by SDS-PAGE and western blot analysis with anti-UBE2O, anti-GAPDH and anti-C1QBP antibodies. Samples from individual mice are shown in the panels on the left (20 µg protein loaded per mouse) and grouped samples are shown in panels on the right (15 µg protein loaded per group). There was no difference in C1QBP levels between the groups, but there was a reduction in UBE2O levels in the infected IECs compared to uninfected IECs. **C)** C57BL/6 mice (n=4-5) were infected with ICC169 (WT) or Δ*nleG1* Δ*nleG7* Δ*nleG8 (*ΔΔΔ*nleG) C. rodentium* and colonisation was monitored by enumerating viable bacteria shed in the stool. *This data is also presented in Figure 4.24.* **D)** IECs were isolated from the colon of mice sacrificed at day 8 post-infection. IEC cell lysates were evaluated for levels of UBE2O and C1QBP (alongside a GAPDH control) by SDS-PAGE and western blot analysis with anti-UBE2O, anti-GAPDH and anti-C1QBP antibodies. Samples from individual mice are shown in the panels on the left (20 µg protein loaded per mouse) and grouped samples are shown in panels on the right (15 µg protein loaded per group). There was no difference in C1QBP levels between the groups, but there was a reduction in UBE2O levels in the infected IECs compared to uninfected IECs. The decrease in UBE2O levels was less pronounced in IECs infected with ΔΔΔ*nleG* compared to WT *C. rodentium*.

5.14. NleG expression does not affect degradation of IκBα in response to Poly (I:C) transfection

In a final attempt to see if NleG1 might target a pathway that C1QBP is involved in, a functional assay was performed where cells were transfected with Poly (I:C). This triggers the degradation of IκBα, with downstream activation of NF-κB responses. C1QBP was previously shown to inhibit such responses (Xu *et al.*, 2009). Cells with and without doxycycline stimulation were subjected to Poly (I:C) transfection. Cell lysates were then analysed by western blot. However, Poly (I:C) transfection triggered degradation of IκBα independent of NleG1 expression (**Figure 5.26**), indicating that it is unlikely that NleG1 bears a role in down-regulating NF-κB dependent interferon responses.

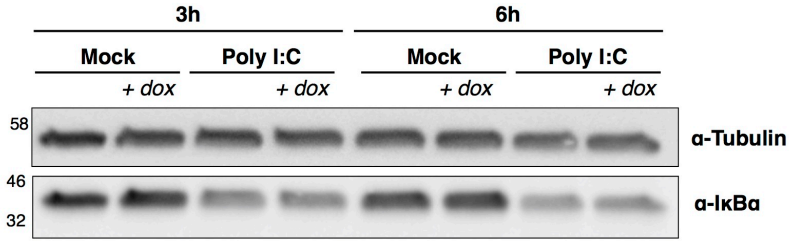


Figure 5.26| NleG1 expression does not affect degradation of IκBα in response to Poly (I:C) transfection in HCT116 cells. Stably transduced HCT116 cells used in this study were grown in cell culture media with or without µM doxycycline (Dox) for 24 hours prior to transfection for regulated doxycycline-induced expression of HA-NleG1. Cells were either transfected with Poly (I:C) or mock-transfected for 3 or 6 hours, and cell lysates were subsequently analysed by SDS-PAGE and immunoblotting. An anti-tubulin antibody was used as a loading control (top panel). An anti-IkBa antibody was used to show degradation of IκBα in response to Poly (I:C) transfection regardless of the induced expressed of NleG1 (bottom panel). Representative of two biological repeats.

5.15. Discussion

Post-translational modifications are indispensable modulators of protein function during infection and studying them by proteomics holds great potential for uncovering mechanisms mediating the progression and spread of bacterial infections. Proteomics-based approaches have been pivotal in furthering our understanding of the ubiquitin system and technical advances in mass spectrometry-based methodologies have permitted identification of substrates of bacterial as well as eukaryotic ubiquitin E3 ligases. For example, stable isotype labelling by amino acids in cell culture (SILAC) coupled to quantitative mass spectrometry have found targets of the human transmembrane E3 ligase MARCH9 (Hör *et al.*, 2009). Likewise, SILAC coupled to selective enrichment of ubiquitin-modified proteins and quantitative mass spectrometry dissected the dynamics of the global host ubiquitinome in response to *Salmonella* infection and identified substrates of the bacterial E3 ligase SopA (Fiskin *et al.*, 2016). Other quantitative MS-based

proteomics, such as the TMT-labelling approach discussed in **Chapter 4**, can also be integrated for studying the host-pathogen interaction during bacterial infection (Jean Beltran *et al.*, 2017).

In this chapter, a co-IP and mass spectrometry-based screen identified putative host interacting partners of NleG1 and NleG8. Ligase-null NleG variants were used to pull down interacting partners, thereby circumventing the problem of E3 ligase-substrate interactions typically being transient. Furthermore, the screen was performed in two different cell lines to add confidence to the hits (O'Connor & Huibregtse, 2017). Nevertheless, challenges posed upon this type of experiment concern all stages of the workflow: sample preparation, data gathering and, importantly, the functional interpretation of data (Schmidt, Forne & Imhof, 2014). Common contaminants are inevitably acquired during sample preparation, and must therefore be manually removed from the LFQ data sets or simply disregarded during further analysis. A comparison between control and test samples generated a list of proteins significantly enriched in pull-down samples from NleG-expressing cultured cells, some of which were indeed common contaminants (Mellacheruvu *et al.*, 2013). Owing to the high incidence of such contaminants as well as the relatively low number of hits, our data sets were not well suited for hierarchical clustering of proteins based on functional terms to identity over-representation of specific GO-terms. Nonetheless, it was interesting to note that a proportion of the proteins enriched across the NleG samples compared to control samples had reported functions concerning host antiviral immunity. For example, TRIM proteins are commonly associated with antiviral immune cell signalling (Rakebrandt *et al.*, 2014), and ZAP1 (which requires TRIM25 for its activity) binds viral mRNA to repress their translation (Zheng *et al.*, 2017; Li *et al.*, 2017). The notion that A/E pathogens might to some extent translocate bacterial DNA through the T3SS prompts the speculation that the host elicits immune responses directly in response to the T3SS-translocated bacterial nucleic acids. However, these proteins were not reproducibly enriched for specific NleGs in the screen, and the identification of certain antiviral proteins depended on the cell background rather than the NleG tested. For example, MOV10 (an IFN-inducible RNA helicase shown to enhance IFN regulatory factor 3 (IRF3)-mediated transcription of IFN; Cuevas *et al.*, 2016) and Oas3 (an IFN-induced protein that activates ribonuclease L, involved in degrading viral RNA; Li *et al.*, 2016) were enriched in iBMMs expressing NleG1 or NleG8, but never in HCT116 cells.

Following the hypothesis that sequence variability between NleG homologs reflect a difference in host substrate or binding partner (while overall a shared pathway might still be targeted), we

focused on putative interactors specifically identified for one NleG. The assumption that NleGs have different binding partners had been strengthened by previously observed differences in intracellular localization during transfection of cultured cells and in conferring yeast cytotoxicity (discussed in **Chapter 3**). We prioritised validating putative interactors of NleG1, as expression levels of NleG1 had been higher than for NleG8 (thereby giving more confidence to the obtained data) and a larger number of proteins had been found to be significantly enriched. A second biological repeat of the NleG1 screen in HCT116 cells revealed that three proteins were reproducibly enriched in NleG1-containing samples across all repeats, while not being enriched in any NleG8-containing samples. These included UBE2O and C1QBP, which had both been implicated in cellular pathways relevant to bacterial infection, and efforts were dedicated to validating these putative interactors.

While UBE2O and C1QBP interacted with both WT and E3 ligase-null NleG1 in repeated IPs (as performed in the original screen), a recurring hurdle in validating interactions in this project was the lack of a robust negative control. Although the HA-tag control had been pivotal for the mass spectrometry screen, we were unable to verify expression levels of HA upon doxycycline induction due its small size, and a negative control with comparable expression of the transgene would have been useful. Such a control, with no suspected overlap with NleGs in terms of interacting partners, would help delineate true NleG interactions from background noise – a problem that remains inherent to mass spectrometry methodologies. Both NleG-family and non-NleG family controls would benefit such efforts. For example, it would be interesting to see if C1QBP and UBE2O co-immunoprecipitated with NleG members such as StoD as well as with other bacterial E3 ligases (RING-type, HECT-type and NELs).

Purified recombinant His-NleG1 could not pull down C1QBP from HeLa cell lysate, giving a first indication that C1QBP might not be a true target of NleG1. Secondly, C1QBP did not interact with any of the *C. rodentium* NleGs by direct yeast two-hybrid. Although we observed ubiquitination of C1QBP by NleG1 in cell-free ubiquitination assays *in vitro*, suggesting that these proteins can form an interaction sufficient for ubiquitination of C1QBP to occur in this assay, this was not specific to NleG1: NleG8 and StoD also ubiquitinated C1QBP here. The HECT-type bacterial E3 ligase NleL did not, but without a non-NleG family RING-type E3 ligase at hand we could not discern between NleG family specific-ubiquitination of C1QBP and RING/U-box-type E3 ligase-specific ubiquitination of C1QBP. More importantly, we were unable to show that NleG1 ubiquitinated C1QBP in cultured cells, which would be expected if C1QBP were the true

substrate of NleG1. Further, the abundance of C1QBP remained unchanged in *C. rodentium*-infected enterocytes (*in vivo* infection) and HeLa cells (*in vitro* infection), regardless of the presence of NleGs. One function of C1QBP is down-regulation of MAVS-dependent IFN responses (where C1QBP binds MAVS to disrupt downstream antiviral signalling), but we did not find that C1QBP overexpression in HCT116 cells altered degradation of IκBα in response to Poly (I:C) transfection. Thus, in this project we did could confirm that NleG1 and C1QBP interact with functional relevance to *C. rodentium*. Taken together with reports of C1QBP interacting with a hundred different proteins by IP-coupled mass spectrometry (Zhang *et al.*, 2013a) and its presence in an online repository for suspected false positives in such experiments (Mellacheruvu *et al.*, 2013), this cast doubts on the validity of C1QBP as a true target of NleG1.

We could not confirm that UBE2O and NleG1 interacted in reciprocal IP experiments, where FLAG-UBE2O was expressed in NleG1- or NleG1$_{P163K}$-expressing cells. Moreover, there was no ubiquitination of UBE2O in HCT116 cells by NleG1, as assessed by TUBE-IP and western blot analysis. During the course of this project, two papers were published that outlined a role for UBE2O in protein quality control in the cell. The first showed that UBE2O recognises misfolded, unassembled or mislocalised proteins and targets them for degradation (Yanagitani, Juszkiewicz & Hegde, 2017). The second paper showed that UBE2O is important in large-scale proteaome remodelling during reticulocyte maturation (Nguyen *et al.*, 2017). Therefore, we ensured that UBE2O was not targeting NleG1 rather than *vice versa*, as NleG1 could potentially be recognised as an aberrant protein in HCT116 cells. However, we did not see obvious ubiquitination of NleG1 or NleG1$_{P163K}$ by UBE2O, and their expression in cultured cells did not impact on the abundance of either protein. *In vitro* infection of HeLa cells using *C. rodentium* also did not change UBE2O abundance in the cell. *In vivo*, on the other hand, endogenous UBE2O levels in intestinal epithelial cells were dramatically reduced in response to *C. rodentium* infection, confirming proteomics data from our lab suggesting that UBE2O levels decline with infection. *In vivo* infection with a full *nleG* knockout mutant displayed an intermediate phenotype in terms of UBE2O degradation, but further complementation experiments must be performed to ensure that this was due specifically to the lack of NleG proteins, rather than to the potential slight colonisation attenuation discussed in **Chapter 4**. Such experiments would help shed light on whether an interaction between NleG1 and UBE2O occurs in intestinal epithelial cells during *C. rodentium* infection.

Intriguingly, UBE2O has a paralog in yeast, BRUCE, that was recovered from a Y2H screen that used N-terminal NleG1 as the bait (discussed in **Chapter 3**). The interaction with BRUCE could not be verified by subsequent direct two-hybrid and was therefore not considered further. Taken together with the identification of UBE2O in the MS screen in this chapter, however, it may still reflect the ability of NleG1 to associate with proteins that resemble UBE2O/BRUCE. Notably, these proteins both have RING-type E3 ligase domains. When considering all the proteins enriched for NleGs across the mass spectrometry repeats in this chapter, it is interesting to note that many RING E3 ligase proteins were found, including the TRIM proteins and components of larger multi-protein RING-E3 complexes. This might reflect NleG biology, hinting at a preference for binding RING-domain proteins. SopA in *Salmonella*, for instance, specifically targets the E3 ligases TRIM56 and TRIM65 during infection (Fiskin *et al.*, 2017). On the other hand, it might simply reflect the tendency of RING proteins to form RING-RING interactions, without hinting at specific substrates during infection. A study demonstrated the requirement of RING dimerization for the E3 activity of TRIM25 and TRIM32 (Koliopoulos *et al.*, 2016), and therefore putative interactions between NleGs and TRIMs might stem from TRIMs attempting to dimerise with proteins with structurally similar RING domains. Alternatively, the NleG RING domain itself might be prone RING-RING dimerization in the cell, and therefore bind other proteins with RING domains. This should be kept in mind if analysing any further putative hits that are E3 ligases, and validation experiments should map interactions, for example by yeast two-hybrid, to test whether they are due to RING-RING interactions.

While our validation experiments did not validate UBE2O or C1QBP as true interacting partners of NleG1, further validations of the mass spectrometry data could be performed. For instance, it would be interesting to test if there is any biological relevance to the enrichment of calprotectin components that were found in both NleG1- and NleG8-containing samples. As calprotectin is a known antimicrobial protein that akin Lcn2 is triggered during enteric infection, it could potentially be targeted by bacterial effector proteins (Simard, Girard & Tessier, 2010).

Although this chapter highlights difficulties not just in elucidating partners for NleG but within the E3 ligase field of research, it presents a cellular toolkit that could be used together with both new- and established –omics methods to provide novel information about ubiquitination during *C. rodentium* infection. Importantly, the cell lines described in this chapter could still be used in a number of ways to characterise the NleGs. For instance, they could be used to perform SILAC and ubiquitination-enriching experiments to compare ubiquitinatomes between samples, as has

been done for other effector proteins (Fiskin *et al.*, 2016). This approach allowed analysis of the global ubiquitinome during bacterial infection, and during *Salmonella* infection the most targeted pathways concerned actin rearrangement, NF-κB signalling, autophagy pathways and RHO GTPase systems. Being able to define ubiquitinomes has been a major advance in the ubiquitin field, where an approach called diGly proteomics recognises a diglycine (diGly) remnant on the side chain of ubiquitinated lysine residues following a trypsinisation step. This can be enriched for using antibodies that specifically bind the diGly remnant (Kim *et al.*, 2011). Proteins can then be identified by mass spectrometry, and experiments can be further enhanced by the addition of specific inhibitors (such as proteasome-, NEDD8-, and de-ubiquitinase-inhibitors). With this approach, even subtle changes in ubiquitinomes can be detected. If applied to *C. rodentium* infection or the cell lines developed here, it could help dissect ubiquitin dynamics in response to *C. rodentium* infection and identify substrates of NleG1.

Another way of using our cell lines to identify targets of NleG1 is coupling pull-downs with a ubiquitination reaction, which has been useful in finding targets of some E3 ligases (O'Connor & Huibregtse, 2017). In such experiments, IPs are performed as in this chapter, followed by the addition of the necessary components for ubiquitination to occur directly on beads. FLAG-tagged ubiquitin is used to allow further FLAG-IPs to enrich for ubiquitinated proteins that can be identified my mass spectrometry. For instance, Claspin was identified as a substrate of the human E3 complex $SCF^{B\text{-}TrCP2}$ in this manner (Peschiaroli *et al.*, 2006).

Additional ways of finding E3 targets include using Ubiquitin-Activated Interaction Traps (UBAITS), engineered to covalently capture substrates of E3 ligases, as the E3 ligase of interest is fused (via a flexible linker) C-terminally to a ubiquitin moiety (O'Connor *et al.*, 2015). Alternatively, protein microarrays can be used, where they are loaded with thousands of tagged proteins can be incubate with fluorescently labeled ubiquitin as well as E1, E2, ATP and the E3 of interest for ubiquitination to occurs, which can be detected using a fluorescent scanner.

A clear understanding of the interactions between bacterial E3 ligases and host proteins could hold therapeutic advantages. While identifying biologically relevant host targets is a difficult task, this chapter highlights options for future investigations into finding targets of the *C. rodentium* NleGs.

Chapter 6: Final discussion

6. Final discussion

Infectious diarrhoea poses a vast burden on human health, with pathogenic *E. coli* contributing significantly to morbidity and mortality across the globe. The two A/E pathogens EPEC and EHEC are amongst these and their infection strategy involves a series of intricate interactions between bacterial and host factors, modulating a range of host responses that determine the outcome of the infection. A battery of T3-secreted effector proteins is essential in the host-pathogen cross-talk, and an increased understanding of how they affect the host could drive the development of novel therapeutics and help disease prevention. An important avenue within the T3SS field, particularly considering the growing threat of antimicrobial resistance, is indeed identifying and developing compounds that target T3SS biology to treat infection (Charro & Mota, 2015; Beckham & Roe, 2014). Efforts in developing such anti-virulence strategies would be aided by a more complete understanding of the individual and concerted contributions of T3SS effector proteins to the virulence of A/E pathogens. The project presented in this thesis set out to decipher the role of the NleG proteins during A/E infection. It focused on NleG homologs found within *C. rodentium*, a natural mouse pathogen routinely used to study A/E infection *in vivo* due to its close relatedness to human A/E pathogens. Beyond being a fundamental tool in studying A/E pathogenesis, the *C. rodentium* model is also employed to study enteric mucosal immune responses during colitis and the involvement of the microbiota in protection against enteric infection, which is a rapidly expanding field (Keeney *et al.*, 2014; Collins *et al.*, 2014; Mullineaux-Sanders *et al.*, 2017).

The NleG family of T3SS effector proteins is the largest described family of non-LEE encoded effector proteins, with homologs in A/E pathogens as well as in typhoidal (but not non-typhoidal) *Salmonella* spp. (Petty *et al.*, 2010). In EHEC, at least 8 NleG homologs are secreted by the T3SS, and the NleGs thus constitute a significant proportion of the EHEC effector protein repertoire (Tobe *et al.*, 2006). Prior to this work, the structure of NleG2.3 from EHEC was solved, which revealed structural similarity of the conserved NleG C-terminus to eukaryotic RING/U-box-type ubiquitin E3 ligase domains, despite no sequence homology (Wu *et al.*, 2010). This prompted the hypothesis that A/E pathogens translocate NleG proteins to exploit the host ubiquitin code to favour their pathogenesis. However, there were no reports of NleG substrates or target cell processes during infection and while the biochemical activity of EHEC NleGs had been studied, the only investigation of the *C. rodentium* NleGs was bioinformatical. Considering the strength of the *C. rodentium* model in studying T3SS effector protein contribution *in vivo*,

this project sought to specifically characterize the *C. rodentium* NleGs. Three overarching aims were defined to achieve this: (1) analysing the NleGs by *in vitro* methods (focusing on confirming their translocation into eukaryotic cells, their putative E3 ligase enzymatic activity and examining their localisation within the host); (2) assessing their importance *in vivo* using the murine infection model; and (3) performing screens to identify their interacting partners or substrates and to characterise putative interactions.

The key efforts, findings and the future experiments they call for or motivate are discussed at length in the discussions at the ends of **Chapter 3**, **Chapter 4** and **Chapter 5**. To summarise the findings presented **Chapter 3**, we demonstrated for the first time that the *C. rodentium* NleG proteins NleG1, NleG7 and NleG8 can be translocated into eukaryotic cells via the T3SS. By characterizing them biochemically, we showed that NleG1 and NleG8 harbour E3 ligase activity reliant on a conserved proline residue (residue P163 in NleG1) that is involved in the interaction between the E3 ligase and E2 enzymes. Further support for the hypothesis that the NleG family evolved to exploit the ubiquitin code comes from immunofluorescence microscopy that revealed that NleG7 and NleG8 cause specific relocalisation of endogenous ubiquitin when expressed ectopically in cultured eukaryotic cells. Both NleG7 and NleG8 co-localised with ubiquitin in such experiments and could also be shown to associate with or be modified by ubiquitin by yeast two-hybrid methodology. A clear difference between the *C. rodentium* NleG proteins concerned their cytotoxicity in yeast, as only NleG8 was shown to be cytotoxic and, importantly, this cytotoxicity was dependent on its E3 ligase activity. Together, these findings support the notion that NleG proteins hijack or disrupt host ubiquitin-dependent pathways: a theme that is emerging as increasingly important in the context of bacterial infection. The eukaryote-specific ubiquitin system is a target for many bacterial effector proteins and E3 ligase mimicry is observed in an array of Gram-negative pathogens, including clinically important and T3SS-dependent strains of S*. flexneri* and *S. enterica* (Ashida & Sasakawa, 2017; Maculins *et al.*, 2016b). Interestingly, bacterial manipulation of this system extends beyond structural mimicry of eukaryotic proteins, as novel classes of E3 ligases that are structurally distinct from eukaryotic E3 ligases have been discovered (Bhogaraju *et al.*, 2016; Singer *et al.*, 2013; Kubori, Hyakutake & Nagai, 2008). Further yet, the SidE family of T4SS effector proteins in *L. pneumophila* even perform non-canonical ubiquitination never observed for eukaryotic E3 ligases, which permits ubiquitin attachment independent of the conventional ubiquitin cycle machinery (Qiu *et al.*, 2016). Aided by technical advances in methods to study ubiquitination as well as host-pathogen interactions *in vivo* and *in vitro*, the molecular mechanisms that pathogens have evolved to intricately exploit the ubiquitin

system are rapidly being uncovered. As such, pharmacological inhibition of bacterial E3 ligase activity should be considered as a novel strategy to treat infections (Rossi *et al.*, 2014; Maculins *et al.*, 2016b) and could be combined with high-throughput screens for identifying E3 ligase inhibitors (Maculins *et al.*, 2016a).

