Characterization of the host-mediated modification of bacterial effector NleA

By
Lindsay Burns

Department of Microbiology and Immunology
McGill University, Montreal, Quebec, Canada

July 2024

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

ABSTRACT

Escherichia coli is a ubiquitous member of the human intestinal microbiota. Although commensal *E. coli* and its human host generally coexist with mutual benefits in digestion and immunity, diarrheagenic *E. coli* are pathogens that have acquired virulence factors that permit them to cause a wide range of enteric or diarrheal diseases in healthy individuals, resulting in more than 500 million illnesses and nearly 200,000 deaths each year. Enteropathogenic and enterohaemorrhagic *Escherichia coli* (EPEC and EHEC) are gastrointestinal pathogens responsible for severe diarrheal illness. EPEC and EHEC form "attaching and effacing" lesions during colonization and, upon adherence, inject proteins directly into host intestinal cells via the type III secretion system (T3SS). Injected bacterial proteins have a variety of functions, typically alter host cell biology to favour survival and/or replication of the pathogen. Non-LEE-encoded effector A (NleA) is a T3SS-injected effector of EPEC, EHEC, and the related mouse pathogen *Citrobacter rodentium*. *In vivo* experiments have shown that NleA is necessary for pathogenesis, however, the mechanism by which NleA contributes to disease remains largely unknown. I have determined that following translocation into host cells, a serine- and threonine-rich region of NleA is modified by host-mediated mucin-type O-linked glycosylation. Importantly, this region was absent in several EHEC outbreak clinical isolates. When expressed in *C. rodentium,* a non-modifiable variant of NleA was indistinguishable from wildtype NleA in an acute mortality model but conferred a modest increase in persistence over the course of infection in mixed infections in C57BL/6J mice. My data also suggests that this modification may confer a selective disadvantage to the bacteria during *in vivo* infection. The modification of NleA occurs in the host cell secretory pathway, where I have identified the interaction of NleA with the Sec24 subunit of the COPII vesicle coat as a necessary step for the glycosylation of NleA. Preliminary *in vitro* work indicates

that the presence of the modification may be detrimental to NleA's stability within the host cell. Altogether, this work has characterized NleA as the first known example of a bacterial effector being modified by host-mediated O-linked glycosylation and contributes to a growing understanding of the range of interactions possible between bacterial effector proteins and host cells.

RÉSUMÉ

Escherichia coli est un membre omniprésent du microbiote intestinal humain. Bien que *l'E. coli* commensal et son hôte humain coexistent généralement avec des avantages mutuels en matière de digestion et d'immunité, les *E. coli* diarrhéiques sont des agents pathogènes qui ont acquis des facteurs de virulence leur permettant de provoquer un large éventail de maladies entériques ou diarrhéiques chez les individus sains, ce qui entraîne plus de 500 millions de maladies et près de 200 000 décès chaque année. Les *Escherichia coli* entéropathogènes et entérohémorragiques (EPEC et EHEC) sont des agents pathogènes gastro-intestinaux responsables de maladies diarrhéiques graves. Les EPEC et les EHEC forment des lésions « attachantes et effaçantes » pendant la colonisation et, après l'adhésion, injectent des protéines directement dans les cellules intestinales de l'hôte par le système de sécrétion de type III (T3SS). Les protéines bactériennes injectées ont des fonctions diverses, mais modifient habituellement la biologie des cellules de l'hôte pour favoriser la survie et/ou la réplication de l'agent pathogène. L'effecteur A non codé par le LEE (NleA) est un effecteur injecté par le T3SS des EPEC, des EHEC et de l'agent pathogène apparenté *Citrobacter rodentium* chez la souris. Des expériences *in vivo* ont montré que NleA est nécessaire à la pathogenèse, mais le mécanisme par lequel NleA contribue à la maladie reste largement inconnu. J'ai déterminé qu'après la translocation dans les cellules hôtes, une région riche en sérine et en thréonine de NleA est modifiée par une glycosylation liée en O de type mucine médiée par l'hôte. Il est important de noter que cette région était absente dans plusieurs isolats cliniques de l'épidémie de EHEC. Lorsqu'elle est exprimée dans *C. rodentium*, une variante non modifiable de NleA ne se distingue pas de NleA de type sauvage dans un modèle de mortalité aiguë, mais elle confère une augmentation modeste de la persistance au cours de l'infection dans des infections mixtes chez des souris C57BL/6J. Mes données suggèrent également que cette

modification puisse conférer un désavantage sélectif à la bactérie lors d'une infection *in vivo*. La modification de NleA se produit dans la voie sécrétoire de la cellule hôte, où j'ai identifié l'interaction de NleA avec la sous-unité Sec24 de l'enveloppe de la vésicule COPII comme une étape nécessaire à la glycosylation de NleA. Des travaux préliminaires *in vitro* indiquent que la présence de la modification peut être compromettre à la stabilité de NleA dans la cellule hôte. Dans l'ensemble, ces travaux ont permis de caractériser la NleA comme le premier exemple connu d'un effecteur bactérien modifié par une glycosylation O médiée par l'hôte et contribuent à une compréhension croissante de la gamme d'interactions possibles entre les protéines effectrices bactériennes et les cellules hôtes.

ACKNOWLEDGEMENTS

Don't adventures ever have an end? I suppose not.
Someone else always has to carry on the story.
- J.R.R. Tolkien

I would first like to thank my supervisor, Dr. Samantha Gruenheid, for believing in me to carry on the story. I am deeply appreciative of her patience and enthusiasm in guiding me on this adventure and creating a research environment that encourages creativity and freedom in which I could fail, endlessly. I am extremely fortunate to have found such an incredible mentor so early on in my research career. I am very grateful to my advisory committee, Dr. Ajitha Thanabalasuriar and Dr. Jesse Shapiro, for their ardent encouragement, as well as all the collaborators involved in this work for their contributions and guidance.

Thank you to past and present members of the Gruenheid lab from whom I have learned so much, both in and outside the lab. Their passion, enthusiasm, and knowledge are inspiring. An especially sincere thank you to the labbies – Jessica, Tyler, Elena, and Brendan. I did not envision my PhD beginning with a global pandemic and lab shutdown, but their levity and reassurance made it all feel a little less lonely. Many thanks to Jacqueline Sung for editing my abstract. Thank you to Lei Zhu for her animal expertise and movie recommendations, and Dr. Alexandra Kazanova for more help, experimental and emotional, than could reasonably be listed here. I would also like to acknowledge the unsung lab members: the laboratory mice that were part of this work. Throughout scientific history, lab animals have made the ultimate sacrifice for the betterment of human life, and for that we are all eternally indebted.

Thank you to my friends Kirsten, Hayley, Rebecca, Tiffany, Kayla, Dilhan, and many others that have impacted my time at McGill, namely through late-night debriefs, questionable

driving, and all-you-can-eat anything. To my family and feels-like-family – Anita, Eric, Mary-Kay, and Peter, thank you for your incredible kindness and support even when you didn't know I needed it most. A heartfelt thank you to my parents Janet and Mark for their unconditional love and support. They have gifted me with a curiosity for questions and the stubbornness to get answers. Their simple request that I always just try my best has kept me going. Thank you to my brother Matthew for the lengthy discussions about nothing at all. Thank you to my cat Lola for cuddles and keeping my ego in check. And finally, thank you to "my" Sam, who has so patiently waited the past six years to carry on our story. I can't possibly put into a few short words how much your love means to me. The next adventure awaits with you.

CONTRIBUTION OF AUTHORS

The work presented in this manuscript-based thesis was published, or is in preparation for publication, as detailed below. Titles for unpublished chapters may change based on experiments pending and in progress.

Chapter 1

The literature review was written by LB and edited by SG.

Chapter 2

Burns L, Giannakopoulou N, Zhu L, Xu YZ, Khan RH, Bekal S, Schurr E, Schmeing TM, Gruenheid S. The bacterial virulence factor NleA undergoes host-mediated O-linked glycosylation. Molecular Microbiology. 2023;119(2):161-173.

LB, NG, and SG conceptualized and designed the experiments. LB, NG, and LZ performed the experiments and analyzed the data. RHK and SB contributed to the work by providing EHEC outbreak strain *nleA* sequences. TMS performed predictive structure modelling and generated the model figure. SG supervised the project. The manuscript was written by LB and edited by SG and TMS.

Chapter 3

Burns L, Le Mauff F, Gruenheid, S. Direct evidence of host-mediated glycosylation of NleA and its dependence on interaction with the COPII complex. Gut Microbes. 2024;16(1).

LB and SG conceptualized and designed the experiments. LB performed the experiments and analyzed the data. FLM provided technical support. SG supervised the project. The manuscript was written by LB and edited by FLM and SG.

Chapter 4

Burns L, Kazanova A, Gruenheid, S. Host-mediated glycosylation of bacterial effector NleA may affect its stability inside host cells. *Manuscript in preparation*.

LB and AK conceptualized the experiments, performed the experiments, and analyzed the data. AK performed flow cytometry analyses. SG supervised the project. The manuscript was written by LB and edited by AK and SG.

Chapter 5

The discussion and conclusion were written by LB and edited by SG.

CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

The work presented here contributes novel findings to our understanding of host-pathogen interactions and the role bacterial virulence factors play in this evolutionary arms race.

Chapter 2

Although a shift in mobility by Western blot of NleA had been observed since at least 2007, this study was the first to characterize this shift as the modification of NleA by host-mediated O-linked glycosylation. This study is the first to describe O-linked glycosylation as a host-mediated post-translational modification of a bacterial effector protein. We identified a serine- and threonine-rich region of the NleA protein sequence that, when deleted, abrogates modification of NleA and predicted where this region exists in structural models of NleA. We demonstrated the first evidence of this region being variable in human clinical isolates. Finally, we showed that although glycosylation of NleA does not seem to be crucial for virulence *in vivo*, we did observe a subtle persistence effect of bacteria expressing non-modifiable NleA.

Chapter 3

Prior to the publication of this study, we were only able to obtain indirect evidence of NleA's O-glycosylation. This study presents the first direct evidence of mucin-type O-glycans bound to NleA using lectin blotting. We also determined that the modification of NleA is dependent on its interaction with Sec24, a subunit of the COPII vesicle coat involved in anterograde trafficking, providing new insight into interaction between NleA and the host secretory pathway.

Chapter 4

We demonstrate a novel technique for investigating host-effector interactions using detection of intracellular effector proteins by flow cytometry to identify host cells as positive or negative for injection by the T3SS. Using this technique, we show that non-modifiable NleA may be more stable inside the host cell than O-glycosylated NleA.

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF TABLES

Chapter 1

Chapter 2

LIST OF ABBREVIATIONS

A/E	Attaching and effacing
aEPEC	Atypical EPEC
AMP	Antimicrobial peptide
AQP	Aquaporin
Asn	Asparagine
BFP	Bundle-forming pilus
CFU	Colony forming unit
CHO	Chinese hamster ovary
Cif	Cycle inhibiting factor
COPII	Coat protein complex 2
DAEC	Diffusely adhering *Escherichia coli*
DEPEC	Dog enteropathogenic *Escherichia coli*
DM	Double mutant
DMEM	Dulbecco's modified eagle medium
DPI	Days post-infection
DRA	Down-regulated in adenoma
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EHEC	Enterohaemorrhagic *Escherichia coli*
EPEC	Enteropathogenic *Escherichia coli*
ER	Endoplasmic reticulum
ETEC	Enterotoxigenic *Escherichia coli*
FBS	Fetal bovine serum
Fuc	Fucose
Gal	Galactose
GalNAc	N-acetyl galactosamine
GALNT	GalNAc transferase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
Gb_3	Globotriaosylceramide
GlcNAc	N-acetyl glucosamine
gRNA	Guide ribonucleic acid
HC	Hemorrhagic colitis
HUS	Hemolytic uremic syndrome
ICAM-1	Intercellular adhesion molecule 1
ICE	Integrative conjugative element
IL	Interleukin
ILC3	Type 3 innate lymphoid cell
INF	Interferon
IP	Immunoprecipitation
kb	Kilobase
kDa	Kilodalton
KO	Knockout
LB	Luria Bertani

LCN2	Lipocalin 2
LEE	Locus of enterocyte effacement
Ler	LEE-encoded regulator
LOD	Limit of detection
LPS	Lipopolysaccharide
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight
Man	Mannose
MAP	Mitogen-associated protein
Map	Mitochondrial-associated protein
Mb	Megabase
MFI	Mean fluorescence intensity
MGE	Mobile genetic element
MOI	Multiplicity of infection
MPEC	Mouse pathogenic *Escherichia coli*
MS	Mass spectrometry
MUC2; Muc2	Mucin-2
MyD88	Myeloid differentiation primary-response protein 88
NF-κB	Nuclear factor κ B
NFDM	Non-fat dried milk
NHE3	Sodium-hydrogen exchanger 3
Nle	Non-LEE-encoded effector
NleA	Non-LEE-encoded effector A
NLR	NOD-like receptor
OGA	O-GlcNAcase
OGT	O-GlcNAc transferase
PAI	Pathogenicity island
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pEAF	EPEC adherence factor plasmid
PEPEC	Porcine enteropathogenic *Escherichia coli*
Per	Plasmid-encoded regulator
PTM	Post-translational modification
REPEC	Rabbit enteropathogenic *Escherichia coli*
RIPA	Radioimmunoprecipitation assay
Rspo2	R-spondin 2
SCV	*Salmonella*-containing vacuole
SDS	Sodium dodecyl-sulfate
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
S.E.M.	Standard error of the mean
Ser	Serine
SGLT-1	Sodium-D-glucose transporter 1
sgRNA	Single guide ribonucleic acid
STEC	Shiga toxin-producing *Escherichia coli*
Stx	Shiga toxin
SUMO	Small ubiquitin-like modifier
T3SS	Type III secretion system

TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween
tEPEC	Typical EPEC
Thr	Threonine
Tir	Translocated intimin receptor
TJ	Tight junction
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TRAF2	TNF receptor associated factor 2
TRIF	Tir-domain-containing adapter protein inducing interferon β
Ub	Ubiquitin
UDP	Uridine diphosphate
VVL	*Vicia villosa* lectin
WT	Wildtype

PREFACE TO CHAPTER 1

Host-pathogen interactions are a complex balance of accidental or obligatory encounters, temporary or persistent associations, and subtle or acute molecular and cellular responses that ultimately result in a binary outcome: health or disease. Using injected effectors, pathogenic bacteria can tip the scales in their favour. This thesis investigates a major effector of the attaching and effacing pathogens: non-LEE-encoded effector A (see **Chapter 1.3.2**). **Chapter 1** provides a literature review of the field to date in four relevant areas of research: A/E pathogens and their characteristics (**Chapter 1.1**), host defenses against intestinal infection (**Chapter 1.2**), pathogenesis of A/E bacteria (**Chapter 1.3**), and post-translational protein modifications and their importance to effector functionality (**Chapter 1.4**).

CHAPTER 1: **Introduction and Literature Review**

1.1 Attaching and effacing pathogens

Escherichia coli is a ubiquitous and highly versatile organism. It can survive in a multitude of environments, including soil, water, food, and the human intestine. Indeed, *E. coli* colonizes the gastrointestinal (GI) tract of infants within hours of birth and remains the most abundant facultative anaerobe throughout life. Although commensal *E. coli* and its human host generally coexist with mutual benefits in digestion and immunity, there are several highly adapted pathotypes of *E. coli* that have acquired virulence attributes. These virulence attributes, encoded on mobile genetic elements, confer an increased ability to cause disease. Pathogenic *E. coli* is broadly classified as either diarrheagenic or extraintestinal. Diarrheagenic *E. coli* are gastrointestinal pathogens that have acquired virulence factors that permit them to cause a wide range of enteric diseases in healthy individuals, resulting in more than 500 million illnesses and nearly 200,000 deaths each year [1].

The attaching and effacing (A/E) pathogens are a prominent subgroup of diarrheagenic *E. coli* characterized by their formation of distinct lesions during intestinal colonization. The histopathology of A/E pathogens is characterized by intimate bacterial attachment to intestinal epithelial cells, effacement of the microvilli architecture, and formation of actin-rich pedestals atop which the bacteria adhere (**Figure 1**) [2,3]. The A/E classification consists of human pathogens Enteropathogenic *E. coli* (EPEC), Enterohaemorrhagic *E. coli* (EHEC), and *Escherichia albertii*, as well as animal pathogens rabbit EPEC (REPEC), porcine EPEC (PEPEC), dog EPEC (DEPEC), and the murine pathogen *Citrobacter rodentium* [4].

1.1.1 Enteropathogenic *Escherichia coli*

In 1945, Bray described EPEC as the first pathotype of *E. coli* to be associated with human disease [5]. Today, EPEC remains a significant cause of potentially fatal infant diarrhea in developing countries, accounting for 30-40% of all cases [3]. EPEC are non-invasive diarrheagenic *E. coli* that are subdivided into two groups: typical EPEC (tEPEC) and atypical EPEC (aEPEC) based on the presence of the EPEC adherence factor plasmid (pEAF) in tEPEC and its absence in aEPEC [6]. The pEAF encodes the bundle-forming pilus (BFP), a type IV adhesin important for initial localized adherence to the epithelium, further described below [7–10]. However, genetic analyses suggest that EPEC lineages are incredibly diverse, with varied combinations of virulence elements from several independent loss and acquisition events, indicating that the tEPEC and aEPEC classification may be a gross oversimplification [11,12].

As with other diarrheagenic *E. coli* strains, EPEC is transmitted via the fecal-oral route, with contaminated hands, foods, or fomites serving as vehicles [13]. Although rare, outbreaks of EPEC among adults seem to occur due to ingestion of contaminated food and water, however no specific environmental reservoir has been identified as a source of infection [3]. Epidemiologic studies across several countries have shown high asymptomatic carriage with some studies finding 17-20% of healthy infants under two years old shed EPEC in their stool (reviewed in [13]). In symptomatic individuals, EPEC can be isolated from stool up to two weeks following the cessation of symptoms [14]. Humans are the only known reservoir for tEPEC, with symptomatic or asymptomatic children and asymptomatic adults as the primary carriers [13,15,16]. In contrast, aEPEC strains have been isolated from both human and animal sources, including rabbits, dogs, monkeys, and sheep [17,18]. Many animal EPEC species, including REPEC, PEPEC, and DEPEC, are clonally related to human EPEC. These animal pathogens share many virulence properties and

mechanisms to cause analogous diseases in their respective animal hosts, such as formation of the characteristic pedestal lesions and symptoms including profuse watery diarrhea [4,18,19].

EPEC targets to the proximal small intestine where it causes acute or persistent diarrhea especially in infants under two years old. Although acute infections causing watery diarrhea are more common, persistent cases lasting more than two weeks are associated with complications such as dehydration and malnutrition and often require hospitalization [3,19]. The onset of disease is rapid, with symptoms appearing just a few hours following the ingestion of bacteria [20,21]. Human volunteer studies indicate that an infectious dose of 10^8-10^{10} colony forming units (CFU) is sufficient to cause infection in adults, but the actual dose required to cause disease during natural infection is unknown, and presumed to be much lower in infants [22]. The precise mechanism of diarrhea production by EPEC is not fully understood, but is likely a combinatorial effect of several processes, including decreased absorption due to microvilli effacement, bacterial effector disruption of water and ion transport channels, and impaired barrier function leading to increased intestinal permeability [4,23–29]. Thus, the primary goal of treatment for EPEC diarrhea is to prevent dehydration. In most cases, EPEC-induced diarrhea is self-limiting and can be effectively treated with oral rehydration therapy [3]. Persistent infections may require antibiotic administration, however multiple antibiotic resistance is common in EPEC [30–32].

1.1.2 Enterohaemorrhagic *Escherichia coli*

EHEC is a highly infectious A/E pathogen and one of several strains characterized as Shiga toxin-producing *E. coli* (STEC) [33]. STEC are strains that produce the phage-encoded Shiga toxin (Stx; also called verocytotoxin), a potent bacterial toxin originally described in *Shigella dysenteriae* type 1 [34]. Although there are more than 200 STEC serotypes, most do not contain

the locus of enterocyte effacement (LEE) pathogenicity island (PAI) and are not associated with human disease [35]. Thus, STEC is used as a general term for any *E. coli* strain that produces Stx, while EHEC denotes the subset of Stx-producing strains that also contain the LEE. Among the 400 serogroups of STEC that infect humans, EHEC O157:H7 is the most commonly isolated in North America [36]. In addition to O157, several non-O157 serogroups have emerged as causative agents of severe illness and death in humans. Six serogroups: O26, O45, O103, O111, O121, and O145, collectively known as the "big six," have been estimated to account for more than 80% of non-O157 STEC infections [37].

Cattle were identified as a primary reservoir of EHEC following the first outbreaks associated with consumption of undercooked contaminated hamburgers [38]. Cattle are asymptomatic carriers of EHEC and do not suffer from haemorrhagic diarrhea due to the absence of Stx receptors [39–43], making it impossible to visually identify infected animals. Therefore, transmission of EHEC to humans may occur via consumption of contaminated meat and dairy products (foodborne), contamination of crops or drinking water by infected animal feces (waterborne), or by direct person-to-person transmission [44]. The high infectivity of EHEC is a result of its pervasiveness in agriculture, ability to survive for months in the environment, and low infectious dose – fewer than 100 CFU [3,45]. Outbreaks from EHEC are sporadic but primarily occur in developed countries, in contrast to those caused by EPEC [33,46,47]. Despite industry and public health efforts to limit contamination of food and water, the incidence of disease attributed to EHEC has been steady since its discovery in 1984 [48,49].

EHEC colonizes the colon where it causes a spectrum of disease severity, ranging from asymptomatic to fatal. Patients initially experience self-limiting watery diarrhea that may resolve in about one week [3]. Approximately 90%, however, will progress to haemmorrhagic colitis (HC),

and up to 15% of those will develop hemolytic uremic syndrome (HUS) [3,50]. HC may become so severe that some patients report diarrhea that is "all blood and no stool" [38]. HUS is characterized by concurrent hemolytic anemia, thrombocytopenia, and acute kidney injury that can progress to renal failure and death. The tendency to develop worsened disease is affected by several risk factors, namely the age of the host and the strain of EHEC. HUS incidence is highest in patients under 5 years and over 65 years of age [47,51,52], with mortality rates of 5% and up to 90%, respectively [3,53]. In addition, progression to HUS is more likely in those infected with the EHEC O157:H7 strain than non-O157 serotypes, due to disease determinants that are not fully elucidated [54–60].

EHEC produces either or both of two immunologically distinct Stx types: Stx1 and Stx2. Stx1 is nearly identical to the prototype Shiga toxin of *S. dysenteriae* type 1. In comparison, Stx2 shares only 55% amino acid homology to Stx1 and cannot be neutralized by the same antibodies but is associated with greater disease severity [60,61]. Stx is among the most potent biologic substances known given its toxicity to cells at picomolar concentrations [62]. EHEC *stx* genes are encoded within antibiotic-inducible resident lambdoid prophages [63,64], and toxin secretion is achieved via phage-mediated bacterial lysis [65]. Stx is an AB_5 toxin, composed of an A subunit covalently bound to a pentamer of B subunits [62,66,67]. StxB mediates binding to the globotriaosylceramide (Gb_3) cellular receptor on the surface of target cells, leading to subsequent internalization by receptor-mediated endocytosis [68–70]. StxA possesses highly specific RNA N-glycosidase activity that cleaves an adenine residue from the 28S rRNA of the 60S ribosomal subunit, halting protein synthesis and ultimately leading to cell death [62]. Stx secreted in the colon appears to be translocated intact into the bloodstream, however, the mechanism by which it does so is not well characterized. In circulation, the toxin binds to polymorphonuclear leukocytes and

is transported by these cells to various organs, particularly affecting the kidneys [71,72]. The Gb_3 receptor is highly expressed in the glomerular endothelium [73], where Stx damages glomerular endothelial cells leading to reduced filtration and eventual renal failure [74]. In addition to kidney damage, other organs may be affected. Neurological involvement is common, occurring in approximately 25% of HUS patients, in addition to reported pancreatic, gastrointestinal, ocular, cardiac, and pulmonary effects [75–78].

Supportive therapy remains the primary treatment option for EHEC infection. Affected organs in humans and experimental animals demonstrate significant vascular damage, with the most severe cases requiring blood transfusion and dialysis therapy [79]. Administration of antibiotics for EHEC infection treatment is controversial. Worsening of disease is hypothesized to occur following antibiotic treatment either due to bacterial cell death and subsequent mass Stx release, or dysbiosis in the commensal microbiota that allows Stx to freely bind the epithelium. A contrasting hypothesis proposes that earlier destruction of the pathogen leads to reduced Stx excretion and decreased severity or propensity to develop disease. Indeed, meta-analyses of combined cohorts have reached opposing conclusions [80–83].