Efforts to identify inhibitors of bacterial E3 ligases would benefit from a detailed understanding of specific E3 ligase and substrate interactions, and finding the host targets of bacterial E3 ligases is therefore of high importance – as attempted for the *C. rodentium* NleG proteins in this thesis. Uncovering the cellular targets of the NleGs would undoubtedly shed light on the role of NleGs during infection. Despite our best attempts, however, we could not confirm any interacting partners or substrates of NleG1, NleG7 or NleG8, likely due to the inherent challenges in finding substrates of E3 ligase proteins. We performed both yeast two-hybrid screening (in **Chapter 3**) and a mass spectrometry based screen (in **Chapter 5**), both of which have been previously used to identify substrates of E3 ligase proteins (Lomma *et al.*, 2010; Bex *et al.*, 2007; Coyaud *et al.*, 2015). In the final week of writing this thesis, a report was published that revealed interacting partners of the two EHEC NleG homologs NleG2.3 and NleG5.1 (Valleau *et al.*, 2018). The authors used affinity-purified NleG proteins immobilised on magnetic beads to capture interacting partners from eukaryotic cell lysate (U937 cells) and identified them by MS, highlighting the potential of the MS-based methodologies attempted in our work. They found that NleG2.3 bound the host protein hexokinase 2 (HK2), which is involved in glucose metabolism (including the Warburg effect) and in the inhibition of apoptosis (Pastorino, Shulga & Hoek, 2002; Mathupala, Ko & Pedersen, 2009; Roberts & Miyamoto, 2015). NleG5.1 on the other hand interacted with MED15, a component of the Mediator complex involved in transcriptional regulation (Allen & Taatjes, 2015). They verified the interactions by direct yeast two-hybrid and by transfection and co-IP experiments, similar to the approaches adopted in our work. Likewise, they used both WT and enzymatically deficient NleG variants to investigate the interactions. They also performed site-directed mutagenesis and generated chimeric NleG constructs to show that the N-termini of NleGs – as initially hypothesised – indeed confer binding to different host proteins. Binding of the target proteins in transfected cells resulted in their poly-ubiquitination and proteasomal degradation, which is the first conclusive evidence that NleGs target host proteins for degradation. Moreover, the study showed specific degradation of the NleG substrates in cells infected with Δ*nleG* EPEC strains expressing epitope-tagged EHEC NleG variants from a plasmid.

Importantly, Valleau and colleagues also solved the structure of N-terminal and full-length NleG5.1 to reveal a previously unreported N-terminal α/β domain required for substrate binding. The hydrophobic core of this domain was conserved across the NleG family and is likely to form the common target interface that together with variable surface residues creates a modifiable platform to direct functionally diverse host proteins for degradation. Thus, similar to what is seen in *Shigella* spp. (Ashida & Sasakawa, 2016), EHEC has an expanded arsenal of functionally diverse E3 ligase effector proteins. The NleGs studied by Valleau and colleagues do not have close homologs in *C. rodentium* and there is no overlap between the interacting partners they identify and the putative interacting partners we found in **Chapters 3** and **5**. Nevertheless, their work supports the idea that the *C. rodentium* NleGs also have different target proteins in their host and that they may modulate different pathways, especially as they localize to different subcellular compartments akin the EHEC NleG homologs (Valleau *et al.*, 2018). Ultimately, the report validates many of the approaches we attempted in our work and emphasises the potential for continued investigations into the effects of NleG-translocation into eukaryotic cells. Further, it motivates *in vivo* investigations of the importance of NleGs and of the downstream consequences of the degradation of their interacting partners during infection.

To address the role of the NleGs during infection, we employed the *C. rodentium in vivo* model to study the contribution of NleG proteins to the kinetics and histopathological features of *C. rodentium* infection. In **Chapter 4**, we conclusively showed that knocking out *nleG* genes on their own has no effect on the ability of *C. rodentium* to colonise the gut mucosa to WT-levels. The same held true for a double *nleG7 nleG8* deletion mutant. However, a full *nleG* knockout mutant exhibited a slight colonisation attenuation, hinting at some functional redundancy between the *C. rodentium* NleGs. Nevertheless, further work is required to clarify whether this depends on the lack of *nleG* genes and, if true, to elucidate molecular mechanism behind why NleG proteins are required for WT-levels of infection. Complementation experiments are essential, and would also clarify if the accompanied reduction in Lcn2 levels in the colon of *nleG*-null *C. rodentium*-infected mice compared to WT-infected mice is due specifically to the lack of translocated NleGs. If confirmed, this would be the first *in vivo* phenotype attributable to the NleG proteins. Moreover, the complementation experiments could be performed in conjunction with further TMT-10 plex experiments (as described in **Chapter 4**). Such proteomic-based approaches are well-suited for studying E3 ligase proteins, and beyond the NleG project, for understanding the infectious process and host-pathogen interactions to eventually allow the development of novel treatment options (Jean Beltran *et al.*, 2017).

At present, rehydration therapy is the most effective way to manage symptoms of infectious diarrhoea and using antibiotics is not always recommended as this increases the risk of HUS in the case of Shiga toxin-producing *E. coli*. Moreover, antibiotic treatment disturbs the endogenous microflora and its efficacy may both be limited by and contribute to the emergence and spread of multi-drug resistance in enteric bacteria. In a climate of increasing antimicrobial resistance, there are also ethical implications to consider when recommending antibiotic treatments that directly contribute to the overuse of antibiotics. However, targeting the T3SS to treat disease circumvents the above problems, as this targets bacterial virulence rather than viability. The requirement of a T3SS for host colonization by many clinically relevant Gram-negative pathogens and the absence of a T3SS in non-pathogenic members of the microflora make the T3SS a promising target for anti-virulence compounds (Zambelloni *et al.*, 2017). Such compounds could be aimed at structural elements of the T3SS (basal body, translocon pore or needle), at elements that intricately regulate the LEE or at T3SS effector proteins (Charro & Mota, 2015), as discussed for bacterial E3 ligases. An example of a well-studied class of anti-virulence compounds is salicylidene acylhydrazides (SAs), originally identified as potential T3SS inhibitors in *Y. pseudotuberculosis* (Nordfelth *et al.*, 2005; Kauppi *et al.*, 2007). They interact with proteins of the basal body and their ability to block T3SS function has also been demonstrated in *S.* Typhimurium, *Shigella* spp. and EHEC (Negrea *et al.*, 2007; Hudson *et al.*, 2007; Veenendaal, Sundin & Blocker, 2009; Tree *et al.*, 2009). SAs also interact with proteins involved in *E. coli* metabolism, which ultimately down-regulates T3SS gene expression (McShan & De Guzman, 2015). Nevertheless, drawbacks of SAs include their substrate promiscuity, poor solubility in biological buffers and a tendency to hydrolyse under acidic conditions (Wang *et al.*, 2011). Development of novel compounds with higher selectivity and improved stability is therefore underway, including a new class of compounds that target EspD that should be examined further for potential clinical use (Zambelloni *et al.*, 2017). In terms of vaccine development, structural elements of the T3SS could be targeted (O'Ryan *et al.*, 2015). For the Gram-negative pathogen *Pseudomonas syringae*, antibodies against its T3SS tip protein PrcV and its exopolysaccharide PsI have been shown to give protection against pulmonary infection in an *in vivo* mouse model with a multidrug resistant strain (MedImmune patent MED13902). This supports the idea that with continued research on the T3SS, a similar approach could also be used to prevent infection by A/E pathogens.

Advances in studying the impact of effector protein secretion on host cells keeps revealing a fascinating complexity and sophistication to T3SS-dependent pathogenesis. The efforts

described in this thesis provide novel information about the *C. rodentium* NleGs and presents a toolkit of plasmids, *C. rodentium* mutant strains and transgenic eukaryotic cell lines that are well suited for further investigations into the NleG family of effector proteins. Moreover, this work highlights the potential in studying how bacterial effector proteins exploit ubiquitin pathways during infection. Specifically targeting bacterial E3 ligase activity without compromising host E3 ligases is a promising avenue for virulence-targeting antimicrobial drug development, as such therapeutics would not encourage the emergence of drug-resistance bacteria to the same extent as traditional antibiotics. Ultimately, the work performed here will hopefully motivate and direct future efforts to elucidate the mechanism by which NleG proteins contribute to the host-pathogen interactions and A/E pathogenesis.

References

Abella, V., Scotece, M., Conde, J., Gómez, R., *et al*. (2015) The potential of lipocalin-2/NGAL as biomarker for inflammatory and metabolic diseases. *Biomarkers*. [Online] Available from: doi:10.3109/1354750X.2015.1123354.

Abramovitch, R.B., Kim, Y.J., Chen, S., Dickman, M.B., *et al*. (2003) *Pseudomonas* type III effector AvrPtoB induces plant disease susceptibility by inhibition of host programmed cell death. *EMBO Journal*. [Online] Available from: doi:10.1093/emboj/cdg006.

Abt, M.C. & Pamer, E.G. (2014) Commensal bacteria mediated defenses against pathogens. *Current Opinion in Immunology*. [Online]. Available from: doi:10.1016/j.coi.2014.03.003.

Abu-Median, A.B., Van Diemen, P.M., Dziva, F., Vlisidou, I., *et al*. (2006) Functional analysis of lymphostatin homologues in enterohaemorrhagic *Escherichia coli*. *FEMS Microbiology Letters*. [Online] Available from: doi:10.1111/j.1574-6968.2006.00198.x.

Adams, N.L., Byrne, L., Smith, G.A., Elson, R., *et al*. (2016) Shiga toxin-producing *Escherichia coli* O157, England and Wales, 1983-2012. *Emerging Infectious Diseases*. [Online] Available from: doi:10.3201/eid2204.151485.

Agin, T.S., Cantey, J.R., Boedeker, E.C. & Wolf, M.K. (1996) Characterization of the *eaeA* gene from rabbit enteropathogenic *Escherichia coli* strain RDEC-1 and comparison to other *eaeA* genes from bacteria that cause attaching-effacing lesions. *FEMS Microbiology Letters*. [Online] Available from: doi:10.1016/0378-1097(96)00371-0.

Ale, A., Crepin, V.F., Collins, J.W., Constantinou, N., *et al*. (2017) Model of host-pathogen interaction dynamics links in vivo optical imaging and immune responses. *Infect Immun*. [Online] Available from: doi:10.1128/IAI.00606-16.

Alhagamhmad, M.H., Day, A.S., Lemberg, D.A. & Leach, S.T. (2016) An overview of the bacterial contribution to Crohn disease pathogenesis. *Journal of Medical Microbiology*. [Online]. Available from: doi:10.1099/jmm.0.000331.

Allen, B.L. & Taatjes, D.J. (2015) The Mediator complex: A central integrator of transcription. *Nature Reviews Molecular Cell Biology*. [Online]. Available from: doi:10.1038/nrm3951.

Allen, K.J., Rogan, D., Finlay, B.B., Potter, A., *et al*. (2011) Vaccination with type III secreted proteins leads to decreased shedding in calves after experimental infection with *Escherichia coli* O157. *Canadian journal of veterinary research = Revue canadienne de recherche vétérinaire*.

Alsharif, G., Ahmad, S., Islam, M.S., Shah, R., *et al*. (2015) Host attachment and fluid shear are integrated into a mechanical signal regulating virulence in *Escherichia coli* O157:H7. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1422986112.

Alto, N.M., Shao, F., Lazar, C.S., Brost, R.L., *et al*. (2006) Identification of a bacterial type III effector family with G protein mimicry functions. *Cell*. [Online] Available from: doi:10.1016/j.cell.2005.10.031.

Alto, N.M., Weflen, A.W., Rardin, M.J., Yarar, D., *et al*. (2007) The type III effector EspF coordinates membrane trafficking by the spatiotemporal activation of two eukaryotic signaling pathways. *Journal of Cell Biology*. [Online] Available from: doi:10.1083/jcb.200705021.

Anand, A., Rojas, C.M., Tang, Y. & Mysore, K.S. (2012) Several components of SKP1/Cullin/F-box E3 ubiquitin ligase complex and associated factors play a role in *Agrobacterium*-mediated plant transformation. *The New phytologist*. [Online] Available from: doi:10.1111/j.1469-8137.2012.04133.x.

Angot, A., Peeters, N., Lechner, E., Vailleau, F., *et al*. (2006) *Ralstonia solanacearum* requires F-box-like

domain-containing type III effectors to promote disease on several host plants. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0509393103.

Arbeloa, A., Blanco, M., Moreira, F.C., Bulgin, R., *et al*. (2009) Distribution of *espM* and *espT* among enteropathogenic and enterohaemorrhagic *Escherichia coli*. *Journal of Medical Microbiology*. [Online] Available from: doi:10.1099/jmm.0.010231-0.

Arbeloa, A., Bulgin, R.R., Mackenzie, G., Shaw, R.K., *et al*. (2008) Subversion of actin dynamics by EspM effectors of attaching and effacing bacterial pathogens. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01136.x.

Arbeloa, A., Oates, C. V., Marchès, O., Hartland, E.L., *et al.* (2011) Enteropathogenic and enterohemorrhagic *Escherichia coli* type III secretion effector EspV induces radical morphological changes in eukaryotic cells. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.01003-10.

Ardley, H.C. & Robinson, P.A. (2005) E3 ubiquitin ligases. *Essays in Biochemistry*. [Online] Available from: doi:10.1042/EB0410015.

Arena, E.T., Auweter, S.D., Antunes, L.C.M., Vogl, A.W., *et al*. (2011) The deubiquitinase activity of the *Salmonella* pathogenicity island 2 effector, SseI, prevents accumulation of cellular lipid droplets. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.05478-11.

Ashida, H., Kim, M., Schmidt-Supprian, M., Ma, A., *et al*. (2010) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-κB-mediated inflammatory response. *Nature cell biology*. [Online] Available from: doi:10.1038/ncb2006.

Ashida, H., Nakano, H. & Sasakawa, C. (2013) *Shigella* IpaH0722 E3 ubiquitin ligase effector targets TRAF2 to inhibit PKC-NF-κB activity in invaded epithelial cells. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1003409.

Ashida, H. & Sasakawa, C. (2017) Bacterial E3 ligase effectors exploit host ubiquitin systems. *Current Opinion in Microbiology*. [Online]. Available from: doi:10.1016/j.mib.2016.11.001.

Ashida, H. & Sasakawa, C. (2016) *Shigella* IpaH family effectors as a versatile model for studying pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2015.00100.

Barnich, N., Carvalho, F.A., Glasser, A.L., Darcha, C., *et al*. (2007) CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *Journal of Clinical Investigation*. [Online] Available from: doi:10.1172/JCI30504.

Basler, M. (2015) Type VI secretion system: Secretion by a contractile nanomachine. *Philosophical Transactions of the Royal Society B: Biological Sciences*. [Online]. Available from: doi:10.1098/rstb.2015.0021.

Basler, M., Ho, B.T. & Mekalanos, J.J. (2013) Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell*. [Online] Available from: doi:10.1016/j.cell.2013.01.042.

Basler, M. & Mekalanos, J.J. (2012) Type 6 secretion dynamics within and between bacterial cells. *Science*. [Online]. Available from: doi:10.1126/science.1222901.

Basu, R., O'Quinn, D.B., Silberger, D.J., Schoeb, T.R., *et al*. (2012) Th22 cells are an important source of IL-22 for host protection against enteropathogenic bacteria. *Immunity*. [Online] Available from: doi:10.1016/j.immuni.2012.08.024.

Beckham, K.S.H. & Roe, A.J. (2014) From screen to target: insights and approaches for the development of anti-virulence compounds. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from:

doi:10.3389/fcimb.2014.00139.

Van Den Beld, M.J.C. & Reubsaet, F.A.G. (2012) Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*. *European Journal of Clinical Microbiology and Infectious Diseases*. [Online]. Available from: doi:10.1007/s10096-011-1395-7.

Belkaid, Y. & Hand, T.W. (2014) Role of the microbiota in immunity and inflammation. *Cell*. [Online]. Available from: doi:10.1016/j.cell.2014.03.011.

Belzer, C., Liu, Q., Carroll, M.C. & Bry, L. (2011) The role of specific IgG and complement in combating aprimary mucosal infection of the gut epithelium. *European Journal of Microbiology and Immunology*. [Online] Available from: doi:10.1556/EuJMI.1.2011.4.7.

Benz, I. & Schmidt, M.A. (1992) AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesized via a precursor molecule. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.1992.tb00875.x.

Benz, I. & Schmidt, M.A. (1989) Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infection and Immunity*. [Online] Available from: doi:10.1099/00221287-135-3-531.

Berger, C.N., Crepin, V.F., Jepson, M.A., Arbeloa, A., *et al*. (2009a) The mechanisms used by enteropathogenic *Escherichia coli* to control filopodia dynamics. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01254.x.

Berger, C.N., Crepin, V.F., Roumeliotis, T.I., Wright, J.C., *et al*. (2017) *Citrobacter rodentium* subverts ATP flux and cholesterol homeostasis in intestinal epithelial cells *in vivo*. *Cell Metabolism*. [Online] Available from: doi:10.1016/j.cmet.2017.09.003.

Berger, C.N., F. Crepin, V., Baruch, K., Mousnier, A., *et al.* (2012) EspZ of Enteropathogenic and Enterohemorrhagic *Escherichia coli*. *mBio*. [Online] Available from: doi:10.1128/mBio.00317-12.

Berger, C.N., Shaw, R.K., Brown, D.J., Mather, H., *et al*. (2009b) Interaction of *Salmonella enterica* with basil and other salad leaves. *ISME Journal*. [Online] Available from: doi:10.1038/ismej.2008.95.

Berger, T., Togawa, A., Duncan, G.S., Elia, A.J., *et al*. (2006) Lipocalin 2-deficient mice exhibit increased sensitivity to *Escherichia coli* infection but not to ischemia-reperfusion injury. *Proc. Natl. Acad. Sci. U. S. A*. [Online] Available from: doi:0510847103 [pii];10.1073/pnas.0510847103 [doi].

Bergstrom, K.S.B., Guttman, J.A., Rumi, M., Ma, C., *et al.* (2008) Modulation of intestinal goblet cell function during infection by an attaching and effacing bacterial pathogen. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00093-07.

Bern, C., Martines, J., de Zoysa, I. & Glass, R.I. (1992) The magnitude of the global problem of diarrhoeal disease: a ten-year update. *Bulletin of the World Health Organization*.

Bernal-Bayard, J., Cardenal-Muñoz, E. & Ramos-Morales, F. (2010) The *Salmonella* type III secretion effector, *Salmonella* Leucine-rich Repeat Protein (SlrP), targets the human chaperone ERdj3. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M110.100669.

Bernal-Bayard, J. & Ramos-Morales, F. (2009) *Salmonella* type III secretion effector SlrP is an E3 ubiquitin ligase for mammalian thioredoxin. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M109.010363.

Bernassola, F., Karin, M., Ciechanover, A. & Melino, G. (2008) The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell*. [Online]. Available from: doi:10.1016/j.ccr.2008.06.001.

Berry, A.A., Yang, Y., Pakharukova, N., Garnett, J.A., *et al*. (2014) Structural insight into host recognition by aggregative adherence fimbriae of Enteroaggregative *Escherichia coli*. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1004404.

Beuming, T., Skrabanek, L., Niv, M.Y., Mukherjee, P., *et al*. (2005) PDZBase: A protein-protein interaction database for PDZ-domains. *Bioinformatics*. [Online]. Available from: doi:10.1093/bioinformatics/bti098.

Bex, C., Knauth, K., Dambacher, S. & Buchberger, A. (2007) A yeast two-hybrid system reconstituting substrate recognition of the von Hippel-Lindau tumor suppressor protein. *Nucleic Acids Research*. [Online] Available from: doi:10.1093/nar/gkm932.

Bhat, A., Shin, M., Jeong, J.-H., Kim, H.-J., *et al.* (2014) DNA looping-dependent autorepression of *LEE1* P1 promoters by Ler in enteropathogenic *Escherichia coli* (EPEC). *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1322033111.

Bhatt, S., Egan, M., Jenkins, V., Muche, S., *et al*. (2016) The Tip of the iceberg: on the roles of regulatory small RNAs in the vrulence of Enterohemorrhagic and Enteropathogenic *Escherichia coli*. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2016.00105.

Bhavsar, A.P., Brown, N.F., Stoepel, J., Wiermer, M., *et al*. (2013) The *Salmonella* Type III effector SspH2 specifically exploits the NLR co-chaperone activity of SGT1 to subvert immunity. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1003518.

Bhogaraju, S., Kalayil, S., Liu, Y., Bonn, F., *et al*. (2016) Phosphoribosylation of ubiquitin promotes serine ubiquitination and impairs conventional ubiquitination. *Cell*. [Online] Available from: doi:10.1016/j.cell.2016.11.019.

Bieber, D., Ramer, S.W., Wu, C.Y., Murray, W.J., *et al*. (1998) Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science*. [Online]. Available from: doi:10.1126/science.280.5372.2114.

Biemans-Oldehinkel, E., Sal-Man, N., Deng, W., Foster, L.J., *et al*. (2011) Quantitative proteomic analysis reveals formation of an EscL-EscQ-EscN type III complex in enteropathogenic *Escherichia coli*. *Journal of Bacteriology*. [Online]. Available from: doi:10.1128/JB.05235-11.

Blasche, S., Mörtl, M., Steuber, H., Siszler, G., *et al*. (2013) The *E. coli* effector protein NleF is a Caspase inhibitor. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0058937.

Blattner, F.R. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*. [Online] Available from: doi:10.1126/science.277.5331.1453.

Blomstrom, D.C., Fahey, D., Kutny, R., Korant, B.D., *et al.* (1986) Molecular characterization of the interferon-induced 15-kDa protein. Molecular cloning and nucleotide and amino acid sequence. *Journal of Biological Chemistry*.

Boisen, N., Melton-Celsa, A.R., Scheutz, F., O'Brien, A.D., *et al*. (2015) Shiga toxin 2a and Enteroaggregative *Escherichia coli* – a deadly combination. *Gut Microbes*. [Online] Available from: doi:10.1080/19490976.2015.1054591.

Boisen, N., Ruiz-Perez, F., Scheutz, F., Krogfelt, K.A., *et al*. (2009) Short report: high prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative *Escherichia coli*. *The American journal of tropical medicine and hygiene*.

Boisen, N., Struve, C., Scheutz, F., Krogfelt, K.A., *et al*. (2008) New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infection and Immunity*. [Online] Available from:

doi:10.1128/IAI.01646-07.

Boratyn, G.M., Camacho, C., Cooper, P.S., Coulouris, G., *et al*. (2013) BLAST: a more efficient report with usability improvements. *Nucleic acids research*. [Online] Available from: doi:10.1093/nar/gkt282.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., *et al*. (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*. [Online] Available from: doi:10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.

Bray, J. (1945) Isolation of antigenically homogeneous strains of *Bact. coli neapolitanum* from summer diarrhœa of infants. *The Journal of Pathology and Bacteriology*. [Online] 57 (2), 239–247. Available from: doi:10.1002/path.1700570210 [Accessed: 14 September 2018].

Brooks, J.T., Sowers, E.G., Wells, J.G., Greene, K.D., *et al*. (2005) Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. *The Journal of infectious diseases*. [Online] Available from: doi:10.1086/466536.

Bruckner, A., Polge, C., Lentze, N., Auerbach, D., *et al*. (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci*. [Online] Available from: doi:10.3390/ijms10062763.

Bulgin, R., Arbeloa, A., Goulding, D., Dougan, G., *et al*. (2009a) The T3SS effector EspT defines a new category of invasive enteropathogenic *E. coli* (EPEC) which form intracellular actin pedestals. In: *PLoS Pathogens*. [Online]. 2009 p. Available from: doi:10.1371/journal.ppat.1000683.

Bulgin, R., Raymond, B., Garnett, J.A., Frankel, G., *et al*. (2010) Bacterial guanine nucleotide exchange factors SopE-Like and WxxxE effectors. *Infection and Immunity*. [Online]. Available from: doi:10.1128/IAI.01250-09.

Bulgin, R.R., Arbeloa, A., Chung, J.C.S. & Frankel, G. (2009b) EspT triggers formation of lamellipodia and membrane ruffles through activation of Rac-1 and Cdc42. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01248.x.

Burger, R. (2012) EHEC O104:H4 in Germany 2011: Large outbreak of bloody diarrhea and haemolytic uraemic syndrome by Shiga toxin-producing *E. coli* via contaminated food. In: *Improving Food Safety Through a One Health Approach: Workshop Summary*.

Burts, M.L., Williams, W.A., DeBord, K. & Missiakas, D.M. (2005) EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0405620102.

Buttner, D. (2012) Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiology and Molecular Biology Reviews*. [Online] Available from: doi:10.1128/MMBR.05017-11.

Cantey, J.R. & Blake, R.K. (1977) Diarrhea due to *Escherichia coli* in the rabbit: a novel mechanism. *Journal of Infectious Diseases*. [Online] Available from: doi:10.1093/infdis/135.3.454.

Cantey, J.R., Lushbaugh, W.B. & In Man, L.R. (1981) Attachment of bacteria to intestinal epithelial cells in diarrhea caused by *Escherichia coli* strain RDEC-1 in the rabbit: stages and role of capsule. *Journal of Infectious Diseases*. [Online] Available from: doi:10.1093/infdis/143.2.219.

Cantey, J.R. & Moseley, S.L. (1991) HeLa cell adherence, actin aggregation, and invasion by nonenteropathogenic *Escherichia coli* possessing the *eae* gene. *Infection and Immunity*. [Online] Available from: doi:0019-9567/91/113924-06$02.00/0.

Carlson-Banning, K.M. & Sperandio, V. (2018) Enterohemorrhagic *Escherichia coli* outwits hosts through sensing small molecules. *Current Opinion in Microbiology*. [Online]. Available from: doi:10.1016/j.mib.2017.12.002.

Carvalho, F. a, Barnich, N., Sivignon, A., Darcha, C., *et al.* (2009) Crohn's disease adherent-invasive *Escherichia col*i colonize and induce strong gut inflammation in transgenic mice expressing human CEACAM. *The Journal of experimental medicine*. [Online] Available from: doi:10.1084/jem.20090741.

Cassady-Cain, R.L., Blackburn, E.A., Alsarraf, H., Dedic, E., *et al*. (2016) Biophysical characterization and activity of lymphostatin, a multifunctional virulence factor of attaching and effacing *Escherichia coli*. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M115.709600.

Cella, M., Fuchs, A., Vermi, W., Facchetti, F., *et al*. (2009) A human natural killer cell subset provides an innate source of IL-22 for mucosal immunity. *Nature*. [Online] Available from: doi:10.1038/nature07537.

Cepeda-Molero, M., Berger, C.N., Walsham, A.D.S., Ellis, S.J., *et al*. (2017) Attaching and effacing (A/E) lesion formation by enteropathogenic *E. coli* on human intestinal mucosa is dependent on non-LEE effectors. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1006706.

Cha, B. & Donowitz, M. (2008) The epithelial brush border Na+/H+ exchanger NHE3 associates with the actin cytoskeleton by binding to ezrin directly and via PDZ domain-containing Na+/H+ exchanger regulatory factor (NHERF) proteins. *Clinical and Experimental Pharmacology and Physiology*. [Online]. Available from: doi:10.1111/j.1440-1681.2008.04931.x.

Chan, J.M., Bhinder, G., Sham, H.P., Ryz, N., et al. (2013) CD4 + T Cells Drive Goblet Cell Depletion during Citrobacter rodentium Infection. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00655-13.

Charpentier, X. & Oswald, E. (2004) Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 β-lactamase as a new fluorescence-based reporter. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.186.16.5486-5495.2004.

Charro, N. & Mota, L.J. (2015) Approaches targeting the type III secretion system to treat or prevent bacterial infections. *Expert Opinion on Drug Discovery*. [Online] Available from: doi:10.1517/17460441.2015.1019860.

Chassaing, B., Srinivasan, G., Delgado, M.A., Young, A.N., *et al*. (2012) Fecal Lipocalin 2, a Sensitive and Broadly Dynamic Non-Invasive Biomarker for Intestinal Inflammation. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0044328.

Chaudhuri, R.R., Sebaihia, M., Hobman, J.L., Webber, M.A., *et al*. (2010) Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain O42. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0008801.

Chen, F.X., Xie, P., Collings, C.K., Cao, K., *et al*. (2017) PAF1 regulation of promoter-proximal pause release via enhancer activation. *Science*. [Online] Available from: doi:10.1126/science.aan3269.

Chen, L., Li, J., Shirota, M., Ohnishi, K., *et al*. (2018) Involvement of GALA effectors in *Ralstonia solanacearum* disease development towards two host plants. *Acta Microbiologica Sinica*. [Online] Available from: doi:10.13343/j.cnki.wsxb.20170057.

Chen, Y.-B., Jiang, C.-T., Zhang, G.-Q., Wang, J.-S., *et al*. (2009) Increased expression of hyaluronic acid binding protein 1 is correlated with poor prognosis in patients with breast cancer. *Journal of surgical oncology*. [Online] Available from: doi:10.1002/jso.21329.

Chosed, R., Tomchick, D.R., Brautigam, C.A., Mukherjee, S., *et al*. (2007) Structural analysis of *Xanthomonas* XopD provides insights into substrate specificity of ubiquitin-like protein proteases. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M608730200.

Chou, C.C., Lee, K.H., Lai, I.L., Wang, D., *et al*. (2014) AMPK reverses the mesenchymal phenotype of cancer cells by targeting the Akt-MDM2-Foxo3a signaling axis. *Cancer Research*. [Online] Available from: doi:10.1158/0008-5472.CAN-14-0135.

Christie, P.J. (2016) The mosaic type IV secretion systems. *EcoSal Plus*. [Online] Available from: doi:10.1128/ecosalplus.ESP-0020-2015.

Chu, X., Qin, X., Xu, H., Li, L., *et al*. (2013) Structural insights into Paf1 complex assembly and histone binding. *Nucleic Acids Research*. [Online] Available from: doi:10.1093/nar/gkt819.

Claessen, J.H.L., Witte, M.D., Yoder, N.C., Zhu, A.Y., *et al*. (2013) Catch-and-release probes applied to semi-intact cells reveal ubiquitin-specific protease expression in *Chlamydia trachomatis* infection. *ChemBioChem*. [Online] Available from: doi:10.1002/cbic.201200701.

Clarke, M.B., Hughes, D.T., Zhu, C., Boedeker, E.C., *et al*. (2006) The QseC sensor kinase: A bacterial adrenergic receptor. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0604343103.

Cleary, J., Lai, L.C., Shaw, R.K., Straatman-Iwanowska, A., *et al*. (2004) Enteropathogenic *Escherichia coli* (EPEC) adhesion to intestinal epithelial cells: Role of bundle-forming pili (BFP), EspA filaments and intimin. *Microbiology*. [Online] Available from: doi:10.1099/mic.0.26740-0.

Clements, A., Young, J.C., Constantinou, N. & Frankel, G. (2012) Infection strategies of enteric pathogenic *Escherichia coli*. *Gut Microbes*. [Online]. Available from: doi:10.4161/gmic.19182.

Cole, A.J., Clifton-Bligh, R. & Marsh, D.J. (2015) Histone H2B monoubiquitination: Roles to play in human malignancy. *Endocrine-Related Cancer*. [Online]. Available from: doi:10.1530/ERC-14-0185.

Collier-Hyams, L.S., Zeng, H., Sun, J., Tomlinson, A.D., *et al*. (2002) Cutting edge: *Salmonella* AvrA effector inhibits the key proinflammatory, anti-apoptotic NF-κB Pathway. *The Journal of Immunology*. [Online] Available from: doi:10.4049/jimmunol.169.6.2846.

Collins, J.W., Keeney, K.M., Crepin, V.F., Rathinam, V.A.K., *et al*. (2014) *Citrobacter rodentium*: Infection, inflammation and the microbiota. *Nature Reviews Microbiology*. [Online]. Available from: doi:10.1038/nrmicro3315.

Connolly, J.P.R., Brett Finlay, B. & Roe, A.J. (2015) From ingestion to colonization: The influence of the host environment on regulation of the LEE encoded type III secretion system in enterohaemorrhagic *Escherichia coli. Frontiers in Microbiology*. [Online]. Available from: doi:10.3389/fmicb.2015.00568.

Cordero-Alba, M., García-Gómez, J.J., Aguilera-Herce, J. & Ramos-Morales, F. (2016) Proteomic insight into the effects of the *Salmonella* ubiquitin ligase SlrP on host cells. *Biochemical and Biophysical Research Communications*. [Online] Available from: doi:10.1016/j.bbrc.2016.03.014.

Costa, T.R.D., Felisberto-Rodrigues, C., Meir, A., Prevost, M.S., *et al*. (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*. [Online] Available from: doi:10.1038/nrmicro3456.

Coyaud, E., Mis, M., Laurent, E.M.N., Dunham, W.H., *et al*. (2015) BioID-based identification of Skp Cullin F-box (SCF) $^{\beta\text{-TrCP1/2}}$ E3 ligase substrates. *Molecular & Cellular Proteomics*. [Online] Available from: doi:10.1074/mcp.M114.045658.

Cramer, E.P., Dahl, S.L., Rozell, B., Knudsen, K.J., *et al*. (2017) Lipocalin-2 from both myeloid cells and the epithelium combats *Klebsiella pneumoniae* lung infection in mice. *Blood*. [Online] Available from: doi:10.1182/blood-2016-11-753434.

Creasey, E.A., Delahay, R.M., Daniell, S.J. & Frankel, G. (2003) Yeast two-hybrid system survey of interactions between LEE-encoded proteins of enteropathogenic *Escherichia coli*. *Microbiology*. [Online]. Available from: doi:10.1099/mic.0.26355-0.

Crepin, V.F., Collins, J.W., Habibzay, M. & Frankel, G. (2016) *Citrobacter rodentium* mouse model of bacterial infection. *Nature Protocols*. [Online] Available from: doi:10.1038/nprot.2016.100.

Crepin, V.F., Habibzay, M., Glegola-Madejska, I., Guenot, M., *et al.* (2015) Tir triggers expression of CXCL1 in enterocytes and neutrophil recruitment during *Citrobacter rodentium* infection. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00291-15.

Croxen, M.A. & Finlay, B.B. (2010) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews Microbiology*. [Online]. Available from: doi:10.1038/nrmicro2265.

Cuevas, R.A., Ghosh, A., Wallerath, C., Hornung, V., *et al*. (2016) MOV10 provides antiviral activity against RNA vruses by enhancing RIG-I–MAVS-independent IFN induction. *The Journal of Immunology*. [Online] Available from: doi:10.4049/jimmunol.1501359.

Cui, J., Yao, Q., Li, S., Ding, X., *et al.* (2010) Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family. *Science*. [Online] Available from: doi:10.1126/science.1193844.

Curciarello, R., Steele, A., Cooper, D., MacDonald, T.T., *et al*. (2014) The role of galectin-1 and galectin-3 in the mucosal immune response to *Citrobacter rodentium* infection. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0107933.

Czeczulin, J.R., Balepur, S., Hicks, S., Phillips, A., *et al*. (1997) Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infection and Immunity*. [Online] Available from: doi:65(10):4135-45.

Dahan, S., Knutton, S., Shaw, R.K., Crepin, V.F., *et al.* (2004) Transcriptome of enterohemorrhagic *Escherichia coli* O157 adhering to eukaryotic plasma membranes. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.72.9.5452-5459.2004.

Daniell, S.J., Takahashi, N., Wilson, R., Friedberg, D., *et al*. (2001) The filamentous type III secretion translocon of enteropathogenic *Escherichia coli*. *Cellular microbiology*. [Online] Available from: doi:168 [pii].

Das, S. & Chaudhuri, K. (2003) Identification of a unique IAHP (IcmF associated homologous proteins) cluster in *Vibrio cholerae* and other proteobacteria through *in silico* analysis. *In silico biology*. [Online] Available from: doi:2003030025 [pii].

Dean-Nystrom, E.A., Bosworth, B.T., Cray, W.C. & Moon, H.W. (1997) Pathogenicity of *Escherichia coli* O157:H7 in the intestines of neonatal calves. *Infection and Immunity*.

Dean-Nystrom, E.A., Gansheroff, L.J., Mills, M., Moon, H.W., *et al*. (2002) Vaccination of pregnant dams with intimin(O157) protects suckling piglets from *Escherichia coli* O157:H7 infection. *Infection and immunity*. [Online] Available from: doi:10.1128/IAI.70.5.2414-2418.2002.

Dean, P., Maresca, M., Sch ller, S., Phillips, A.D., *et al*. (2006) Potent diarrheagenic mechanism mediated by the cooperative action of three enteropathogenic *Escherichia coli*-injected effector proteins. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0509451103.

Dean, P., Scott, J.A., Knox, A.A., Quitard, S., *et al*. (2010) The enteropathogenic *E. coli* effector EspF targets and

disrupts the nucleolus by a process regulated by mitochondrial dysfunction. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1000961.

Delepelaire, P. (2004) Type I secretion in Gram-negative bacteria. *Biochimica et Biophysica Acta - Molecular Cell Research*. [Online]. Available from: doi:10.1016/j.bbamcr.2004.05.001.

Dembitzer, F.R., Kinoshita, Y., Burstein, D., Phelps, R.G., *et al*. (2012) gC1qR expression in normal and pathologic human tissues: differential expression in tissues of epithelial and mesenchymal origin. *Journal of Histochemistry and Cytochemistry*. [Online] Available from: doi:10.1369/0022155412440882.

Deng, W., Li, Y., Hardwidge, P.R., Frey, E.A., *et al*. (2005) Regulation of type III secretion hierarchy of translocators and effectors in attaching and effacing bacterial pathogens. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.73.4.2135-2146.2005.

Deng, W., Puente, J.L., Gruenheid, S., Li, Y., *et al*. (2004) Dissecting virulence: Systematic and functional analyses of a pathogenicity island. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0400326101.

Deng, W., Vallance, B.A., Li, Y., Puente, J.L., *et al*. (2003) *Citrobacter rodentium* translocated intimin receptor (Tir) is an essential virulence factor needed for actin condensation, intestinal colonization and colonic hyperplasia in mice. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2003.03429.x.

Deng, W., Yu, H.B., Li, Y. & Finlay, B.B. (2015) SepD/SepL-dependent secretion signals of the type III secretion system translocator proteins in enteropathogenic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.02401-14.

Deretic, V., Jiang, S. & Dupont, N. (2012) Autophagy intersections with conventional and unconventional secretion in tissue development, remodeling and inflammation. *Trends in Cell Biology*. [Online]. Available from: doi:10.1016/j.tcb.2012.04.008.

Diao, J., Zhang, Y., Huibregtse, J.M., Zhou, D., *et al.* (2008) Crystal structure of SopA, a *Salmonella* effector protein mimicking a eukaryotic ubiquitin ligase. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb1346.

Diepold, A., Wiesand, U., Amstutz, M. & Cornelis, G.R. (2012) Assembly of the *Yersinia* injectisome: The missing pieces. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2012.08146.x.

Diepold, A., Wiesand, U. & Cornelis, G.R. (2011) The assembly of the export apparatus (YscR,S,T,U,V) of the *Yersinia* type III secretion apparatus occurs independently of other structural components and involves the formation of an YscV oligomer. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2011.07830.x.

Dong, N., Liu, L. & Shao, F. (2010) A bacterial effector targets host DH-PH domain RhoGEFs and antagonizes macrophage phagocytosis. *EMBO Journal*. [Online] Available from: doi:10.1038/emboj.2010.33.

Donowitz, M., Cha, B., Zachos, N.C., Brett, C.L., *et al.* (2005) NHERF family and NHE3 regulation. *Journal of Physiology*. [Online] Available from: doi:10.1113/jphysiol.2005.090399.

Duda, D.M., Scott, D.C., Calabrese, M.F., Zimmerman, E.S., *et al*. (2011) Structural regulation of cullin-RING ubiquitin ligase complexes. *Current Opinion in Structural Biology*. [Online]. Available from: doi:10.1016/j.sbi.2011.01.003.

Dümmler, A., Lawrence, A.M. & de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microbial Cell Factories*. [Online] Available

from: doi:10.1186/1475-2859-4-34.

Dupont, N., Jiang, S., Pilli, M., Ornatowski, W., *et al*. (2011) Autophagy-based unconventional secretory pathway for extracellular delivery of IL-1?? *EMBO Journal*. [Online] Available from: doi:10.1038/emboj.2011.398.

Dutta, P.R., Cappello, R., Navarro-García, F. & Nataro, J.P. (2002) Functional comparison of serine protease autotransporters of Enterobacteriaceae. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.70.12.7105-7113.2002.

Echtenkamp, F., Deng, W., Wickham, M.E., Vazquez, A., *et al*. (2008) Characterization of the NleF effector protein from attaching and effacing bacterial pathogens. *FEMS Microbiology Letters*. [Online] Available from: doi:10.1111/j.1574-6968.2008.01088.x.

Eggleton, P., Lieu, T.S., Zappi, E.G., Sastry, K., *et al*. (1994) Calreticulin is released from activated neutrophils and binds to C1q and mannan-binding protein. *Clinical Immunology and Immunopathology*. [Online] Available from: doi:10.1006/clin.1994.1160.

Ehrbar, K., Friebel, A., Miller, S.I. & Hardt, W.D. (2003) Role of the *Salmonella* Pathogenicity Island 1 (SPI-1) protein InvB in type III secretion of SopE and SopE2, two *Salmonella* effector proteins encoded outside of SPI-1. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.185.23.6950-6967.2003.

Elliott, S.J., Krejany, E.O., Mellies, J.L., Robins-Browne, R.M., *et al.* (2001) EspG, a novel type III system-secreted protein from enteropathogenic *Escherichia coli* with similarities to virA of *Shigella flexneri*. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.69.6.4027-4033.2001.

Elsinghorst, E.A. & Weitz, J.A. (1994) Epithelial cell invasion and adherence directed by the enterotoxigenic *Escherichia coli* tib locus is associated with a 104-kilodalton outer membrane protein. *Infection and Immunity*.

Ensminger, A.W. & Isberg, R.R. (2010) E3 ubiquitin ligase activity and targeting of BAT3 by multiple *Legionella pneumophila* translocated substrates. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00344-10.

Erickson, J.L., Adlung, N., Lampe, C., Bonas, U., *et al*. (2018) The *Xanthomonas* effector XopL uncovers the role of microtubules in stromule extension and dynamics in *Nicotiana benthamiana*. *Plant Journal*. [Online] Available from: doi:10.1111/tpj.13813.

Etienne-Manneville, S. (2004) Cdc42 - the centre of polarity. *Journal of Cell Science*. [Online] Available from: doi:10.1242/jcs.01115.

Fairbrother, J.M., Nadeau, É. & Gyles, C.L. (2005) *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Animal Health Research Reviews*. [Online] Available from: doi:10.1079/AHR2005105.

Faubert, B., Boily, G., Izreig, S., Griss, T., et al. (2013) AMPK is a negative regulator of the warburg effect and suppresses tumor growth in vivo. *Cell Metabolism*. [Online] Available from: doi:10.1016/j.cmet.2012.12.001.

Feldsine, P., Lienau, A.H., Shah, K., Immermann, A., *et al.* (2016) Comparison of assurance GDS® MPX ID for Top STEC with reference culture methods for the detection of *E. coli* Top 6 STEC; Direct Confirmation of Top 6 STEC from isolation plates and determination of equivalence of PickPen® and FSIS OctoMACS concentration p. *Journal of AOAC International*. [Online] Available from: doi:10.5740/jaoacint.15-0261.

Fernandez-Prada, C.M., Hoover, D.L., Tall, B.D., Hartman, A.B., *et al*. (2000) *Shigella flexneri* IpaH7.8 facilitates

escape of virulent bacteria from the endocytic vacuoles of mouse and human macrophages. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.68.6.3608-3619.2000.