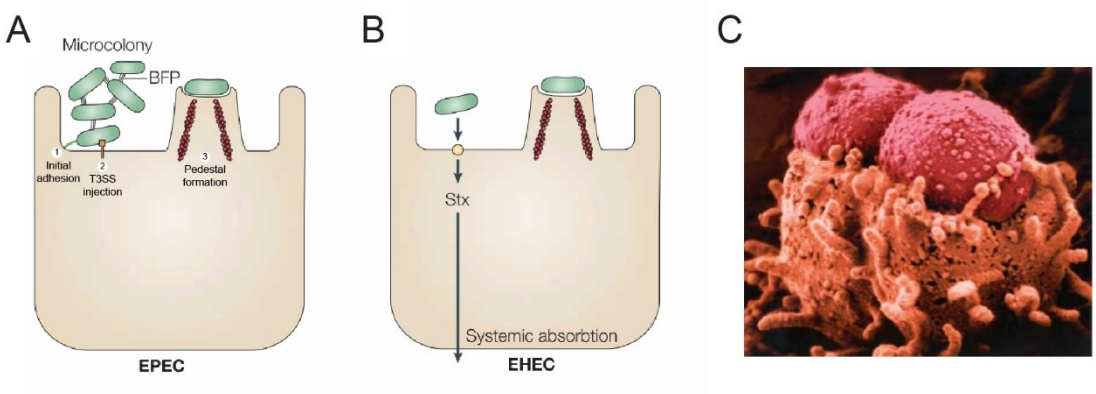

Figure 1. EPEC and EHEC form a characteristic A/E lesion histopathology. (A) EPEC initially adhere to the small intestine epithelium via the BFP, followed by T3SS injection of bacterial effectors that result in actin-rich pedestal formation. (B) EHEC also induce A/E lesion

formation in the colon. EHEC-translocated Stx disseminates systemically which leads to potentially life-threatening complications. (C) Scanning electron microscopy of A/E lesions (orange) atop which bacteria (red) are intimately adhered. Figure modified from [33].

1.1.3 *Escherichia albertii*

Escherichia albertii is an emerging, albeit underappreciated, A/E pathogen, causing outbreaks of gastroenteritis in humans [84]. *E. albertii* is a close relative of *E. coli* and frequently misidentified as EPEC or EHEC due to shared virulence and genetic traits, namely the ability to form A/E lesions. *E. albertii* is challenging to classify using traditional methods due to the very limited number of characterized strains that show non-universal results with multilocus sequence typing or biochemical approaches [85]. As a result, distinguishing *E. albertii* from *E. coli* has largely relied on whole genome sequencing [85]. Consequently, *E. albertii* remains relatively unrecognized despite several reports demonstrating a strong correlation between *E. albertii* and disease outbreaks, some of which had been incorrectly attributed to EPEC and EHEC [86–90]. Although it was first isolated from infants, *E. albertii* infections have been reported in children, adults, and elderly [86,91,92]. Symptoms of *E. albertii* infection include gastroenteritis, dehydration, abdominal distension, vomiting, and fever [84].

In addition to humans, *E. albertii* causes disease in multiple species of birds [93], that therefore may act as a primary reservoir. Independent investigations in the United States, Japan, and China recovered *E. albertii* from processed chickens, suggesting that chicken could be a vehicle for transmissibility of *E. albertii* to humans [94–96]. Besides chicken, *E. albertii* has also been isolated from other processed meat products, including duck, mutton, and pork [96]. In addition, *E. albertii* has been isolated from several contaminated water sources [90,97]. Altogether, these findings implicate *E. albertii* as a food- and water-borne pathogen.

Since many *E. albertii* isolates are uncharacterized, it is challenging to determine the clinical relevance, prevalence, reservoirs, transmission routes, and epidemiology of the species. However, an increasing amount of sequence data on complete or draft genomes of many *E. albertii* strains have provided valuable information about the potential virulence and diversity of this pathogen. These analyses show that *E. albertii* carries a large collection of virulence-related genes, many of which are shared with other A/E pathogens. For example, a small subset of *E. albertii* isolates harbor the *stx* genes that encode Shiga toxin, typically a genetic hallmark of EHEC [92,98].

1.1.4 *Citrobacter rodentium*

Citrobacter rodentium is a natural mouse pathogen that shares several pathogenic characteristics and mechanisms with EPEC and EHEC. *C. rodentium* was initially identified as *C. freundii* biotype 480 and mouse pathogenic *E. coli* (MPEC) by two independent groups in Boston and Tokyo that isolated the pathogen from severe outbreaks of diarrhea in their mouse colonies [99,100]. Genetic analyses of these pathogens showed significant DNA relatedness, suggesting a common source, and the two strains were reassigned to a single species: *C. rodentium* [101]. *C. rodentium* shares 67% homology with the genomes of EPEC and EHEC, including several virulence determinants that indicate *C. rodentium*, EPEC, and EHEC acquired a common host infection strategy via convergent evolution [102].

Like EPEC and EHEC, *C. rodentium* is an extracellular pathogen that colonizes the lumen of the mouse gut mucosa by the formation of A/E lesions on the apical surfaces of enterocytes (**Figure 2**) [103,104]. *C. rodentium* infection produces lesions that are indistinguishable from those formed by EPEC and EHEC in humans [104]. *C. rodentium* colonizes the intestine, residing predominantly in the cecum and colon [105]. Mice infected with *C. rodentium* develop diarrhea,

colitis, and hyperplasia of the intestinal epithelium, which are all shared symptoms of human infections with EPEC and EHEC [106]. *C. rodentium*-mediated colonic hyperplasia is highly infectious, spreading rapidly between mice due to a reduced infectious dose and lack of requirement for initial colonization of the cecal patch in secondary hosts [107]. Notably, mice are naturally resistant to EPEC and EHEC infection, with poor colonization in the absence of antibiotic pretreatment and no colonic pathology [106,108]. Thus, due to the similarities and challenges outlined above, *C. rodentium* is widely used as a robust model for studying molecular mechanisms of A/E pathogenesis *in vivo*.

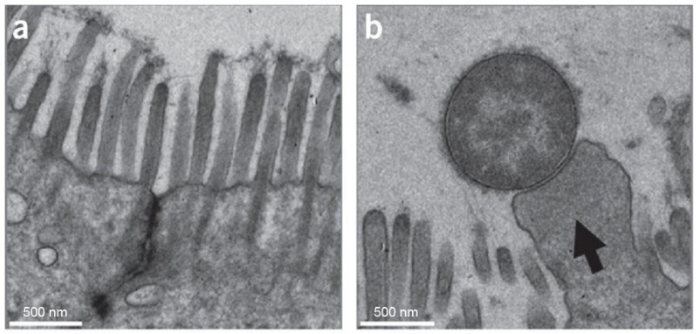

Figure 2. *C. rodentium* induces A/E lesion formation on mucosal surfaces. Transmission electron microscopy of C57BL/6 mouse colonic epithelium. (A) Uninfected epithelial cells have intact brush border microvilli architecture. (B) *C. rodentium*-infected epithelial cells have characteristic A/E pedestal-like lesions with effacement of the microvilli and actin accumulation (arrow) below adherent *C. rodentium*. Figure modified from [109].

1.1.4.1 *C. rodentium* as a mouse model of A/E pathogen infection

In this model, infection with *C. rodentium* is achieved by oral gavage of 10^8-10^9 CFU directly into the GI tract. Although a majority of the bacterial inoculum will die in the stomach, remaining *C. rodentium* colonize the cecum during the first day of infection, progressing to the distal colon 2-3 days post-infection [105]. Bacterial levels in the colon peak at approximately 10^9 organisms between days 5-14 post-infection, comprising 1-3% of the total intestinal microbiota

[105,106,110–112]. Overgrowth of *C. rodentium* in the colon results in pronounced dysbiosis, consequently reducing the overall abundance and diversity of the commensal microbiota [110]. The pattern of *C. rodentium* clearance parallels colonization: the cecum is cleared first, followed by clearance of the colon 21-28 days post-infection [106].

The severity of *C. rodentium* infection is largely dependent on mouse genetic background. The majority of immunocompetent strains, including C57BL/6, NIH Swiss, and Balb/c, are relatively resistant to clinical disease, developing mild self-limiting colitis upon *C. rodentium* infection [106]. Genetically susceptible mouse strains, such as C3H/HeJ, display severe diarrhea and colon inflammation, and ultimately succumb to infection [106]. Genetic analyses have implicated the *Cri1* genetic locus as a common genetic cause of susceptibility [113,114]. Specifically, the *Cri1* locus has been shown to regulate *R-spondin 2* (*Rspo2*) expression [115]. During homeostasis, *Rspo2*-induced activation of Wnt signaling is restricted to the proliferative compartment of the colonic crypt, playing a key role in maintaining the stem cell niche and functioning as an epithelial repair response [116]. In C57BL/6 mice, *Rspo2* expression is tightly regulated to maintain Wnt signaling to increase epithelial cell proliferation, prompting sloughing of epithelial cells and clearance of the pathogen [117,118]. Comparatively, C3H/HeJ mice have a genetic haplotype upstream of *Rspo2* that drives excessive *Rspo2* expression and robust Wnt activation, resulting in increased proliferation and decreased differentiation of intestinal epithelial cells [115,117]. Interestingly, *Cri1*-mediated susceptibility is independent of the ability of the pathogen to colonize the gut or dysregulation of the immune response [115,119]. In addition to genetic determinants, the composition of the intestinal microbiota has been shown to influence susceptibility to *C. rodentium* infection [120].

1.2 Host defenses against intestinal infection

Although the human body is constantly exposed to microorganisms, infectious diseases are fortunately quite rare. This is due to several physical and chemical barriers that protect the host against commensal and pathogenic bacteria alike.

Invading enteric pathogens encounter a major obstacle in the form of the trillions of microorganisms that occupy the mammalian intestine. Collectively known as the gut microbiota, resident bacteria, viruses, fungi, archaea, and protozoa contribute several benefits to the host including digestion of nutrients and development of the immune system [121]. Furthermore, gut symbionts form a stable community that oppose invasion by exogenous bacteria, a function termed colonization resistance. Recognized since at least the 1950s, colonization resistance is multifaceted: involving inhibition of initial infection, improved tolerance to established infection, and promotion of infection clearance [122,123]. It has been shown that germ-free and certain antibiotic-treated mice are more susceptible to *C. rodentium*, indicating the key role the microbiota plays in affecting infection outcome [124,125]. Altering the composition of the microbiota also influences host susceptibility to infection. Pre- or co-administration of probiotics during the early stages of *C. rodentium* infection was shown to ameliorate barrier dysfunction, colonic pathology, and pathogen colonization [126]. In addition, transfer of the microbiota from genetically resistant to susceptible mouse strains has been shown to confer protection against *C. rodentium*-induced colitis [120,127].

At the epithelial interface, the mucus layer provides a physical barrier between host and environment, serving as the site for interactions between invading pathogens, the commensal microbiota, and the host immune system [128–130]. The mucus architecture is variable along the GI tract due to the differing functions of the epithelium required at these organs (**Figure 3A**). The

small intestine: the primary site of food digestion, has a single permeable mucus layer that permits nutrient absorption by enterocytes (**Figure 3B**) [131]. In the cecum, the mucus layer becomes discontinuous, predominantly covering the bottom of epithelial crypts while leaving the surface of the crypts exposed for immune antigen sampling and pathogen colonization. The colon produces a majority of the mucus in the GI tract, increasing in thickness from the proximal to the distal colon to reduce mechanical damage from movement of fecal content and protect from increasing bacterial load [129,132,133]. The colonic mucus is organized in two layers: an inner, stratified mucus layer that is firmly adhered to epithelial cells and approximately 50 μm thick and an outer, nonattached layer of variable thickness (**Figure 3C**) [129,134]. The normal intestinal microbiota inhabits the outer mucus layer without triggering an inflammatory response [133–135]. The inner layer, however, is virtually sterile and impenetrable to the microbiota [133]. Both mucus layers are primarily composed of the protein Mucin-2 (MUC2) secreted by goblet cells, a specific secretory subtype of intestinal epithelial cell [136,137]. MUC2 is an extensively O-glycosylated protein that, following secretion, unfolds into net-like sheets [138]. The numerous O-glycans of MUC2 serve as a bacterial nutrient source and attachment site [139,140]. Mice deficient in Muc2 are highly susceptible to *C. rodentium* infection and have a 10-100 fold increase in pathogen burden compared to infected wild-type (WT) mice [141]. Although the mechanisms invading bacteria use to surpass the inner mucus layer and reach the epithelium are not well characterized, it is believed that they employ a combination strategy of secreted proteases, flagellar motility, and exploitation of commensal mucin degradation [141–149].

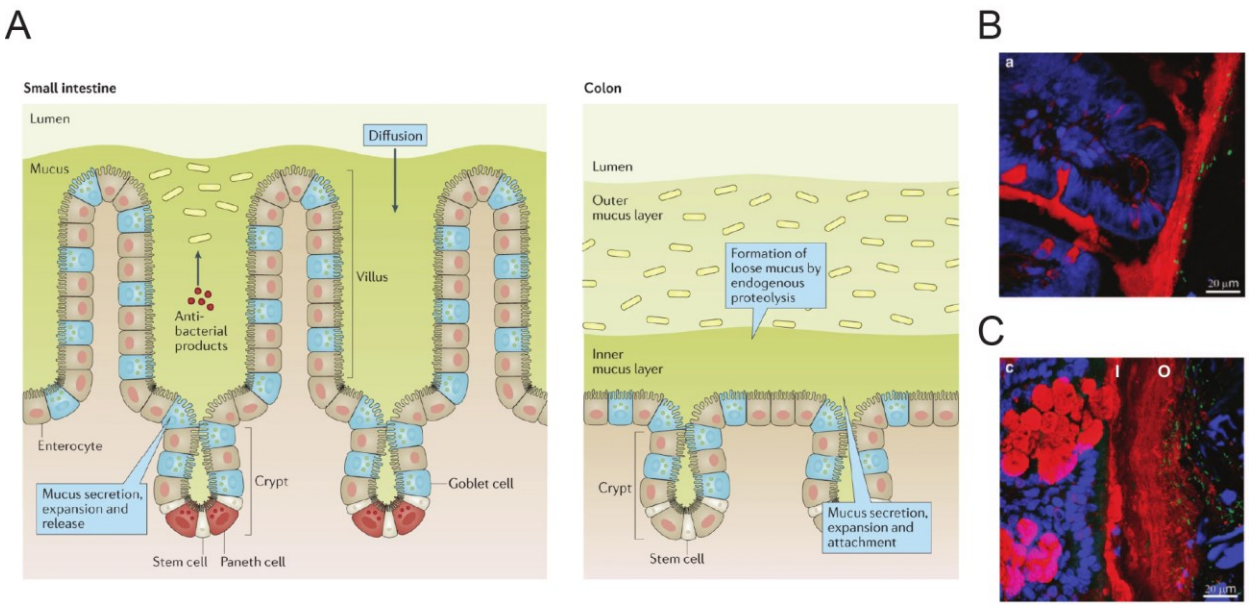

Figure 3. The gastrointestinal tract is protected by mucus. (A) The small intestinal mucus is permeable but contains antibacterial mediators to limit penetration by bacteria. In the colon, the mucus is comprised of an inner stratified layer attached to the epithelium and an outer loose layer. The inner layer is typically impenetrable to bacteria while the outer layer can serve as a habitat for resident bacteria. Fluorescence microscopy of (B) small intestine and (C) colon tissue from WT C57BL/6 mice shows the structure of Muc2-comprising mucus (red) and the microbiota (green) at these sites. Cell nuclei are shown in blue. The inner (I) and outer (O) mucus layers can be observed in the colon. Figures modified from [150,151].

Considering much of the work characterizing the host response to A/E pathogen infection has been done in the *C. rodentium* mouse model, the following will focus on the murine immune response. The cycle of *C. rodentium* infection in C57BL/6 mice, that develop mild disease, is divided into four defined phases: cecal establishment, colonic expansion, steady-state shedding, and clearance (**Figure 4**). The resident microbiome and mucus described above act as primary barriers to prevent bacterial establishment in the first phase of infection. In subsequent phases, both the innate and adaptive immune systems are responsible for eliminating bacteria that breach the mucus and pass the epithelial cell barrier. Extracellular immune surveillance by Toll-like receptors (TLRs) on the surface of epithelial and myeloid cells initially recognize *C. rodentium* [152–154]. TLR2 and TLR4 detect peptidoglycan and lipopolysaccharide (LPS), respectively, in

the bacterial cell wall [152,155]. Binding of specific ligands to TLRs recruits the adapter proteins myeloid differentiation primary-response protein 88 (MyD88) and TIR-domain-containing adapter protein inducing interferon β (TRIF) [156]. MyD88 and TRIF recruitment results in subsequent activation of the nuclear factor-κB (NF-κB) and mitogen-activated protein (MAP) kinase signaling cascades [156]. NF-κB signaling triggers a strong inflammatory response, with the recruitment of dendritic cells, macrophages, and neutrophils, and production of pro-inflammatory cytokines interleukin (IL)-6, IL-12, IL-23, and tumor necrosis factor alpha (TNF-α) [152,153,157]. Activation of TLR2 is not required for *C. rodentium* clearance, however, TLR2-deficient mice display severe colonic pathology, rapid weight loss, and accelerated mortality [152]. Likewise, TLR4-mediated responses to *C. rodentium* are not protective of the host, and in fact promote early colonization of the colon and can exacerbate tissue damage during infection [155].

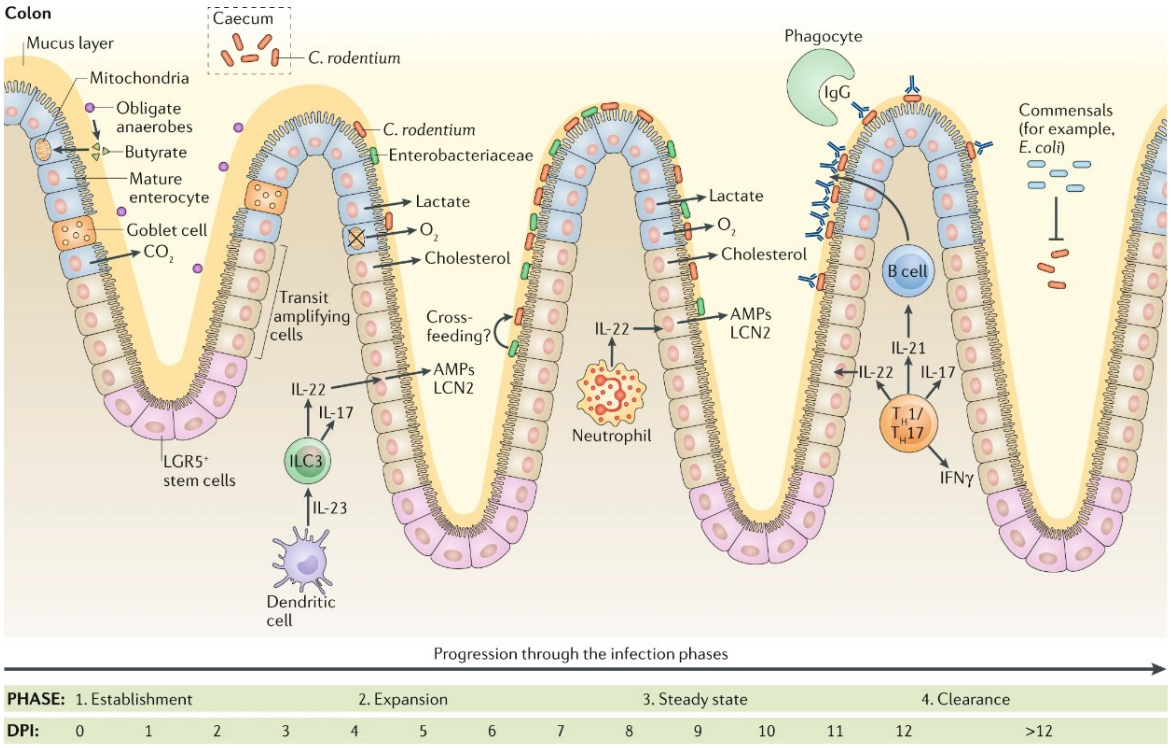

Figure 4. Phases of *C. rodentium* infection and host immune response in a resistant mouse model. In the establishment phase, 1-3 days post-infection (DPI), a small proportion of inoculated *C. rodentium* colonizes the cecum, where it adapts to the gastrointestinal microenvironment. In the colon, invading *C. rodentium* encounters commensal bacteria, producing metabolites including butyrate, and mucus that protect the epithelium. During the expansion phase, *C. rodentium* is first observed sparsely adhered to epithelial cells, followed by rapid proliferation. Increased oxygenation and secretion of lactate into the lumen results from downregulation of mitochondrial oxidative phosphorylation and switch to aerobic glycolysis. Concurrently, cholesterol biosynthesis and excretion into the lumen is increased. These alterations are associated with reduced tissue-associated obligate anaerobes and expansion of facultative anaerobes, including *C. rodentium*. Dendritic cells secrete IL-23 which induces ILC3s to secrete proinflammatory cytokines IL-22 and IL-17. IL-22 signals intestinal epithelial cells to secrete antimicrobial peptides (AMPs) as well as lipocalin 2 (LCN2) and calprotectin (not shown). 8-12 DPI, the steady-state phase is characterized by continued *C. rodentium* colonization of the colonic mucosa. Tissue neutrophils secrete IL-22, further inducing AMP and LCN2 secretion. In the clearance phase, Th1 and Th17 effector cells secrete IL-22, IL-21, IFNγ, and IL-17. IL-21 promotes B cells to secrete IgG antibodies. IgG recognize adhered *C. rodentium* which are subsequently engulfed by phagocytes. Figure reproduced from [158].

Intracellularly, nucleotide-binding oligomerization domain (NOD) proteins NOD1 and NOD2 within epithelial cells recognize bacterial cell wall components to promote *C. rodentium* clearance by inducing T helper type 17 (Th17) and Th1 responses in the gut [159,160]. During early infection, surveilling resident mononuclear phagocytes secrete IL-23 which stimulates secretion of IL-17 and IL-22 by type 3 innate lymphoid cells (ILC3s), aiding in neutrophil recruitment and secretion of antimicrobial peptides, respectively [159,161–164]. Epithelial cells, macrophages, and dendritic cells also express members of the intracellular NOD-like receptor (NLR) protein family, including NLRP3 which assembles into a multiprotein complex called the inflammasome [165]. NLRP3 inflammasome formation is initiated by LPS-mediated activation of caspase 11, leading to the induction of caspase 1 and subsequent cleavage of pro-inflammatory cytokines IL-1β and IL-18 [165–167]. Caspase 1-deficient mice infected with *C. rodentium* have increased bacterial loads, more severe colonic pathology, and rapid weight loss, while NLRP3-deficient mice display the same effects albeit milder [168]. Therefore, although caspase 1-mediated immune responses are critical for host resistance to *C. rodentium* infection, NLRP3 may not be the only inflammasome pathway involved.

The host adaptive immune response to intestinal infection is characterized by $CD4^+$ T cells and B cells. C57BL/6 mice depleted of $CD4^+$ T cells or B cells, but not $CD8^+$ T cells, are hypersusceptible to *C. rodentium* infection, develop severe colitis, and have impaired bacterial clearance resulting in systemic dissemination [169,170]. Once recruited, $CD4^+$ T cells become the main producers of IL-22, IL-17A, and IFNγ [159,171–173]. Although T cells are important for bacterial clearance, they are not solely responsible for eradicating *C. rodentium*, which is mostly found extracellularly at the epithelial surface and in the intestinal lumen. B cells are crucial for clearance of *C. rodentium*, since mice lacking B cells do not clear the bacteria or decrease bacterial

burden over time [169,174] Notably, pathogen-specific IgG antibodies, but not IgA or IgM, are required for pathogen clearance [174,175]. IgG recognizes bacterial virulence factors in the intestinal lumen, resulting in selective eradication of virulent pathogens via neutrophil engulfment [176].

1.3 Pathogenesis of A/E bacteria

The study of host-pathogen interactions is a lesson in the following refrain from pioneering microbiologist Abigail Salyers: never underestimate an adversary that has a 3.5-billion-year head start [177]. Bacteria and their human hosts share a deep evolutionary history, forming strong interdependencies since the first animals diverged from their protistan ancestors more than 700 million years ago [178]. Since then, various genetic adaptations and ecological drivers have enabled microorganisms that were formerly benign or even beneficial to become pathogens capable of harming humans.

Virulence is the relative capacity of bacteria to cause damage to a host. Virulence is conferred in the form of genetic attributes termed virulence factors. These are bacterially secreted, membrane-associated, or cytosolic molecules that assist in colonization of a host and are classified based on function as adherence factors, invasion factors, capsular and surface components, endotoxins, and exotoxins. These factors can be acquired via mobile genetic elements (MGEs) that are encoded on plasmids or are integrated at chromosomal sites as prophages, integrative conjugative elements (ICEs), or PAIs and are transferred between bacterial cells by conjugation, transformation, or transduction (**Figure 5**) (reviewed in [179]). MGEs often carry signatures that differentiate them from the rest of the genome such as flanking direct repeats, presence of integrase genes, and a different G+C content [180,181]. In addition to acquiring extraneous DNA from these

sources, commensal *E. coli* can undergo deletions, point mutations, and other DNA rearrangements that contribute to virulence [33].

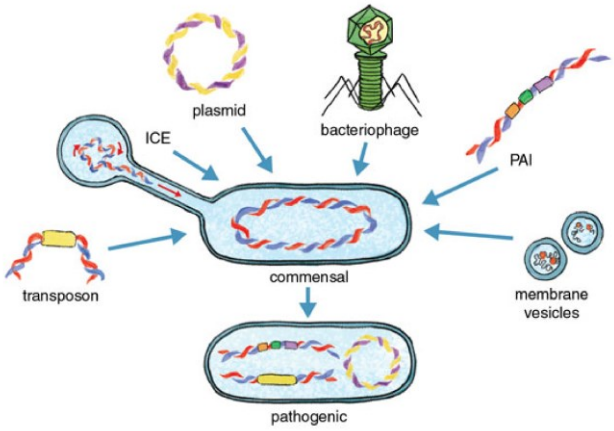

Figure 5. Mobile genetic elements contribute to the evolution of pathogenic bacteria. Mobile genetic elements including transposons, ICEs, plasmids, prophages, PAIs, and DNA within membrane vesicles confer pathogenicity to commensal bacteria. Additions, deletions, point mutations, and rearrangements caused by these elements give rise to pathogenic bacteria capable of causing disease. Figure modified from [182].