Finley, D., Ulrich, H.D., Sommer, T. & Kaiser, P. (2012) The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. *Genetics*. [Online] Available from: doi:10.1534/genetics.112.140467.

Fischbach, M.A. & Sonnenburg, J.L. (2011) Eating for two: How metabolism establishes interspecies interactions in the gut. *Cell Host and Microbe*. [Online]. Available from: doi:10.1016/j.chom.2011.10.002.

Fiskin, E., Bhogaraju, S., Herhaus, L., Kalayil, S., *et al*. (2017) Structural basis for the recognition and degradation of host TRIM proteins by *Salmonella* effector SopA. *Nature Communications*. [Online] Available from: doi:10.1038/ncomms14004.

Fiskin, E., Bionda, T., Dikic, I. & Behrends, C. (2016) Global analysis of host and bacterial ubiquitinome in Response to *Salmonella* Typhimurium Infection. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2016.04.015.

Fitzhenry, R., Dahan, S., Torres, A.G., Chong, Y., *et al*. (2006) Long polar fimbriae and tissue tropism in *Escherichia coli* O157:H7. *Microbes and Infection*. [Online] Available from: doi:10.1016/j.micinf.2006.02.012.

Fookes, M., Schroeder, G.N., Langridge, G.C., Blondel, C.J., *et al*. (2011) *Salmonella bongori* provides insights into the evolution of the *salmonellae*. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1002191.

Fraser, M.E., Fujinaga, M., Cherney, M.M., Melton-Celsa, A.R., *et al.* (2004) Structure of shiga toxin type 2 (Stx2) from *Escherichia coli* O157:H7. *The Journal of biological chemistry*. [Online] Available from: doi:10.1074/jbc.M401939200.

Fritz, J.H., Rojas, O.L., Simard, N., McCarthy, D.D., *et al*. (2011) Acquisition of a multifunctional IgA+ plasma cell phenotype in the gut. *Nature*. [Online] Available from: doi:10.1038/nature10698.

Fu, P., Zhang, X., Jin, M., Xu, L., *et al*. (2013) Complex structure of OspI and Ubc13: the molecular basis of Ubc13 deamidation and convergence of bacterial and sost E2 recognition. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1003322.

Furniss, R.C.D. & Clements, A. (2018) Regulation of the locus of enterocyte effacement in attaching and effacing pathogens. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.00336-17.

Furniss, R.C.D., Slater, S., Frankel, G. & Clements, A. (2016) Enterohaemorrhagic *E. coli* modulates an ARF6:Rab35 signaling axis to prevent recycling endosome maturation during infection. *Journal of Molecular Biology*. [Online] Available from: doi:10.1016/j.jmb.2016.05.023.

Furtado, A.R., Essid, M., Perrinet, S., Balañá, M.E., *et al*. (2013) The chlamydial OTU domain-containing protein ChlaOTU is an early type III secretion effector targeting ubiquitin and NDP52. *Cellular Microbiology*. [Online] Available from: doi:10.1111/cmi.12171.

Gaastra, W. & Svennerholm, A.M. (1996) Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends in Microbiology*. [Online]. Available from: doi:10.1016/0966-842X(96)10068-8.

Gallo, R.L. & Hooper, L. V. (2012) Epithelial antimicrobial defence of the skin and intestine. *Nature Reviews Immunology*. [Online]. Available from: doi:10.1038/nri3228.

Gao, X., Wang, X., Pham, T.H., Feuerbacher, L.A., *et al*. (2013) NleB, a bacterial effector with glycosyltransferase activity, targets GADPH function to inhibit NF-κB activation. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2012.11.010.

García-Angulo, V.A., Deng, W., Thomas, N.A., Finlay, B.B., *et al*. (2008) Regulation of expression and secretion of NleH, a new non-locus of enterocyte effacement-encoded effector in *Citrobacter rodentium*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01602-07.

García-Gómez, E., Espinosa, N., de la Mora, J., Dreyfus, G., *et al*. (2011) The muramidase EtgA from enteropathogenic *Escherichia coli* is required for efficient type III secretion. *Microbiology*. [Online] Available from: doi:10.1099/mic.0.045617-0.

Garmendia, J., Phillips, A.D., Carlier, M.-F., Chong, Y., *et al*. (2004) TccP is an enterohaemorrhagic *Escherichia coli* O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2004.00459.x.

Gauthier, A. & Finlay, B.B. (2003) Translocated intimin receptor and its chaperone interact with ATPase of the type III secretion apparatus of enteropathogenic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.185.23.6747-6755.2003.

Gaytán, M.O., Martínez-Santos, V.I., Soto, E. & González-Pedrajo, B. (2016) Type Three Secretion System in Attaching and Effacing Pathogens. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2016.00129.

Gaytán, M.O., Monjarás Feria, J., Soto, E., Espinosa, N., *et al*. (2018) Novel insights into the mechanism of SepL-mediated control of effector secretion in enteropathogenic *Escherichia coli. MicrobiologyOpen*. [Online] Available from: doi:10.1002/mbo3.571.

Geddes, K., Rubino, S.J., Magalhaes, J.G., Streutker, C., *et al*. (2011) Identification of an innate T helper type 17 response to intestinal bacterial pathogens. *Nature Medicine*. [Online] Available from: doi:10.1038/nm.2391.

Gee, H.Y., Noh, S.H., Tang, B.L., Kim, K.H., *et al*. (2011) Rescue of Δf508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell*. [Online] Available from: doi:10.1016/j.cell.2011.07.021.

Germane, K.L., Ohi, R., Goldberg, M.B. & Spiller, B.W. (2008) Structural and functional studies indicate that *Shigella* VirA is not a protease and does not directly destabilize microtubules. *Biochemistry*. [Online] Available from: doi:10.1021/bi801533k.

Ghavami, S., Eshragi, M., Ande, S.R., Chazin, W.J., *et al*. (2010) S100A8/A9 induces autophagy and apoptosis via ROS-mediated cross-talk between mitochondria and lysosomes that involves BNIP3. *Cell Research*. [Online] Available from: doi:10.1038/cr.2009.129.

Ghosh, I., Chowdhury, A.R., Rajeswari, M.R. & Datta, K. (2004) Differential expression of Hyaluronic Acid Binding protein 1 (HABP1)/P32/C1QBP during progression of epidermal carcinoma. *Molecular and Cellular Biochemistry*. [Online] Available from: doi:10.1023/B:MCBI.0000049362.04033.ea.

Gibson, D.L., Ma, C., Bergstrom, K.S.B., Huang, J.T., *et al*. (2008) MyD88 signalling plays a critical role in host defence by controlling pathogen burden and promoting epithelial cell homeostasis during *Citrobacter rodentium*-induced colitis. *Cellular microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2007.01071.x.

Giniger, E., Varnum, S.M. & Ptashne, M. (1985) Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell*. [Online] Available from: doi:10.1016/0092-8674(85)90336-8.

Giogha, C., Wong Fok Lung, T., Mühlen, S., Pearson, J.S., *et al*. (2015) Substrate recognition by the zinc metalloprotease effector NleC from enteropathogenic *Escherichia coli*. *Cellular Microbiology*. [Online] Available from: doi:10.1111/cmi.12469.

Girôn, J.A., Ho, A.S.Y. & Schoolnik, G.K. (1991) An inducible bundle-forming pilus of enteropathogenic *Escherichia coli. Science*. [Online] Available from: doi:10.1126/science.1683004.

Giuliani, F., Grieve, A. & Rabouille, C. (2011) Unconventional secretion: A stress on GRASP. *Current Opinion in Cell Biology*. [Online]. Available from: doi:10.1016/j.ceb.2011.04.005.

Glasser, A.L., Boudeau, J., Barnich, N., Perruchot, M.H., *et al*. (2001) Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.69.9.5529-5537.2001.

Green, E.R. & Mecsas, J. (2016) Bacterial secretion systems: An overview. *Microbiology spectrum*. [Online] Available from: doi:10.1128/microbiolspec.VMBF-0012-2015.

Grijpstra, J., Arenas, J., Rutten, L. & Tommassen, J. (2013) Autotransporter secretion: Varying on a theme. *Research in Microbiology*. [Online] Available from: doi:10.1016/j.resmic.2013.03.010.

Grishin, A.M., Cherney, M., Anderson, D.H., Phanse, S., *et a*l. (2014) NleH Defines a new family of bacterial effector kinases. *Structure*. [Online] Available from: doi:10.1016/j.str.2013.11.006.

Grivennikov, S., Karin, E., Terzic, J., Mucida, D., *et al*. (2009) IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell*. [Online] Available from: doi:10.1016/j.ccr.2009.01.001.

Gruenheid, S., Sekirov, I., Thomas, N.A., Deng, W., *et al*. (2004) Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic *Escherichia coli* O157:H7. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2003.03911.x.

Gueguen, E. & Cascales, E. (2013) Promoter swapping unveils the role of the *Citrobacter rodentium* CTS1 type VI secretion system in interbacterial competition. *Applied and Environmental Microbiology*. [Online] Available from: doi:10.1128/AEM.02504-12.

Gujral, T., Kumar, A., Priyamvada, S., Saksena, S., *et al*. (2015) Mechanisms of DRA recycling in intestinal epithelial cells: effect of enteropathogenic *E. coli. American Journal of Physiology-Cell Physiology*. [Online] Available from: doi:10.1152/ajpcell.00107.2015.

Guttman, J.A., Samji, F.N., Li, Y., Deng, W., *et al.* (2007) Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2006.00773.x.

Gwinn, D.M., Shackelford, D.B., Egan, D.F., Mihaylova, M.M., *et al*. (2008) AMPK Phosphorylation of Raptor Mediates a Metabolic Checkpoint. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2008.03.003.

Hachani, A., Wood, T.E. & Filloux, A. (2016) Type VI secretion and anti-host effectors. *Current Opinion in Microbiology*. [Online]. Available from: doi:10.1016/j.mib.2015.11.006.

Hao, R., Qiu, S., Wang, Y., Yang, G., *et al*. (2012) Quinolone-resistant *Escherichia coli* O127a:K63 serotype with an extended-spectrum-beta-lactamase phenotype from a food poisoning outbreak in China. *Journal of Clinical Microbiology*. [Online] Available from: doi:10.1128/JCM.00276-12.

Haraga, A. & Miller, S.I. (2006) A *Salmonella* type III secretion effector interacts with the mammalian serine/threonine protein kinase PKN1. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2005.00670.x.

Hardwidge, P.R., Deng, W., Vallance, B.A., Rodriguez-Escudero, I., *et al*. (2005) Modulation of host cytoskeleton function by the enteropathogenic *Escherichia coli* and *Citrobacter rodentium* effector protein EspG.

Infection and Immunity. [Online] Available from: doi:10.1128/IAI.73.5.2586-2594.2005.

Hartland, E.L., Batchelor, M., Delahay, R.M., Hale, C., *et al.* (1999) Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.1999.01338.x.

Hayakawa, S., Shiratori, S., Yamato, H., Kameyama, T., *et al*. (2011) ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. *Nature Immunology*. [Online] Available from: doi:10.1038/ni.1963.

Heise, T. & Dersch, P. (2006) Identification of a domain in *Yersinia* virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. *Proceedings of the National Academy of Sciences of the United States of America*. [Online] Available from: doi:10.1073/pnas.0507749103.

Hemrajani, C., Marches, O., Wiles, S., Girard, F., *et al*. (2008) Role of NleH, a type III secreted effector from attaching and effacing pathogens, in colonization of the bovine, ovine, and murine gut. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00742-08.

Herhaus, L. & Dikic, I. (2017) Regulation of *Salmonella*-host cell interactions via the ubiquitin system. *International Journal of Medical Microbiology*. [Online] Available from: doi:10.1016/j.ijmm.2017.11.003.

Herrero, M., De Lorenzo, V. & Timmis, K.N. (1990) Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/jb.172.11.6557-6567.1990.

Hershko, A. & Ciechanover, A. (1998) The ubiquitin system. *Annual review of biochemistry*. [Online] Available from: doi:10.1146/annurev.biochem.67.1.425.

Higgins, L.M., Frankel, G., Douce, G., Dougan, G., *et al*. (1999) *Citrobacter rodentium* infection in mice elicits a mucosal Th1 cytokine response and lesions similar to those in murine inflammatory bowel disease. *Infection and Immunity*.

Hjerpe, R., Hjerpe, R., Aillet, F., Aillet, F., *et al*. (2009) Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO reports*. [Online] Available from: doi:10.1038/embor.2009.192.

Hodges, K., Alto, N.M., Ramaswamy, K., Dudeja, P.K., *et al*. (2008) The enteropathogenic *Escherichia coli* effector protein EspF decreases sodium hydrogen exchanger 3 activity. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01163.x.

Hodges, K. & Gill, R. (2010) Infectious diarrhea: Cellular and molecular mechanisms. *Gut Microbes*. [Online]. Available from: doi:10.4161/gmic.1.1.11036.

Hodgkinson, J.L., Horsley, A., Stabat, D., Simon, M., *et al*. (2009) Three-dimensional reconstruction of the *Shigella* T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.1599.

Hoffman, M.A., Menge, C., Casey, T.A., Laegreid, W., *et al.* (2006) Bovine immune response to Shiga-toxigenic *Escherichia coli* O157:H7. *Clinical and Vaccine Immunology*. [Online] Available from: doi:10.1128/CVI.00205-06.

Holden, V.I., Lenio, S., Kuick, R., Ramakrishnan, S.K., *et al*. (2014) Bacterial siderophores that evade or overwhelm lipocalin 2 induce hypoxia inducible factor 1α and proinflammatory cytokine secretion in cultured respiratory epithelial cells. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.01849-14.

Holmes, A., Mühlen, S., Roe, A.J. & Dean, P. (2010) The EspF effector, a bacterial pathogen's swiss army knife. *Infection and Immunity*. [Online]. Available from: doi:10.1128/IAI.00635-10.

Hör, S., Ziv, T., Admon, A. & Lehner, P.J. (2009) Stable isotope labeling by amino acids in cell culture and differential plasma membrane proteome quantitation identify new substrates for the MARCH9 transmembrane E3 ligase. *Molecular & cellular proteomics : MCP*. [Online] Available from: doi:10.1074/mcp.M900174-MCP200.

Horne, C., Vallance, B.A., Deng, W. & Finlay, B.B. (2002) Current progress in enteropathogenic and enterohemorrhagic *Escherichia coli* vaccines. *Expert Review of Vaccines*. [Online]. Available from: doi:10.1586/14760584.1.4.483.

Hospenthal, M.K., Freund, S.M.V. & Komander, D. (2013) Assembly, analysis and architecture of atypical ubiquitin chains. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.2547.

Hsu, F., Luo, X., Qiu, J., Teng, Y.-B., *et al*. (2014) The *Legionella* effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1402605111.

Hsu, Y., Jubelin, G., Taieb, F., Nougayrède, J.-P., *et al*. (2008) Structure of the cyclomodulin Cif from pathogenic *Escherichia coli*. *Journal of molecular biology*. [Online] Available from: doi:10.1016/j.jmb.2008.09.051.

Hu, J. & Torres, A.G. (2015) Enteropathogenic *Escherichia coli*: foe or innocent bystander? *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. [Online] Available from: doi:10.1016/j.cmi.2015.01.015.

Huang, Z., Sutton, S.E., Wallenfang, A.J., Orchard, R.C., *et al*. (2009) Structural insights into host GTPase isoform selection by a family of bacterial GEF mimics. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.1647.

Hudson, D.L., Layton, A.N., Field, T.R., Bowen, A.J., *et al.* (2007) Inhibition of type III secretion in *Salmonella enterica* serovar Typhimurium by small-molecule inhibitors. *Antimicrobial Agents and Chemotherapy*. [Online] Available from: doi:10.1128/AAC.01492-06.

Hughes, D.T., Clarke, M.B., Yamamoto, K., Rasko, D.A., *et al*. (2009) The QseC adrenergic signaling cascade in enterohemorrhagic *E. coli* (EHEC). *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1000553.

Humphries, R.M. & Linscott, A.J. (2015) Laboratory diagnosis of bacterial gastroenteritis. *Clinical Microbiology Reviews*. [Online] Available from: doi:10.1128/CMR.00073-14.

Ibuki, T., Imada, K., Minamino, T., Kato, T., *et al.* (2011) Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.1977.

Ide, T., Laarmann, S., Greune, L., Schillers, H., *et al.* (2001) Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic *Escherichia coli*. *Cellular Microbiology*. [Online] Available from: doi:10.1046/j.1462-5822.2001.00146.x.

Ide, T., Michgehl, S., Knappstein, S., Heusipp, G., *et al*. (2003) Differential modulation by Ca2+ of type III secretion of diffusely adhering enteropathogenic *Escherichia coli*. *Infect. Immun.* [Online] Available from: doi:10.1128/iai.71.4.1725-1732.2003.

Iguchi, A., Thomson, N.R., Ogura, Y., Saunders, D., *et al.* (2009) Complete genome sequence and comparative

genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01238-08.

Imada, K., Minamino, T., Tahara, A. & Namba, K. (2007) Structural similarity between the flagellar type III ATPase FliI and F1-ATPase subunits. *Proceedings of the National Academy of Sciences of the United States of America*. [Online] 104 (2), 485–490. Available from: doi:10.1073/pnas.0608090104 [Accessed: 6 May 2014].

Inman, L.R. & Cantey, J.R. (1984) Peyer's patch lymphoid follicle epithelial adherence of a rabbit enteropathogenic *Escherichia coli* (strain RDEC-1). Role of plasmid-mediated pili in initial adherence. *Journal of Clinical Investigation*. [Online] Available from: doi:10.1172/JCI111423.

Inman, L.R. & Cantey, J.R. (1983) Specific adherence of *Escherichia coli* (strain RDEC-1) to membranous (M) cells of the Peyer's patch in *Escherichia coli* diarrhea in the rabbit. *Journal of Clinical Investigation*. [Online] Available from: doi:10.1172/JCI110737.

Inoue, J., Nishiumi, S., Fujishima, Y., Masuda, A., *et al*. (2012) Autophagy in the intestinal epithelium regulates *Citrobacter rodentium* infection. *Archives of Biochemistry and Biophysics*. [Online] Available from: doi:10.1016/j.abb.2012.03.019.

Itahana, K. & Zhang, Y. (2008) Mitochondrial p32 is a critical mediator of ARF-induced apoptosis. *Cancer Cell*. [Online] Available from: doi:10.1016/j.ccr.2008.04.002.

Itoh, Y., Nagano, I., Kunishima, M. & Ezaki, T. (1997) Laboratory investigation of enteroaggregative *Escherichia coli* O untypeable:H10 associated with a massive outbreak of gastrointestinal illness. *Journal of Clinical Microbiology*.

Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., *et al*. (2009) Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. [Online] Available from: doi:10.1016/j.cell.2009.09.033.

Ivanov, S.S. & Roy, C.R. (2009) Modulation of ubiquitin dynamics and suppression of DALIS formation by the *Legionella pneumophila* Dot/Icm system. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01251.x.

Iwobi, A., Heesemann, J., Garcia, E., Igwe, E., *et al*. (2003) Novel virulence-associated type II secretion system unique to high-pathogenicity *Yersinia enterocolitica*. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.71.4.1872-1879.2003.

Izoré, T., Job, V. & Dessen, A. (2011) Biogenesis, regulation, and targeting of the type III secretion system. *Structure*. [Online]. Available from: doi:10.1016/j.str.2011.03.015.

Jaehning, J.A. (2010) The Paf1 complex: Platform or player in RNA polymerase II transcription? *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. [Online]. Available from: doi:10.1016/j.bbagrm.2010.01.001.

Jean Beltran, P.M., Federspiel, J.D., Sheng, X. & Cristea, I.M. (2017) Proteomics and integrative omic approaches for understanding host–pathogen interactions and infectious diseases. *Molecular Systems Biology*. [Online] Available from: doi:10.15252/msb.20167062.

De Jong, M.F., Liu, Z., Chen, D. & Alto, N.M. (2016) *Shigella flexneri* suppresses NF-κB activation by inhibiting linear ubiquitin chain ligation. *Nature Microbiology*. [Online] Available from: doi:10.1038/nmicrobiol.2016.84.

Journet, L., Agrain, C., Broz, P. & Cornelis, G.R. (2003) The needle length of bacterial injectisomes is determined by a molecular ruler. *Science*. [Online] Available from: doi:10.1126/science.1091422.

Journet, L. & Cascales, E. (2016) The type VI secretion system in *Escherichia coli* and related species. *EcoSal Plus*. [Online] Available from: doi:10.1128/ecosalplus.ESP-0009-2015.

Kamada, N., Seo, S.-U., Chen, G.Y. & Núñez, G. (2013) Role of the gut microbiota in immunity and inflammatory disease. *Nature Reviews Immunology*. [Online] Available from: doi:10.1038/nri3430.

Kamal, A. & Datta, K. (2006) Upregulation of hyaluronan binding protein 1 (HABP1/p32/gC1qR) is associated with Cisplatin induced apoptosis. *Apoptosis*. [Online] Available from: doi:10.1007/s10495-006-5396-4.

Kamanova, J., Sun, H., Lara-Tejero, M. & Galán, J.E. (2016) The *Salmonella* effector protein SopA modulates innate immune responses by targeting TRIM E3 ligase family members. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1005552.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., *et al*. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*. [Online] Available from: doi:10.1093/nar/gkv1070.

Kang, E., Zhou, G., Yousefi, M., Cayrol, R., *et al*. (2018) Loss of disease tolerance during *Citrobacter rodentium* infection is associated with impaired epithelial differentiation and hyperactivation of T cell responses. *Scientific Reports*. [Online] Available from: doi:10.1038/s41598-017-17386-y.

Kaniga, K., Delor, I. & Cornelis, G.R. (1991) A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene*. [Online] Available from: doi:10.1016/0378-1119(91)90599-7.

Kaniuk, N.A., Lam, G.Y., Ma, C., Galindo-Mata, E., *et al*. (2011) *Citrobacter rodentium* infection induces MyD88-dependent formation of ubiquitinated protein aggregates in the intestinal epithelium. *Journal of Innate Immunity*. [Online] Available from: doi:10.1159/000320644.