However, acquiring a single virulence-associated gene is insufficient to convert commensal *E. coli* to pathogenic *E. coli*, as *Stx* has been shown to be readily transmitted via gene transduction to commensal strains without conferring virulence [33]. Instead, a combination of genes encoding toxins, adhesion and invasion factors, as well as other functions are necessary to make *E. coli* pathogenic. Furthermore, a genome comparison of EHEC with the non-pathogenic *E. coli* K-12 laboratory strain showed a 4.1 megabase (Mb) shared genetic backbone, with an additional 1.34 Mb of DNA found only in EHEC and 0.53 Mb of DNA found only in K-12 [183]. This observation reinforces that pathogen emergence likely involves both a gain of virulence factor-encoding DNA as well as a loss of genes that are detrimental to a pathogenic lifestyle.

Attachment of A/E pathogens and lesion formation involves several virulence factors encoded on MGEs that function in three stages: initial adherence of bacteria to epithelial cells, translocation of bacterial effectors into host cells via the type III secretion system (T3SS), followed by intimate adherence of bacteria and pedestal formation [21]. The initial attachment of EPEC to enterocytes in the small intestine involves BFP, encoded on the pEAF present in tEPEC strains [8,9]. BFP are rope-like fimbriae that interact with receptors on host cell surface or with other EPEC bacteria to form microcolonies for localized adherence [9,184–186]. Thus, BFP are important adhesins for EPEC pathogenicity, as human volunteers infected with EPEC lacking the pEAF plasmid were at lower risk for developing diarrhea [187]. EHEC and *C. rodentium* do not possess the pEAF but do have homologous gene clusters, *hcp* in EHEC and *cfc* in *C. rodentium*, that encode type IV pili functionally similar to BFP [188,189].

1.3.1 Locus of enterocyte effacement

The genes required for A/E lesion formation are encoded on a 35 kilobase (kb) PAI termed the locus of enterocyte effacement (LEE) [190]. The LEE is divided into five polycistronic operons designated *LEE1-LEE5* encoding a T3SS, several effector proteins, chaperones, and gene regulators [191–193]. LEE homologs are present in all A/E pathogens and, although not identical, are highly conserved in gene repertoire and organization, suggesting a common origin [191,194–197]. Similar to other PAIs, the LEE has a lower G+C content than that of the *E. coli* chromosome as a whole [191,194,195]. Sequence comparison analysis has shown that genes encoding the T3SS have greater than 95% identity among A/E pathogens, however several LEE-encoded genes demonstrate greater variability than would be expected for clonal divergence of *E. coli* strains [194], many of which encode proteins that are involved in interactions with the host. The

contribution of the LEE to the characteristic A/E phenotype was confirmed when the LEE from EPEC strain E2348/69 introduced into non-pathogenic *E. coli* K-12 enabled the formation of A/E lesions [198]. When the EHEC LEE was introduced, however, the A/E phenotype could not be reproduced [199], therefore suggesting that the LEE adapts to different genetic backgrounds and host specificities. Thus, it is hypothesized that A/E pathogens individually acquired the LEE PAI via horizontal gene transfer, however it is not clear where the LEE originated, the exact mechanism of transfer, and when this transfer occurred [200].

A/E pathogens have complex regulatory networks, integrating many host- and microbiota-derived signals to sense the gastrointestinal environment and control expression of the LEE genes. Such signals include microbiota-liberated sugars and metabolites in the gut lumen, molecular oxygen at the gut epithelium, and host hormones, among others [201]. More than 40 regulators of LEE transcription have been described, many of which converge on the master LEE regulator, LEE-encoded regulator (Ler), encoded in the *LEE1* operon [201,202]. Production of Ler activates expression of *LEE2-LEE5* to produce other LEE-encoded proteins [202,203]. In EPEC, but not EHEC, Ler expression is positively regulated by plasmid-encoded regulator (Per), encoded on the pEAF [204]. Ler-mediated expression of the LEE is fine-tuned by the LEE-encoded positive and negative regulators GrlA and GrlR, respectively [205]. GrlA binds directly to the promoter of *LEE1*, driving the expression of *ler* and subsequently the entire LEE [206,207]. Conversely, GrlR inhibits this activation by binding directly to GrlA and preventing its interaction with DNA [208]. In addition to the chemical signals mentioned above, physical cues can regulate LEE expression [209]. Notably, bacterial attachment to host cells mediated by LEE- and non-LEE-encoded adhesins increases GrlA activity, further activating the *LEE1* promoter [210]. This effect is further amplified in response to laterally acting fluid flow forces at the epithelial surface [210].

1.3.1.1 Type 3 secretion system

T3SSs are highly conserved protein secretion machinery widely found in gram-negative bacteria, including human and plant pathogens of the *Salmonella*, *Shigella*, *Yersinia*, *Pseudomonas*, and *E. coli* genera [211]. The T3SS is a large, multiring protein complex, referred to as an injectisome, that spans the inner and outer bacterial membranes, with a hollow needle-like filament terminating in a translocation pore that is inserted into the host cell plasma membrane (**Figure 6**) [212]. The T3SS is critical for virulence, functioning as a molecular syringe to inject bacterial proteins directly into the target host cell cytoplasm [211].

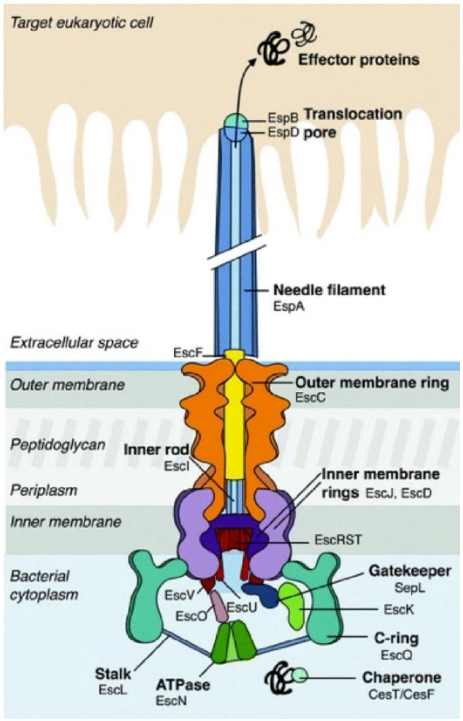

Figure 6. Schematic representation of the A/E pathogen T3SS. The T3SS consists of a basal body, export apparatus, needle complex, and several cytoplasmic components. The basal body is comprised of inner (purple) and outer (orange) membrane rings. The needle complex has an inner rod (yellow), needle filament (blue) and translocation pore through which bacterial effector proteins are delivered across both bacterial and host membranes directly into the cytosol of host cells. A cytoplasmic ATPase (green) powers injectisome assembly and drives translocation of effector proteins. Figure reproduced from [213].

Pathogens secrete a large number of proteins via the T3SS which are largely categorized as effectors and translocators [212]. Effectors are proteins that modulate host cell biology and subvert host defense mechanisms to facilitate bacterial colonization and disease development upon injection into host cells [214]. Translocators are required for the injection of effectors by assembling the injectisome complex [214]. In comparison to translocators which are highly conserved, secreted effectors vary in number and function among bacterial species, reflecting unique repertoires depending on pathogenesis and host niche [215].

To ensure effector proteins are efficiently injected into host enterocytes and not the extracellular milieu, T3SS secretion is highly hierarchical. Assembly of the EPEC and EHEC T3SS follows an "inside out" process, beginning with formation of the inner rings and export apparatus enabling secretion of the outer ring and subsequently needle complex components [215,216]. Secreted EspA self-polymerizes to form the filament over the EscF needle structure, resulting in a hollow tube with an inner diameter of 2-3 nm and varying in length [217–219]. Upon contact with the target host cell, translocators EspB and EspD are inserted into the plasma membrane to form the translocation pore [219,220]. To prevent aberrant secretion of effectors before the injectisome is fully assembled, the LEE-encoded proteins SepD and SepL function as a translocation "checkpoint." Although it is unclear how the T3SS distinguishes translocators from effectors, mutations in either *sepD* or *sepL* abolish secretion of translocators and results in hypersecretion of effectors [192,214]. The ATPase EscN binds to the type III secretion signal of substrates and their chaperones, mediating the unfolding of effectors and driving translocation through the injectisome [221–224]. In EPEC, deletion of *escN* resulted in complete abrogation of type III secretion [225].

Like components of the injectisome, effectors have been shown to be hierarchically translocated into host cells [226,227]. In general, effectors contain an approximately 20 amino acid

N-terminal signal sequence that is necessary and sufficient to target them for type III secretion [228–231]. Additionally, a chaperone-binding domain is located downstream of the N-terminal secretion signal [232]. The LEE encodes eight chaperone proteins that are categorized based on their binding substrates: effectors, translocators, or needle components, and have been shown to be critical for *C. rodentium* virulence in mice [205,212,233–235]. Chaperones function to promote type III secretion by increasing protein stability, mediating substrate recognition and targeting, preventing premature oligomerization or nonspecific interactions, and enforcing proper secretion order [217,218,222,231–241].

1.3.1.2 LEE-encoded effectors

A/E pathogens share seven effectors encoded in the LEE: translocated intimin receptor (Tir), mitochondrial-associated protein (Map), EspB, EspF, EspG, EspH, and EspZ [246]. Most of the LEE-encoded effectors, except EspZ, have important functions in destabilizing the physiology of the intestinal epithelium, triggering cytoskeletal reorganization, and inducing cytotoxicity and electrolyte imbalance that ultimately leads to diarrhea [192,247]. The contributions to pathogen virulence and characterized functions of the LEE-encoded effectors are denoted in **Table 1**.

Table 1: LEE-encoded effectors

Effector	Deletion effect on virulence in mice	Characterized function(s) during infection
Tir	Complete attenuation [205]	Bacterial adhesion, actin polymerization, TJ disruption, cell detachment, microvilli effacement, immune suppression, and others [248–256]
Map	Slight attenuation [205]	TJ disruption, filopodia formation, mitochondrial dysfunction, microvilli effacement, SGLT-1 inactivation [25,29,257–259]
EspB*	Complete attenuation [205]	Inhibits phagocytosis, microvilli effacement, actin disruption, translocation pore formation [260–263]
EspF	Defect in early colonization [205]	Promotes apoptosis, TJ disruption, microvilli effacement, SGLT-1 inactivation, mitochondrial dysfunction, pedestal maturation, membrane remodeling, and others [264–267]
EspG family	Defect in early colonization [205]	Microtubule disruption, TJ disruption, stress fiber formation, inhibition of protein secretion, and others [268–272]
EspH	Slight attenuation [205]	Modulates actin dynamics, cytoskeleton disruption [273,274]
EspZ	Significantly decreased colonization [205]	Control of effector translocation, inhibits cell death [275,276]

*EspB is classified as both a translocator and an effector.

The rapid onset of diarrhea during infection is likely due to cooperative function of Tir, Map, and EspF which inhibit the sodium-D-glucose transporter 1 (SGLT-1), a major mechanism of glucose absorption in the small intestine [25]. Map and EspF also inhibit sodium absorption by the sodium-hydrogen exchanger 3 (NHE3) [266,277]. EspG1/2 prevents membrane association of the down-regulated in adenoma (DRA) Cl^-/OH^- exchanger found on intestinal epithelial cells and EspF and EspG induce mislocalization of aquaporins (AQP), resulting in reduced chlorine uptake and water absorption [24,26]. The combined inhibition of these transporters results in accumulation of solutes in the lumen of the GI tract and subsequent substantial water loss from the

mucosa. Furthermore, Tir, Map, EspB, and EspF induce effacement of the microvilli, reducing the absorptive surface area of the epithelium and exacerbating diarrhea [260].

Map and EspF have been shown to function synergistically to disrupt tight junctions (TJs) between cells of the epithelium [29]. EspG induces microtubule disruption also contributing to TJ disruption [272]. TJs are intercellular junctional complexes that mechanically connect adjacent epithelial cells, forming a regulatable barrier between the luminal and serosal sides of the epithelial monolayer [278]. TJs maintain epithelial cell polarity by restricting free movement of apical and basolateral components, thereby aiding in the directional movement of water, electrolytes, and nutrients [278]. Thus, the disruption of TJs increases intestinal permeability and contributes to A/E pathogen-induced diarrhea [35].

LEE-encoded effectors are also involved in the formation of pedestals that are associated with the characteristic A/E lesion. Tir, the first effector to be translocated by the T3SS, is inserted as a hairpin-like structure into the host cell plasma membrane [248,279–282]. The extracellular domain of Tir binds intimin, an integral component of the bacterial cell wall that is also encoded by the LEE [280,281]. This intimate binding interaction triggers actin nucleation and subsequent pedestal formation. The intracellular N- and C-terminal cytoplasmic domains of Tir interact with several focal adhesion and cytoskeletal proteins, linking the extracellular bacterium to the host cell cytoskeleton [283,284]. These interactions lead to the formation of actin-rich pedestals beneath adherent A/E bacteria. EHEC relies on the delivery of an additional T3SS-translocated effector protein, EspFU, to mediate binding to the host cytoskeleton [285–287]. In addition, EspG triggers the formation of actin stress fibers and the destruction of microtubule networks beneath adherent bacteria in intestinal epithelial cells [288,289]. EspH localizes to the host cell membrane and modulates the host actin cytoskeleton structure, affecting pedestal formation [273]. However,

EspH does not play a critical role in *C. rodentium* infection, as mutant strains exhibit only slight attenuation [205,290].

Both Map and EspF target to the host mitochondria: organelles that play a crucial role in regulating programmed cell death (apoptosis). Map and EspF disrupt mitochondrial membrane potential, triggering apoptosis [264,291,292]. EspZ is a transmembrane protein that targets the plasma and mitochondrial membranes of infected host cells [276,293]. EspZ is the second most abundant effector protein after Tir, and also the second effector translocated into the host cell by EPEC, suggesting it may have a key role during infection [227,294]. Interestingly, EPEC Δ*espZ* have impaired effector and increased translocator secretion, indicating that EspZ is involved in controlling translocation through an unknown mechanism [275]. In addition, EspZ has been shown to protect against EPEC-induced cell death [276]. Although the exact process by which EspZ blocks cell death remains uncharacterized, a hypothesized mechanism involves inhibition of the translocation of pro-apoptotic effectors Map and EspF [295]. This is one of several examples of contradictory effector functions, highlighting the importance of highly coordinated spatiotemporal effector translocation.

1.3.2 Non-LEE-encoded effectors

T3SS-translocated effectors encoded outside the LEE have been identified in all A/E pathogens. Complete genome sequencing of EPEC and EHEC in the 2000s led to the identification of more than 400 and 200 predicted effector protein sequences in EPEC and EHEC, respectively [296,297]. Of these, 21 EPEC and 39 EHEC sequences were confirmed to be T3SS-secreted proteins [296,297]. In *C. rodentium*, 23 additional effector genes are located outside of the LEE [102]. Non-LEE-encoded (Nle) effector genes are found within several prophage and integrative

element PAIs, with some effectors having duplicated genes or paralogs encoded on more than on PAI [102,296,297]. Nle effectors have diverse, often compensatory, functions. **Table 2** presents several characterized non-LEE-encoded effectors, their presence in A/E pathogens, contribution to bacterial virulence, and described functions during infection.

Table 2: Characterized non-LEE-encoded effectors

Effector	Gene present in: EPEC O127:H6	EHEC O157:H7	*C. rodentium*	Deletion effect on virulence in mice	Characterized function(s) during infection
NleA/ EspI	+	+	+	Complete attenuation [205]	Inhibits protein secretion, inhibits inflammasome activation [298,299]
NleB family	+	+	+	Reduced colonization and pathology [300] Severe attenuation [301]	Inhibits apoptosis [302,303]
NleC	+	+	+	No effect [300]	Inhibits NF-κB signaling [304–306]
NleD	+	+	+	No effect [300]	Inhibits MAPK signaling [304]
NleE	+	+	+	No effect [300] Severe attenuation [307]	Inhibits NF-κB signaling; inhibits DNA repair [308,309]
NleF	+	+	+	Reduced bacterial fitness [310]	Inhibits apoptosis [311]
NleG family	+	+	+	Unknown	Ubiquitin ligase [312,313]
NleH family	+	+	+	Moderate colonization defect [314]	Inhibits cell death and NF-κB signaling; pro-inflammatory [315–317]
Cif	+	–*	–	Unknown	Promotes cell cycle arrest and actin stress fiber formation [318]
EspFU/ TccP	–	+	–	Moderate colonization defect in rabbits [319]	Actin polymerization, intimate bacterial adhesion [285,320–322]
EspJ	+	+	+	Increased bacterial colonization [323]	Inhibits phagocytosis [324–326]
EspL family	+	+	+	Moderate colonization defect [327]	Inhibits necroptosis and inflammasome [327]
EspM family	+	+	+	Moderate attenuation [328]	Disrupts cell-cell junctions, promotes stress fiber formation, inhibits pedestal formation [329,330]

*Cif is present in some EHEC strains, but not EHEC O157:H7.

A/E pathogens inject multiple Nle effector proteins into the host cell that specifically target innate immune factors. It must first be acknowledged that A/E pathogens induce both pro- and anti-inflammatory responses. Several studies have shown variable, often conflicting, effector activity in the inhibition of inflammatory signaling. Therefore, elucidating the exact contribution of individual effectors to provoking an immune response seems to be highly dependent on the context of infection. This section will present the current accepted functions of Nle effectors in modulating the immune response, but these discrepancies, and the often-limited validation of effector function *in vivo*, should be noted.

In general, host cells respond to A/E pathogen infection by increasing NF-κB p50/p65 heterodimer binding affinity to DNA to initiate expression of genes involved in inflammation and cytokine production [331]. NleE inhibits activation of NF-κB by preventing activation of IKKβ and degradation of the NF-κB inhibitor IκB [332]. This activity by NleE is enhanced by, but not dependent on, NleB [332]. NleC and NleD are also potent inhibitors of the NF-κB and MAPK innate signaling networks via their proteolytic activities. NleC targets the NF-κB p65 subunit for degradation, while NleD cleaves the MAPK enzymes JNK and p38 [304,305,333,334]. Neither *nleC* nor *nleD C. rodentium* deletion mutants were attenuated in mice [300]. However, mice infected with a *C. rodentium* Δ*nleC* mutant did have increased pathology, suggesting that NleC assists in reducing the severity of colitis during infection [334]. Considering the complexity of innate immune signaling pathways, it is not surprising that A/E pathogens have evolved mechanisms to target specific pathways at multiple levels. Although NleC and NleE have apparent functional redundancy, secreted levels of pro-inflammatory cytokine IL-8 were significantly higher upon infection with a double *nleC*/*nleE* deletion mutant than either single deletion mutant, indicating that NleC and NleE act synergistically to inhibit NF-κB-mediated IL-8 production [304–

306,334,335]. During infection, NleB functions as a translocated N-acetyl-D-glucosamine (O-GlcNAc) transferase and is believed to inhibit NF-κB activation by modifying host glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [336]. NleB-mediated GAPDH O-GlcNAcylation prevents its interaction with TNF receptor associated factor 2 (TRAF2), a protein required for TNF-α-mediated NF-κB activation [336].

Apoptosis is a normal homeostatic process to maintain cell populations in tissues. The gastrointestinal epithelium has a cell high turnover rate, involving a cycle of proliferation, differentiation, and apoptosis [337]. During infection, however, induction of cell death is often a strategy used by the host to remove infected cells along with any associated bacteria, accelerating clearance of a colonizing pathogen. To overcome this, A/E pathogens encode effectors with anti-apoptotic activity to promote cell survival. Despite the presence of pro-apoptotic factors, namely Map and EspF, EPEC-infected cells do not develop late apoptotic symptoms [292,338,339]. NleH1/2 block caspase-3 activation, nuclear condensation, and membrane blebbing to promote cell survival [316]. NleH1/2 possess kinase activity that is essential for the inhibition of NF-κB, however the targets of this phosphorylation remain to be identified [317,340]. Additionally, cycle inhibiting factor (Cif) is as an Nle effector that blocks intestinal cell proliferation by causing cell cycle arrest during interphase [318,341].

1.3.2.1 Non-LEE-encoded effector A

In 2004, non-LEE-encoded effector A (NleA) was the first A/E pathogen effector described to be encoded on a PAI separate from the LEE [342]. Concurrently, the same protein was separately identified and named EspI [290], but for simplicity will only be referred to as NleA henceforth. Initial *in vivo* mortality experiments demonstrated that NleA is required for virulence [342].

Genetically susceptible C3H/HeJ mice infected with *C. rodentium* Δ*nleA* have mild symptoms but ultimately survive infection, in contrast to mice infected with WT *C. rodentium* that display severe pathology and high rates of mortality [342]. In addition, the gene coding for NleA is absent from non-pathogenic *E. coli* and is preferentially found in strains associated with outbreaks of human disease [342]. NleA is translocated into host cells by the T3SS [342] and, using real-time translocation analysis, was found to be the third most efficiently translocated EPEC effector after Tir and EspZ [294]. Furthermore, Tir, NleA, and EspZ were determined to be the only effectors essential for *C. rodentium* colonization in mice [343]. Altogether, these findings indicate NleA contributes significantly to virulence, however the complete function of NleA over the course of infection is not well-characterized.

Several host cell pathways subject to modulation by NleA have been identified (**Figure 7**). Following translocation, NleA localizes to the secretory pathway, where it binds directly to the Sec24 subunit of the coat complex type II (COPII) protein complex to disrupt anterograde vesicular trafficking between the endoplasmic reticulum (ER) and Golgi apparatus [299,342]. Sec24 functions as the primary cargo-determining subunit of the COPII coat [344]. NleA's interaction with Sec24 was shown to prevent both COPII vesicle budding and cargo packaging in a dose-dependent manner *in vitro* [299]. Mutant NleA with diminished binding to Sec24 did not impede COPII trafficking *in vitro* and *C. rodentium* expressing this mutant was completely avirulent *in vivo* [345]. NleA is also involved in disruption of TJs between epithelial cells, resulting in increased barrier permeability and fecal water content in infected mice [346,347]. NleA mediates TJ disruption through its inhibition of COPII-mediated protein trafficking, preventing renewal of tight junction proteins during infection [346]. Functional cooperation of NleA, Map, and EspF is required for complete disruption of TJ integrity, in a mechanism that is not yet

understood [29,348]. Finally, NleA has been found to interact with ubiquitinated NLRP3 in macrophages and epithelial cells, preventing its de-ubiquitination and subsequent NLRP3 inflammasome assembly [298]. NleA inhibition of inflammasome maturation resulted in subdued IL-1β secretion, an essential inflammatory cytokine involved in host defenses against enteric pathogens [298].

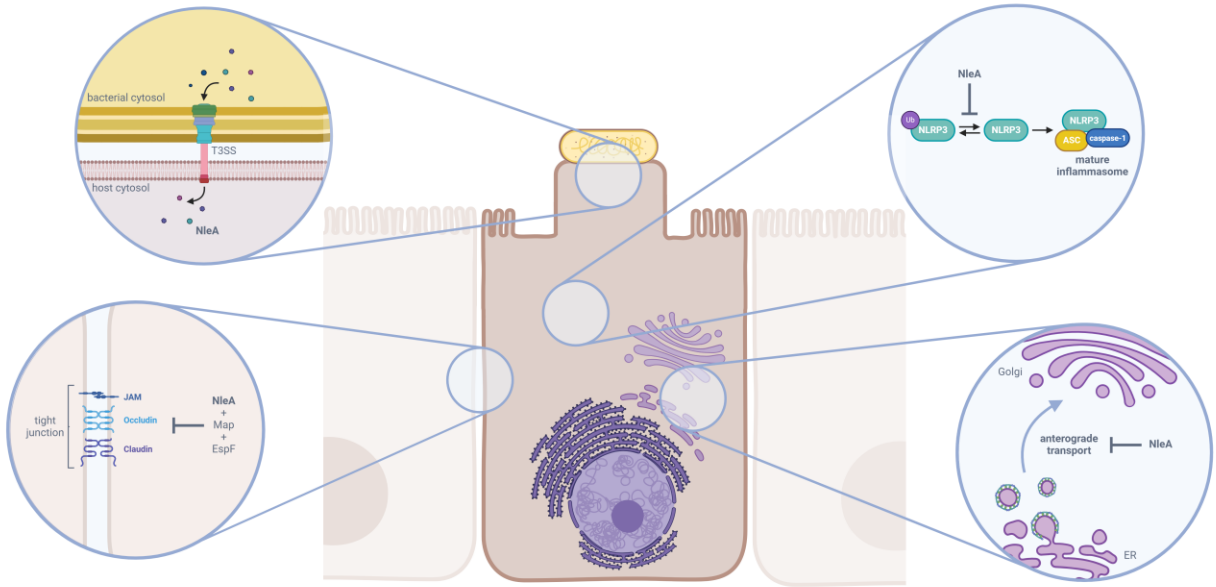

Figure 7. NleA has several characterized functions upon translocation into the host cell. Schematic of localizations of NleA within the host epithelial cell. NleA is translocated by the T3SS (top left). It has been shown to disrupt tight junctions between epithelial cells (bottom left) primarily via its inhibition of anterograde protein trafficking by interacting with the COPII vesicle coat (bottom right). In the cytoplasm, NleA prevents de-ubiquitination of NLRP3 to inhibit inflammasome activation (top right). Created with BioRender.com.