Kaper, J.B., Nataro, J.P. & Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nature reviews. Microbiology*. [Online] Available from: doi:10.1038/nrmicro818.

Kaser, a, Lee, a H., Franke, a, Glickman, J.N., *et al*. (2008) XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell*. [Online] Available from: doi:10.1016/j.cell.2008.07.021.

Kauppi, A.M., Andersson, C.D., Norberg, H.A., Sundin, C., *et al*. (2007) Inhibitors of type III secretion in *Yersinia*: Design, synthesis and multivariate QSAR of 2-arylsulfonylamino-benzanilides. *Bioorganic and Medicinal Chemistry*. [Online] Available from: doi:10.1016/j.bmc.2007.07.047.

Kawahara, T., Kuwano, Y., Teshima-Kondo, S., Takeya, R., *et al*. (2004) Role of nicotinamide adenine dinucleotide phosphate oxidase 1 in oxidative burst response to Toll-like receptor 5 signaling in large intestinal epithelial cells. *Journal of immunology (Baltimore, Md. : 1950)*. [Online] Available from: doi:10.4049/jimmunol.172.5.3051.

Kayagaki, N., Warming, S., Lamkanfi, M., Walle, L. Vande, *et al*. (2011) Non-canonical inflammasome activation targets caspase-11. *Nature*. [Online] Available from: doi:10.1038/nature10558.

Kayagaki, N., Wong, M.T., Stowe, I.B., Ramani, S.R., *et al.* (2013) Noncanonical inflammasome activation by intracellular LPS independent of TLR4. *Science*. [Online] Available from: doi:10.1126/science.1240248.

Keeney, K.M., Yurist-Doutsch, S., Arrieta, M.-C. & Finlay, B.B. (2014) Effects of antibiotics on human microbiota and subsequent disease. *Annual Review of Microbiology*. [Online] Available from: doi:10.1146/annurev-micro-091313-103456.

Kenny, B. (1999) Phosphorylation of tyrosine 474 of the enteropathogenic *Escherichia coli* (EPEC) Tir receptor molecule is essential for actin nucleating activity and is preceded by additional host modifications.

Molecular Microbiology. [Online] Available from: doi:10.1046/j.1365-2958.1999.01265.x.

Kenny, B., Abe, a, Stein, M. & Finlay, B.B. (1997) Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. *Infection and immunity*.

Kenny, B., Ellis, S., Leard, A.D., Warawa, J., *et al.* (2002) Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic *Escherichia coli* effector molecules. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2002.02952.x.

Kerns, J.A., Emerman, M. & Malik, H.S. (2008) Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genetics*. [Online] Available from: doi:10.1371/journal.pgen.0040021.

Keszei, A.F.A. & Sicheri, F. (2017) Mechanism of catalysis, E2 recognition, and autoinhibition for the IpaH family of bacterial E3 ubiquitin ligases. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1611595114.

Keszei, A.F.A., Tang, X., McCormick, C., Zeqiraj, E., *et al.* (2014) Structure of an SspH1-PKN1 complex reveals the basis for host substrate recognition and mechanism of activation for a bacterial E3 ubiquitin ligase. *Molecular and Cellular Biology*. [Online] Available from: doi:10.1128/MCB.01360-13.

Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M., *et al.* (2010) The ARRIVE guidelines Animal Research : Reporting In Vivo Experiments. *Journal of pharmacology & pharmacotherapeutics*. [Online] Available from: doi:10.4103/0976-500X.72351.

Kim, D.W., Lenzen, G., Page, A.-L., Legrain, P., *et al.* (2005) The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0504466102.

Kim, J.-G., Taylor, K.W., Hotson, A., Keegan, M., *et al*. (2008) XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in xanthomonas-infected tomato leaves. *THE PLANT CELL ONLINE*. [Online] Available from: doi:10.1105/tpc.108.058529.

Kim, J. & Gibson, G. (2010) Insights from GWAS into the quantitative genetics of transcription in humans. *Genetics Research*. [Online] Available from: doi:10.1017/S001667231000056X.

Kim, J., Thanabalasuriar, A., Chaworth-Musters, T., Fromme, J.C., *et al*. (2007) The bacterial virulence factor NleA inhibits cellular protein secretion by disrupting mammalian COPII function. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2007.07.010.

Kim, J.G., Stork, W. & Mudgett, M.B. (2013) *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2013.01.006.

Kim, K., Kim, M.J., Kim, K.H., Ahn, S.A., *et al*. (2017) C1QBP is upregulated in colon cancer and binds to apolipoprotein A-I. *Experimental and Therapeutic Medicine*. [Online] Available from: doi:10.3892/etm.2017.4249.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., *et al*. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2011.08.025.

Kimmitt, P.T., Harwood, C.R. & Barer, M.R. (2000) Toxin gene expression by shiga toxin-producing *Escherichia coli*: the role of antibiotics and the bacterial SOS response. *Emerging infectious diseases*. [Online] Available from: doi:10.3201/eid0605.000503.

Klapproth, J.M.A., Sasaki, M., Sherman, M., Babbin, B., *et al*. (2005) *Citrobacter rodentium lifA/efa1* is essential

for colonic colonization and crypt cell hyperplasia in vivo. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.73.3.1441-1451.2005.

Klapproth, J.M.A., Scaletsky, I.C.A., McNamara, B.P., Lai, L.C., *et al*. (2000) A large toxin from pathogenic *Escherichia coli* strains that inhibits lymphocyte activation. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.68.4.2148-2155.2000.

Klink, B.U., Barden, S., Heidler, T. V., Borchers, C., *et al*. (2010) Structure of *Shigella* IpgB2 in complex with human RhoA: Implications for the mechanism of bacterial guanine nucleotide exchange factor mimicry. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M110.107953.

Koike, A., Arai, S., Yamada, S., Nagae, A., *et al.* (2012) Dynamic mobility of immunological cells expressing S100A8 and S100A9 *in vivo*: A variety of functional roles of the two proteins as regulators in acute inflammatory reaction. *Inflammation*. [Online] Available from: doi:10.1007/s10753-011-9330-8.

Koliopoulos, M.G., Esposito, D., Christodoulou, E., Taylor, I.A., *et al*. (2016) Functional role of TRIM E 3 ligase oligomerization and regulation of catalytic activity. *The EMBO journal*. [Online] Available from: doi:10.15252/embj.201593741.

Komander, D. (2009) The emerging complexity of protein ubiquitination. *Biochemical Society Transactions*. [Online] Available from: doi:10.1042/BST0370937.

Komander, D. & Rape, M. (2012) The ubiquitin code. *Annual Review of Biochemistry*. [Online] Available from: doi:10.1146/annurev-biochem-060310-170328.

Konowalchuk, J., Speirs, J.I. & Stavric, S. (1977) Vero response to a cytotoxin of *Escherichia coli*. *Infection and Immunity*.

Koroleva, E.P., Halperin, S., Gubernatorova, E.O., Macho-Fernandez, E., *et al*. (2015) *Citrobacter rodentium*-induced colitis: A robust model to study mucosal immune responses in the gut. *Journal of Immunological Methods*. [Online] Available from: doi:10.1016/j.jim.2015.02.003.

Korotkov, K. V., Sandkvist, M. & Hol, W.G.J. (2012) The type II secretion system: Biogenesis, molecular architecture and mechanism. *Nature Reviews Microbiology*. [Online]. Available from: doi:10.1038/nrmicro2762.

Kosek, M., Bern, C. & Guerrant, R.L. (2003) The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bulletin of the World Health Organization*. [Online] Available from: doi:S0042-96862003000300010 [pii].

Kotloff, K.L., Nataro, J.P., Blackwelder, W.C., Nasrin, D., et al. (2013) Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *The Lancet*. [Online] Available from: doi:10.1016/S0140-6736(13)60844-2.

Krause, G., Frank, C., Gilsdorf, A., Mielke, M., *et al.* (2013) The 2011 HUS epidemic in Germany. Challenges for disease control: What should be improved? *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*. [Online] Available from: doi:10.1007/s00103-012-1585-1.

Kruglov, A. a, Grivennikov, S.I., Kuprash, D. V, Winsauer, C., *et al*. (2013) Nonredundant function of soluble LTα3 produced by innate lymphoid cells in intestinal homeostasis. *Science*. [Online] Available from: doi:10.1126/science.1243364.

Kubori, T., Hyakutake, A. & Nagai, H. (2008) *Legionella* translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-

2958.2008.06124.x.

Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., *et al.* (1998) Supramolecular structure of the *Salmonella* Typhimurium type III protein secretion system. *Science (New York, N.Y.)*. [Online] Available from: doi:10.1126/science.280.5363.602.

Kubori, T., Shinzawa, N., Kanuka, H. & Nagai, H. (2010) *Legionella* metaeffector exploits host proteasome to temporally regulate cognate effector. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1001216.

Kudryashev, M., Stenta, M., Schmelz, S., Amstutz, M., *et al*. (2013) *In situ* structural analysis of the *Yersinia enterocolitica* injectisome. *eLife*. [Online] 2, e00792. Available from: doi:10.7554/eLife.00792 [Accessed: 6 May 2014].

Kulathu, Y. & Komander, D. (2012) Atypical ubiquitylation-the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nature Reviews Molecular Cell Biology*. [Online]. Available from: doi:10.1038/nrm3394.

Kunzi, M.S. & Pitha, P.M. (2003) Interferon targeted genes in host defense. *Autoimmunity*. [Online]. Available from: doi:10.1080/08916930310001605855.

Kuo, a, Zhong, C., Lane, W.S. & Derynck, R. (2000) Transmembrane transforming growth factor-α tethers to the PDZ domain-containing, Golgi membrane-associated protein p59/GRASP55. *The EMBO journal*. [Online] Available from: doi:10.1093/emboj/19.23.6427.

De la Cruz, M.A., Morgan, J.K., Ares, M.A., Yáñez-Santos, J.A., *et al*. (2016) The two-component system CpxRA negatively regulates the Locus of Enterocyte Effacement of enterohemorrhagic *Escherichia coli* involving σ32 and Lon protease. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2016.00011.

Lamprecht, G. & Seidler, U. (2006) The emerging role of PDZ adapter proteins for regulation of intestinal ion transport. *AJP: Gastrointestinal and Liver Physiology*. [Online] Available from: doi:10.1152/ajpgi.00135.2006.

Lanata, C.F., Fischer-Walker, C.L., Olascoaga, A.C., Torres, C.X., *et al*. (2013) Global causes of diarrheal disease mortality in children. *PloS one*. [Online] Available from: doi:10.1371/journal.pone.0072788.

Lara-Tejero, M., Kato, J., Wagner, S., Liu, X., *et al*. (2011) A sorting platform determines the order of protein secretion in bacterial type III systems. *Science (New York, N.Y.)*. [Online] 331 (6021), 1188–1191. Available from: doi:10.1126/science.1201476 [Accessed: 28 March 2014].

Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J.M., *et al*. (2005) BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2005.04907.x.

Lee, H.J. & Zheng, J.J. (2010) PDZ domains and their binding partners: Structure, specificity, and modification. *Cell Communication and Signaling*. [Online]. Available from: doi:10.1186/1478-811X-8-8.

Lee, P.-C. & Rietsch, A. (2015) Fueling type III secretion. *Trends in Microbiology*. [Online] Available from: doi:10.1016/j.tim.2015.01.012.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., *et al*. (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0813360106.

Leimbach, A., Hacker, J. & Dobrindt, U. (2013) *E. coli* as an all-rounder: The thin line between commensalism

and pathogenicity. *Current Topics in Microbiology and Immunology*. [Online] Available from: doi:10.1007/82_2012_303.

Leininger, E., Roberts, M., Kenimer, J.G., Charles, I.G., *et al*. (1991) Pertactin, an Arg-Gly-Asp-containing Bordetella pertussis surface protein that promotes adherence of mammalian cells. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.88.2.345.

Lelouard, H., Ferrand, V., Marguet, D., Bania, J., *et al.* (2004) Dendritic cell aggresome-like induced structures are dedicated areas for ubiquitination and storage of newly synthesized defective proteins. *Journal of Cell Biology*. [Online] Available from: doi:10.1083/jcb.200312073.

Lelouard, H., Gatti, E., Cappello, F., Gresser, O., *et al.* (2002) Transient aggregation of ubiquitinated proteins during dendritic cell maturation. *Nature*. [Online] Available from: doi:10.1038/417177a.

Lesser, C.F. & Miller, S.I. (2001) Expression of microbial virulence proteins in *Saccharomyces cerevisiae* models mammalian infection. *EMBO Journal*. [Online] Available from: doi:10.1093/emboj/20.8.1840.

Levi, S.K. & Glick, B.S. (2007) GRASPing Unconventional Secretion. *Cell*. [Online]. Available from: doi:10.1016/j.cell.2007.07.030.

Levine, M.M., Kotloff, K.L., Nataro, J.P. & Muhsen, K. (2012) The Global Enteric Multicenter Study (GEMS): Impetus, rationale, and genesis. *Clinical Infectious Diseases*. [Online] Available from: doi:10.1093/cid/cis761.

Levine, M.M., Nalin, D.R., Hornick, R.B., Bergquist, E.J., *et al*. (1978) *Escherichia coli* strains that cause diarrhœa but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *The Lancet*. [Online] Available from: doi:10.1016/S0140-6736(78)90299-4.

Li, J., Poulikakos, P.I., Dai, Z., Testa, J.R., *et al*. (2007) Protein kinase c phosphorylation disrupts Na+/H+ Exchanger Regulatory Factor 1 autoinhibition and promotes cystic fibrosis transmembrane conductance regulator macromolecular assembly. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M702019200.

Li, M., Rosenshine, I., Yu, H.B., Nadler, C., *et al*. (2006) Identification and characterization of NleI, a new non-LEE-encoded effector of enteropathogenic *Escherichia coli* (EPEC). *Microbes and Infection*. [Online] Available from: doi:10.1016/j.micinf.2006.09.006.

Li, M.M.H., Lau, Z., Cheung, P., Aguilar, E.G., *et al.* (2017) TRIM25 enhances the antiviral action of Zinc-finger Antiviral Protein (ZAP). *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1006145.

Li, Y., Banerjee, S., Wang, Y., Goldstein, S.A., *et al*. (2016) Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1519657113.

Liao, A.P., Petrof, E.O., Kuppireddi, S., Zhao, Y., *et al.* (2008) *Salmonella* type III effector AvrA stabilizes cell tight junctions to inhibit inflammation in intestinal epithelial cells. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0002369.

Lightowlers, R.N. & Chrzanowska-Lightowlers, Z.M.A. (2013) Human pentatricopeptide proteins: Only a few and what do they do? *RNA Biology*. [Online]. Available from: doi:10.4161/rna.24770.

Lilic, M., Vujanac, M. & Stebbins, C.E. (2006) A common structural motif in the binding of virulence factors to bacterial secretion chaperones. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2006.01.026.

Lim, B.L., Reid, K.B.M., Ghebrehiwet, B., Peerschke, E.I.B., *et al.* (1996) The binding protein for globular heads

of complement C1q, gC1qR: Functional expression and characterization as a novel vitronectin binding factor. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.271.43.26739.

Lin, D.Y. -w., Diao, J. & Chen, J. (2012) Crystal structures of two bacterial HECT-like E3 ligases in complex with a human E2 reveal atomic details of pathogen-host interactions. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1115025109.

Lin, D.Y., Diao, J., Zhou, D. & Chen, J. (2011) Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7. *The Journal of biological chemistry*. [Online] Available from: doi:10.1074/jbc.M110.167643.

Lin, Y.H., Doms, A.G., Cheng, E., Kim, B., et al. (2015) Host cell-catalyzed S-palmitoylation mediates Golgi targeting of the *Legionella* ubiquitin ligase GobX. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M115.637397.

Lin, Y.H., Lucas, M., Evans, T.R., Abascal-Palacios, G., *et al*. (2018) RavN is a member of a previously unrecognized group of *Legionella pneumophila* E3 ubiquitin ligases. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1006897.

Lindenthal, C. & Elsinghorst, E.A. (1999) Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli. Infection and Immunity*.

Linstedt, A.D. & Hauri, H.P. (1993) Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Molecular Biology of the Cell*. [Online] Available from: doi:10.1091/mbc.4.7.679.

Lippert, E., Falk, W., Bataille, F., Kaehne, T., *et al*. (2007) Soluble galectin-3 is a strong, colonic epithelial-cell-derived, lamina propria fibroblast-stimulating factor. *Gut*. [Online] Available from: doi:10.1136/gut.2005.081646.

Litvak, Y., Sharon, S., Hyams, M., Zhang, L., *et al.* (2017) Epithelial cells detect functional type III secretion system of enteropathogenic *Escherichia coli* through a novel NF-κB signaling pathway. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1006472.

Liu, C., Heath, L.S. & Turnbough, C.L. (1994) Regulation of pyrBI operon expression in *Escherichia coli* by UTP-sensitive reiterative RNA synthesis during transcriptional initiation. *Genes and Development*. [Online] Available from: doi:10.1101/gad.8.23.2904.

Liu, L., Oza, S., Hogan, D., Perin, J., *et al*. (2015a) Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: An updated systematic analysis. *The Lancet*. [Online] Available from: doi:10.1016/S0140-6736(14)61698-6.

Liu, Z., Kato, A., Oyama, M., Kozuka-Hata, H., *et al*. (2015b) Role of host cell p32 in *Herpes simplex* virus 1 de-envelopment during viral nuclear egress. *Journal of Virology*. [Online] Available from: doi:10.1128/JVI.01220-15.

Lloyd, S.A., Norman, M., Rosqvist, R. & Wolf-Watz, H. (2001) *Yersinia* YopE is targeted for type III secretion by N-terminal, not mRNA, signals. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2001.02271.x.

Loerch, S., Maucuer, A., Manceau, V., Green, M.R., *et al*. (2014) Cancer-relevant splicing factor CAPERα engages the essential splicing factor SF3b155 in a specific ternary complex. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M114.558825.

Lomma, M., Dervins-Ravault, D., Rolando, M., Nora, T., *et al.* (2010) The *Legionella pneumophila* F-box protein

Lpp2082 (AnkB) modulates ubiquitination of the host protein parvin B and promotes intracellular replication. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2010.01467.x.

Lopez, P.J., Marchand, I., Joyce, S.A. & Dreyfus, M. (1999) The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.1999.01465.x.

Loukiadis, E., Nobe, R., Herold, S., Tramuta, C., *et al.* (2008) Distribution, functional expression, and genetic organization of Cif, a phage-encoded type III-secreted effector from enteropathogenic and enterohemorrhagic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.00844-07.

Low, H.H., Gubellini, F., Rivera-Calzada, A., Braun, N., *et al*. (2014) Structure of a type IV secretion system. *Nature*. [Online] Available from: doi:10.1038/nature13081.

Lu, L.S. & Yeh, E.T.H. (2013) Ubiquitin-Like Proteins. In: *Encyclopedia of Biological Chemistry: Second Edition*. [Online]. p. Available from: doi:10.1016/B978-0-12-378630-2.00068-2.

Lubelski, J., Konings, W.N. & Driessen, A.J.M. (2007) Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiology and Molecular Biology Reviews*. [Online] Available from: doi:10.1128/MMBR.00001-07.

Luperchio, S.A., Newman, J. V., Dangler, C.A., Schrenzel, M.D., *et al*. (2000) *Citrobacter rodentium*, the causative agent of transmissible murine colonic hyperplasia, exhibits clonality: Synonymy of *C. rodentium* and mouse-pathogenic *Escherichia coli*. *Journal of Clinical Microbiology*.

Luperchio, S.A. & Schauer, D.B. (2001) Molecular pathogenesis of *Citrobacter rodentium* and transmissible murine colonic hyperplasia. *Microbes and Infection*. [Online]. Available from: doi:10.1016/S1286-4579(01)01387-9.

Lupfer, C.R., Anand, P.K., Liu, Z., Stokes, K.L., *et al.* (2014) Reactive oxygen species regulate caspase-11 expression and activation of the non-canonical NLRP3 Inflammasome during enteric pathogen infection. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1004410.

Lupp, C., Robertson, M.L., Wickham, M.E., Sekirov, I., *et al*. (2007) Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of *Enterobacteriaceae*. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2007.06.010.

Lynn, R.M., O'Brien, S.J., Taylor, C.M., Adak, G.K., *et al.* (2005) Childhood hemolytic uremic syndrome, United Kingdom and Ireland. *Emerging Infectious Diseases*. [Online] Available from: doi:10.3201/eid1104.040833.

Macho-Fernandez, E., Koroleva, E.P., Spencer, C.M., Tighe, M., *et al*. (2015) Lymphotoxin beta receptor signaling limits mucosal damage through driving IL-23 production by epithelial cells. *Mucosal Immunology*. [Online] Available from: doi:10.1038/mi.2014.78.

Mackman, N. & Holland, I.B. (1984) Functional characterization of a cloned haemolysin determinant from *E. coli* of human origin, encoding information for the secretion of a 107K polypeptide. *MGG Molecular & General Genetics*. [Online] Available from: doi:10.1007/BF00334104.

Maculins, T., Carter, N., Dorval, T., Hudson, K., *et al*. (2016a) A generic platform for cellular screening against ubiquitin ligases. *Scientific Reports*. [Online] Available from: doi:10.1038/srep18940.

Maculins, T., Fiskin, E., Bhogaraju, S. & Dikic, I. (2016b) Bacteria-host relationship: Ubiquitin ligases as weapons of invasion. *Cell Research*. [Online]. Available from: doi:10.1038/cr.2016.30.

Magori, S. & Citovsky, V. (2011) *Agrobacterium* counteracts host-induced degradation of its effector F-box protein. *Science Signaling*. [Online] Available from: doi:10.1126/scisignal.2002124.

Mahajan, A., Currie, C.G., Mackie, S., Tree, J., *et al*. (2009) An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157: H7 with bovine intestinal epithelium. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01244.x.

Manta, C., Heupel, E., Radulovic, K., Rossini, V., *et al.* (2013) CX3CR1+ macrophages support IL-22 production by innate lymphoid cells during infection with *Citrobacter rodentium*. *Mucosal Immunology*. [Online] Available from: doi:10.1038/mi.2012.61.

Marchès, O., Batchelor, M., Shaw, R.K., Patel, A., *et al*. (2006) EspF of enteropathogenic *Escherichia coli* binds sorting nexin 9. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.188.8.3110-3115.2006.

Marchès, O., Ledger, T.N., Boury, M., Ohara, M., *et al*. (2003) Enteropathogenic and enterohaemorrhagic *Escherichia coli* deliver a novel effector called Cif, which blocks cell cycle G2/M transition. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2003.03821.x.

Marteyn, B.S., Gazi, A.D. & Sansonetti, P.J. (2012) *Shigella*: A model of virulence regulation *in vivo*. *Gut Microbes*. [Online] Available from: doi:10.4161/gmic.19325.

Martínez-García, E., Aparicio, T., Goñi-Moreno, A., Fraile, S., *et al.* (2015) SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic acids research*. [Online] 43 (Database issue), D1183-9. Available from: doi:10.1093/nar/gku1114 [Accessed: 7 September 2018].

Martinez, E., Schroeder, G.N., Berger, C.N., Lee, S.F., *et al*. (2010) Binding to Na+/H+exchanger regulatory factor 2 (NHERF2) affects trafficking and function of the enteropathogenic *Escherichia coli* type III secretion system effectors Map, EspI and NleH. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2010.01503.x.

Mathupala, S.P., Ko, Y.H. & Pedersen, P.L. (2009) Hexokinase-2 bound to mitochondria: Cancer's stygian link to the 'Warburg effect' and a pivotal target for effective therapy. *Seminars in Cancer Biology*. [Online]. Available from: doi:10.1016/j.semcancer.2008.11.006.

McDowell, G.S. & Philpott, A. (2013) Non-canonical ubiquitylation: Mechanisms and consequences. *International Journal of Biochemistry and Cell Biology*. [Online]. Available from: doi:10.1016/j.biocel.2013.05.026.

Mcewan, W.A., Mallery, D.L., Rhodes, D.A., Trowsdale, J., *et al*. (2011) Intracellular antibody-mediated immunity and the role of TRIM21. *BioEssays*. [Online] Available from: doi:10.1002/bies.201100093.

McGee, A.M. & Baines, C.P. (2011) Complement 1q-binding protein inhibits the mitochondrial permeability transition pore and protects against oxidative stress-induced death. *The Biochemical journal*. [Online] Available from: doi:10.1042/BJ20101431.

McGee, A.M., Douglas, D.L., Liang, Y., Hyder, S.M., *et al*. (2011) The mitochondrial protein C1qbp promotes cell proliferation, migration and resistance to cell death. *Cell cycle (Georgetown, Tex.)*. [Online] Available from: doi:10.4161/cc.10.23.18287.

McNamara, B.P., Koutsouris, A., O'Connell, C.B., Nougayréde, J.P., *et al.* (2001) Translocated EspF protein from enteropathogenic *Escherichia coli* disrupts host intestinal barrier function. *Journal of Clinical Investigation*. [Online] Available from: doi:10.1172/JCI11138.

McShan, A.C. & De Guzman, R.N. (2015) The bacterial type III secretion system as a target for developing new

antibiotics. *Chem Biol Drug Des*. [Online] Available from: doi:10.1111/cbdd.12422.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., *et al.* (2013) The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nature Methods*. [Online] Available from: doi:10.1038/nmeth.2557.

Mesquita, F.S., Thomas, M., Sachse, M., Santos, A.J.M., *et al.* (2012) The *Salmonella* deubiquitinase SseI inhibits selective autophagy of cytosolic aggregates. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1002743.

Miao, E.A., Scherer, C.A., Tsolis, R.M., Kingsley, R.A., *et al.* (1999) *Salmonella* typhimurium leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. *Mol Microbiol*. [Online] Available from: doi:mmi1651 [pii].

Michiels, T. & Cornelis, G.R. (1991) Secretion of hybrid proteins by the *Yersinia* Yop export system. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/jb.173.5.1677-1685.1991.

Minamino, T. & Macnab, R.M. (2000) Domain structure of *Salmonella* FlhB , a flagellar export component responsible for substrate specificity switching. *J Bacteriol* [Online] 182 (17). Available from: doi:10.1128/JB.182.17.4906-4914.2000.Updated.

Misaghi, S., Balsara, Z.R., Catic, A., Spooner, E., *et al.* (2006) *Chlamydia trachomatis*-derived deubiquitinating enzymes in mammalian cells during infection. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2006.05199.x.

Miyasaki, K.T., Bodeau, A.L., Murthy, A.R. & Lehrer, R.I. (1993) *In vitro* antimicrobial activity of the human neutrophil cytosolic S-100 protein complex, calprotectin, against *Capnocytophaga sputigena*. *Journal of dental research*. [Online] Available from: doi:10.1177/00220345930720020801.

Von Moll, L.K. & Cantey, J.R. (1997) Peyer's patch adherence of enteropathogenic *Escherichia coli* strains in rabbits. *Infect Immun*.

Moon, H.W., Whipp, S.C., Argenzio, R.A., Levine, M.M., *et al*. (1983) Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infection and Immunity*.

Morales, F.C., Takahashi, Y., Momin, S., Adams, H., et al. (2007) NHERF1/EBP50 Head-to-tail intramolecular interaction masks association with PDZ domain ligands. *Molecular and Cellular Biology*. [Online] Available from: doi:10.1128/MCB.01372-06.

Moreira, C.G., Russell, R., Mishra, A.A., Narayanan, S., *et al*. (2016) Bacterial adrenergic sensors regulate virulence of enteric pathogens in the gut. *mBio*. [Online] Available from: doi:10.1128/mBio.00826-16.

Morikawa, H., Kim, M., Mimuro, H., Punginelli, C., *et al.* (2010) The bacterial effector Cif interferes with SCF ubiquitin ligase function by inhibiting deneddylation of Cullin1. *Biochemical and Biophysical Research Communications*. [Online] Available from: doi:10.1016/j.bbrc.2010.09.048.

Mortha, A., Chudnovskiy, A., Hashimoto, D., Bogunovic, M., *et al*. (2014) Microbiota-dependent crosstalk between macrophages and ILC3 promotes intestinal homeostasis. *Science*. [Online] Available from: doi:10.1126/science.1249288.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., *et al*. (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*. [Online] Available from: doi:10.1126/science.1128393.

Mukherjee, S., Zheng, H., Derebe, M.G., Callenberg, K.M., *et al.* (2014) Antibacterial membrane attack by a pore-forming intestinal C-type lectin. *Nature*. [Online] Available from: doi:10.1038/nature12729.

Mullineaux-Sanders, C., Collins, J.W., Ruano-Gallego, D., Levy, M., *et al*. (2017) *Citrobacter rodentium* relies on commensals for colonization of the colonic mucosa. *Cell Reports*. [Online] Available from: doi:10.1016/j.celrep.2017.11.086.

Mundy, R., Girard, F., Fitzgerald, A.J. & Frankel, G. (2006) Comparison of colonization dynamics and pathology of mice infected with enteropathogenic *Escherichia coli*, enterohaemorrhagic *E. coli* and *Citrobacter rodentium*. *FEMS Microbiology Letters*. [Online] Available from: doi:10.1111/j.1574-6968.2006.00481.x.

Mundy, R., MacDonald, T.T., Dougan, G., Frankel, G., et al. (2005) *Citrobacter rodentium* of mice and man. *Cellular Microbiology*. [Online]. Available from: doi:10.1111/j.1462-5822.2005.00625.x.

Mundy, R., Schüller, S., Girard, F., Fairbrother, J.M., et al. (2007) Functional studies of intimin *in vivo* and *ex vivo*: Implications for host specificity and tissue tropism. *Microbiology*. [Online] Available from: doi:10.1099/mic.0.2006/003467-0.

Muniesa, M., Hammerl, J.A., Hertwig, S., Appel, B., *et al.* (2012) Shiga toxin-producing *Escherichia coli* O104:H4: A new challenge for microbiology. *Applied and Environmental Microbiology*. [Online]. Available from: doi:10.1128/AEM.00217-12.

Nada, R.A., Shaheen, H.I., Khalil, S.B., Mansour, A., *et al*. (2011) Discovery and phylogenetic analysis of novel members of class b enterotoxigenic *Escherichia coli* adhesive fimbriae. *Journal of Clinical Microbiology*. [Online] Available from: doi:10.1128/JCM.02006-10.

Nadler, C., Baruch, K., Kobi, S., Mills, E., *et al.* (2010) The type III secretion effector NleE inhibits NF-kappaB activation. *PLoS Pathog*. [Online] Available from: doi:10.1371/journal.ppat.1000743.

Nagai, H. & Kubori, T. (2013) Purification and characterization of *Legionella* U-Box-Type E3 ubiquitin ligase. *Methods in Molecular Biology*. [Online] Available from: doi:10.1007/978-1-62703-161-5-21.

Nagai, T., Abe, A. & Sasakawa, C. (2005) Targeting of enteropathogenic *Escherichia coli* EspF to host mitochondria is essential for bacterial pathogenesis: Critical role of the 16th leucine residue in EspF. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M411550200.

Nagy, A., Xu, Y., Bauchan, G.R., Shelton, D.R., *et al*. (2016) Aggregative adherence fimbriae I (AAF/I) mediate colonization of fresh produce and abiotic surface by Shiga toxigenic enteroaggregative *Escherichia coli* O104:H4. *International Journal of Food Microbiology*. [Online] Available from: doi:10.1016/j.ijfoodmicro.2016.04.007.

Nakatani, Y., Yamazaki, M., Chazin, W.J. & Yui, S. (2005) Regulation of S100A8/A9 (calprotectin) binding to tumor cells by zinc ion and its implication for apoptosis-inducing activity. *Mediators of Inflammation*. [Online] Available from: doi:10.1155/MI.2005.280.

Nans, A., Kudryashev, M., Saibil, H.R. & Hayward, R.D. (2015) Structure of a bacterial type III secretion system in contact with a host membrane *in situ*. *Nature Communications*. [Online] Available from: doi:10.1038/ncomms10114.

Natale, P., Brüser, T. & Driessen, A.J.M. (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane - Distinct translocases and mechanisms. *Biochimica et Biophysica Acta - Biomembranes*. [Online]. Available from: doi:10.1016/j.bbamem.2007.07.015.

Nataro, J.P. & Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clinical microbiology reviews*.

Nataro, J.P., Kaper, J.B., Robins-Browne, R., Prado, V., *et al.* (1987) Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatric Infectious Disease Journal*. [Online] Available from: doi:10.1097/00006454-198709000-00008.

Nataro, J.P., Yikang, D., Cookson, S., Cravioto, A., *et al*. (1995) Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated. *Journal of Infectious Diseases*. [Online] Available from: doi:10.1093/infdis/171.2.465.

Nataro, J.P., Yikang, D., Yingkang, D. & Walker, K. (1994) AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli. Journal of Bacteriology*. [Online] Available from: doi:10.1128/jb.176.15.4691-4699.1994.

Le Negrate, G., Faustin, B., Welsh, K., Loeffler, M., *et al*. (2008a) *Salmonella* secreted factor L deubiquitinase of *Salmonella* Typhimurium inhibits NF- B, suppresses IκBα ubiquitination and modulates innate immune responses. *The Journal of Immunology*. [Online] Available from: doi:10.4049/jimmunol.180.7.5045.

Le Negrate, G., Krieg, A., Faustin, B., Loeffler, M., *et al*. (2008b) ChlaDub1 of *Chlamydia trachomatis* suppresses NF-κB activation and inhibits IκBα ubiquitination and degradation. *Cellular microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2008.01178.x.

Negrea, A., Bjur, E., Ygberg, S.E., Elofsson, M., *et al.* (2007) Salicylidene acylhydrazides that affect type III protein secretion in *Salmonella enterica* serovar typhimurium. *Antimicrobial Agents and Chemotherapy*. [Online] Available from: doi:10.1128/AAC.00223-07.

Nguyen, A.T., Prado, M.A., Schmidt, P.J., Sendamarai, A.K., *et al*. (2017) UBE2O remodels the proteome during terminal erythroid differentiation. *Science*. [Online] Available from: doi:10.1126/science.aan0218.

Nguyen, R.N., Taylor, L.S., Tauschek, M. & Robins-Browne, R.M. (2006) Atypical enteropathogenic *Escherichia coli* infection and prolonged diarrhea in children. *Emerging Infectious Diseases*. [Online] Available from: doi:10.3201/eid1204.051112.

Nguyen, Y. & Sperandio, V. (2012) Enterohemorrhagic *E. coli* (EHEC) pathogenesis. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2012.00090.

de Nisco, N.J., Rivera-Cancel, G. & Orth, K. (2018) The biochemistry of sensing: Enteric pathogens regulate type III secretion in response to environmental and host cues. *mBio*. [Online]. Available from: doi:10.1128/mBio.02122-17.

Njoroge, J.W., Gruber, C. & Sperandio, V. (2013) The interacting Cra and KdpE regulators are involved in the expression of multiple virulence factors in enterohemorrhagic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.02252-12.

Njoroge, J.W., Nguyen, Y., Curtis, M.M., Moreira, C.G., *et al*. (2012) Virulence meets metabolism: Cra and KdpE gene regulation in enterohemorrhagic *Escherichia coli*. *mBio*. [Online] Available from: doi:10.1128/mBio.00280-12.

Nordfelth, R., Kauppi, A.M., Norberg, H.A., Wolf-Watz, H., *et al.* (2005) Small-molecule inhibitors specifically targeting type III secretion. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.73.5.3104-3114.2005.

Nougayrède, J.P. & Donnenberg, M.S. (2004) Enteropathogenic *Escherichia coli* EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2004.00421.x.

O'Connell, C.B., Creasey, E.A., Knutton, S., Elliott, S., *et al.* (2004) SepL, a protein required for enteropathogenic *Escherichia coli* type III translocation, interacts with secretion component SepD. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2004.04101.x.

O'Connor, H.F. & Huibregtse, J.M. (2017) Enzyme–substrate relationships in the ubiquitin system: approaches

for identifying substrates of ubiquitin ligases. *Cellular and Molecular Life Sciences*. [Online]. Available from: doi:10.1007/s00018-017-2529-6.

O'Connor, H.F., Lyon, N., Leung, J.W., Agarwal, P., *et al*. (2015) Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO reports*. [Online] Available from: doi:10.15252/embr.201540620.

O'Ryan, M., Vidal, R., del Canto, F., Salazar, J.C., *et al*. (2015) Vaccines for viral and bacterial pathogens causing acute gastroenteritis: Part I: Overview, vaccines for enteric viruses and *Vibrio cholerae*. *Human vaccines & immunotherapeutics*. [Online] Available from: doi:10.1080/21645515.2015.1011019.

Ochoa, T.J., Barletta, F., Contreras, C. & Mercado, E. (2008) New insights into the epidemiology of enteropathogenic *Escherichia col*i infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. [Online]. Available from: doi:10.1016/j.trstmh.2008.03.017.

Ogino, T., Ohno, R., Sekiya, K., Kuwae, A., et al. (2006) Assembly of the type III secretion apparatus of enteropathogenic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.188.8.2801-2811.2006.

Ogura, Y., Ooka, T., Iguchi, A., Toh, H., et al. (2009) Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0903585106.

Okuda, J., Toyotome, T., Kataoka, N., Ohno, M., *et al.* (2005) *Shigella* effector IpaH9.8 binds to a splicing factor U2AF35to modulate host immune responses. *Biochemical and Biophysical Research Communications*. [Online] Available from: doi:10.1016/j.bbrc.2005.05.145.

Van Oss, S.B., Cucinotta, C.E. & Arndt, K.M. (2017) Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. *Trends in Biochemical Sciences*. [Online]. Available from: doi:10.1016/j.tibs.2017.08.003.

Pacheco, A.R., Munera, D., Waldor, M.K., Sperandio, V., *et al.* (2012) Fucose sensing regulates bacterial intestinal colonization. *Nature*. [Online]. Available from: doi:10.1038/nature11623.

Pallen, M.J., Beatson, S.A. & Bailey, C.M. (2005) Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: A Darwinian perpective. *FEMS Microbiology Reviews*. [Online]. Available from: doi:10.1016/j.femsre.2005.01.001.

Pallett, M.A., Berger, C.N., Pearson, J.S., Hartland, E.L., *et al*. (2014) The type III secretion effector NleF of enteropathogenic *Escherichia coli* activates NF-κB early during infection. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.02131-14.

Pastorino, J.G., Shulga, N. & Hoek, J.B. (2002) Mitochondrial binding of hexokinase II inhibits Bax-induced Cytochrome C release and apoptosis. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M109950200.

Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., *et al*. (2008) Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science*. [Online] Available from: doi:10.1126/science.1163885.

Pearson, J.S., Giogha, C., Ong, S.Y., Kennedy, C.L., *et al*. (2013) A type III effector antagonises death receptor signalling during bacterial gut infection. *Nature*. [Online] Available from: doi:10.1038/nature12524.

Peiffer, I., Servin, A.L. & Bernet-Camard, M.F. (1998) Piracy of decay-accelerating factor (CD55) signal transduction by the diffusely adhering strain *Escherichia coli* C1845 promotes cytoskeletal F- actin rearrangements in cultured human intestinal INT407 cells. *Infection and Immunity*.

Penno, C., Sansonetti, P. & Parsot, C. (2005) Frameshifting by transcriptional slippage is involved in production of MxiE, the transcription activator regulated by the activity of the type III secretion apparatus in *Shigella flexneri*. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2004.04530.x.

Peschiaroli, A., Dorrello, N.V., Guardavaccaro, D., Venere, M., *et al*. (2006) SCFβTrCP-mediated degradation of Claspin regulates recovery from the DNA replication checkpoint response. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2006.06.013.

Petty, N.K., Bulgin, R., Crepin, V.F., Cerdeño-Tárraga, A.M., *et al*. (2010) The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01144-09.

Philpott, D.J., Edgeworth, J.D. & Sansonetti, P.J. (2000) The pathogenesis of *Shigella flexneri* infection: Lessons from *in vitro* and *in vivo* studies. *Philosophical Transactions of the Royal Society B: Biological Sciences*. [Online] Available from: doi:10.1098/rstb.2000.0599.

Pickert, G., Neufert, C., Leppkes, M., Zheng, Y., *et al*. (2009) STAT3 links IL-22 signaling in intestinal epithelial cells to mucosal wound healing. *The Journal of Experimental Medicine*. [Online] Available from: doi:10.1084/jem.20082683.

Piscatelli, H., Kotkar, S.A., McBee, M.E., Muthupalani, S., *et al*. (2011) The EHEC type III effector NleL is an E3 ubiquitin ligase that modulates pedestal formation. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0019331.

Pollard, D.J., Berger, C.N., So, E.C., Yu, L., *et al*. (2018) Broad-spectrum regulation of nonreceptor tyrosine kinases by the bacterial ADP-ribosyltransferase EspJ. *mBio*. [Online] Available from: doi:10.1128/mBio.00170-18.

Popa, C., Coll, N.S., Valls, M. & Sessa, G. (2016) Yeast as a heterologous model system to uncover type III effector function. *PLoS Pathogens*. [Online]. Available from: doi:10.1371/journal.ppat.1005360.

Portaliou, A.G., Tsolis, K.C., Loos, M.S., Balabanidou, V., *et al*. (2017) Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic *Escherichia coli*. *The EMBO Journal*. [Online] Available from: doi:10.15252/embj.201797515.

Potter, A.A., Klashinsky, S., Li, Y., Frey, E., *et al.* (2004) Decreased shedding of *Escherichia coli* O157:H7 by cattle following vaccination with type III secreted proteins. *Vaccine*. [Online] Available from: doi:10.1016/j.vaccine.2003.08.007.

Pradel, N., Santini, C.L., Ye, C.Y., Fevat, L., *et al*. (2003) Influence of *tat* mutations on the ribose-binding protein translocation in *Escherichia coli*. *Biochemical and Biophysical Research Communications*. [Online] Available from: doi:10.1016/S0006-291X(03)01043-X.

Pruneda, J.N., Smith, F.D., Daurie, A., Swaney, D.L., *et al*. (2014) E2~Ub conjugates regulate the kinase activity of *Shigella* effector OspG during pathogenesis. *EMBO Journal*. [Online] Available from: doi:10.1002/embj.201386386.

Pym, A.S., Brodin, P., Majlessi, L., Brosch, R., *et al*. (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nature Medicine*. [Online] Available from: doi:10.1038/nm859.

Qadri, F., Svennerholm, A.-M., Faruque, A.S.G. & Sack, R.B. (2005) Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clinical microbiology reviews*. [Online] Available from: doi:10.1128/CMR.18.3.465-483.2005.

Qiu, J. & Luo, Z.Q. (2017) *Legionella* and *Coxiella* effectors: Strength in diversity and activity. *Nature Reviews Microbiology*. [Online]. Available from: doi:10.1038/nrmicro.2017.67.

Qiu, J., Sheedlo, M.J., Yu, K., Tan, Y., *et al*. (2016) Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature*. [Online] Available from: doi:10.1038/nature17657.

Quaile, A.T., Urbanus, M.L., Stogios, P.J., Nocek, B., *et al.* (2015) Molecular characterization of LubX: Functional divergence of the U-Box fold by *Legionella pneumophila*. *Structure*. [Online] Available from: doi:10.1016/j.str.2015.05.020.

Quezada, C.M., Hicks, S.W., Galan, J.E. & Stebbins, C.E. (2009) A family of *Salmonella* virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0811058106.

Radics, J., Königsmaier, L. & Marlovits, T.C. (2014) Structure of a pathogenic type 3 secretion system in action. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.2722.

Rajagopala, S. V., Sikorski, P., Caufield, J.H., Tovchigrechko, A., *et al*. (2012) Studying protein complexes by the yeast two-hybrid system. *Methods*. [Online] Available from: doi:10.1016/j.ymeth.2012.07.015.

Rakebrandt, N., Lentes, S., Neumann, H., James, L.C., *et al.* (2014) Antibody- and TRIM21-dependent intracellular restriction of *Salmonella enterica*. *Pathogens and Disease*. [Online] Available from: doi:10.1111/2049-632X.12192.