Injected effectors form robust intracellular networks that can withstand substantial contractions while maintaining pathogenicity and host adaptation [343]. In general, effectors hijack cellular processes in a coordinated manner to alter host cell biology to favour pathogen survival and replication. Effectors enact these functions through a variety of mechanisms,

including altering epithelial cell function to induce water loss and diarrhea, modulating inflammatory and immune responses to avoid detection or clearance, and promoting or inhibiting host cell survival [349]. The function of an effector can be dependent on the activity of another, for example EHEC actin polymerization by Tir is dependent on interaction with the adapter protein EspFU [285,287]. Effectors can also have complementary functions. NleC and NleE both function as inhibitors of the NF-κB signaling cascade. In addition, many effectors from bacterial pathogens of both plants and animals undergo host-mediated post-translational modifications (PTMs) following translocation that regulate effector function, including enzymatic activity, or aid in subcellular localization (reviewed in [350,351]).

1.4 Post-translational effector protein modification

PTMs are chemical modifications of a polypeptide chain that occur after DNA has been transcribed into RNA and then translated into protein. These chemical alterations influence the properties of the modified residues and nearby polypeptide regions, including protein net charge, conformation, binding properties and, ultimately, function. PTMs extend the 21 proteinogenic amino acid repertoire of eukaryotes by producing more than 400 different types of modifications, vastly increasing the functional diversity of the proteome [352]. Importantly, most known PTMs are substoichiometric, meaning they do not occur on all molecules of a given protein within the cell [353]. This further enables the cell to quickly adapt by increasing or decreasing the amount of modified protein without requiring time and energy intensive protein synthesis or degradation. The majority of PTMs are reversible and dynamic; they are transiently added or removed from a polypeptide chain by specialized enzymes in response to cell stimuli. Others are irreversible or occur spontaneously. Large-scale protein modification studies in eukaryotes have identified tens

of thousands of predicted modification sites [353]. In contrast, modifications of bacterial proteins are much less abundant but more diverse, making their analysis challenging. Further limited is the identification and functional characterization of host-mediated PTM of bacterial proteins. The host-specific covalent modifications of injected bacterial proteins described to date include phosphorylation, ubiquitination, SUMOylation, and lipidation [351].

Phosphorylation is the process by which specific kinases attach negatively charged phosphate groups to a phosphorylatable amino acid residue of their substrate. These residues include histidine, serine, threonine, and tyrosine. The *Helicobacter pylori* injected effector CagA is modified by host cytosolic kinases [354]. Tyrosine phosphorylation of CagA results in its activation and subsequent changes related to cytoskeletal rearrangements and induction of cell elongation [355,356]. Following its delivery into host cells, EPEC Tir is phosphorylated on two serine residues, S434 and S463, and two tyrosine residues, Y454 and Y474 [282,357,358]. This host-mediated phosphorylation has been shown to be critical for EPEC Tir's actin signaling function *in vitro* [282]. In contrast, EHEC Tir lacks Y474 or an equivalent tyrosine and generates pedestals independently of phosphorylation [359–361]. EPEC expressing EHEC Tir failed to generate actin-rich pedestals unless Y474 and surrounding amino acids are substituted into the protein sequence [362,363], thereby indicating that EPEC and EHEC have evolved different mechanisms for A/E lesion formation. In addition, serine phosphorylation of EPEC Tir at S434 and S463 leads to, in contrast to the CagA-induced phenotype, inhibition of cell elongation [364,365].

Ubiquitination is a PTM involving the addition of a small ubiquitin (Ub) molecule on lysine residues, and rarely cysteine or serine, in the N-terminus of the protein substrate [366]. Although attachment of a single Ub can affect localization and activity of a target protein [367], the Ub

subunit can itself be modified by addition of Ub, forming a poly-Ub chain, which can target a protein for degradation by the host proteasome [366]. Because ubiquitination is involved in many cellular functions including protein degradation, the cell cycle, vesicle trafficking, and immune responses, some bacterial effectors have evolved to exploit this by binding to or modifying host ubiquitination components while other effectors are subjected to ubiquitination [368,369]. *Salmonella* Typhimurium T3SS-translocated effectors SopA, SopB/SigD, SopE, and SptP are ubiquitinated following their delivery into the host cell [370]. Time-dependent ubiquitination of SopE and SptP lead to their degradation by the host protease machinery [371]. Host mediated ubiquitination of SopB, however, serves as a non-proteolytic signal and rather contributes to its intracellular localization and effector function [370,372]. SopB ubiquitination is required for it to relocate from the plasma membrane to the *Salmonella*-containing vacuole (SCV) [373]. Delivery of SopB to the SCV by the host-ubiquitin machinery concentrates SopB activity at this location, ultimately allowing *Salmonella* to escape degradation by the lysosome [373,374]. SopA is both modified by ubiquitin and functions as a ubiquitin ligase [375,376]. Although ubiquitination of SopA leads to its proteasomal degradation, it also functions as a signal regulating *Salmonella* escape into the host cell cytosol where it can rapidly multiply and disseminate [375].

SUMOylation is target substrate modification of one or more lysine residues by covalent attachment of a member of the small ubiquitin-like modifier (SUMO) family of proteins [377]. Although many pathogenic bacteria negatively regulate the host SUMOylation system, there are few cases of SUMOylated effectors and none described in enteric or A/E pathogens [351,378,379]. AmpA of the zoonotic pathogen *Anaplasma phagocytophilum* is poly-SUMOylated to promote intracellular bacterial survival [380].

Lipidation is an important PTM in which lipid moieties are attached to proteins. Lipidation drastically increases protein hydrophobicity, resulting in changes to conformation, stability, membrane association, binding affinity, and localization [381]. Prenylation, one type of lipidation, is the irreversible addition of farnesyl and geranylgeranyl isoprene groups to specific cysteine residues within a C-terminal CaaX motif [381]. Prenylation of *Salmonella* Typhimurium effector SifA by geranylgeranyl addition was shown to be required for maintenance of the membrane surrounding the SCV [382]. Although modification by prenylation has not been identified for any A/E pathogen effectors, *in silico* analyses have determined it may be a conserved mechanism for effector modification among animal pathogens [383]. A screen of microbial genomes for C-terminal CaaX-motif-containing proteins identified 56 predicted prenylated proteins, 3 of which were identified as EHEC translocated Nle effectors from the NleB and NleG families [383].

1.4.1 Glycosylation

Glycosylation is a ubiquitous and remarkably diverse PTM; in addition to forming important structural features, glycoconjugates modulate or mediate many physiological and pathophysiological cell functions [384]. Unlike the genome or proteome, the glycome is produced in a non-templated manner and reflects a given cell's unique gene expression [385]. Glycosylated proteins are found on the cell surface, abundantly in the nucleus and cytoplasm, and as secreted macromolecules [385]. Glycosylation refers to the addition of sugars to proteins to create a vast number of unique glycan structures that modulate protein function. Multiple enzymatic site preferences and the use of α or β conjugations create further diversity in precisely where and how these sugars are linked to each other, generating the potential for 10^{12} possible branched glycan structures [386].

In eukaryotic cells, glycosylation occurs primarily in the ER and Golgi, with terminal processing mainly confined to the cis-, medial-, and trans-Golgi compartments [385]. Import of proteins into the ER is typically co-translational: the growing polypeptide chain is threaded across the ER membrane via the Sec61 channel as it is synthesized by the ribosome [387]. Additionally, proteins synthesized in the cytoplasm can be post-translationally transported into the ER by accessory proteins that feed the polypeptide chain into the Sec61 complex and drive translocation [388]. Protein glycosylation occurs in two major forms: N-linked and O-linked (**Figure 8**) [389]. N-linked glycosylation consists of the attachment of sugars to the nitrogen atom of asparagine (Asn) [389]. O-linked glycosylation is the attachment of sugars to the hydroxy groups of primarily serine (Ser) or threonine (Thr) residues [389].

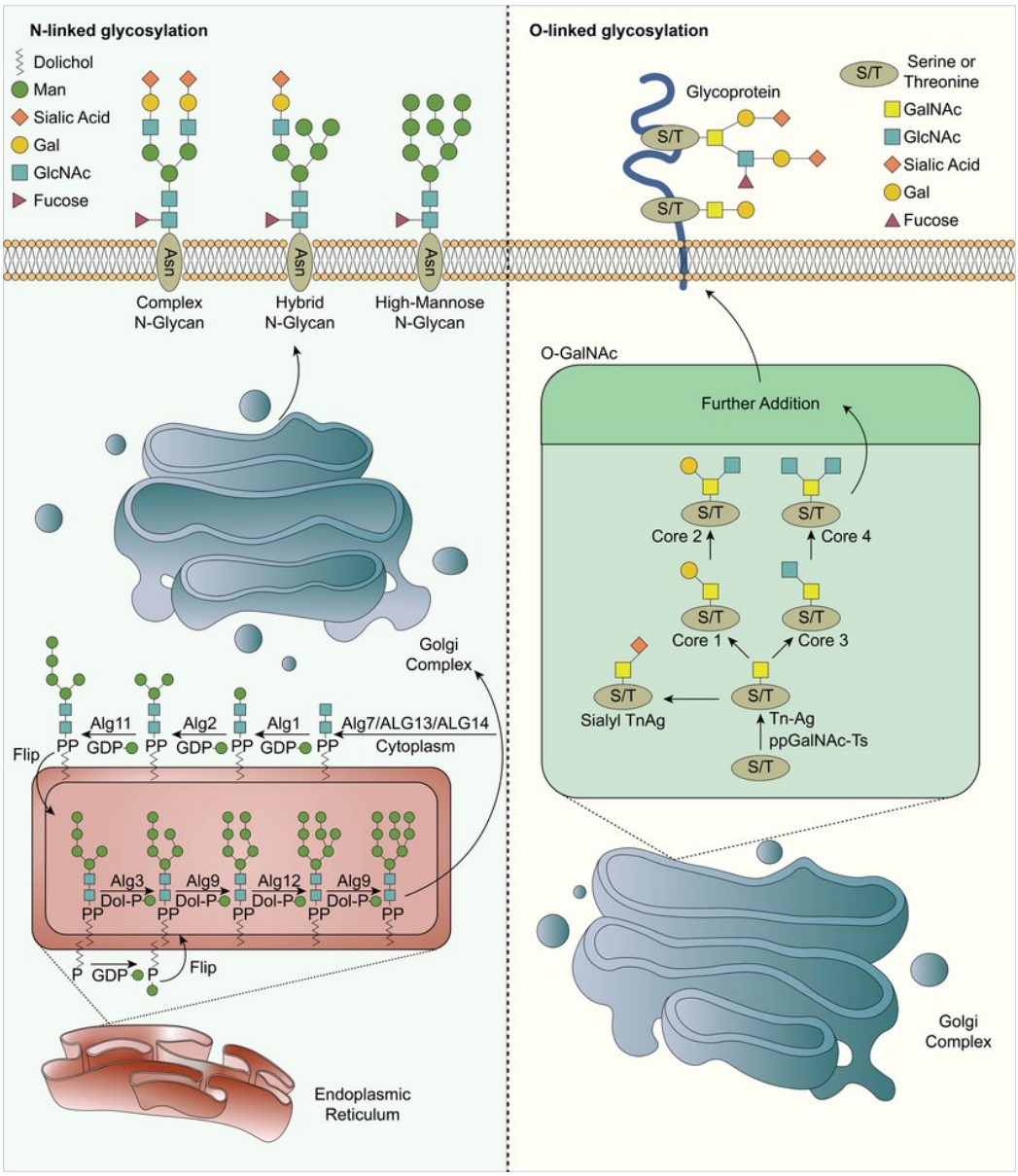

Figure 8. Eukaryotic subcellular organization of glycosylation. Protein glycosylation takes place in the ER and Golgi, nucleus, cytoplasm, and mitochondria. In the cytoplasm and nucleus, OGT performs O-GlcNAcylation. N-glycosylation begins with synthesis of a lipid-linked oligosaccharide linked to dolichol. By the action of invertase, this precursor undergoes transbilayer translocation across the ER membrane. Following further sugar additions in the ER lumen, the oligosaccharide is transferred to an asparagine residue of a newly synthesized protein. N-glycoproteins leave the ER and enter the Golgi complex, where they are additionally modified to form complex N-glycoproteins that are ultimately transported to the cell membrane. GalNAc-type O-glycosylation is initiated by GALNTs in the Golgi. Following the addition of an initial GalNAc residue to serine or threonine, further glycan processing, including core extension, elongation, and capping, occur as the protein traffics through the Golgi. Figure reproduced from [390].

N-linked glycosylation is preceded by synthesis of a preformed oligosaccharide precursor in the cytosol, linked to the ER membrane by the carrier lipid molecule dolichol [391]. The dolichol-linked oligosaccharide is then flipped to face the ER lumen followed by transfer *en bloc* to a protein substrate [392]. Transfer of the precursor is mediated by the oligosaccharyl transferase in a single enzymatic step to an asparagine residue of the Asn-X-Ser/Thr consensus sequence [392]. The nascent glycoprotein undergoes further processing in the ER before being transported by COPII-coated transport vesicles to the Golgi for maturation [385]. N-linked glycans share a common core region consisting of two N-acetylglucosamine (GlcNAc) and three mannose (Man) residues, however, this core structure can be further modified to generate diversity among N-glycans. Glycosyltransferases and glycosidases add or remove sugar residues, respectively, in a multi-step process that is controlled by substrate availability, enzyme activity, gene transcription levels, and enzyme location within the Golgi [385]. This process ultimately results in three broad classes of N-linked oligosaccharides: complex, high-mannose, and hybrid (**Figure 8**) [391]. The terminal region of complex N-glycans contains GlcNAc, galactose (Gal), fucose (Fuc), and sialic acid sugars [391]. In contrast, high-mannose N-glycans are not fully trimmed and contain additional Man residues [391]. Hybrid N-glycans have a combination of complex and high-mannose branches [391].

O-linked glycosylation typically begins with a GlcNAc or GalNAc sugar linked to a Ser or Thr residue [393]. GalNAc-linked glycans, often called mucin-type O-glycans, are abundantly found on extracellular and secreted glycoproteins, including mucins [393]. As described earlier, mucins form a crucial barrier to protect the gut epithelium from external stress, microbial infection, and self-recognition by the immune system (**Chapter 1.2**). Mucin-type O-glycan synthesis is initiated by polypeptide GalNAc transferases (GALNTs) in the Golgi [393]. These enzymes are

sequentially and functionally conserved across species and are differentially expressed over tissue and time, indicating strict regulation [394–396]. There are up to 20 different known GALNT isoforms that differ in their specificity for amino acid motifs which determines how and where O-glycans are attached to the protein [396]. However, O-linked glycosylation lacks a known amino acid consensus motif and GALNTs often have promiscuous binding affinity, allowing for some diversity among produced carbohydrate structures [385]. Following initial attachment of GalNAc to Ser/Thr of the completely folded protein, further non-templated sequential addition of sugars occurs as the protein moves through the cis-, medial-, and trans-Golgi compartments [385]. Addition of GalNAc to Ser/Thr forms with an α linkage forms the Tn antigen, which can be extended by Gal, GlcNAc, or GalNAc to produce four major core structures (cores 1-4) (**Figure 8**) [397]. Extension of O-GalNAc with Gal forms the core 1 O-glycan, also called the Thomsen-Friedenreich (T) antigen, the most common O-GalNAc glycan [393]. Core structures are further modified to form complex glycans with variable termini of sialic acid, GalNAc, or Fuc [397]. In general, O-linked glycans are smaller but more diverse than N-linked glycans [396].

GlcNAc-linked O-glycans are typically found on intracellular glycoproteins present in the nucleus, mitochondria, and cytoplasm [385]. Unlike mucin-type O-glycans, addition and removal of GlcNAc is mediated by O-linked GlcNAc transferases (OGTs) and O-linked GlcNAcases (OGAs), respectively [398]. These enzymes, located in the nucleus and cytoplasm, enable rapid addition and removal of GlcNAc from protein substrates, and this process is thought to regulate many cellular functions, including cell metabolism [385].

The functions of glycosylated proteins are nearly as diverse as their structures. Many immune receptors expressed on the surface of innate and adaptive immune cells recognize glycan-containing molecules on the surface of microorganisms, including LPS, peptidoglycan teichoic

acids, and capsular polysaccharides [385]. Moreover, glycosylated adhesion molecules on the surface of endothelial cells are crucial for the trafficking and recruitment of leukocytes to sites of tissue injury [385]. Glycoproteins are also important for B cell and T cell differentiation, functioning as cell surface and secreted proteins, contributing to cell-cell interactions, and recognizing glycan-containing antigens [385,399]. For example, CD43 and CD45 are N- and O-glycosylated proteins abundantly expressed on the surface of T cells that regulate multiple T cell functions including cellular migration, T cell receptor signaling, cell survival, and apoptosis [400]. Glycosylation of immunoglobulins, secreted proteins of B cells, impacts antibody function depending on the branching structure and terminal sugars of the glycan. In fact, glycosylation can determine whether an antibody is pro-inflammatory, for instance IgG with galactose-deficient N-linked glycosylation, or anti-inflammatory, such as IgG with sialylated N-linked glycosylation [401,402].

Historically, glycosylation was believed to be restricted to eukaryotes, however, both N- and O-linked glycosylation pathways have been well-established in commensal and pathogenic bacteria [403]. Glycoproteins have been identified in: *P. aeruginosa* [404,405], *Neisseria meningitides* [406], *Haemophilus influenzae* [407], *Campylobacter jejuni* [408], *Mycobacterium tuberculosis* [409], *Streptococcus parasanguis* [410], diffusely adhering *E. coli* (DAEC) [411], and enterotoxigenic *E. coli* (ETEC) [412]. Most bacterially glycosylated proteins are not secreted into the extracellular milieu, instead remaining attached to the cell [413]. Flagellar components are the most well-studied bacterial glycoproteins, where glycosylation has been shown to be important for assembly, motility, host cell recognition, and adhesion to epithelial cells [414]. Bacterial pathogens possess an arsenal of secreted toxins and effector proteins that function to modulate host cell biology to facilitate infection. As more bacterial effector functions are characterized, there is

emerging evidence of effectors that bear glycosyltransferase activity to add glycans to host molecules. To date, only one effector of A/E pathogens, NleB, has been identified as a glycosyltransferase that modifies host proteins of the death receptor signaling pathway. Therefore, while the ability of bacteria to modify their own or host proteins by glycosylation has been established, at the outset of this thesis project a translocated bacterial effector protein glycosylated by the host had never been described.

PREFACE TO CHAPTER 2

NleA is an essential effector for A/E pathogen virulence. Since its identification 20 years ago, several functions of NleA have been characterized. However, the complete mechanism of NleA in the host cell remains to be elucidated. In **Chapter 2**, we demonstrate that NleA undergoes an apparent mobility shift between the bacterially secreted form and NleA present in the host cell lysate resolved by SDS-PAGE. This shift indicates that NleA is post-translationally modified upon translocation into the host. Using a cell line with a reversible defect in glycosylation, we have obtained evidence that NleA is modified by O-linked glycosylation within the host cell secretory pathway.

CHAPTER 2: **The bacterial virulence factor NleA undergoes host-mediated O-linked glycosylation**

2.1 Abstract

Enterohaemorrhagic and enteropathogenic *Escherichia coli* (EHEC and EPEC) are gastrointestinal pathogens responsible for severe diarrheal illness. EHEC and EPEC form "attaching and effacing" lesions during colonization and, upon adherence, inject proteins directly into host intestinal cells via the type III secretion system (T3SS). Injected bacterial proteins have a variety of functions but generally alter host cell biology to favor survival and/or replication of the pathogen. Non-LEE-encoded effector A (NleA) is a T3SS-injected effector of EHEC, EPEC, and the related mouse pathogen *Citrobacter rodentium*. Studies in mouse models indicate that NleA has an important role in bacterial virulence. However, the mechanism by which NleA contributes to disease remains unknown. We have determined that following translocation into host cells, a serine and threonine- rich region of NleA is modified by host-mediated mucin-type O-linked glycosylation. Surprisingly, this region was not present in several clinical EHEC isolates. When expressed in *C. rodentium*, a non-modifiable variant of NleA was indistinguishable from wildtype NleA in an acute mortality model but conferred a modest increase in persistence over the course of infection in mixed infections in C57BL/6J mice. This is the first known example of a bacterial effector being modified by host-mediated O-linked glycosylation. Our data also suggests that this modification may confer a selective disadvantage to the bacteria during *in vivo* infection.

2.2 Introduction

Escherichia coli is a ubiquitous member of the human intestinal microbiota that generally coexists with mutual benefits in digestion and immunity [33]. However, enterohaemorrhagic *E. coli* (EHEC) and enteropathogenic *E. coli* (EPEC) are pathogenic strains found in contaminated food and water that, when ingested, can infect the host gastrointestinal tract to cause potentially lethal diarrheal illness [3]. EHEC, EPEC, and the related murine pathogen *Citrobacter rodentium* belong to a group of gram-negative bacteria characterized by their ability to form "attaching and effacing" (A/E) lesions on the surface of host cells during intestinal colonization [3,415,416]. A/E lesion development occurs through localized destruction of the normal epithelial microvilli architecture and sub-sequent cytoskeletal rearrangements for the formation of an actin-rich pedestal structure, beneath the attached bacteria [33,415]. The genes required for A/E lesion formation are clustered in the Locus for Enterocyte Effacement (LEE) which contains genes encoding a type III secretion system (T3SS) and multiple translocated effector proteins [416].

The T3SS is found exclusively in gram-negative bacteria, where it functions as a "molecular syringe" to deliver effector proteins across bacterial and host membranes directly into the cytosol of host cells [191,194,195,205]. However, while the genes encoding type III system components are conserved between bacterial species, the repertoire of translocated effectors tends to be unique to each pathogen [211]. EHEC, EPEC, and *C. rodentium* utilize an arsenal of virulence factors that have a variety of functions, but typically alter host cell biology to facilitate pathogen colonization of host tissue, suppression of the host immune system, or otherwise favor pathogen survival and/or replication inside the host. In addition, bacterial effector proteins may undergo several types of host-mediated post-translational modifications (PTM) following translocation that

may trigger effector enzymatic activity, or aid in proper subcellular location targeting and functional regulation.

Several translocated bacterial effectors modified by host-mediated covalent modifications have already been described. After translocation, EPEC Tir, a LEE-encoded effector secreted by the T3SS essential for colonization, is tyrosine phosphorylated by host cell kinases upon its integration into the intestinal epithelial plasma membrane [282,417–419]. Phosphorylation of EPEC Tir is necessary for efficient actin pedestal formation and thereby intimate bacterial attachment by EPEC, indicating functional regulation by this host-mediated modification [248,282]. The *Salmonella* type III effector, SopB, is a membrane-associated inositol polyphosphate phosphatase that is ubiquitinated by host machinery following translocation [420]. Although post-translational modification of proteins by ubiquitin has primarily been associated with protein degradation, ubiquitination of SopB regulates its enzymatic activity at the plasma membrane and intracellular localization rather than its intracellular stability [374,421].

One of the first-characterized T3SS-translocated non-LEE encoded effectors of EHEC, EPEC, and related pathogens is non-LEE- encoded effector A (NleA; also known as EspI). The gene encoding NleA is absent from non-pathogenic *E. coli* and is preferentially found in those strains associated with outbreaks of human disease [422,423]. In addition, it has been shown that NleA is present in *C. rodentium*, the murine model of EHEC and EPEC infections, and is required for virulence during infection of mice [290,342,424]. Genetically hyper-susceptible strains of mice infected with wild-type *C. rodentium* suffer high rates of mortality following infection. In comparison, mice infected with mutant *C. rodentium* in which the *nleA* gene is disrupted display mild symptoms but ultimately survive. In C57BL/6J mice, the *nleA* gene is also required for high-level bacterial colonization and associated histopathology [342]. Although this indicates a clear

effect of NleA on *C. rodentium* virulence, the functional role of NleA during infection is still unknown. NleA associates tightly with host cell membranes, although bioinformatic prediction of transmembrane domains yields inconclusive results [342]. Following translocation into host cells, NleA predominantly localizes to the host secretory pathway, colocalizing with markers of the Golgi apparatus [342,425,426] and has been shown to interact with COPII [299,342]. Among the dozens of effectors injected by EHEC, EPEC, and *C. rodentium*, NleA was noted as one of four effectors essential to bacterial pathogenicity [343].

In this study, we show that NleA undergoes an apparent size shift following translocation from the bacterium into the host cell, provide several lines of evidence consistent with the modification of NleA by O-linked glycosylation, and investigate the implications of the modification using analysis of human outbreak strains and *in vivo* infection models.

2.3 Results and Discussion

NleA is modified by host-mediated O-glycosylation

We first noted a significant mobility change of NleA between bacterially-secreted proteins and those found within infected cells, resulting in a larger apparent size of host-translocated NleA on SDS-PAGE gels (**Figure 1A**). The bacterially secreted form of the protein is present at approximately 55 kDa, close to its predicted molecular weight of 48 kDa [342], while in the host cell lysate a band is observed at a higher apparent molecular weight (~60-65 kDa), suggesting a PTM of NleA inside the host cell. We assessed the phosphorylation of NleA by infecting HeLa cells with wild-type EPEC and incubating the lysates with λ phosphatase enzyme, which cleaves phosphates from serine, threonine, and tyrosine residues. We observed no change in apparent size of NleA after treatment with λ phosphatase (**Figure 1B**). In contrast, Tir, which has been previously

reported to be phosphorylated upon host translocation [282], was used as a positive control and showed a decrease in apparent size after enzyme treatment (**Figure 1B**). To assess the possibility of modification by ubiquitination, NleA was immunoprecipitated from Caco2/TC7 cell lysates infected with wild-type EPEC. There was an enrichment of NleA in the immunoprecipitate lysate, however, we detected no ubiquitin (**Figure 1C**). Therefore, neither phosphorylation nor ubiquitination could account for the host-mediated modification of the NleA protein.

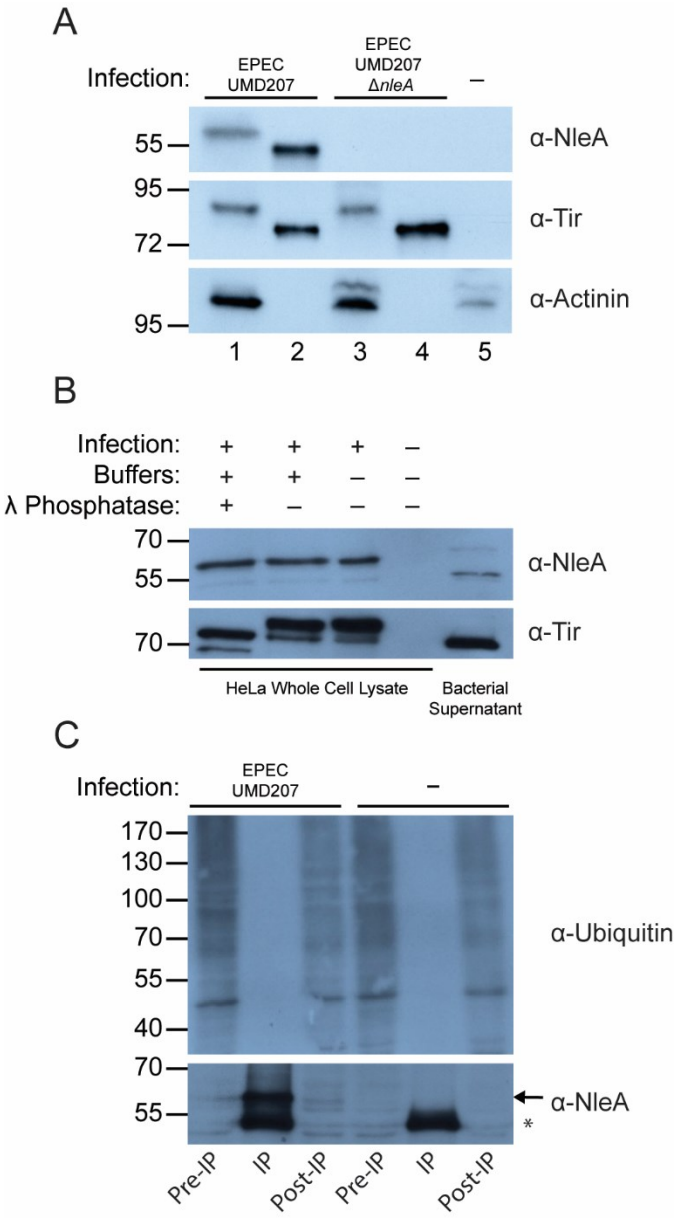

Figure 1. NleA undergoes a mobility shift upon translocation into host cells. (a) Western blot analysis of HeLa cell lysate infected with EPEC UMD207 (lane 1), EPEC UMD207 Δ*nleA* (lane 3), or uninfected (lane 5). Bacterial proteins secreted by indicated strain were concentrated from culture supernatant by trichloroacetic acid precipitation and separated (lanes 2 and 4). Blots were probed with anti-NleA, anti-Tir, and anti-Actinin antibodies. (b) Western blot analysis of HeLa cells infected with EPEC UMD207 and treated with lambda protein phosphatase. Cell lysates were probed with anti-NleA and anti-Tir antibody antibodies. (c) Western blot analysis of immunoprecipitate from Caco2/TC7 cells infected with wild-type EPEC probed with anti-Ubiquitin and anti-NleA antibodies. Lanes 1 and 4 depict the pre-immunoprecipitation (pre-IP), lanes 2 and 5 the immunoprecipitated (IP), and lanes 3 and 6 the post-immunoprecipitation (post-IP) lysates of infected and uninfected Caco2/TC7 cells, respectively. The Ig heavy chain is indicated with an asterisk (*), and NleA is indicated by an arrow. Migration of molecular weight markers (kDa) is indicated on the left of each blot.

To further investigate the type of modification underlying the apparent size shift of the protein, a modified cell line with a reversible defect in glycosylation was used. Glycosylation, in which sugar moieties are glycosidically linked to specific amino acids of a protein structure, plays a significant role in protein folding, targeted transport, cellular localization, and activity [427]. Like phosphorylation and ubiquitination, glycosylation is a dynamic and reversible process. Under standard cell culture conditions, glucose is the only sugar directly provided in growth media. Wild-type Chinese Hamster Ovary (CHO) K1 cells maintain the ability to convert glucose to the required nucleotide sugars for normal glycosylation to occur. Mutant CHO LDLD cells, however, are defective in the synthesis of N-linked, O-linked, and lipid-linked glycoconjugates due to the absence of the UDP-Gal/UDP-GalNAc 4-epimerase, resulting in the inability to catalyze the reversible isomerizations of UDP-Glucose and UDP-N-acetylglucosamine to UDP-Galactose and UDP-N-acetylgalactosamine, respectively [428]. The mutant LDLD cells still possess sugar salvage pathways, however, and so these defects can be reversed by supplementing galactose and N-acetylgalactosamine (GalNAc) directly to the culture media [428].

To determine whether glycoconjugate addition could explain the size shift of translocated NleA, CHO K1 and LDLD cells were infected with EPEC UMD207 with or without the addition of both galactose and GalNAc to the culture media. The UMD207 strain of EPEC was used as it is deficient in host cell adherence, but still functional for T3SS-mediated protein translocation [429], allowing us to wash away bacteria prior to preparing cell lysates. This strain, therefore, enables the analysis of only translocated bacterial protein, without or with minimal interference from NleA in remaining attached bacteria. Host-translocated NleA was visualized by Western blotting of the cell lysates. In the infected CHO K1 cells, two reactive bands were present: one at 55 kDa, corresponding to the expected size of unmodified NleA, and one at approximately 65 kDa,

similar to the previously-observed size of the host-modified protein (**Figure 2A**). Both these bands were present whether or not the media was supplemented with sugars. In the infected CHO LDLD cells, however, the appearance of NleA differed depending on whether the sugars were added to the growth media or not. Only the 55 kDa band was present without sugar supplementation. Upon supplementation with galactose and GalNAc, both the 55 and 65 kDa bands were observable. Since glycoconjugate addition is independent of sugar supplementation in the CHO K1 cells but dependent on the addition of galactose and GalNAc in the CHO LDLD cells, these results are consistent with the modification of NleA by glycoconjugate addition.

To further explore the potential modification of NleA by glycoconjugates, membrane preparations from cells infected with EPEC UMD207 were treated with a mix of deglycosylation enzymes that remove all N-linked, as well as simple O-linked glycans. This treatment did not decrease the apparent molecular weight of host-associated NleA but did decrease the apparent molecular weight of the positive control, fetuin (**Figure 2B,C**). To further characterize the modification, we examined if supplementing CHO K1 and LDLD culture media prior to infection with each individual sugar, galactose, and GalNAc, both, or neither affected the molecular weight of NleA. As before, the appearance of NleA in the CHO K1 cells was independent of sugars supplemented to the media (**Figure 2D**). In the CHO LDLD cells only a band at 55 kDa was present in the absence of sugars or when galactose alone was added (**Figure 1D**). When GalNAc alone was added, the apparent molecular weight of NleA increased, although not to the same extent as when both galactose and GalNAc were supplemented to the media (**Figure 2D**).

O-glycosylation is classified based on the attachment of the glycan to the oxygen atom of the hydroxyl group of a serine (Ser) or threonine (Thr) residue and can be further subclassified based on the initial sugar attached to the protein and the subsequent sugar structures added. O-

glycans are most commonly linked to the polypeptide via GalNAc, a subtype of O-glycosylation referred to as mucin-type, and can vary in length from a single GalNAc residue to structures of more than 20 sugars [393]. Given that we saw a restoration in the size shift of NleA when only GalNAc was added to the culture media, we posit that this implicates mucin-type O-linked glycosylation as the host-cell mediated modification of the bacterial protein upon translocation into the host cell. Mucin-type O-linked glycosylation primarily occurs in the endoplasmic reticulum (ER) and Golgi apparatus, where NleA has previously been shown to localize by immunofluorescence imaging [342,425,426].

Furthermore, the deglycosylation enzyme mix that did not re-move the modification on the NleA protein (**Figure 2B**) is known to be ineffective on mucin-like substrates. It should also be noted that when CHO LDLD cells are cultured with both galactose and GalNAc, the NleA apparent molecular weight is slightly larger than when only GalNAc is available, consistent with a restriction on the size of glycan structure with only GalNAc as a donor.

Unlike N-linked glycosylation, O-linked glycosylation does not have a consensus peptide sequence, which can make predicting specific sites of modification exceedingly difficult. We used the neural-network trained NetOGlyc 4.0 algorithm [430] to predict mucin-type O-glycosylation sites in the EPEC NleA protein sequence (**Figure 2E**). There are 22 individual sites predicted to be modified by O-linked glycosylation, concentrated in a particularly serine- and threonine-rich portion of the EPEC NleA protein sequence. However, little is known about interactions between initial and subsequent sites of O-glycosylation, and the capability of predictive technology to precisely annotate substrate specificity of single GalNAc transferases remains limited. As a result, although NetOGlyc predicts specific sites of modification, the safe interpretation of a positive prediction is that the local region of the protein in which that residue is present is more likely to

be O-glycosylated [430]. Therefore, it is strongly suggested that NleA possesses at least one site of O-linked glycosylation in this serine- and threonine-rich region.

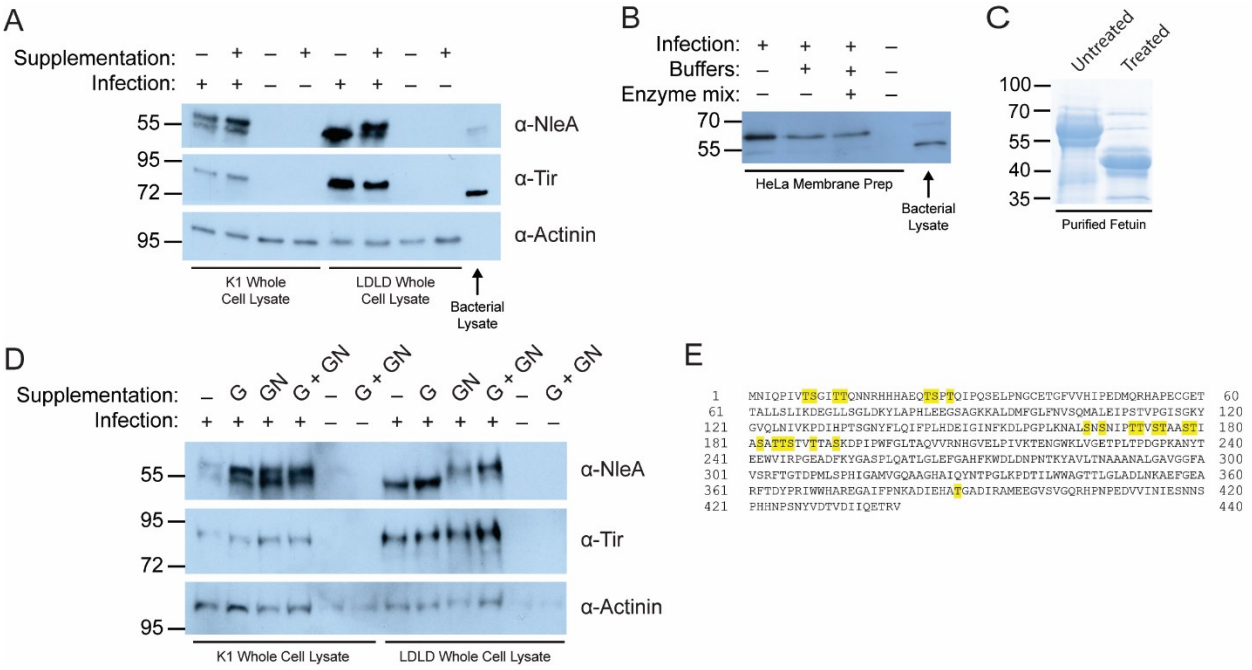

Figure 2. NleA modification is reversibly abrogated in glycosylation-deficient CHO LDLD cells. (A) Western blot analysis of CHO K1 and LDLD cell lysates infected with EPEC UMD207 probed with anti-NleA, anti-Tir, and anti-Actinin antibodies. Cells were cultured in media without (−) or with galactose and GalNAc sugar supplementation (+). (B) Western blot analysis of membrane fraction from HeLa cells infected with EPEC UMD207 and treated with deglycosylation enzyme mix. Membrane was probed with anti-NleA antibody. Lysates were incubated with enzyme mix for 4 h at 37°C. Deglycosylation enzyme mix contains PNGase F, O-glycosidase, α2-3,6,8 neuraminidase, β1-4 galactosidase, and β-N-acetylglucosaminidase enzymes. (C) Coomassie-stained SDS- PAGE of fetuin protein treated with deglycosylation enzyme mix. Purified protein was incubated with enzyme mix for 4 h at 37°C. (D) Western blot analysis of CHO K1 and LDLD cell lysates infected with EPEC UMD207. Blots were probed with anti-NleA, anti-Tir, and anti-Actinin antibodies. Cells were cultured in media without (−), with galactose (G), with N-acetylgalactosamine (GN), or with both galactose and N-acetylgalactosamine (G + GN) sugar supplementation. Migration of molecular weight markers (kDa) is indicated on the left of each blot. (E) NetOGlyc 4.0 algorithm predicted O-glycosylation sites (highlighted) of EPEC NleA.

NleA is one of many bacterial effectors injected by A/E pathogens into host cells following adherence. Recently, the nature in which relationships between these effectors form a complex network has been emphasized, including functional dependencies between certain effectors [343]. Of note, NleB is a T3SS-translocated effector of A/E pathogens that acts a translocated N-acetyl-D-glucosamine (O-GlcNAc)-transferase inside host cells [300,336]. To determine whether the modification of NleA depended on the activity NleB or any other translocated effectors, we transfected a plasmid encoding enhanced green fluorescent protein-tagged NleA (NleA-EGFP) into uninfected LDLD cells with and without sugar supplementation. In the absence of sugar supplementation, NleA-EGFP was not modified (**Figure 3A**). In comparison, in transfected cells with sugar supplementation, two bands were present at the expected sizes of the unmodified and modified forms of NleA-EGFP (**Figure 3A**). No NleA-EGFP was detected in the mock transfection conditions or in the DNA-negative controls. Therefore, NleA is modified in host cells even in the absence of any other translocated bacterial effectors.

Disruption of the host secretory pathway affects NleA modification

The canonical ER-Golgi pathway of glycan biosynthesis in eukaryotic cells involves co-translational protein translocation into the ER where proteins are folded and modified, followed by subsequent passage through the Golgi apparatus for further modification and traffic to their various destinations. However, an additional glycosylation pathway also occurs in the cytoplasm, primarily via O-GlcNAc transferase (OGT)-catalyzed addition of a single N-acetylglucosamine to Ser or Thr residues. Although the extent of the apparent molecular weight shift of host-modified NleA, as well as the requirement of GalNAc for modification to occur in LDLD cells argue against the implication of the OGT pathway in NleA glycosylation, we sought to formally rule out this

possibility. We used CRISPR-Cas9 to generate an OGT knockout HEK293 cell line and infected these cells with EPEC UMD207. We observed that NleA translocated into either WT HEK293- or HEK293 OGT KO-infected cells displayed an increased apparent molecular weight compared to the bacterial lysate (**Figure 3B**). Thus, OGT is not implicated in the mobility shift of NleA following its translocation into host cells.

Next, we investigated the implication of the host cell secretory pathway in the modification of NleA with brefeldin A and monensin treatment in infected CHO LDLD cells with and without sugar supplementation. Brefeldin A treatment induces rapid disassembly of the Golgi apparatus [431–434] and blocks protein transport into cell compartments post-Golgi [432,435]. The retrograde transport of components back to the ER results in nonselective separation of Golgi enzymes from those located in the trans-Golgi network. Therefore, some glycans may remain truncated due to an uncoupling of core structures from later glycosylation reactions. Comparatively, treatment with monensin, a sodium-hydrogen ionophore, interrupts intra-Golgi protein trafficking and inhibits glycan processing by neutralizing the Golgi luminal pH, impairing the function of many glycosylation enzymes [436–438].

As expected, in the absence of sugars only a band at 55 kDa was present regardless of brefeldin A and monensin treatment (**Figure 3C**). In lysate from cells cultured in the presence of sugar supplementation untreated with the protein transport inhibitors, two NleA bands were observed corresponding to unmodified and modified protein (**Figure 3C**). Following treatment with brefeldin A alone, there was a striking reduction in the modified form of the protein. Upon monensin treatment, an intermediate NleA protein size was apparent in addition to the unmodified form. The molecular weight decreased further when brefeldin A and monensin were added in combination, although not to the same extent as when no sugars were supplemented. This

intermediary modified form mimics what was observed when GalNAc alone was supplemented to the culture media, indicating that some degree of modification may be occurring even in the presence of protein transport inhibitors, albeit truncated.

Tir, which was used as a control due to modification independent of the secretory pathway, showed the expected pattern of modification (**Figure 3C**). Together, these results implicate the host cell secretory pathway as the site of NleA modification.

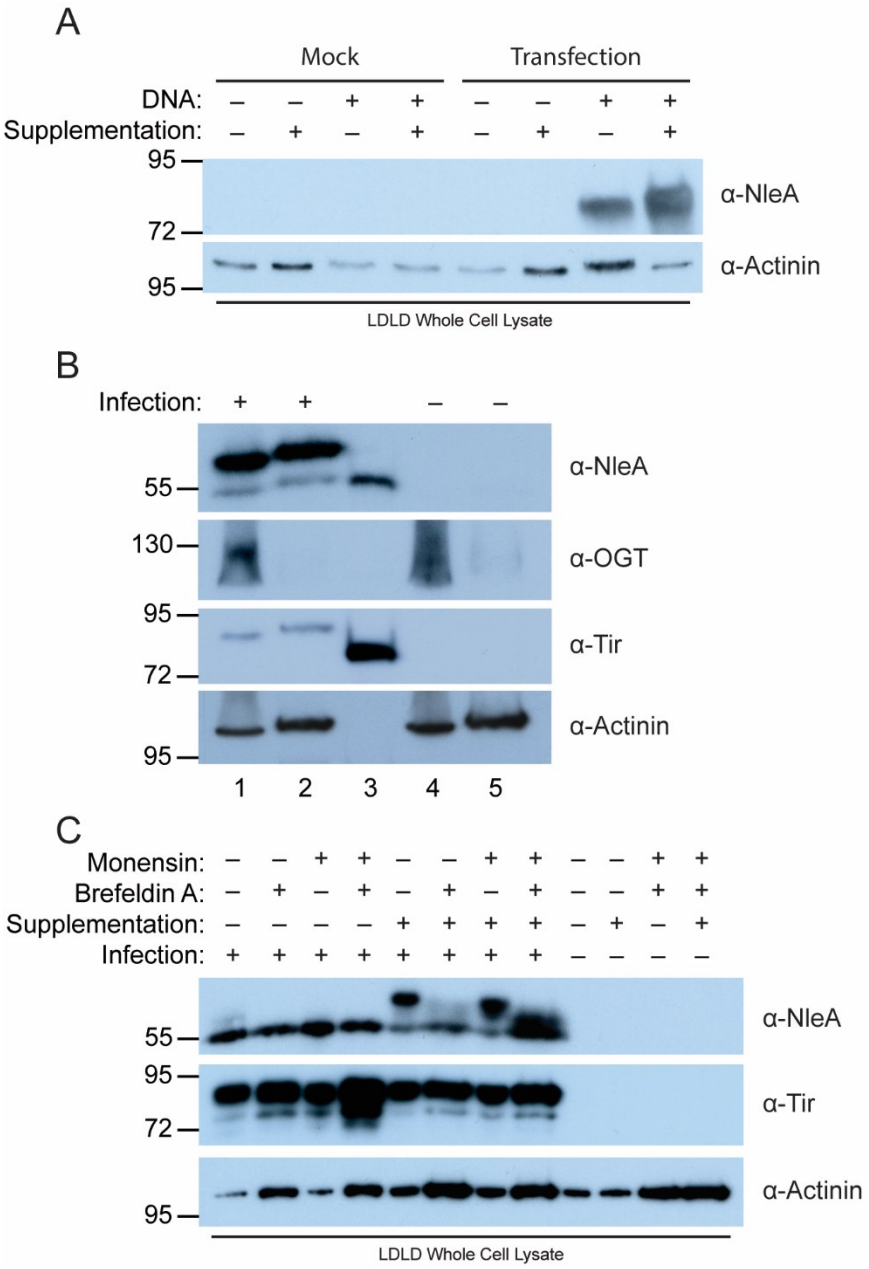

Figure 3. NleA is modified by the host secretory pathway. (A) Western blot analysis of CHO LDLD cells transfected with NleA-EGFP. Cells were cultured in media without (−) or with galactose and GalNAc sugar supplementation (+). Cells were incubated for 24 h with mock or transfection conditions. Blots were probed with anti-NleA and anti-Actinin antibodies. (B) Western blot analysis of wild-type HEK293 (lane 1) or HEK293 OGT knockout (lane 2) cell lysates infected with EPEC UMD207, uninfected cells (lanes 4, 5), and bacterial lysate collected from culture supernatant (lane 3). Blots were probed with anti-NleA, anti-OGT, anti-Tir, and anti-Actinin antibodies. (C) Western blot analysis of CHO LDLD cells treated with brefeldin A and monensin, infected with EPEC UMD207. Cells were cultured in media without (−) or with galactose and GalNAc sugar supplementation (+). Migration of molecular weight markers (kDa) is indicated on the left of each blot.

The suspected region of NleA modification is variable

Based on the concentration of predicted O-glycosylation sites, we investigated whether the serine- and threonine-rich region of the NleA sequence was conserved between isolates. Protein sequences of NleA from *C. rodentium*, EPEC, and EHEC acquired from NCBI, and from EHEC human outbreak strains collected by the Institut National de Santé Publique du Québec (INSPQ) were aligned using the Clustal Omega Multiple Sequence Alignment tool [439].

Sequences showed considerable variability in the region encompassing the serine- and threonine-rich portion of the protein predicted to be modified by O-glycosylation (**Figure 4A**). Surprisingly, several of the strains from infected patients exhibited a partial or complete deletion of this region.

To gain the understanding of where this region exists within the structure of NleA, we used RoseTTAFold [440] for protein structure prediction. The resulting models were somewhat low-confidence (0.45 confidence value), because there are few sequences of NleA homologs and no NleA structure has been determined experimentally. However, the models clearly indicate NleA possesses a central beta barrel, surrounded by one coiled-coil and several other helices and loops (**Figure 4B**). The serine- and threonine-rich region is an unstructured loop segment between two strands of the beta barrel. That the region is a protruding loop is consistent with it being exposed for modification and suggests that its deletion would not greatly affect the overall structure of NleA.

Using these natural variants as well as the structure and glycosylation site predictions, we designed and constructed a variant of *C. rodentium* NleA protein, in which this serine- and threonine-rich region is deleted. *C. rodentium* NleA was chosen for these studies in order to align with our ultimate goal of assessing any effects on virulence *in vivo*.

To test the prediction that the serine- and threonine-rich region of NleA was the site of host-mediated modification, we first assessed whether the variant lacking this region appeared to be modified in host cells, using EPEC UMD207 as a delivery system as described above. CHO K1 and LDLD cells were infected with EPEC UMD207 Δ*nleA* + wild-type *C. rodentium nleA* or EPEC UMD207 Δ*nleA* + *C. rodentium nleA*Δ169-183 without sugar supplementation to the culture media. Consistent with the previous results, translocation of wild-type *C. rodentium* NleA into CHO K1 cells, but not CHO LDLD cells, led to an increase in its apparent molecular weight, with the appearance of an upper band present at approximately 60-65 kDa. In contrast, NleAΔ169-183 did not appear to increase in apparent molecular weight upon translocation into CHO K1 or LDLD cells, consistent with this region being critical for host-mediated modification (**Figure 4C**). Thus, modification within host cells is conserved in the NleA protein of *C. rodentium*, and this modification is abrogated when the serine- and threonine-rich region is deleted.