Rasko, D.A., Rosovitz, M.J., Myers, G.S.A., Mongodin, E.F., *et al*. (2008) The pangenome structure of *Escherichia coli* - Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *JOURNAL OF BACTERIOLOGY*. [Online] Available from: doi:10.1128/JB.00619-08.

Reichow, S.L., Korotkov, K. V., Hol, W.G.J. & Gonen, T. (2010) Structure of the cholera toxin secretion channel in its closed state. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.1910.

Remigi, P., Anisimova, M., Guidot, A., Genin, S., et al. (2011) Functional diversification of the GALA type III effector family contributes to *Ralstonia solanacearum* adaptation on different plant hosts. *New Phytologist*. [Online] Available from: doi:10.1111/j.1469-8137.2011.03854.x.

Rendon, M.A., Saldana, Z., Erdem, A.L., Monteiro-Neto, V., *et al*. (2007) Commensal and pathogenic *Escherichia coli* use a common pilus adherence factor for epithelial cell colonization. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0704104104.

Rhodes, D.A., Ihrke, G., Reinicke, A.T., Malcherek, G., *et al*. (2002) The 52 000 MW Ro/SS-A autoantigen in Sjögren's syndrome/systemic lupus erythematosus (Ro52) is an interferon-γ inducible tripartite motif protein associated with membrane proximal structures. *Immunology*. [Online] Available from: doi:10.1046/j.1365-2567.2002.01417.x.

Rhodes, D.A. & Isenberg, D.A. (2017) TRIM21 and the Function of Antibodies inside Cells. *Trends in Immunology*. [Online]. Available from: doi:10.1016/j.it.2017.07.005.

Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., *et al*. (1983) Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *The New England journal of medicine*. [Online] Available from: doi:10.1056/NEJM198303243081203.

Ritchie, J.M. & Waldor, M.K. (2005) The locus of enterocyte effacement-encoded effector proteins all promote enterohemorrhagic *Escherichia coli* pathogenicity in infant rabbits. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.73.3.1466-1474.2005.

Roberts, D.J. & Miyamoto, S. (2015) Hexokinase II integrates energy metabolism and cellular protection: Akting on mitochondria and TORCing to autophagy. *Cell Death and Differentiation*. [Online]. Available from: doi:10.1038/cdd.2014.173.

Robins-Browne, R.M., Holt, K.E., Ingle, D.J., Hocking, D.M., *et al*. (2016) Are *Escherichia coli* pathotypes still relevant in the era of whole-genome sequencing? *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2016.00141.

Roghi, C., Jones, L., Gratian, M., English, W.R., *et al.* (2010) Golgi reassembly stacking protein 55 interacts with membrane-type (MT) 1-matrix metalloprotease (MMP) and furin and plays a role in the activation of the MT1-MMP zymogen. *FEBS Journal*. [Online] Available from: doi:10.1111/j.1742-4658.2010.07723.x.

Rohde, J.R., Breitkreutz, A., Chenal, A., Sansonetti, P.J., *et al*. (2007) Type III secretion effectors of the IpaH family are E3 ubiquitin ligases. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2007.02.002.

Rolando, M., Sanulli, S., Rusniok, C., Gomez-Valero, L., *et al*. (2013) *Legionella pneumophila* effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2013.03.004.

Rolhion, N., Hofman, P. & Darfeuille-Michaud, A. (2011) The endoplasmic reticulum stress response chaperone: Gp96, a host receptor for Crohn disease-associated adherent-invasive *Escherichia coli*. *Gut microbes*. [Online]. Available from: doi:10.4161/gmic.2.2.15725.

Romo-Castillo, M., Andrade, A., Espinosa, N., Feria, J.M., *et al.* (2014) EscO, a functional and structural analog of the flagellar FliJ protein, is a positive regulator of EscN ATPase activity of the enteropathogenic *Escherichia coli* injectisome. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01551-14.

Rosebrock, T.R., Zeng, L., Brady, J.J., Abramovitch, R.B., *et al*. (2007) A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature*. [Online] Available from: doi:10.1038/nature05966.

Rosenbaum, J.C., Fredrickson, E.K., Oeser, M.L., Garrett-Engele, C.M., *et al*. (2011) Disorder targets misorder in nuclear quality control degradation: A disordered ubiquitin ligase directly recognizes its misfolded substrates. *Molecular Cell*. [Online] Available from: doi:10.1016/j.molcel.2010.12.004.

Rossi, M., Rotblat, B., Ansell, K., Amelio, I., *et al*. (2014) High throughput screening for inhibitors of the HECT ubiquitin E3 ligase ITCH identifies antidepressant drugs as regulators of autophagy. *Cell Death and Disease*. [Online] Available from: doi:10.1038/cddis.2014.113.

Rossier, O., Dao, J. & Cianciotto, N.P. (2009) A type II secreted RNase of *Legionella pneumophila* facilitates optimal intracellular infection of *Hartmannella vermiformis*. *Microbiology*. [Online] Available from: doi:10.1099/mic.0.023218-0.

Rotin, D. & Kumar, S. (2009) Physiological functions of the HECT family of ubiquitin ligases. *Nature Reviews Molecular Cell Biology*. [Online]. Available from: doi:10.1038/nrm2690.

Roxas, J.L., Wilbur, J.S., Zhang, X., Martinez, G., *et al*. (2012) The enteropathogenic *Escherichia coli*-secreted protein EspZ inhibits host cell apoptosis. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00335-12.

Ruano-Gallego, D., Álvarez, B. & Fernández, L.Á. (2015) Engineering the controlled assembly of filamentous injectisomes in *E. coli* K-12 for protein translocation into mammalian cells. *ACS Synthetic Biology*. [Online] Available from: doi:10.1021/acssynbio.5b00080.

Rubino, S.J., Geddes, K. & Girardin, S.E. (2012) Innate IL-17 and IL-22 responses to enteric bacterial pathogens. *Trends in Immunology*. [Online]. Available from: doi:10.1016/j.it.2012.01.003.

Rubino, S.J., Geddes, K., Magalhaes, J.G., Streutker, C., *et al.* (2013) Constitutive induction of intestinal Tc17 cells in the absence of hematopoietic cell-specific MHC class II expression. *European Journal of Immunology*. [Online] Available from: doi:10.1002/eji.201243028.

Rubinstein, D.B., Stortchevoi, A., Boosalis, M., Ashfaq, R., *et al*. (2004) Receptor for the globular heads of C1q (gC1q-R, p33, hyaluronan-binding protein) is preferentially expressed by adenocarcinoma cells. *International Journal of Cancer*. [Online] Available from: doi:10.1002/ijc.20105.

Russell, A.B., Peterson, S.B. & Mougous, J.D. (2014) Type VI secretion system effectors: Poisons with a purpose. *Nature Reviews Microbiology*. [Online]. Available from: doi:10.1038/nrmicro3185.

Ryckman, C., Vandal, K., Rouleau, P., Talbot, M., *et al.* (2003) Proinflammatory activities of S100: Proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *The Journal of Immunology*. [Online] Available from: doi:10.4049/jimmunol.170.6.3233.

Rytkonen, A., Poh, J., Garmendia, J., Boyle, C., *et al*. (2007) SseL, a *Salmonella* deubiquitinase required for macrophage killing and virulence. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0610095104.

Sakkejha, H., Byrne, L., Lawson, A.J. & Jenkins, C. (2013) An update on the microbiology and epidemiology of enteropathogenic *Escherichia coli* in England 2010-2012. *Journal of Medical Microbiology*. [Online] Available from: doi:10.1099/jmm.0.062380-0.

Sal-Man, N., Deng, W. & Finlay, B.B. (2012) EscI: a crucial component of the type III secretion system forms the inner rod structure in enteropathogenic *Escherichia coli*. *Biochemical Journal*. [Online] Available from: doi:10.1042/BJ20111620.

Samba-Louaka, A., Nougayrède, J.P., Watrin, C., Oswald, E., *et al*. (2009) The enteropathogenic *Escherichia coli* effector Cif induces delayed apoptosis in epithelial cells. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00860-09.

Sanada, T., Kim, M., Mimuro, H., Suzuki, M., et al. (2012) The *Shigella flexneri* effector OspI deamidates UBC13 to dampen the inflammatory response. *Nature*. [Online] Available from: doi:10.1038/nature10894.

Sasaki, K., Gotoh, K., Miake, S., Setoyama, D., *et al*. (2017) p32 is required for appropriate interleukin-6 production upon LPS stimulation and protects mice from endotoxin shock. *EBioMedicine*. [Online] Available from: doi:10.1016/j.ebiom.2017.05.018.

Sason, H., Milgrom, M., Weiss, A.M., Melamed-Book, N., *et al.* (2009) Enteropathogenic *Escherichia coli* subverts phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 3,4,5-trisphosphate upon epithelial cell infection. *Molecular Biology of the Cell*. [Online] Available from: doi:10.1091/mbc.E08-05-0516.

Scaletsky, I.C.A., Silva, M.L.M. & Trabulsi, L.R. (1984) Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. *Infection and Immunity*.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., *et al*. (2012) Fiji: An open-source platform for biological-image analysis. *Nature Methods*. [Online]. Available from: doi:10.1038/nmeth.2019.

Schmidt, A., Forne, I. & Imhof, A. (2014) Bioinformatic analysis of proteomics data. *BMC systems biology*. [Online]. Available from: doi:10.1186/1752-0509-8-S2-S3.

Schraidt, O. & Marlovits, T.C. (2011) Three-dimensional model of *Salmonella*'s needle complex at

subnanometer resolution. *Science*. [Online] Available from: doi:10.1126/science.1199358.

Schüller, S., Chong, Y., Lewin, J., Kenny, B., *et al*. (2007) Tir phosphorylation and Nck/N-WASP recruitment by enteropathogenic and enterohaemorrhagic *Escherichia coli* during *ex vivo* colonization of human intestinal mucosa is different to cell culture models. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2006.00879.x.

Scott, N.E., Giogha, C., Pollock, G.L., Kennedy, C.L., *et al*. (2017) The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M117.805036.

Scully, O.J., Yu, Y., Salim, A., Thike, A.A., *et al*. (2015) Complement component 1, q subcomponent binding protein is a marker for proliferation in breast cancer. *Experimental biology and medicine (Maywood, N.J.)*. [Online] Available from: doi:10.1177/1535370214565075.

Segal, G., Feldman, M. & Zusman, T. (2005) The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*. *FEMS Microbiology Reviews*. [Online]. Available from: doi:10.1016/j.femsre.2004.07.001.

Sekiya, K., Ohishi, M., Ogino, T., Tamano, K., *et al.* (2001) Supermolecular structure of the enteropathogenic *Escherichia coli* type III secretion system and its direct interaction with the EspA-sheath-like structure. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.191378598.

Selyunin, A.S., Reddick, L.E., Weigele, B.A. & Alto, N.M. (2014) Selective protection of an ARF1-GTP signaling axis by a bacterial scaffold induces bidirectional trafficking arrest. *Cell Reports*. [Online] Available from: doi:10.1016/j.celrep.2014.01.040.

Selyunin, A.S., Sutton, S.E., Weigele, B.A., Reddick, L.E., *et al.* (2011) The assembly of a GTPase-kinase signalling complex by a bacterial catalytic scaffold. *Nature*. [Online] Available from: doi:10.1038/nature09593.

Servin, A.L. (2014) Pathogenesis of human diffusely adhering *Escherichia coli* expressing Afa/Dr adhesins (Afa/Dr DAEC): Current insights and future challenges. *Clinical Microbiology Reviews*. [Online] Available from: doi:10.1128/CMR.00036-14.

Seyedarabi, A. (2010) The orchestrators of *Shigella flexneri* infection. In: *Crystallography Reviews*. [Online]. 2010 p. Available from: doi:10.1080/08893110903563732.

Seyedarabi, A., Sullivan, J.A., Sasakawa, C. & Pickersgill, R.W. (2010) A disulfide driven domain swap switches off the activity of *Shigella* IpaH9.8 E3 ligase. *FEBS Letters*. [Online] Available from: doi:10.1016/j.febslet.2010.09.006.

Shah, S. & Briken, V. (2016) Modular Organization of the ESX-5 Secretion System in Mycobacterium tuberculosis. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2016.00049.

Shanks, J., Burtnick, M.N., Brett, P.J., Waag, D.M., *et al*. (2009) *Burkholderia mallei tssM* encodes a putative deubiquitinase that is secreted and expressed inside infected RAW 264.7 murine macrophagest. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.01339-08.

Shaw, R.K., Smollett, K., Cleary, J., Garmendia, J., *et al*. (2005) Enteropathogenic *Escherichia coli* type III effectors EspG and EspG2 disrupt the microtubule network of intestinal epithelial cells. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.73.7.4385-4390.2005.

Sheikh, J., Hicks, S., Dall'Agnol, M., Phillips, A.D., *et al.* (2001) Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-

2958.2001.02512.x.

Shema, E., Tirosh, I., Aylon, Y., Huang, J., *et al*. (2008) The histone H2B-specific ubiquitin ligase RNF20 / hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes and Development*. [Online] Available from: doi:10.1101/gad.1703008.as.

Sheng, X., You, Q., Zhu, H., Chang, Z., *et al*. (2017) Bacterial effector NleL promotes enterohemorrhagic *E. coli*-induced attaching and effacing lesions by ubiquitylating and inactivating JNK. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1006534.

Shenolikar, S., Voltz, J.W., Cunningham, R. & Weinman, E.J. (2004) Regulation of ion transport by the NHERF family of PDZ proteins. *Physiology (Bethesda, Md.)*. [Online] Available from: doi:10.1152/physiol.00020.2004.

Shi, H., Fang, W., Liu, M. & Fu, D. (2017) Complement component 1, q subcomponent binding protein (C1QBP) in lipid rafts mediates hepatic metastasis of pancreatic cancer by regulating IGF-1/IGF-1R signaling. *International Journal of Cancer*. [Online] Available from: doi:10.1002/ijc.30831.

Shin, M. (2017) The mechanism underlying Ler-mediated alleviation of gene repression by H-NS. *Biochemical and Biophysical Research Communications*. [Online] Available from: doi:10.1016/j.bbrc.2016.12.132.

Shiomi, H., Masuda, A., Nishiumi, S., Nishida, M., *et al*. (2010) Gamma interferon produced by antigen-specific CD4+ T cells regulates the mucosal immune responses to *Citrobacter rodentium* infection. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.01343-09.

Shorter, J., Watson, R., Giannakou, M.E., Clarke, M., *et al*. (1999) GRASP55, a second mammalian GRASP protein involved in the stacking of Golgi cisternae in a cell-free system. *EMBO Journal*. [Online] Available from: doi:10.1093/emboj/18.18.4949.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*. [Online] Available from: doi:10.1038/msb.2011.75.

Silverman, J.M., Brunet, Y.R., Cascales, E. & Mougous, J.D. (2012) Structure and regulation of the type VI secretion system. *Annu Rev Microbiol*. [Online] Available from: doi:10.1146/annurev-micro-121809-151619.

Simard, J.C., Girard, D. & Tessier, P.A. (2010) Induction of neutrophil degranulation by S100A9 via a MAPK-dependent mechanism. *Journal of Leukocyte Biology*. [Online] Available from: doi:10.1189/jlb.1009676.

Simmons, C.P., Clare, S., Ghaem-Maghami, M., Uren, T.K., *et al*. (2003) Central role for B lymphocytes and CD4+ T cells in immunity to infection by the attaching and effacing pathogen *Citrobacter rodentium*. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.71.9.5077-5086.2003.

Simovitch, M., Sason, H., Cohen, S., Zahavi, E.E., *et al.* (2010) EspM inhibits pedestal formation by enterohaemorrhagic *Escherichia coli* and enteropathogenic *E. coli* and disrupts the architecture of a polarized epithelial monolayer. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2009.01410.x.

Simpson, N., Shaw, R., Crepin, V.F., Mundy, R., *et al*. (2006) The enteropathogenic *Escherichia coli* type III secretion system effector Map binds EBP50/NHERF1: Implication for cell signalling and diarrhoea. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2006.05109.x.

Singer, A.U., Schulze, S., Skarina, T., Xu, X., et al. (2013) A pathogen type III effector with a novel E3 ubiquitin ligase architecture. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1003121.

Slater, S.L., Sågfors, A.M., Pollard, D.J., Ruano-Gallego, D., *et al*. (2018) *The Type III Secretion System of Pathogenic Escherichia coli*. In: [Online]. Springer, Berlin, Heidelberg. pp. 1–22. Available from: doi:10.1007/82_2018_116 [Accessed: 17 September 2018].

Smit, J.J. & Sixma, T.K. (2014) RBR E3-ligases at work. *EMBO Reports*. [Online]. Available from: doi:10.1002/embr.201338166.

Smith, K., Humphreys, D., Hume, P.J. & Koronakis, V. (2010) Enteropathogenic *Escherichia coli* Recruits the Cellular Inositol Phosphatase SHIP2 to Regulate Actin-Pedestal Formation. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2009.12.004.

Smits, A.H. & Vermeulen, M. (2016) Characterizing Protein–Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. *Trends in Biotechnology*. [Online]. Available from: doi:10.1016/j.tibtech.2016.02.014.

Snyder, J.D. & Merson, M.H. (1982) The magnitude of the global problem of acute diarrhoeal disease: A review of active surveillance data. *Bulletin of the World Health Organization*.

Song, T., Li, K., Zhou, W., Zhou, J., et al. (2017) A type III effector NleF from EHEC inhibits epithelial inflammatory cell death by targeting Caspase-4. *BioMed Research International*. [Online] Available from: doi:10.1155/2017/4101745.

Soni, M. & Mondal, K.K. (2018) *Xanthomonas axonopodis pv. punicae* uses XopL effector to suppress pomegranate immunity. *Journal of Integrative Plant Biology*. [Online] Available from: doi:10.1111/jipb.12615.

Spratt, D.E., Walden, H. & Shaw, G.S. (2014) RBR E3 ubiquitin ligases: new structures, new insights, new questions. *The Biochemical journal*. [Online] Available from: doi:10.1042/BJ20140006.

Stebbins, C.E. & Galán, J.E. (2001) Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature*. [Online] Available from: doi:10.1038/35102073.

Studier, F.W. & Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*. [Online] Available from: doi:10.1016/0022-2836(86)90385-2.

Sugimoto, K., Ogawa, A., Mizoguchi, E., Shimomura, Y., *et al.* (2008) IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis. *The Journal of clinical investigation*. [Online] Available from: doi:10.1172/JCI33194.

Suzuki, S., Mimuro, H., Kim, M., Ogawa, M., *et al*. (2014) *Shigella* IpaH7.8 E3 ubiquitin ligase targets glomulin and activates inflammasomes to demolish macrophages. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1324021111.

Tachikawa, T., Seo, G., Nakazawa, M., Sueyoshi, M., *et al*. (1998) Estimation of probiotics by infection model of infant rabbit with enterohemorrhagic *Escherichia coli* O157:H7. *Kansenshogaku zasshi. The Journal of the Japanese Association for Infectious Diseases*. [Online] Available from: doi:10.11150/kansenshogakuzasshi1970.72.1300.

El Tahir, Y. & Skurnik, M. (2001) YadA, the multifaceted *Yersinia* adhesin. *International Journal of Medical Microbiology*. [Online]. Available from: doi:10.1078/1438-4221-00119.

Taieb, F., Nougayrède, J.P., Watrin, C., Samba-Louaka, A., *et al*. (2006) *Escherichia coli* cyclomodulin Cif induces G2arrest of the host cell cycle without activation of the DNA-damage checkpoint-signalling pathway. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2006.00757.x.

Takagi, K., Kim, M., Sasakawa, C. & Mizushima, T. (2016) Crystal structure of the substrate-recognition domain of the *Shigella* E3 ligase IpaH9.8. *Acta Crystallographica Section:F Structural Biology Communications*. [Online] Available from: doi:10.1107/S2053230X16002715.

Tan, K.S., Chen, Y., Lim, Y.-C., Tan, G.-Y.G., *et al.* (2010) Suppression of host innate immune response by *Burkholderia pseudomallei* through the virulence factor TssM. *Journal of immunology (Baltimore, Md. : 1950)*. [Online] Available from: doi:10.4049/jimmunol.0902663.

Tarabykina, S., Griffiths, T.R.L., Tulchinsky, E., Mellon, J.K., *et al*. (2007) Metastasis-associated protein S100A4: spotlight on its role in cell migration. *Current cancer drug targets*. [Online] Available from: doi:10.2174/156800907780618329.

Thomas, J., Stafford, G.P. & Hughes, C. (2004) Docking of cytosolic chaperone-substrate complexes at the membrane ATPase during flagellar type III protein export. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0307223101.

Thomas, N.A., Deng, W., Puente, J.L., Frey, E.A., *et al.* (2005) CesT is a multi-effector chaperone and recruitment factor required for the efficient type III secretion of both LEE- and non-LEE-encoded effectors of enteropathogenic *Escherichia coli*. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2005.04802.x.

Thomas, N.A., Ma, I., Prasad, M.E. & Rafuse, C. (2012) Expanded roles for multicargo and class 1B effector chaperones in type III secretion. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.00406-12.

Thomas, S., Holland, I.B. & Schmitt, L. (2014) The Type 1 secretion pathway - The hemolysin system and beyond. *Biochimica et Biophysica Acta - Molecular Cell Research*. [Online]. Available from: doi:10.1016/j.bbamcr.2013.09.017.

Thurston, T.L.M., Wandel, M.P., Von Muhlinen, N., Foeglein, Á., *et al*. (2012) Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion. *Nature*. [Online] Available from: doi:10.1038/nature10744.

Tobe, T., Beatson, S.A., Taniguchi, H., Abe, H., *et al*. (2006) An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.0604891103.

Tokunaga, F. & Iwai, K. (2012) LUBAC, a novel ubiquitin ligase for linear ubiquitination, is crucial for inflammation and immune responses. *Microbes and Infection*. [Online]. Available from: doi:10.1016/j.micinf.2012.01.011.

De Torres, M., Mansfield, J.W., Grabov, N., Brown, I.R., *et al.* (2006) *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in Arabidopsis. *Plant Journal*. [Online] Available from: doi:10.1111/j.1365-313X.2006.02798.x.