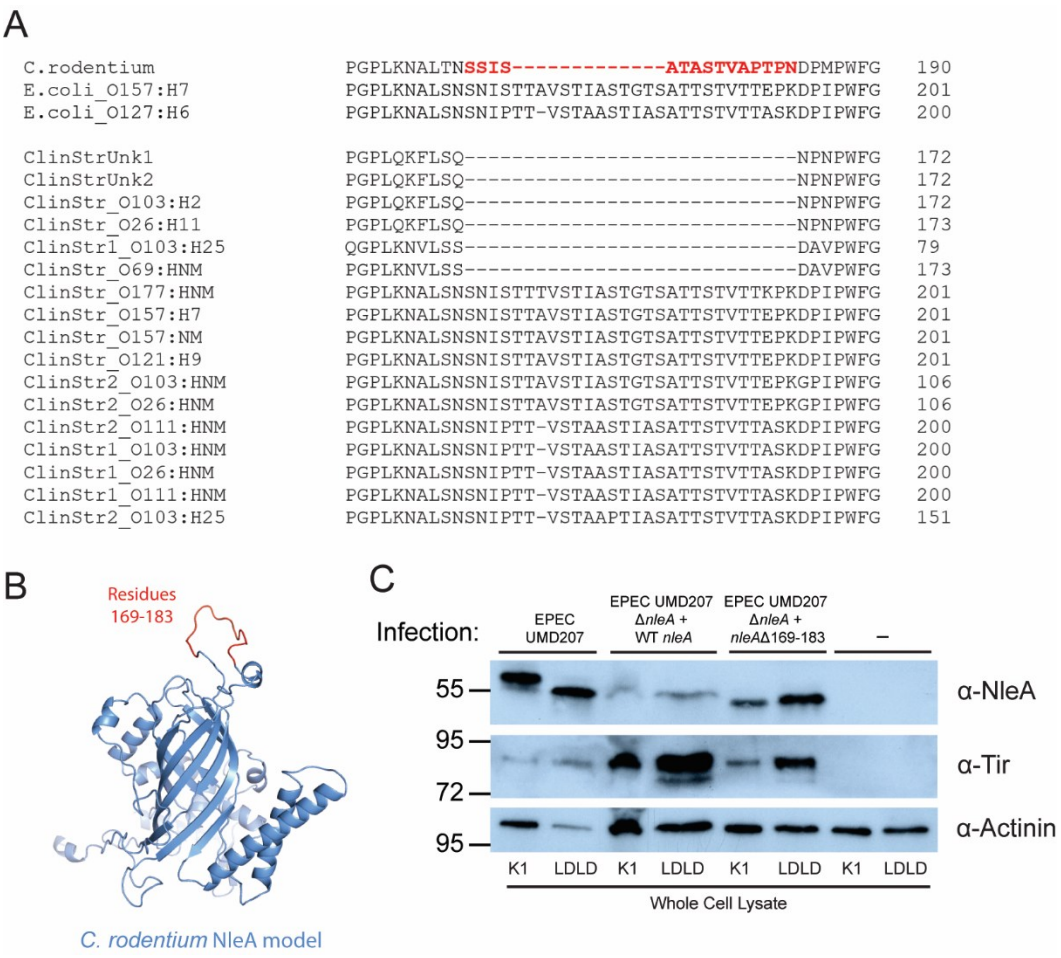

Figure 4. The NleA region predicted to be modified is variable. (A) Multiple protein sequence alignment of NleA from homologs identified via BLAST search and EHEC clinical strains courtesy of INSPQ. Sequences were aligned using Clustal Omega. Absent residues are represented by a dash (−). (B) Structural model of *C. rodentium* NleA with residues 169-183 colored in red. (C) Western blot analysis of CHO K1 and LDLD cell lysates infected with EPEC UMD207 ΔnleA+ wild-type *C. rodentium nleA* or EPEC UMD207 Δ*nleA* + *C. rodentium nleA*Δ169-183. Blots were probed with anti-NleA, anti-Tir, and anti-Actinin antibodies. Migration of molecular weight markers (kDa) is indicated on the left.

Modification of NleA is not required for virulence in susceptible mice

NleA has previously been shown to be absolutely required for pathogen virulence in the *C. rodentium*-infection model in both susceptible and resistant mice [342]. Mice infected with wild-type *C. rodentium* develop diarrhea, colitis, and hyperplasia of the intestinal epithelium, all shared symptoms of EHEC and EPEC infections in humans. In contrast, mice infected with *C. rodentium* lacking NleA develop only mild symptoms and do not succumb to infection [342]. To analyze the role of NleA's host-mediated modification in bacterial virulence, we infected susceptible C3H/HeJ mice with wild-type *C. rodentium* (WT), *C. rodentium* Δ*nleA*, *C. rodentium* Δ*nleA* + wild-type *nleA* (Δ*nleA* + WT *nleA*), or *C. rodentium* Δ*nleA* + non-modifiable mutant *nleA* (Δ*nleA* + *nleA*Δ169-183) and monitored mortality over the course of infection. At day 6 post-infection, all mice were colonized, with a trend for lower colonization in mice infected with *C. rodentium* Δ*nleA* (**Figure 5A**). All infected mice, except for those infected with *C. rodentium* Δ*nleA*, succumbed to infection between 8-10 days post-infection with no significant differences in mortality between the groups (**Figure 5B**). As previously described, mice infected with *C. rodentium* Δ*nleA* survived. Thus, overall, deletion of the modified region of NleA did not visibly alter virulence in this infection model. This result is consistent with the finding that some human outbreak strains naturally lack this region.

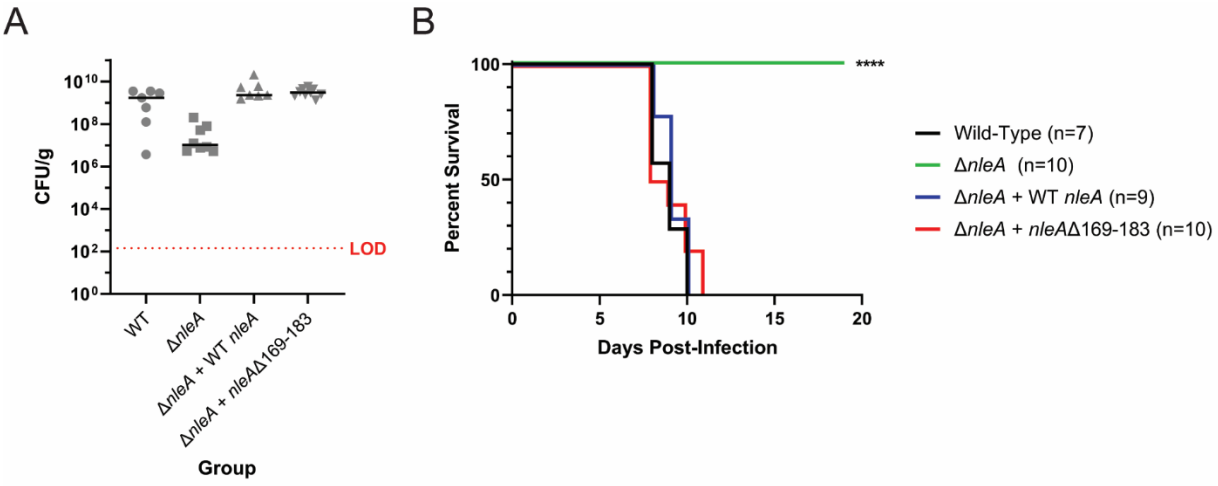

Figure 5. Mutated NleA does not alter virulence *in vivo*. Female C3H/HeJ mice were infected by oral gavage with wild-type *C. rodentium*, *C. rodentium* Δ*nleA*, *C. rodentium* Δ*nleA*+ mutant *nleA* or *C. rodentium* Δ*nleA* + wild-type *nleA* and (A) fecal bacterial burden was assessed at day 6 post-infection by diluting and plating on selective MacConkey agar and counting CFU. Each mouse in the experiment is represented by a single point. The median of each group is indicated on the graph as a horizontal line. The limit of detection (LOD) is indicated by a dashed red line. (B) Survival curve for infected mice was monitored 20 days post-infection. Percentage of the starting number of mice in each group that were viable on each day is shown. Data are pooled from two individual experiments with the number of mice per group indicated in the legend. The log-rank (Mantel-Cox) method was used to determine statistical significance (****, $p < 0.0001$).

Glycan structures are bulky and hydrophilic, and often contribute to protein solubility and increased stability against proteolysis. However, the distinct role of glycans, especially O-linked structures, obtaining correct protein folding is ambiguous. While there is evidence for protein misfolding and aggregation in the absence of glycans, others have shown that the elimination of some or all glycans had no effect on folding [441,442]. This implies that some glycosylation sites may be more important than others for folding and that glycan-induced conformation effects are more than likely local in nature. The finding that the non-modifiable protein confers full virulence in susceptible mice provides evidence that this protein is likely to be properly folded and translocated during *in vivo* infection.

Mutant NleA-complemented strains outcompete wild-type-complemented

To test for a more subtle effect on virulence, we employed the competitive index assay in C57BL/6J mice, which undergo a self-limiting infection with wild-type *C. rodentium*. Mice were infected with a 1:1 ratio of *C. rodentium* Δ*nleA* + WT *nleA*: *C. rodentium* Δ*nleA* + *nleA*Δ169-183 and the ratio of the two strains within the inoculum and within each animal was assessed at early, peak, and late timepoints post-infection by colony PCR. An index of 1 indicates that wild-type and mutant strains are present in equal proportions, while an index less than 1 means that the wild-type population outcompeted the mutant population and conversely with an index greater than 1, the mutant outcompeted the wild-type strain. Surprisingly, the mutant-complemented strain, lacking the serine- and threonine-rich region, persisted longer than the WT-complemented strain, resulting in a median index significantly greater than 1 at day 26 post-infection (**Figure 6A**). Likewise, the percent of total colonies skewed toward mutant-complemented at peak, day 12, and late, day 26,

infection timepoints, despite total bacterial burden not differing markedly from typically observed CFU (**Figure 6B,C**).

These results indicate that the absence of the host-mediated modification of NleA may be somewhat favorable for the bacteria, or suggests that the contraction of this region may be associated with some kind of modest gain-of-function, which is consistent with the clinical isolate data. An intriguing hypothesis is that NleA is modified by the host to serve as a signal to the immune system of an active infection, and that the deletion of this region, as observed in human outbreak strains, is an adaptation of the bacteria to prolong clearance, consistent with superior fitness of the mutant observed in our competition assay. Among their repertoire of functions, glycans can serve as recognition markers as well as modulate immune responses [443,444]. The consequences of the absence of the glycosylation of NleA on immune interplay have not yet been explored.

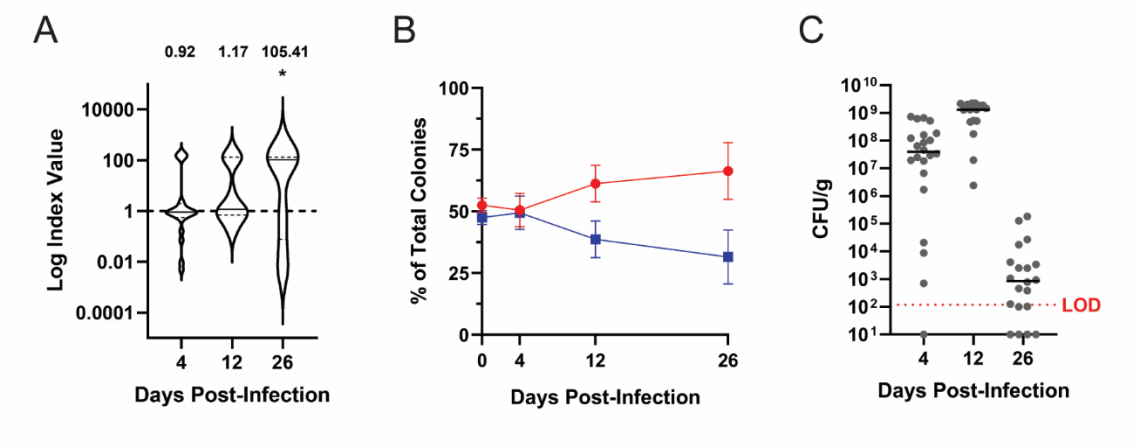

Figure 6. Bacteria expressing mutant NleA outcompete those with wild-type NleA. C57BL/6J mice (n = 22) were infected by oral gavage with a 1:1 mixture of *C. rodentium* Δ*nleA* + *nleA*Δ 169-183:*C. rodentium* Δ*nleA* + wild-type *nleA*. Data were pooled from four independent experiments. (A) Competitive index of B6 mice at various timepoints post-infection. The competitive index is defined as the mutant-to-wild type ratio within the output sample divided by the mutant-to-wild type ratio within the input (inoculum). The index was assessed at days 4, 12, and 26 post-infection by diluting and plating on selective MacConkey agar supplemented with kanamycin. 100 colonies were selected for polymerase chain reaction and run on a 1% agarose gel and the ratio of mutant:WT was determined. The median at each time point is indicated by a solid horizontal line. P values were determined via one sample t-test at a 95% confidence interval. (B) Change in percent of total colonies identified as of *C. rodentium* Δ*nleA* + *nleA*Δ169-183 (red) or *C. rodentium* Δ*nleA* + WT *nleA* (blue) at various timepoints post- infection. Data are represented as mean ± S.E.M. (C) CFU/g over time for each group of mice infected with 1:1 v:v of *C. rodentium* Δ*nleA* + WT *nleA* and *C. rodentium* Δ*nleA* + *nleA*Δ169-183. The median of each time point is indicated by a horizontal line. The limit of detection (LOD) is indicated by a dashed red line.

Our results indicate that the bacterial virulence factor NleA is modified by host-mediated O-linked glycosylation upon translocation into the host cell during infection. To our knowledge, this is the first example of a bacterial effector protein undergoing this type of modification inside the host cell. Our findings are, however, reminiscent of another bacterial effector modified inside host cells. Although Tir has been studied extensively *in vitro*, with epithelial cell culture demonstrating that tyrosine phosphorylation of EPEC Tir is necessary for efficient A/E lesion formation, it has been challenging to attribute the significance of this modification to Tir functionality in mouse models and *in vitro* organ culture. In intestinal explants, Tir phosphorylation-deficient mutants were still able to colonize, and tyrosine phosphorylation modification of Tir is not required for colonization, A/E lesion formation, or crypt hyperplasia in the mouse gut, despite being necessary *in vitro* [417,445]. Nonetheless, we are intrigued by the potential role of this modification given the importance of glycosylation in intestinal biology.

One important remaining question for future investigation is how the serine- and threonine-rich region of NleA gains access to the enzymes involved in glycosylation, which are located within the lumen of the secretory pathway. Notably, when NleA was used as bait for affinity purification of effector-host protein interactions from HeLa cell lysates, GalNAc transferase 7 (GALNT7) was identified with intermediate confidence [446]. GALNT7 is part of a large subfamily of enzymes residing in the Golgi apparatus that control the initiation step of mucin-type O-linked protein glycosylation and transfer of GalNAc to Ser and Thr amino acid residues. However, this putative interaction was not validated by the authors. We have previously shown that in addition to localizing to the secretory pathway, NleA interacts with the Sec24 component of the COPII complex involved in protein trafficking from the ER to the Golgi [299,346]. It is unclear whether Sec24 interaction of NleA precedes or follows its modification. Together these

studies lend credence to NleA being in the vicinity of and potentially interacting with glycosyltransferase enzymes. Intriguingly, Sec24 is part of a coat complex on the outside of transport vesicles, whereas the glycosylation enzymes reside within the lumen of the secretory pathway. This suggests that NleA must either span, cross, or disrupt the secretory pathway membranes. NleA was previously shown to be tightly associated with host membranes, resistant to extraction with high salt and high pH [342] although the nature of this association and the topology of NleA at the membrane remain open questions.

Altogether, this study contributes to a growing understanding of the range of interactions possible between bacterial effector proteins and host cells. Further investigation into the mechanism of NleA function, how it is modified in host cells, and how this modification modulates its function promises to provide new insights into host-microbe interactions.

2.4 Experimental Procedures

Bacterial strains, plasmids, and growth conditions

Bacterial strains, antibodies, plasmids, and oligonucleotides used in this study are listed in **Table S1**. Bacterial strains were routinely cultured at 220 rpm at 37°C in Luria Bertani (LB) agar or broth (1% (wt/vol) tryptone, 0.5% (wt/vol) yeast extract, 1% (wt/vol) NaCl). When appropriate, LB was supplemented with chloramphenicol (Cm; 30 μg/ml), kanamycin (Kan; 50 μg/ml), or ampicillin (Amp; 100 μg/ml).To make EPEC UMD207 Δ*nleA*, sacB gene-based allelic exchange [429] was used to generate an in-frame deletion mutant of *nleA* using the vector pRE112 with primers EPECDHKF1 and EPECDHKR2 (**Table S1**). This mutant was verified by PCR reaction spanning the deletion site, and by Western blot analysis with anti-NleA antiserum.

The complemented EPEC UMD207 Δ*nleA* + WT *nleA* and *Citrobacter rodentium* Δ*nleA* + WT *nleA* strains were created by PCR amplifying the entire *nleA* sequence and flanking region using primers EPECDHKF1 and EPECDHKR2 (**Table S1**) and, after sequence verification, was subcloned into the low-copy number vector pWSK129.

To make EPEC UMD207 Δ*nleA* + *nleA*Δ169-183 and *Citrobacter rodentium* Δ*nleA* + *nleA*Δ169-183, the *nleA* gene with a partial internal deletion was commercially synthesized and inserted into the pBluescript II SK (+) vector and transformed into chemically competent Subcloning Efficiency DH5α cells (ThermoFisher). pWSK129 plasmid containing the WT *nleA* gene and pBluescript II SK (+) containing the mutated *nleA* gene were digested with *BglII* and *SphI*. The resulting pWSK129 vector was ligated to the pBluescript II SK (+) digestion product encompassing the 45 base pair internal deletion. The partial deletion was verified by Western blotting whole cell lysates with polyclonal anti-NleA antiserum, and by sequencing with primers NleA_1_F and NleA_4_R.

Secreted proteins

Overnight cultures were sub-cultured 1:40 M9+ media (KH_2PO_4-$7H_2O$, NaCl, Na_2HPO_4, NH_4Cl, 1 M $MgCl_2$, 5% casamino acids, 20% glucose, $NaHCO_3$) in a 6-well plate and incubated for 6 h at 37°C, 5% CO_2. Bacteria were removed by centrifugation (16,000 × g, 1 min). The supernatant was transferred into a new tube and centrifuged 16,000 × g, 2 min. Supernatant proteins were precipitated by the addition of trichloroacetic acid (10% [vol/vol]) for 1 h on ice. Precipitated proteins were pelleted by centrifugation at 4°C (16,000 × g, 30 min). The supernatant was removed, and 1 ml of ice-cold acetone was added to each tube. The samples were stored at −80°C overnight. The next day, samples were centrifuged at 4°C (16,000 × g, 30 min). The

supernatant was removed, and the pellets were resuspended in 5X Laemmli buffer (50% glycerol, 0.3 M Tris pH 6.8, 5% [wt/vol] SDS, 10% β-mercaptoethanol, bromophenol blue).

Cell culture

HeLa (ATCC no. CCL2) and Caco2/TC7 [447] cells were cultured in Dulbecco's minimal Eagle's medium (DMEM; Wisent) supplemented with 2 mM glutamine and 5% and 10% (v/v) heat-inactivated fetal calf serum (Wisent). Chinese Hamster Ovary (CHO) cells were cultured in Ham's F-12 medium (Wisent) with 5% inactivated fetal calf serum and 1% L-glutamine. Cells were grown at 37°C in a humidified atmosphere of 5% CO_2. 48 h before infection of CHO cells, dishes were washed with phosphate-buffered saline (PBS) once and changed to Ham's F-12 media supplemented with 1% inactivated fetal calf serum with or without galactose (20 μM) and N-acetylgalactosamine (200 μM) sugar supplementation. For each sample, at least one confluent 100 mm dish of cells were infected with EPEC UMD207 Δ*nleA* + WT *nleA* or EPEC UMD207 Δ*nleA* + *nleA*Δ169-183 using an initial multiplicity of infection of 1:100 and incubated for 4 h at 37°C in a 5% CO2 atmosphere. EPEC strain UMD207 lacks intimin and the bundle-forming pilus and therefore delivers bacterial effector proteins without stable bacterial adherence [429]. Following infection, cells were washed with sterile ice-cold PBS and fractionated by re-suspension in 0.5 ml of radioimmunoprecipitation assay (RIPA) lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% (v/v) NP-40, 0.5% (w/v) sodium deoxycholate, 1 mM EDTA, 0.1% (w/v) SDS, 1 mM NaF, 100 μM Na_3VO_4) supplemented with cOmplete protease inhibitor cocktail (Roche). Cell suspensions were sonicated and centrifuged at maximum speed for 15 min at 4°C. The supernatant containing the whole cell lysate was made up to 1X Laemmli buffer using a 5X stock for resolution by SDS-PAGE.

For treatment of cells with brefeldin A and monensin, Brefeldin A Solution (Biolegend) and Monensin Solution (Biolegend) were added at 1X concentration to the culture media 2 h prior to infection and incubated at 37°C, 5% CO_2. Cells were infected with EPEC UMD207 in the presence of brefeldin A and monensin, and whole cell lysates were collected for resolution by SDS-PAGE as outlined above.

λ phosphatase treatment

The λ phosphatase treatment was performed as per the manufacturer's instructions (New England Biolabs). Briefly, HeLa cells were infected with EPEC UMD207. Cells were harvested and lysed in B150 buffer without phosphatase inhibitors (20 mM Tris-HCl pH 8.0, 150 mM KCl, 5 mM MgCl2, 10% glycerol, 0.1% NP-40, protease inhibitors). Lysates were incubated with or without enzyme for 2 h at 30°C. The samples were run on an SDS-PAGE gel, transferred onto a PVDF membrane, and immunoblotted.

Immunoprecipitation and anti-ubiquitin blotting

Confluent Caco2/TC7 cells seeded in 10 cm dishes were infected with 100 µl of overnight culture of EPEC UMD207. Approximately 5 h post-infection, the cells were washed and harvested with cold PBS+/+ (Wisent). The cells were centrifuged at 4°C, 1500 × g for 5 min. Cell pellets were lysed with lysis buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1% NP-40, 3 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM Na_3VO_4, 10 mM NaF, 50 mM $Na_4P_2O_7$, protease inhibitors), by rotating for 30 min at 4°C. The samples were sonicated three times, 15 s at 60% amplitude in a bath sonicator. Lysed cells were centrifuged at 4°C, 16,000 × g for 15 min. The supernatant was transferred to a clean Eppendorf tube, containing 40 µl of protein G Plus (Santa Cruz; prewashed in lysis buffer)

per sample, and incubated for 1 h on the labquake at 4°C. Following incubation, the samples were centrifuged at 4°C, 1500 × g for 5 min. 15 µl of 5X Laemmli was added to 75 µl of the supernatant, boiled for 5 min, and stored as the pre-IP lysate. NHS-activated sepharose beads (GE Healthcare Life Sciences) coupled to rat anti-NleA affinity purified antibody were prepared as per manufacturer's instructions and added to the remaining supernatant, followed by incubation for 2 h on a labquake at 4°C. Following incubation, the samples were centrifuged at 4°C, 1500 × g, and 75 µl of supernatant was stored as the post-IP lysate as described above. The beads were washed with lysis buffer, each wash consisting of 10 min on the labquake at 4°C followed by centrifugation at 4°C, 1500 ×g for 5 min. Beads were resuspended in 100 µl of 5X Laemmli sample buffer to elute the protein and boiled for 5 min. All samples were then run on an SDS-PAGE gel, transferred to a PVDF membrane, and immunoblotted for total ubiquitin.

Deglycosylation assay

HeLa cells were infected with 100 µl of overnight culture of wild-type EPEC for approximately 4-5 h. The cells were harvested in PBS+/+. Cells were centrifuged at 1000 × g for 5 min at 4°C. The supernatant was aspirated, and the pellets were resuspended in 1.5 ml homogenization buffer (250 mM sucrose, 3 mM pH 7.4 imidazole, 0.5 mM EDTA, 1 mM VO_4, 1 mM NaF). Samples were centrifuged at 3000 × g for 10 min at 4°C. The supernatant was aspirated and the samples were resuspended in 300 µl of homogenization buffer. The samples were mechanically lysed by passing through a 22-gauge needle and centrifuged at 3000 × g for 15 min at 4°C. The supernatant was transferred to an ultracentrifuge tube, and centrifuged at 41,000 × g in a TL100 Beckman centrifuge in TLS55 rotor for 20 min at 4°C. The supernatant was aspirated, and the pellet (membrane fraction), was resuspended in 125 µl of dH_2O. The deglycosylation assay

was performed as per the manufacturer's instructions (New England Biolabs, P6039). Briefly, 2 µl of 10X Glycoprotein Denaturing Buffer was added to 18 µl of each sample and denatured at 100°C for 10 min. The samples were cooled on ice and centrifuged at maximum speed for 10 s. Following centrifugation, 5 µl of 10X GlycoBuffer 2 was added to each sample, followed by 5 µl of Deglycosylation Enzyme Cocktail or dH_2O. The contents were mixed gently and incubated at 37°C for 4 h. Following incubation, samples were run on an SDS-PAGE gel and immunoblotted with an anti-NleA antibody. Fetuin was processed as a positive control, run on an SDS-PAGE gel, and stained by Coomassie.

Transfection

LDLD cells were plated 2.5×10^6 cells per 10 cm dish 1 day prior to transfection so that the cells were approximately 80% confluent on the day of transfection. Cells were transiently transfected with 2µg of DNA using 25 µl FuGENE HD Transfection Reagent (Promega) according to the manufacturer's instructions. Transfected cells were briefly washed with sterile ice-cold PBS to remove residual media. Cells were harvested 24h post-transfection by scraping into 1 ml cold PBS, centrifuged at 1500 rpm for 5 min at 4°C and resuspended in 0.5 ml RIPA buffer. Cell suspensions were sonicated and centrifuged at maximum speed for 15 min at 4°C. The supernatant containing the whole cell lysate was made up to 1X Laemmli buffer using a 5X stock for resolution by SDS-PAGE.

Generation of HEK293 OGT knockout cell line

Synthesis of gRNAs: The Alt-R CRISPR-Cas9 sgRNA for generation of OGT knockout was synthesized by IDT (gRNA target sequence: CATCGATGGTTATATTAACC).

Electroporation: 1 day prior to transfection, HEK293 cells were split into a new flask with fresh growth medium such that the cells reach 70%-90% confluent the following day. On the day of electroporation, single cells were prepared, counted, and appropriate amounts of cells (1×10^5 cells per transfection) were transferred to a 1.5 ml microcentrifuge tube. The cells were washed once with PBS by centrifugation at 500 × g for 5 min. At the same time as the preparation of cells for electroporation, 2μg Cas9 protein and 400ng sgRNA were mixed in 10 μl of resuspension buffer R and incubated at room temperature for 10 min, after which 12.3 pmol of Alt-R Cas9 electroporation enhancer (IDT) was added. Prepared cells were re-suspended in the buffer R containing Cas9-gRNA complex and electroporation enhancer, and cell mixture was then transferred into a 10 μl Neon tip with Neon pipette and electroporation were performed using the parameters as following: pulse voltage 1100V, pulse width 20ms and pulse number 2. After electroporation, cells were added into 1 ml prewarmed growth medium in a 24-well plate and cultured for 3 days.

T7EI assay: Genomic DNA was extracted from HEK293 cells transfected with Cas9-gRNA and was then PCR amplified with primers flanking the gRNA target region (Fwd: ACACTTGTCGCCTTTTCCAGA; Rev: GACCCATTATCCACCATTCCTTG). The amplification was carried out with AmpliTaq Gold 360 master mix (ThermoFisher), using the following cycling condition: 95°C for 10 min for initial denaturation; 40 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 35 s; and a final extension at 72°C for 7 min. The PCR product was used for the assay of the success of genome editing by T7EI assay using the ALT- R genome editing detection kit (IDT).

Single-cell clone analysis: single cells were prepared, counted, and serially diluted to 2×10^4, 5×10^2, and 5 cells/ml. Then, 200 μl of 5 cells/ml was dispensed to each well of a 96-well

plate using a multichannel pipette. Plates had been incubated at 37°C in a 5% CO_2 incubator until single cells grew into colonies.

Screening KO: Cell lysates were prepared from single-cell clones and Western blot analyses were used to screen knockout clones.

Immunoblot analysis

Samples for Western blot analysis were resolved by mini-Protean pre-cast gels (7.5% polyacrylamide; BioRad) at 140V for 75 min and transferred electrophoretically at 10V for 30 min onto an activated PVDF membrane. Immunoblots were blocked in 5% non-fat dried milk (NFDM) in TBS, pH 7.2, containing 0.1% Tween 20 (TBST) over-night at 4°C. Membranes were then incubated with primary antibody in 5% NFDM TBST overnight at 4°C before being washed five times in TBST and then incubated with horseradish peroxidase-conjugated secondary antibody for 45 min at room temperature. Membranes were washed as described above and visualized by developing in Immobilon Western ECL Substrate (Millipore), followed by exposure to autoradiography film (Denville). Primary and secondary antibodies used are listed in **Table S1**.

Animal studies

This study was conducted in accordance with the recommendations of the Canadian Council on Animal Care. The protocol was approved by the McGill University Animal Care Committee. Five-week-old female C3H/HeJ and C57BL/6J mice (Jackson Laboratory) were housed in a specific-pathogen-free facility at McGill University. Wild-type or mutant *C. rodentium* DBS100 strains were grown overnight in 3 ml LB broth, 220 rpm, at 37°C, and 100 µl of the cultures was used to infect mice by oral gavage, containing 2-3 × 10^8 CFU. The infectious dose

was confirmed by plating of serial dilutions. For survival experiments, the highly susceptible C3H/HeJ mice were monitored daily over the course of infection and were euthanized if any of the following clinical endpoints were met: 20% body weight loss, hunching, and shaking, inactivity, or body condition score <2 [448]. To determine bacterial colonization, fecal pellets were homogenized in PBS using a MagNA Lyser (Roche) and serially plated on MacConkey agar (Difco Laboratories). Plates with colonies between 30 and 300 were enumerated, and *C. rodentium* was identified by its distinctive morphology on the selective medium. For competitive index experiments, equal volumes of the inoculum of *C. rodentium* Δ*nleA* + WT *nleA* and *C. rodentium* Δ*nleA* + *nleA*Δ169-183 were mixed, and 100 µl of the mixture was used to infect five-week-old female C57BL/6J mice by oral gavage. At the indicated times throughout the course of infection, fecal pellets were collected, homogenized in PBS as described, and serially diluted before being plated on MacConkey agar. After overnight incubation, individual colonies were patched onto LB plates supplemented with Kan and, the following day each clone was genotyped by PCR using primers NleA_1_F and NleA_3_R flanking the deleted region of the *nleA*Δ169-183 mutant strain. 100 colonies were analyzed per mouse per timepoint, and three to five mice were analyzed per experiment. PCR products were visualized on 1% agarose gels containing ethidium bromide. The competitive index was determined as the output ratio (mutant:wild type) divided by the input ratio (mutant:wild type).

Structural modeling

C. rodentium NleA was modeled by providing its sequence (NCBI accession code WP_012904700.1) to the web-based program RoseTTAfold [440]. The five models each had a

confidence score of 0.45, and had average per atom error estimates of 6.8-8.0 Å. The models were overall similar to one another and model 1 is shown in **Figure 4B**.

Statistical analysis

Data analyses were performed using GraphPad Prism v9.0 software. Statistical comparisons were carried out using tests described in the figure legends with a $p < 0.05$ considered significant.

2.5 Acknowledgements

This work was supported by a National Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2014-05119) awarded to S.G. L.B. was supported by Fonds de Recherche Quebec Santé (FRQS) Training Awards. Sequencing was performed at the Genome Quebec Innovation Centre. We thank Drs. Alexandra Kazanova and Hicham Bessaiah for their technical assistance. We thank Dr. Tracy Raivio (University of Alberta) for the EPEC NH4 strain, as well as Drs. Kirk Bergstrom (University of British Columbia) and Heidi McBride (McGill University) for their helpful discussions.

2.6 Supplementary Information

Table S1

A. Oligonucleotides used in this study

EPECDHKF1	GGTACCCCAGCAGCAGTGAAACAATAAT	IDT
EPECDHKR2	GAGCTCTGTGCGTAACAGATTGTGGAC	IDT
NleA_1_F	GCTCTAGAGTCAACCGAACATACATTCCGGAA	IDT
NleA_2_R	GGTTAACGCATTTTTTAATGGGCCAGGAAG	IDT
NleA_3_F	GTTAACAAACGACCCAATGCCATGGTTTGGATT	IDT
NleA_3_R	AATCCAAACCATGGCATTGGGTCGTT	IDT
NleA_4_R	GGGGTACCTTATGATGTGGTGAATTGCCACC	IDT
T7EI_Fwd	ACACTTGTCGCCTTTTCCAGA	IDT
T7EI_Rev	GACCCATTATCCACCATTCCTTG	IDT

B. Plasmids used in this study

pBluescript II SK (+)	Agilent Technologies
pWSK129	[449]
pRE112	[450]
pEGFP-C1	Clontech Laboratories

C. Antibodies used in this study

Rat anti-NleA	[342]; 1:10,000
Mouse anti-Tir	[361]; 1:10,000
Mouse anti-Ubiquitin	Santa Cruz; 1:1000
Rabbit anti-Actinin	Cell Signaling; 1:1000
Rabbit anti-GAPDH	Cell Signaling; 1:10,000
Rabbit anti-OGT	Abcam, 1:5000

D. Bacterial strains used in this study

EPEC UMD207	[429]
EPEC UMD207 Δ*nleA*	This study
EPEC UMD207 Δ*nleA* + WT *nleA*	This study
EPEC UMD207 Δ*nleA* + *nleA*Δ169-183	This study
Wild-type EPEC E2348/69	[22]
EPEC Δ*nleA*	[299]
Citrobacter rodentium DBS100 wild-type	[103]
Citrobacter rodentium Δ*nleA*	[342]
Citrobacter rodentium Δ*nleA* + WT *nleA*	This study
Citrobacter rodentium Δ*nleA* + *nleA*Δ169-183	This study

PREFACE TO CHAPTER 3

In **Chapter 2**, we presented evidence that NleA is modified by mucin-type O-linked glycosylation and identified the host secretory pathway as the subcellular location of this modification. While the indirect evidence presented was highly suggestive of the type of post-translational modification NleA undergoes in the host cell, in **Chapter 3** we sought to obtain direct evidence of NleA O-glycosylation. We also identified an interaction between NleA and a host secretory pathway protein that is required for glycosylation of NleA to occur.

CHAPTER 3: **Direct evidence of host-mediated glycosylation of NleA and its dependence on interaction with the COPII complex**

3.1 Abstract

Non-LEE-encoded Effector A (NleA) is a type III secreted effector protein of enterohaemorrhagic and enteropathogenic *Escherichia coli* as well as the related mouse pathogen *Citrobacter rodentium*. NleA translocation into host cells is essential for virulence. We previously published several lines of evidence indicating that NleA is modified by host-mediated mucin-type O-linked glycosylation, the first example of a bacterial effector protein modified in this way. In this study, we use lectins to provide direct evidence for the modification of NleA by O-linked glycosylation and determine that the interaction of NleA with the COPII complex is necessary for this modification to occur.

3.2 Introduction

Enteropathogenic *Escherichia coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC) are foodborne gastrointestinal pathogens that account for significant morbidity and mortality worldwide. EPEC infection can cause profuse watery diarrhea and is a primary cause of mortality in children under five years old [1]. Those afflicted with EHEC infection also suffer from diarrhea, which in some cases can progress to life-threatening hemolytic uremic syndrome. EPEC, EHEC, and the closely related mouse pathogen *Citrobacter rodentium* are aptly named attaching and effacing pathogens due to their distinct ability to tightly adhere to the host epithelium, efface the absorptive microvilli architecture, and rearrange the cytoskeleton resulting in the formation pedestal-like lesions beneath attached bacteria [2,3,415]. Development of A/E lesions is conferred

by genes encoded in the Locus for Enterocyte Effacement (LEE) [191,195,198] which also encodes a type III secretion system (T3SS) that is essential for pathogenicity [193,451]. EPEC and EHEC utilize the T3SS to inject effector proteins that subvert and modulate host processes to promote bacterial survival and replication within the host.

Continued identification of novel effectors has shown that the effector repertoire of A/E pathogens is much larger than previously thought, with a complex functional network. Although the LEE encodes a small number of effector proteins, effectors encoded outside the LEE can also be translocated by the T3SS. Non-LEE-encoded effector A (NleA; also called EspI) is one of the first identified T3SS-translocated effectors located within a pathogenicity island separate from the LEE [342]. NleA is absent from strains of non-pathogenic *E. coli* but preferentially found in strains associated with human disease [422,423]. Although NleA's function has been strongly implicated in bacterial virulence *in vivo* [290,342,424], the mechanism underlying its strong effect during infection is not well characterized. In our recent publication, we noted an increase in apparent molecular weight of NleA upon its translocation into host cells and presented several lines of evidence that this mobility shift is specifically due to host-mediated mucin-type O-linked glycosylation of NleA [452]. Treatment of cells with brefeldin A and monensin, which inhibit Golgi trafficking, limited the apparent size shift of the NleA protein, implicating a role for the host cell secretory pathway in the modification of NleA.

The Golgi apparatus is the primary site of O-glycosylation and functions as the main sorting hub of the eukaryotic secretory pathway. Transport of proteins along this pathway begins at the endoplasmic reticulum (ER), where proteins are co-translationally translocated into the ER to undergo folding, modification, and quality control measures. After, proteins are segregated and exported from the ER via coat protein complex II (COPII) transport vesicles to either the ER-Golgi

intermediate compartment or Golgi for further targeting to their ultimate intra- or extracellular destinations. Notably, NleA has previously been shown to localize to the secretory pathway, colocalizing with markers of the Golgi apparatus [342,425,426], and the COPII complex has been identified as a binding partner of NleA [299,344]. NleA possesses no known Golgi targeting motifs, but a site-directed mutagenesis study of NleA identified a complex motif on NleA that when mutated, completely abrogated Sec24 binding ($NleA_{DM\Delta IIQ}$) [345]. Although this mutant was shown to be properly folded, it did not localize to the Golgi, and did not confer virulence in an *in vivo* mouse model.

In this addendum, we build on our previously published results to provide direct evidence for O-linked glycosylation of NleA and, using the non-Sec24-binding mutant, implicate COPII interaction as necessary for modification of NleA.

3.3 Results

NleA is O-glycosylated

Previous work has determined that the post-translational modification of NleA consisted of O-glycan additions to residues located between amino acid 169 and 183 and was not attributable to N-linked glycosylation, ubiquitination, or phosphorylation. [452]. To obtain further evidence of NleA's host-mediated O-glycosylation, we infected Chinese Hamster Ovary (CHO) ldlD cells. These cells lack the UDP-Gal/UDP-GalNAc 4-epimerase enzyme, preventing the synthesis of the O-glycan precursors UDP-galactose and UDP-N-acetylgalactosamine (UDP-GalNAc) [428]. However, sugar salvage pathways allow the cell to uptake galactose and GalNAc from the environment, restoring O-glycosylation mechanisms [453]. As we previously showed, when non-supplemented CHO ldlD cells were infected with EPEC UMD207 expressing FLAG-tagged NleA,

the apparent molecular weight of NleA was slightly above 55kDa. In contrast, the apparent molecular weight of NleA increased only when the cells were supplemented with GalNAc, and was further increased when GalNAc and galactose were supplemented (**Figure 1**, middle and bottom panels). The requirement of GalNAc for the mobility shift of NleA and its extension post-galactose supplementation implicates that NleA acquired mucin-type O-linked glycosylation under the form of Tn antigens extending further to core-1 O-glycans, or T antigen [453].

Next, we used lectins, proteins that bind specific glycans, to probe immunoprecipitated NleA from the conditions described above to directly assess the presence of carbohydrate additions on NleA. When probed with *Vicia villosa* lectin (VVL), which preferentially binds to a single α- or β-linked terminal GalNAc, a prominent band at the same size as NleA was revealed in the sample supplemented with GalNAc alone (**Figure 1**, top panel). Immunoprecipated NleA protein from cells without sugar supplementation or with only galactose had no discernible bands when probed with VVL (**Figure 1**). Interestingly, when galactose was used in combination with GalNAc, the VVL signal previously observed disappeared (**Figure 1**). Given that VVL binds only terminal GalNAc residues, the absence of signal with the dual complementation Gal+GalNAc therefore suggests that NleA O-glycosylation extends further than the Tn antigen towards a core-1 O-glycan structure which cannot be recognized by VVL [454].

No signal was observed when probed with lectins recognizing other core glycans (data not shown). Thus, we have obtained direct evidence of the initial addition of GalNAc to NleA, implicating mucin-type O-linked glycosylation as the host-mediated modification. Furthermore, the loss of VVL reactivity, combined with the additional increase in apparent molecular weight when both sugars are supplemented, strongly suggests that the Tn antigens are further processed in the Golgi apparatus creating core-1 based O-glycan structures.

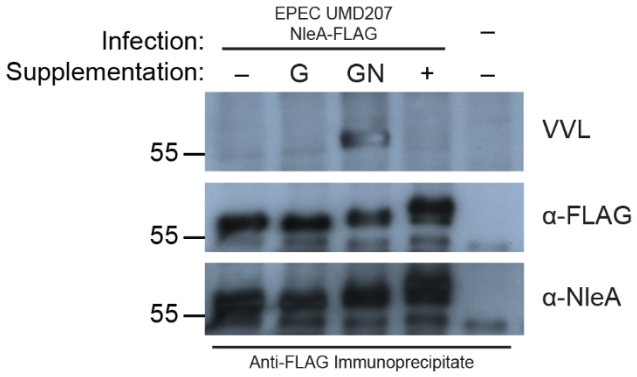

Figure 1. NleA is modified by O-linked glycosylation. Western blot analysis of immunoprecipitate from CHO ldlD cells infected with EPEC UMD207 NleA-FLAG probed with *Vicia villosa* lectin (VVL), and anti-FLAG, and anti-NleA antibodies. Cells were cultured in media without (-), with galactose (G), with GalNAc (GN), or with both galactose and GalNAc (+) sugar supplementation. Migration of molecular weight markers (kDa) is indicated on the left.

Modification of NleA is dependent on binding to Sec24

Mucin-type O-linked glycosylation occurs within the secretory pathway of host cells [455]. Considering that NleA localizes to the secretory pathway and binds the Sec24 component of the COPII vesicle coat, we took advantage of a mutant $NleA_{DM\Delta IIQ}$ protein previously described to have significantly diminished interaction with Sec24 and loss of localization to the secretory pathway [345] to assess whether this impacted the host-mediated modification of NleA. HeLa cells were infected with bacteria expressing WT NleA or $NleA_{DM\Delta IIQ}$ and the apparent molecular weight of NleA was analyzed by Western blot (**Figure 2**). As expected, WT NleA displayed a significant mobility shift from its bacterial lysate size of 55 kDa (lane 1) to approximately 60-65 kDa in infected host cell lysates (lane 2). However, no such shift was seen in cells infected with bacteria expressing $NleA_{DM\Delta IIQ}$, where the bacterial and host-associated NleA protein was visible at 55 kDa (lane 3 and 4, respectively). As a control for bacterial protein translocation, blots were assessed for another T3SS-translocated protein, Tir, which is known to undergo a mobility shift due to host-mediated phosphorylation, and the host protein α-actinin was assessed as a control for equal loading of cell lysates. Together, this data indicates that NleA's interaction with the Sec24 component of the COPII complex seems to be important for its host-mediated post-translational modification to occur.

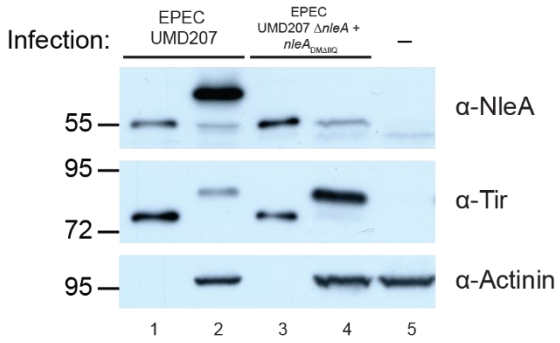

Figure 2. NleA deficient in Sec24 binding does not get modified. Western blot analysis of HeLa cell lysate infected with EPEC UMD207 expressing WT NleA (lane 2), $NleA_{DM\Delta IIQ}$ (lane 4), or uninfected (lane 5). Bacterial lysates from the indicated strain are present in lanes 1 and 3. Blots were probed with anti-NleA, anti-Tir, and anti-Actinin antibodies. Migration of molecular weight makers (kDa) is indicated on the left.

3.4 Concluding Remarks

In this addendum, we have built on our previous findings which indirectly implicated mucin-type O-linked glycosylation in the host-mediated modification of NleA by now providing direct evidence for this modification. Given the pattern of lectin binding to NleA when cells were cultured in media supplemented with both galactose and GalNAc, as well as the size of the mobility shift, we suspect that the final glycan addition on NleA is more complex and more extensive than the Tn antigen form of a single carbohydrate. Future work could narrow down which amino acids are substituted, and which glycans are present using single amino acid replacement of highly predicted sites and mass spectrometry technologies.

One question that remains is how NleA, an exogenous protein that is translocated into host cells fully translated, encounters the O-glycan synthetic machinery involving at least polypeptide GalNAc-transferases, C1GalT-1 and Cosmc [455–457]. The interaction of NleA with host secretory pathway proteins was not investigated here, as we and others have identified only Sec23 and Sec24 as NleA binding partners of the COPII complex [299,446]. Given that we showed NleA binding Sec24 is upstream of its host-mediated glycosylation, we might propose that the NleA-Sec24 interaction could aid in its localization to the secretory pathway via the anterograde trafficking vesicles. However, it is also possible that NleA follows a path of eukaryotic proteins known to enter the ER via post-translational translocation. Although eukaryotic proteins are typically translocated into or across the ER membrane co-translationally, several pathways for post-translational translocation have been described (reviewed in [388]). In this process, specific fully synthesized proteins are transported from the cytoplasm into the lumen of the ER via accessory proteins that feed a polypeptide chain into the pore and drive translocation. It may not be surprising, then, that NleA is secreted into the cytoplasm of the host cell where it is trafficked

to the protein secretory pathway, despite NleA having no identified host chaperone binding partners or Golgi-targeting motifs [342]. While NleA signal has been observed to colocalize with markers of the secretory pathway during infection [342,425,426], $NleA_{DM\Delta IIQ}$ instead showed diffuse, sustained cytosolic localization within the cell [345]. Therefore, Sec24 binding is important for NleA modification and localization, but we cannot be sure which event precedes the other. Future studies are warranted to further dissect trafficking and localization of NleA over the course of infection to determine a timeline and trajectory of its pathway through the host cell.

In vivo studies demonstrate that NleA has an essential role in bacterial virulence. However, the mechanism by which NleA provokes disease, and the involvement of its O-linked glycosylation in this function, is not fully clear. We have previously shown that *C. rodentium* expressing a non-modifiable NleA mutant were indistinguishable from WT in an acute mortality mouse model, but did have a modest increase in persistence in a mixed infection model [452]. Thus, the consequence of the host-mediated modification of NleA may be more relevant to host response to infection rather than to bacterial pathogenicity. We anticipate that future studies will provide clarity on the significance of NleA O-glycosylation on host immune modulation.

Overall, the results in this addendum augment our understanding of the interaction between NleA and host during infection, providing additional evidence of the protein secretory pathway as the locale for its host-mediated O-linked glycosylation.

PREFACE TO CHAPTER 4

In vivo studies demonstrate that NleA has an essential role in bacterial virulence. However, the mechanism by which NleA provokes disease, and the involvement of its O-linked glycosylation in this function, is not fully clear. In **Chapter 2**, we showed that *C. rodentium* expressing a non-modifiable NleA mutant were indistinguishable from WT in an acute mortality mouse model but did have a modest increase in persistence in a mixed infection model. Thus, the consequence of the host-mediated modification of NleA may be more relevant to host response to infection rather than to bacterial pathogenicity. The work in **Chapter 4** focuses on assessing the consequence of NleA glycosylation on protein stability and modulation of host cell responses. We developed a technique to determine whether bacterial effectors have been delivered into infected cells and, using this method to track intracellular NleA signal, we have observed a potential stability effect when NleA remains unglycosylated.

CHAPTER 4: **Host-mediated O-glycosylation of bacterial effector NleA may affect its stability inside host cells**

4.1 Abstract

Attaching and effacing (A/E) pathogens, including enteropathogenic and enterohaemorrhagic *E. coli* (EPEC and EHEC), utilize injected effector proteins to alter host cell biology to conditions favorable for the bacteria. Non-LEE-encoded effector A (NleA) is a crucial effector for A/E pathogen virulence and has been shown to localize to the host secretory pathway where it undergoes host-mediated O-linked glycosylation. *In vivo* studies indicate, however, that the contribution of NleA to bacterial virulence is not dependent on this post-translational modification, and that glycosylation of NleA may confer a selective disadvantage to bacterial persistence. Therefore, we aimed to determine the importance of glycosylation to NleA protein stability and the modulation of host pathways in response to infection. We generated epitope-tagged WT and non-modifiable NleA proteins for the detection of intracellular NleA protein by flow cytometry. Using this system, we performed infection time courses in which we found that glycosylated NleA may have reduced stability in the host cell. This is supported by band densitometry quantification demonstrating that unglycosylated NleA appears more stable in the host cell.

4.2 Introduction

Enteropathogenic and enterohaemorrhagic *E. coli* (EPEC and EHEC) are gram-negative extracellular pathogens that infect the human intestine and cause diarrheal disease [33]. *Citrobacter rodentium* is a mouse-specific pathogen that is widely used as a murine model of EPEC and EHEC infections [106]. EPEC, EHEC, and *C. rodentium* are members of the group of bacteria

known as attaching and effacing (A/E) pathogens. Despite slight differences in disease manifestation, all A/E pathogens form distinct lesions on the host epithelium during infection. These pathogens attach to the host epithelium, efface the local microvilli, and form actin-rich pedestals beneath adherent bacteria [2,3]. A/E pathogens also possess a type 3 secretion system (T3SS), forming a channel that traverses both inner and outer bacterial membranes terminating in a pore in the host cell plasma membrane [212]. Thus, the T3SS functions as a system for the delivery of effector proteins directly from the bacteria into the cytosol of a target cell [211]. Each A/E pathogen has a unique repertoire of effector proteins, which alter cellular pathways to promote bacterial survival and replication. There is a growing appreciation of the complex network of interconnectivity between bacterial effectors and their functions, thus, it can be difficult to determine the contribution of a single effector to the overall changes in host cell biology in response to infection. Indeed, many studies of effector function have relied on single effector gene deletions, which typically exhibit no *in vivo* colonization phenotypes [458].

One effector, Non-LEE-encoded effector A (NleA), is crucial for virulence of A/E pathogens [342]. NleA is translocated into the host cell via the T3SS, where it localizes to the secretory pathway and inhibits cellular protein trafficking [299,342]. We have also demonstrated that NleA is modified by host-mediated mucin-type O-linked glycosylation in this pathway [452,459]. Despite clear implications of NleA on virulence, whereby its deletion renders bacteria avirulent, we observed only subtle effects when glycosylation of NleA is impeded [452]. In addition, glycosylation of NleA seemed to be largely deleterious to bacterial persistence *in vivo* [452]. Therefore, we suspect that NleA's glycosylation may be important for its trafficking or stability in the host cell, and functionally may be more involved in altering host response to infection rather than bacterial pathogenicity. The contribution of NleA to host response modulation

has been difficult to determine due to the very limited quantities of protein delivered into host cells during infection and the presumed few injected target cells. As a result, we aimed to develop a technique to characterize the contribution of NleA by detecting changes in cell responses between those that were injected with NleA versus cells that remained uninjected. We also wanted to further characterize whether glycosylation of NleA is involved in altering these responses and what effect it has on NleA protein stability and trafficking inside the host cell. To do this, we created an epitope-tagged form of WT NleA and non-glycosylatable mutant NleAΔ169-183 with C-terminal FLAG, a short, hydrophilic peptide of eight amino acids: DYKDDDDK. This tag allowed us to specifically recognize intracellular NleA-FLAG or NleAΔ169-183-FLAG by flow cytometry and Western blot using anti-FLAG antibody for more precise analysis of cellular responses and NleA stability based on glycosylation and host cell injection status.