Tozzoli, R., Grande, L., Michelacci, V., Ranieri, P., *et al*. (2014) Shiga toxin-converting phages and the emergence of new pathogenic *Escherichia coli*: a world in motion. *Frontiers in Cellular and Infection Microbiology*. [Online] Available from: doi:10.3389/fcimb.2014.00080.

Trabulsi, L., Keller, R. & Tardelli Gomes, T. (2002) Typical and atypical enteropathogenic *Escherichia coli*. *Emerging Infectious Diseases*. [Online] Available from: doi:10.3201/eid0805.010385.

Trachtman, H. (2013) HUS and TTP in children. *Pediatric Clinics of North America*. [Online]. Available from: doi:10.1016/j.pcl.2013.08.007.

Tree, J.J., Wang, D., McInally, C., Mahajan, A., *et al*. (2009) Characterization of the effects of salicylidene acylhydrazide compounds on type III secretion in *Escherichia coli* O157:H7. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00562-09.

Tsolis, R.M., Townsend, S.M., Miao, E.A., Miller, S.I., *et al*. (1999) Identification of a putative *Salmonella enterica* serotype Typhimurium host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. *Infection and Immunity*.

Tu, X., Nisan, I., Yona, C., Hanski, E., *et al*. (2003) EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic *Escherichia coli*. *Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.2003.03329.x.

Tumanov, A. V., Koroleva, E.P., Guo, X., Wang, Y., *et al*. (2011) Lymphotoxin controls the IL-22 protection pathway in gut innate lymphoid cells during mucosal pathogen challenge. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2011.06.002.

Tye, A.J., Ghebrehiwet, B., Guo, N., Sastry, K.N., *et al.* (2001) The human gC1qR/p32 gene, C1qBP. Genomic organization and promoter analysis. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M009064200.

Tzipori, S., Montanaro, J., Robins-Browne, R.M., Vial, P., *et al*. (1992) Studies with enteroaggregative *Escherichia coli* in the gnotobiotic piglet gastroenteritis model. *Infection and Immunity*.

Tzipori, S., Robins-Browne, R.M., Gonis, G., Hayes, J., *et al.* (1985) Enteropathogenic *Escherichia col*i enteritis: Evaluation of the gnotobiotic piglet as a model of human infection. *Gut*. [Online] Available from: doi:10.1136/gut.26.6.570.

Tzipori, S., Wachsmuth, I.K., Chapman, C., Birden, R., *et al.* (1986) The pathogenesis of hemorrhagic colitis caused by *Escherichia coli* O157:H7 in gnotobiotic piglets. *The Journal of infectious diseases*. [Online] 154 (4), 712–716. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3528323 [Accessed: 17 September 2018].

Umanski, T., Rosenshine, I. & Friedberg, D. (2002) Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. *Microbiology (Reading, England)*. [Online] Available from: doi:10.1099/00221287-148-9-2735.

Vallance, B.A., Deng, W., Knodler, L.A. & Finlay, B.B. (2002) Mice lacking T and B lymphocytes develop transient colitis and crypt hyperplasia yet suffer impaired bacterial clearance during *Citrobacter rodentium* infection. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.70.4.2070-2081.2002.

Valleau, D., Little, D.J., Borek, D., Skarina, T., *et al.* (2018) Functional diversification of the NleG effector family in enterohemorrhagic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. [Online] 201718350. Available from: doi:10.1073/pnas.1718350115 [Accessed: 18 September 2018].

Vasseur, P., Vallet-Gely, I., Soscia, C., Genin, S., *et al*. (2005) The *pel* genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. *Microbiology*. [Online] Available from: doi:10.1099/mic.0.27410-0.

Veenendaal, A.K.J., Sundin, C. & Blocker, A.J. (2009) Small-molecule type III secretion system inhibitors block assembly of the *Shigella* type III secreton. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01004-08.

Velásquez, J.C., Hidalgo, A.A., Villagra, N., Santiviago, C.A., *et al*. (2016) SPI-9 of *Salmonella enterica* serovar Typhi is constituted by an operon positively regulated by RpoS and contributes to adherence to epithelial

cells in culture. *Microbiology (United Kingdom)*. [Online] Available from: doi:10.1099/mic.0.000319.

Vermeulen, M. & Timmers, H.M. (2010) Grasping trimethylation of histone H3 at lysine 4. *Epigenomics*. [Online] Available from: doi:10.2217/epi.10.11.

Viala, J.P.M., Mochegova, S.N., Meyer-Morse, N. & Portnoy, D.A. (2008) A bacterial pore-forming toxin forms aggregates in cells that resemble those associated with neurodegenerative diseases. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2007.01100.x.

Vijay-kumar, S., Bugg, C.E. & Cook, W.J. (1987) Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology*. [Online] Available from: doi:10.1016/0022-2836(87)90679-6.

Vila, I.K., Yao, Y., Kim, G., Xia, W., *et al.* (2017) A UBE2O-AMPKα2 Axis that Promotes Tumor Initiation and Progression Offers Opportunities for Therapy. *Cancer Cell*. [Online] Available from: doi:10.1016/j.ccell.2017.01.003.

Viljanen, M.K., Peltola, T., Kuistila, M., Huovinen, P., *et al.* (1990) Outbreak of diarrhoea due to *Escherichia coli* O111:B4 in schoolchildren and adults: association of Vi antigen-like reactivity. *The Lancet*. [Online] Available from: doi:10.1016/0140-6736(90)92337-H.

Wachter, C., Beinke, C., Mattes, M. & Schmidt, M.A. (1999) Insertion of EspD into epithelial target cell membranes by infecting enteropathogenic *Escherichia coli. Molecular Microbiology*. [Online] Available from: doi:10.1046/j.1365-2958.1999.01303.x.

Wallden, K., Rivera-Calzada, A. & Waksman, G. (2010) Type IV secretion systems: Versatility and diversity in function. *Cellular Microbiology*. [Online]. Available from: doi:10.1111/j.1462-5822.2010.01499.x.

Walterspiel, J.N., Morrow, A.L., Cleary, T.G. & Ashkenazi, S. (1992) Effect of subinhibitory concentrations of antibiotics on extracellular shiga-like toxin I. *Infection*. [Online] Available from: doi:10.1007/BF01704889.

Wang, D., Roe, A.J., McAteer, S., Shipston, M.J., *et al*. (2008) Hierarchal type III secretion of translocators and effectors from *Escherichia coli* O157:H7 requires the carboxy terminus of SepL that binds to Tir. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2008.06377.x.

Wang, D., Zetterström, C.E., Gabrielsen, M., Beckham, K.S.H., *et al.* (2011) Identification of bacterial target proteins for the salicylidene acylhydrazide Class of virulence-blocking compounds. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M111.233858.

Wang, F., Jiang, Z., Li, Y., He, X., *et al.* (2013) *Shigella flexneri* T3SS effector IpaH4.5 modulates the host inflammatory response via interaction with NF-κB p65 protein. *Cellular Microbiology*. [Online] Available from: doi:10.1111/cmi.12052.

Wang, J., Song, Y., Liu, T., Shi, Q., *et al*. (2015) Elevated expression of HABP1 is a novel prognostic indicator in triple-negative breast cancers. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. [Online] Available from: doi:10.1007/s13277-015-3131-x.

Wang, X., Tu, F., Zhu, Y. & Gao, G. (2012) Zinc-finger antiviral protein inhibits XMRV infection. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0039159.

Wang, Y., Koroleva, E.P., Kruglov, A.A., Kuprash, D. V., *et al.* (2010) Lymphotoxin beta receptor signaling in intestinal epithelial cells orchestrates innate immune responses against mucosal bacterial infection. *Immunity*. [Online] Available from: doi:10.1016/j.immuni.2010.02.011.

Wang, Y., Yang, Y., Wu, S., Pan, S., *et al*. (2014) P32 is a novel target for viral protein icp34.5 of herpes simplex virus type 1 and facilitates viral nuclear egress. *Journal of Biological Chemistry*. [Online] Available from: doi:10.1074/jbc.M114.603845.

Watkinson, R.E., McEwan, W.A., Tam, J.C.H., Vaysburd, M., *et al.* (2015) TRIM21 promotes cGAS and RIG-I sensing of viral genomes during Infection by antibody-opsonized virus. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1005253.

Wei, C., Wang, Y., Du, Z., Guan, K., *et al.* (2016) The *Yersinia* Type III secretion effector YopM Is an E3 ubiquitin ligase that induced necrotic cell death by targeting NLRP3. *Cell death & disease*. [Online] Available from: doi:10.1038/cddis.2016.413.

Welch, R.A., Burland, V., Plunkett, G., Redford, P., *et al*. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. [Online] Available from: doi:10.1073/pnas.252529799.

Welinder-Olsson, C. & Kaijser, B. (2005) Enterohemorrhagic *Escherichia coli* (EHEC). *Scandinavian Journal of Infectious Diseases*. [Online]. Available from: doi:10.1080/00365540510038523.

Wells, T.J., Sherlock, O., Rivas, L., Mahajan, A., *et al.* (2008) EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environmental Microbiology*. [Online] Available from: doi:10.1111/j.1462-2920.2007.01479.x.

West, R.W., Yocum, R.R. & Ptashne, M. (1984) *Saccharomyces cerevisiae* GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASG. *Molecular and cellular biology*. [Online] Available from: doi:10.1128/MCB.4.11.2467.Updated.

Whale, A.D., Hernandes, R.T., Ooka, T., Beutin, L., *et al*. (2007) TccP2-mediated subversion of actin dynamics by EPEC 2 - A distinct evolutionary lineage of enteropathogenic *Escherichia coli. Microbiology*. [Online] Available from: doi:10.1099/mic.0.2006/004325-0.

Wickham, M.E., Lupp, C., Vázquez, A., Mascarenhas, M., *et al.* (2007) *Citrobacter rodentium* virulence in mice associates with bacterial load and the type III effector NleE. *Microbes and Infection*. [Online] Available from: doi:10.1016/j.micinf.2006.12.016.

Wilbur, J.S., Byrd, W., Ramamurthy, S., Ledvina, H.E., *et al*. (2015) The secreted effector protein EspZ is essential for virulence of rabbit enteropathogenic *Escherichia coli*. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.02876-14.

Wiles, S., Clare, S., Harker, J., Huett, A., *et al.* (2004) Organ specificity, colonization and clearance dynamics *in vivo* following oral challenges with the murine pathogen *Citrobacter rodentium*. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2004.00414.x.

Wiles, S., Dougan, G. & Frankel, G. (2005) Emergence of a 'hyperinfectious' bacterial state after passage of *Citrobacter rodentium* through the host gastrointestinal tract. *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2005.00544.x.

Williams, S.G., Varcoe, L.T., Attridge, S.R. & Manning, P.A. (1996) *Vibrio cholerae* Hcp, a secreted protein coregulated with HlyA. *Infection and Immunity*.

Wlodarska, M., Thaiss, C.A., Nowarski, R., Henao-Mejia, J., *et al*. (2014) NLRP6 inflammasome orchestrates the colonic host-microbial interface by regulating goblet cell mucus secretion. *Cell*. [Online] Available from: doi:10.1016/j.cell.2014.01.026.

Wolf, M.K. (1997) Occurrence, distribution, and associations of O and H serogroups, colonization factor antigens, and toxins of enterotoxigenic *Escherichia coli*. *Clinical Microbiology Reviews*.

Wong, A.R.C., Clements, A., Raymond, B., Crepin, V.F., *et al*. (2012a) The interplay between the *Escherichia coli* Rho guanine nucleotide exchange factor effectors and the mammalian RHOGEF inhibitor EspH. *mBio*.

[Online] Available from: doi:10.1128/mBio.00250-11.

Wong, A.R.C., Pearson, J.S., Bright, M.D., Munera, D., *et al.* (2011) Enteropathogenic and enterohaemorrhagic *Escherichia coli*: Even more subversive elements. *Molecular Microbiology*. [Online]. Available from: doi:10.1111/j.1365-2958.2011.07661.x.

Wong, A.R.C., Raymond, B., Collins, J.W., Crepin, V.F., *et al.* (2012b) The enteropathogenic *E. coli* effector EspH promotes actin pedestal formation and elongation via WASP-interacting protein (WIP). *Cellular Microbiology*. [Online] Available from: doi:10.1111/j.1462-5822.2012.01778.x.

Wong, C.S., Jelacic, S., Habeeb, R.L., Watkins, S.L., *et al.* (2000) The risk of the hemolytic-uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. *The New England journal of medicine*. [Online] Available from: doi:10.1056/NEJM200006293422601.

Wood, M.W., Jones, M.A., Watson, P.R., Siber, A.M., *et al.* (2000) The secreted effector protein of *Salmonella* dublin, SopA, is translocated into eukaryotic cells and influences the induction of enteritis. *Cellular Microbiology*. [Online] Available from: doi:10.1046/j.1462-5822.2000.00054.x.

Wood, S.E., Jin, J. & Lloyd, S.A. (2008) YscP and YscU switch the substrate specificity of the *Yersinia* type III secretion system by regulating export of the inner rod protein YscI. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.00328-08.

Wu, B., Skarina, T., Yee, A., Jobin, M.C., *et al*. (2010) NleG type 3 effectors from enterohaemorrhagic *Escherichia coli* are U-Box E3 ubiquitin ligases. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1000960.

Xiang, Y. & Wang, Y. (2010) GRASP55 and GRASP65 play complementary and essential roles in Golgi cisternal stacking. *Journal of Cell Biology*. [Online] Available from: doi:10.1083/jcb.200907132.

Xicohtencatl-Cortes, J., Monteiro-Neto, V., Saldaña, Z., Ledesma, M.A., *et al.* (2009) The type 4 pili of enterohemorrhagic *Escherichia coli* O157:H7 are multipurpose structures with pathogenic attributes. *Journal of Bacteriology*. [Online] Available from: doi:10.1128/JB.01306-08.

Xin, D.W., Liao, S., Xie, Z.P., Hann, D.R., *et al.* (2012) Functional analysis of NopM, a novel E3 ubiquitin ligase (NEL) domain effector of *Rhizobium* sp. strain NGR234. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1002707.

Xu, C. & Min, J. (2011) Structure and function of WD40 domain proteins. *Protein and Cell*. [Online]. Available from: doi:10.1007/s13238-011-1018-1.

Xu, L., Xiao, N., Liu, F., Ren, H., *et al*. (2009) Inhibition of RIG-I and MDA5-dependent antiviral response by gC1qR at mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*. [Online] Available from: doi:10.1073/pnas.0811029106.

Yanagitani, K., Juszkiewicz, S. & Hegde, R.S. (2017) UBE2O is a quality control factor for orphans of multiprotein complexes. *Science*. [Online] Available from: doi:10.1126/science.aan0178.

Yao, Q., Cui, J., Wang, J., Li, T., *et al.* (2012) Structural mechanism of ubiquitin and NEDD8 deamidation catalyzed by bacterial effectors that induce macrophage-specific apoptosis. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1210831109.

Yao, Q., Zhang, L., Wan, X., Chen, J., *et al*. (2014) Structure and specificity of the bacterial cysteine methyltransferase effector NleE suggests a novel substrate in human DNA repair pathway. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1004522.

Yatsuyanagi, J., Saito, S., Miyajima, Y., Amano, K.I., *et al.* (2003) Characterization of atypical enteropathogenic

Escherichia coli strains harboring the astA gene that were associated with a waterborne outbreak of diarrhea in Japan. *Journal of Clinical Microbiology*. [Online] Available from: doi:10.1128/JCM.41.5.2033-2039.2003.

Yatsuyanagi, J., Saito, S., Sato, H., Miyajima, Y., *et al*. (2002) Characterization of enteropathogenic and enteroaggregative *Escherichia coli* isolated from diarrheal outbreaks. *Journal of Clinical Microbiology*. [Online] Available from: doi:10.1128/JCM.40.1.294-297.2002.

Ye, Z., Petrof, E.O., Boone, D., Claud, E.C., *et al*. (2007) *Salmonella* effector AvrA regulation of colonic epithelial cell inflammation by deubiquitination. *American Journal of Pathology*. [Online] Available from: doi:10.2353/ajpath.2007.070220.

Yen, H., Ooka, T., Iguchi, A., Hayashi, T., *et al*. (2010) Nlec, a type III secretion protease, compromises NF-kB activation by targeting p65/rela. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1001231.

Yen, H., Sugimoto, N. & Tobe, T. (2015) Enteropathogenic *Escherichia coli* Uses NleA to Inhibit NLRP3 Inflammasome Activation. *PLoS Pathogens*. [Online] Available from: doi:10.1371/journal.ppat.1005121.

Yewdell, J.W. & Nicchitta, C. V. (2006) The DRiP hypothesis decennial: support, controversy, refinement and extension. *Trends in Immunology*. [Online] Available from: doi:10.1016/j.it.2006.06.008.

Yoshida, S., Handa, Y., Suzuki, T., Ogawa, M., *et al*. (2006) Microtubule-severing activity of *Shigella* is pivotal for intercellular spreading. *Science*. [Online] Available from: doi:10.1126/science.1133174.

Young, J.C., Clements, A., Lang, A.E., Garnett, J.A., *et al.* (2014) The *Escherichia coli* effector EspJ blocks Src kinase activity via amidation and ADP ribosylation. *Nature communications*. [Online] Available from: doi:10.1038/ncomms6887.

Zaltsman, A., Krichevsky, A., Loyter, A. & Citovsky, V. (2010) *Agrobacterium* induces expression of a host F-Box protein required for tumorigenicity. *Cell Host and Microbe*. [Online] Available from: doi:10.1016/j.chom.2010.02.009.

Zambelloni, R., Connolly, J.P.R., Huerta Uribe, A., Burgess, K., *et al*. (2017) Novel compounds targeting the enterohemorrhagic *Escherichia coli* type three secretion system reveal insights into mechanisms of secretion inhibition. *Molecular Microbiology*. [Online] Available from: doi:10.1111/mmi.13719.

Zarivach, R., Vuckovic, M., Deng, W., Finlay, B.B., *et al*. (2007) Structural analysis of a prototypical ATPase from the type III secretion system. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb1196.

Zhang, L., Ding, X., Cui, J., Xu, H., *et al*. (2012) Cysteine methylation disrupts ubiquitin-chain sensing in NF-κB activation. *Nature*. [Online] Available from: doi:10.1038/nature10690.

Zhang, X., Zhang, F., Guo, L., Wang, Y., *et al.* (2013a) Interactome analysis reveals that C1QBP (complement component 1, q subcomponent binding protein) is associated with cancer cell chemotaxis and metastasis. *Molecular & cellular proteomics : MCP*. [Online] Available from: doi:10.1074/mcp.M113.029413.

Zhang, X., Zhang, J., Zhang, L., Van Dam, H., *et al*. (2013b) UBE2O negatively regulates TRAF6-mediated NF-κB activation by inhibiting TRAF6 polyubiquitination. *Cell Research*. [Online] Available from: doi:10.1038/cr.2013.21.

Zhang, Y., Higashide, W.M., McCormick, B.A., Chen, J., *et al*. (2006) The inflammation-associated *Salmonella* SopA is a HECT-like E3 ubiquitin ligase. *Molecular Microbiology*. [Online] Available from:

doi:10.1111/j.1365-2958.2006.05407.x.

Zhao, J., Liu, T., Yu, G. & Wang, J. (2015) Overexpression of HABP1 correlated with clinicopathological characteristics and unfavorable prognosis in endometrial cancer. *Tumor Biology*. [Online] Available from: doi:10.1007/s13277-014-2761-8.

Zheng, J. & Leung, K.Y. (2007) Dissection of a type VI secretion system in *Edwardsiella tarda*. *Molecular Microbiology*. [Online] Available from: doi:10.1111/j.1365-2958.2007.05993.x.

Zheng, N. & Shabek, N. (2017) Ubiquitin ligases: Structure, function, and regulation. *Annu. Rev. Biochem*. [Online] Available from: doi:10.1146/annurev-biochem.

Zheng, X., Wang, X., Tu, F., Wang, Q., *et al*. (2017) TRIM25 is required for the antiviral activity of Zinc Finger Antiviral Protein. *Journal of Virology*. [Online] Available from: doi:10.1128/JVI.00088-17.

Zheng, Y., Valdez, P.A., Danilenko, D.M., Hu, Y., *et al*. (2008) Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens. *Nature Medicine*. [Online] Available from: doi:10.1038/nm1720.

Zheng, Z., Wei, C., Guan, K., Yuan, Y., *et al*. (2016) Bacterial E3 ubiquitin ligase IpaH4.5 of *Shigella flexneri* targets TBK1 to dampen the host antibacterial response. *The Journal of Immunology*. [Online] Available from: doi:10.4049/jimmunol.1501045.

Zhou, Y., Dong, N., Hu, L. & Shao, F. (2013) The *Shigella* type three secretion system effector OspG directly and specifically binds to host ubiquitin for activation. *PLoS ONE*. [Online] Available from: doi:10.1371/journal.pone.0057558.

Zhu, Y., Chen, G., Lv, F., Wang, X., *et al*. (2011) Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1101676108.

Zhu, Y., Li, H., Hu, L., Wang, J., *et al*. (2008) Structure of a *Shigella* effector reveals a new class of ubiquitin ligases. *Nature Structural and Molecular Biology*. [Online] Available from: doi:10.1038/nsmb.1517.

Zouhir, S., Bernal-Bayard, J., Cordero-Alba, M., Cardenal-Muñoz, E., *et al*. (2014) The structure of the Slrp–Trx1 complex sheds light on the autoinhibition mechanism of the type III secretion system effectors of the NEL family. *Biochemical Journal*. [Online] Available from: doi:10.1042/BJ20140587.

Zurawski, D. V., Mumy, K.L., Badea, L., Prentice, J.A., *et al.* (2008) The NleE/OspZ family of effector proteins is required for polymorphonuclear transepithelial migration, a characteristic shared by enteropathogenic *Escherichia coli* and *Shigella flexneri* infections. *Infection and Immunity*. [Online] Available from: doi:10.1128/IAI.00684-07.