4.3 Experimental Procedures

Bacterial strains and growth conditions

Bacterial strains were routinely cultured at 220 rpm at 37°C in Luria Bertani (LB) agar or broth (1% (wt/vol) tryptone, 0.5% (wt/vol) yeast extract, 1% (wt/vol) NaCl). When appropriate, LB was supplemented with 100 μg/ml ampicillin. The complemented EPEC UMD207 Δ*nleA* + WT *nleA*-FLAG and EPEC UMD207 Δ*nleA* + *nleA*Δ169-183-FLAG were created as described previously [452].

Cell culture and infection

CMT-93 (ATCC no. CCL-223) were cultured in Dulbecco's minimal Eagle's medium (DMEM; Wisent) supplemented with 10% (v/v) heat-inactivated fetal calf serum (Wisent) and 1

mM sodium pyruvate. Cells were grown at 37°C in a humidified atmosphere of 5% CO_2. For each sample, at least two confluent wells of a 6 well plate were infected with EPEC UMD207 Δ*nleA* + WT *nleA*-FLAG or EPEC UMD207 Δ*nleA* + *nleA*Δ169-183-FLAG using an initial multiplicity of infection (MOI) of 200 and incubated at 37°C in a 5% CO_2 atmosphere.

For Western blot analysis, cells were washed at 2, 3, 4, and 6 hours post-infection with sterile ice-cold PBS, treated with 0.25% trypsin, blocked with DMEM 10% FBS, and fractionated by re-suspension in 0.5 ml of radioimmunoprecipitation assay (RIPA) lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% (v/v) NP-40, 0.5% (w/v) sodium deoxycholate, 1 mM EDTA, 0.1% (w/v) SDS, 1 mM NaF, 100 μM Na_3VO_4) supplemented with cOmplete protease inhibitor cocktail (Roche). Cell suspensions were sonicated and centrifuged at 16,000 × g for 15 min at 4°C. The supernatant containing the whole cell lysate was made up to 1X Laemmli buffer using a 5X stock for resolution by SDS-PAGE. At 6 hours post-infection, the wells for the 8 and 24 hour timepoints were washed with sterile ice-cold PBS and the media was replaced with DMEM 10% FBS supplemented with 10 μg/ml gentamicin and placed back at 37°C, 5% CO_2. Cells were then collected at 8 and 24 hours post-infection and prepared for immunoblot analysis as described above.

For analysis by flow cytometry, infected cells were washed at 6 hours post-infection with sterile ice-cold PBS, and either collected as outlined above or replaced with DMEM 10% FBS supplemented with 10 μg/ml gentamicin and placed back at 37°C, 5% CO_2 for collection at 8 and 24 hours post-infection. To collect, cells were treated with 0.25% trypsin, blocked with DMEM 10% FBS, resuspended in 1X PBS, then processed for flow cytometry.

Flow cytometry

Cells were transferred to black 96-well conical plates for staining. For viability assessment, cells were stained with 1:2000 fixable viability dye eFluor 780 (Invitrogen) in phosphate-buffered saline (PBS) for 15 min on ice, protected from light, then washed with 2% FBS in PBS. Cells were fixed with 2% paraformaldehyde for 20 min at room temperature and subsequently permeabilized using eBioscience Foxp3/Transcription Factor Permeabilization Buffer (Invitrogen) for 10 min on ice. APC- or eFluor 450-conjugated anti-FLAG (BioLegend) and mouse anti-Tir [361] primary antibodies for intracellular markers were diluted 1:250 in permeabilization buffer and incubated with cells for 30 min on ice. This was followed by incubation with Alexa Fluor 488 labeled goat anti-mouse IgG (H+L) secondary antibody (Abcam) diluted 1:500 in permeabilization buffer for 30 min on ice. Samples and controls were acquired using the LSRII Fortessa 5L (BD) and analyzed using FlowJo v10.9 software (BD).

Immunoblot analysis

Samples for Western blot analysis were resolved by 7.5% polyacrylamide gel at 140V for 75 min and transferred electrophoretically at 10V for 30 min onto an activated PVDF membrane. Immunoblots were blocked in 5% non-fat dried milk (NFDM) in TBS, pH 7.2, containing 0.1% Tween 20 (TBST) overnight at 4°C. Membranes were then incubated with primary antibody in 5% NFDM TBST overnight at 4°C before being washed five times in TBST and then incubated with horseradish peroxidase-conjugated secondary antibody for 45 min at room temperature. Membranes were washed as described above and visualized by developing in Immobilon Western ECL Substrate (Millipore), followed by exposure to autoradiography film (Denville). Band densitometry analysis was performed with Fiji software [460].

Statistical analysis

Data analyses were performed using GraphPad Prism v9.0 software. Statistical comparisons were carried out using tests described in the figure legends with a $p<0.05$ considered significant.

4.4 Results

A novel method for determining T3SS injection status of host cells

To determine if FLAG-tagged NleA could be detected by flow cytometry, CMT-93 cells were infected with UMD207 Δ*nleA* expressing WT NleA-FLAG or NleAΔ169-183-FLAG for 6 hours. Both intracellular NleA-FLAG and NleAΔ169-183-FLAG were detectable by flow cytometry 6 hours post-infection (**Figure 1A**). We found that while a small population of cells remained negative for injected effectors 6 hours post-infection, another population of cells had intracellular NleA-FLAG signal that strongly positively correlated to Tir signal (**Figure 1B**).

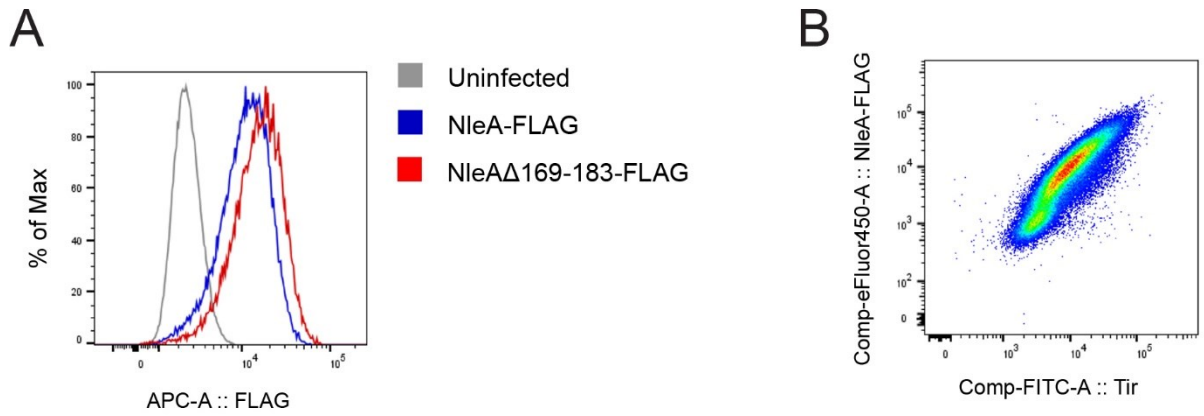

Figure 1. Intracellular FLAG-tagged NleA can be detected by flow cytometry. (A) Histogram of anti-FLAG fluorescence detected intracellularly 6 hours post-infection. (B) Dot plot of anti-Tir:anti-FLAG events gated on live CMT-93 cells 6 hours post-infection.

The modification of NleA may impact its stability

CMT-93 cells were infected with EPEC UMD207 expressing WT NleA-FLAG or NleAΔ169-183-FLAG. Cells were washed and collected according to the schematic in **Figure 2A** to be analyzed by Western blotting and flow cytometry for NleA-FLAG and Tir signal. Cells maintained relatively high viability over the course of the 6 hour infection and following gentamicin treatment (**Figure 2B**). At 6 and 24 hours post-infection, there were significantly fewer CMT-93 cells injected with NleA-FLAG than NleAΔ169-183-FLAG (**Figure 2C**). Additionally, the mean fluorescence intensity (MFI) of FLAG signal was significantly lower at all timepoints tested in FLAG-positive cells infected with EPEC UMD207 expressing WT NleA-FLAG. This likely cannot be attributed to an effector injection defect as Tir signal was comparable at all time points (**Figure 2D**). These findings are further supported by band densitometry analysis of infected CMT-93 whole cell lysates (**Figure 2E**). Band densitometry analysis identified significantly greater normalized NleAΔ169-183-FLAG signal compared to NleA-FLAG at 5, 6, and 8 hours post-infection (**Figure 2F**). Again, there were no significant differences in Tir band intensity at any timepoint.

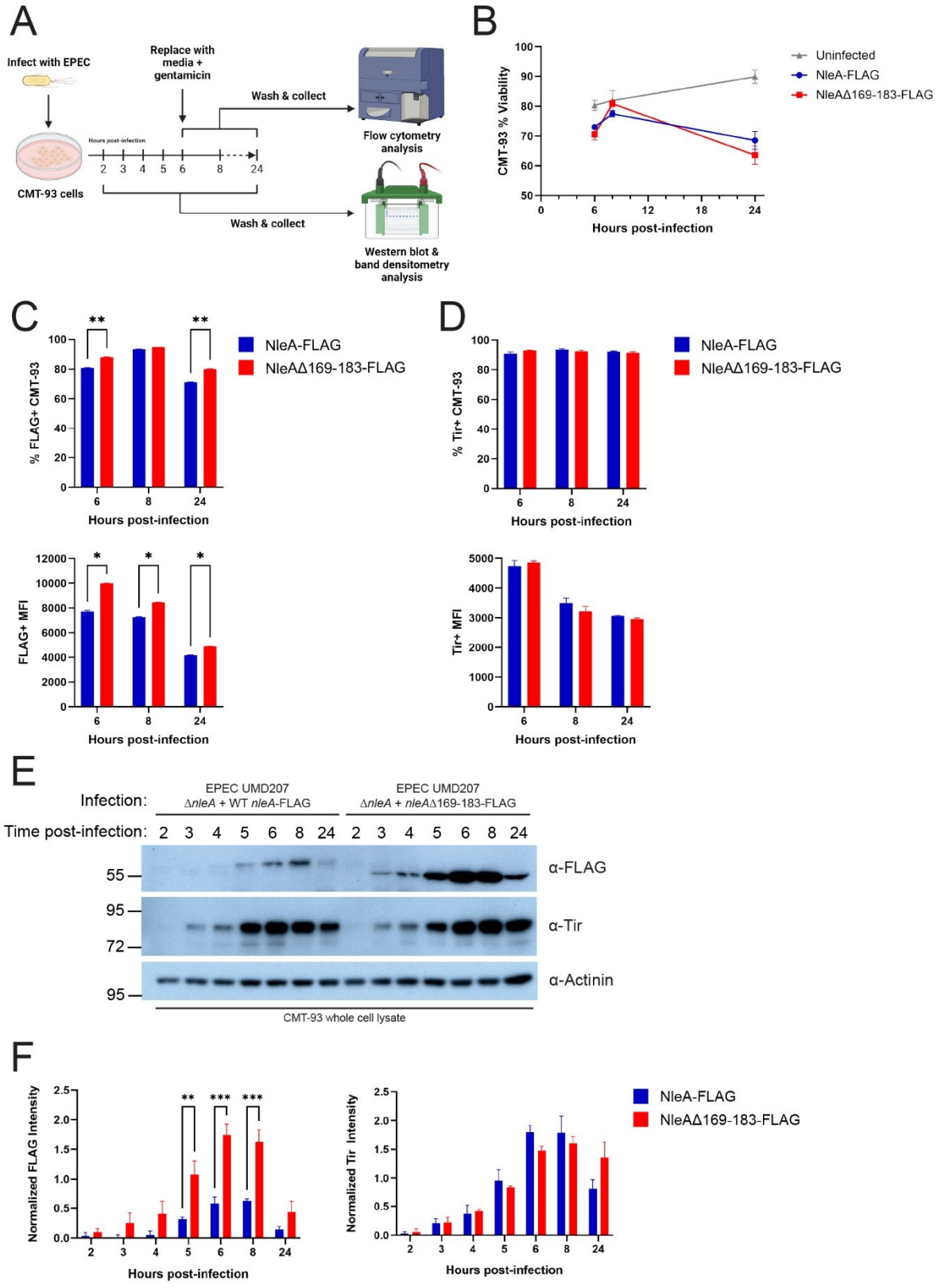

Figure 2. Non-glycosylatable NleA may be more stable in the host cell. (A) Experimental design. Created with BioRender.com. (B) Cell viability post-infection. Percent cells positive for

and MFI of (C) NleA-FLAG and (D) Tir in host cells over time. Data represented as mean ± S.E.M. Pooled technical replicates from one representative experiment of four. (E) Representative Western blot analysis of whole cell lysates from CMT-93 cells infected with EPEC UMD207 NleA-FLAG or NleAΔ169-183-FLAG. Migration of molecular weight makers (kDa) is indicated on the left. (F) Band densitometry analysis of left: FLAG and right: Tir signal normalized to actinin loading control from the same lane. 10 ug of total protein lysate run on three independent blots. Data represented as mean ± S.E.M. Two-way ANOVA with Sidak's post-hoc analysis was used to determine statistical significance (*, $p<0.05$).

4.5 Discussion

Bacterial effector proteins are important virulence determinants, modifying host cell biology to facilitate colonization and the provocation of disease following injection into target host cells. These effectors often have synergistic and overlapping functions that are difficult to fully appreciate with single gene deletions or individual transfection experiments [23,192]. For characterization of NleA function, this is made more difficult by the very limited quantity of protein injected into target cells and the presumed very small number of cells injected during infection *in vivo*. Preliminary *ex vivo* experiments of colonic epithelial cells collected from infected C57BL/6J mice showed fewer than 1% of cells positive for NleA injection at 5 to 12 days post-infection (data not shown). By considering all infected cells, both injected and uninjected, as one condition, we may be further underestimating the amount of effector protein delivered into a host cell and limiting our ability to identify changes in protein stability or cell responses. Therefore, by detecting intracellular effectors to determine a cell's injection status, we can more accurately investigate differences between uninjected, injected, and uninfected cell populations. This therefore allows us to investigate alterations in cellular processes between injected and uninjected cells by combining detection of effectors and various cell and immune markers. By infecting cells with bacteria expressing tagged WT NleA versus a non-modifiable NleA deletion mutant, we anticipate we will identify changes in infected cell biological pathways that are specifically caused by NleA and the effect of NleA modification on these changes.

Here, we applied a combination of effector epitope tagging and flow cytometry to monitor the injection and stability of NleA over the course of and following infection. Full-length WT NleA or non-modifiable NleAΔ169-183 were tagged with C-terminal FLAG. FLAG is a short, eight amino acid peptide that was chosen for its small size to still allow translocation of the proteins

through the T3SS. We detected the presence of intracellular NleA-FLAG or NleAΔ169-183-FLAG at 6 hours post-infection that was not seen in uninfected cells, and used this signal to determine host cell injection status (**Figure 1A**). We also found that NleA signal strongly positively correlated with Tir signal, another injected effector protein (**Figure 1B**). Tir is the first effector protein injected into host cells following bacterial attachment [227], and it was promising to observe co-signaling of NleA with another injected effector. Although we did see a population of cells negative for both Tir and NleA, and very few Tir single positive cells, we did not observe any NleA single positive cells (**Figure 1B**). This suggests that by detecting intracellular epitope-tagged NleA, we can determine host cell effector injection status with relative accuracy.

NleA is known to be required for pathogen virulence, however, our previous findings determined that host-mediated mucin-type O-linked glycosylation of NleA was not required for pathogenicity [452]. Protein glycosylation is an effective way of generating diversity among proteins and modulating their properties due to glycan structural variations. In addition, glycans are bulky hydrophilic molecules that often contribute to high protein solubility and increased stability against denaturation or proteolytic degradation [461]. Therefore, we hypothesized that while glycosylation of NleA may not have an obvious effect on virulence, it may be involved in stabilizing NleA inside the host cell to aid in its trafficking or function. We employed epitope-tagged NleA to examine protein stability during and following infection. CMT-93 cells were infected with EPEC UMD207 expressing WT NleA-FLAG or NleAΔ169-183-FLAG. Cells were washed to remove non-adherent bacteria and collected at 2, 3, 4, and 5 hours post-infection for analysis by Western blot. At 6 hours, cells were washed and either collected for Western blot and flow cytometry or were incubated for up to 24 hours post-infection with culture media containing gentamicin to kill any remaining bacteria and prevent further effector translocation. Again, cells

were washed and collected at 8 and 24 hours post-infection for Western blot and flow cytometry analysis.

We observed comparable cell viability between infected and uninfected conditions over the time course, with a slight decrease in infected cells regardless of the NleA form expressed (**Figure 2B**). Therefore, any perceived differences in injection efficiency or NleA protein amount were not due to cell death between collection timepoints. There was a significantly greater percentage of CMT-93 cells injected with NleAΔ169-183-FLAG compared to NleA-FLAG at both 6 and 24 hours post-infection (**Figure 2C**). Interestingly, there was no difference in the percent CMT-93 cells positive for either form of NleA 8 hours post-infection. However, NleAΔ169-183-FLAG mean fluorescence intensity (MFI) was significantly higher at all three timepoints, suggesting greater protein abundance. These results indicate that NleAΔ169-183-FLAG, which does not get glycosylated, may be more stable inside the host cell. We do not think these differences can be attributed to an increase or decrease in effector injection as a whole, since both the percent of Tir positive cells and Tir MFI remained similar at all timepoints (**Figure 2D**). In addition, NleA structure predictions indicated that the modifiable region was external to the central structure of the protein and that NleAΔ169-183 retained the same core structure as the full-length protein [452]. Therefore, we do not think the observed differences in NleA MFI are due to alterations in the NleA protein structure resulting from deletion of this modified region.

This finding is supported by Western blot quantification in which band densitometry was performed on infected CMT-93 cell lysates to compare FLAG or Tir signal to a housekeeping protein, in this case alpha-actinin, in order to relatively quantify protein signal intensity (**Figure 2E**). There was significantly greater normalized FLAG intensity at 5, 6, and 8 hours post-infection with a trend to be higher at all other timepoints, generally matching what we saw with flow

cytometry (**Figure 2F**). In addition, we saw no major differences in normalized Tir intensity at any point during infection or following cessation.

However, an alternate hypothesis for these preliminary findings is reduced or delayed secretion of WT NleA-FLAG versus NleAΔ169-183-FLAG into the host cell. This is mostly evident by Western blot, in which WT NleA-FLAG is not detected until 5 hours post-infection compared to NleAΔ169-183-FLAG that is prominent at 3 hours post-infection and remains greater in intensity at all other timepoints (**Figure 2E)**. We therefore cannot conclude with certainty from these results that WT NleA-FLAG is lost more quickly, due to a stability defect, rather than due to differences in the amount of protein initially translocated to the cell. Whether the introduced deletion of the serine- and threonine-rich region increases secretion efficiency could be tested in three ways. Firstly, using our FLAG-tagged NleA system we can determine if WT NleA-FLAG has increased signal compared to bacteria with NleAΔ169-183-FLAG within EPEC by flow cytometry and Western blotting of bacterial lysate. Secondly, the stability of NleA could be tested by infecting either CHO ldlD cells deficient in glycosylation or a primary cell system such as colonic organoids from mice that lack the necessary glycosyltransferases for the modification of NleA. Both of these would enable us to monitor NleA stability without mutation of the modified region. Finally, we could treat cells with small molecule inhibitors of GALNTs, such as benzyl-α-GalNAc, which blocks galactosyltransferase activity by acting as a decoy substrate [462]. Again, this method permits analysis of modified versus unmodified NleA stability in the host cell without altering the protein sequence.

These results were largely surprising, as we expected glycosylated NleA to be more stable within the host cell, since this post-translational modification is not typically a marker for protein degradation. However, we have not yet determined whether mucin-type O-glycosylated NleA is

degraded by the host or if its apparent reduced stability compared to non-glycosylatable NleAΔ169-183-FLAG is instead due to an increase in protein secretion. In addition to stability, O-linked glycosylation has well-established roles in protein secretion, processing, and function [463]. Mucin-type O-glycans can be found on many cell surface and secreted proteins where they are important for modulating recognition, adhesion, and communication between cells and their environment [463]. Therefore, future studies will utilize inhibitors of host protein degradation and secretion pathways to determine the fate of NleA once translocated into the host cell and glycosylated. We previously showed that treatment of cells with inhibitors of ER and Golgi protein secretion, brefeldin A and monensin, resulted in a truncated glycan structure added to NleA [452]. While these results identified the host secretory pathway as the site of NleA glycosylation, we could not conclude whether NleA is eventually secreted to the host membrane or extracellularly via this pathway. Inhibitors such as VR23 and bafilomycin A1, which respectively limit proteasomal and lysosomal degradation, are being optimized for use in our time course infection system. We anticipate that by treating CMT-93 cells with these inhibitors prior to infection and detecting intracellular NleA-FLAG, we will determine whether NleA is degraded in either of these pathways.

Besides determining the effect of glycosylation on NleA stability, the usefulness of the technique presented here extends to investigating the contribution of NleA to modulating the host response to infection. However, *nleA* deletion results in complete attenuation of bacterial virulence, making it challenging to identify subtle alterations in cellular pathways by NleA. Other effectors are unable to functionally compensate for the loss of NleA in single gene deletion experiments [343], suggesting that this method is a viable way of investigating NleA function. The technique of separating cells based on effector injection status will allow us to characterize cellular responses

in injected cells, uninjected cells in infected conditions, and uninfected cells. Preliminary phenotyping of the immune response across these cell populations has identified several interesting host pathways that may be affected by characterized and uncharacterized functions of NleA.

CHAPTER 5: **Discussion**

5.1 Overview

The study of host-pathogen interactions is exceedingly complex due to the multifaceted and dynamic nature of host and bacterial mechanisms that are numerous and often intersecting. Effectors are an invaluable bacterial tool for modulation of host cell pathways. Despite their well-established role as host disease determinants, the mechanisms by which many effectors contribute to bacterial pathogenicity remain ill-understood. The effector NleA is required for bacterial virulence and has characterized functions in inhibiting protein trafficking and inflammasome activation. In **Chapter 2**, we characterized the host-mediated modification of NleA as O-linked glycosylation and demonstrated that modification of NleA is not required for its contribution to virulence. In **Chapter 3**, we obtained further evidence of O-glycosylation of NleA and determined that NleA modification is dependent on its interaction with the COPII vesicle coat. In **Chapter 4**, we began to probe the role of glycosylation on stability of NleA within the host cell. We found that non-modifiable NleA may be more stable inside the host cell than O-glycosylated NleA. How these findings fit into the larger context of host-pathogen interactions and what their implications may be are discussed below.

5.2 The bacterial effector NleA is modified by O-glycosylation

While the mobility shift of NleA had been observed for several years, the PTM responsible for this shift and relevance to NleA function remained uncharacterized. In **Chapter 2**, we used the CHO ldlD cell line with a reversible defect in glycosylation to provide indirect evidence for the modification of NleA by mucin-type O-linked glycosylation. We found that modification of NleA was restored upon supplementation of the culture media with GalNAc, although not to the size we

observed when both GalNAc and Gal were supplemented. In **Chapter 3**, we presented further direct evidence of NleA O-glycosylation using specific lectin blotting. The VVL lectin is specific to the Tn antigen, a single GalNAc residue bound to serine or threonine, and does not recognize this motif when it is modified further [454,464]. We observed strong signal solely when infected CHO ldlD cells were cultured with supplemented GalNAc, and not with Gal alone or both sugars. Given that the atomic mass of sugars commonly found in O-glycans ranges from approximately 0.16 to 0.31 kDa and the observed size shift of NleA is about 10 kDa, a modification of this size could contain upwards of 40 sugar residues. Together, these results indicate that NleA is likely glycosylated on more than one serine or threonine, and that this modification extends further than a single GalNAc attached to the protein.

To identify the specific site and structure of the modification of NleA, we sought to use matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). However, after several months of optimization and significant upscaling of our *in vitro* infection system, we were not able to obtain sufficient modified NleA protein to complete MALDI-TOF MS and the preceding purification steps. Collection of glycosylated NleA is substantially hindered by the limited amount of NleA translocated into host cells. Additionally, obtaining infected whole cell lysates and subsequent immunoprecipitation of modified NleA is time intensive and laborious. Together, these barriers limit collection of modified NleA to microgram quantities. We previously performed liquid chromatography mass spectrometry analysis that also proved inconclusive (unpublished). A region of NleA was detected in bacterially secreted NleA that was absent from the mass spectrum of the host-translocated form. This was suggested to be characteristic of heavily O-glycosylated peptides, however this method was not able to confirm this assumption. We later

identified this peptide as the serine- and threonine-rich region predicted to be highly modified by O-linked glycosylation.

The inherent complexity of O-glycans and the diversity of core structures constitute fundamental challenges for the analysis of mucin-type O-glycans [465]. Due to the often isomeric monosaccharides that can be oriented in α and β configurations, the structure of O-linked oligosaccharides can be difficult to distinguish by mass spectrometry alone [466]. In addition, there is no universal enzyme able to cleave all O-glycan core structures, instead requiring harsh chemical extraction, meaning it can be challenging to remove O-glycans while keeping both the glycan and protein intact [466]. With these challenges in mind, mass spectrometry is still a valuable tool for glycan analysis due to its prevalence in proteomics and exceptional sensitivity [466]. In recent years, development of novel ion activation methods, commercialization of ion mobility mass spectrometry, progress in gas-phase ion spectrometry, and advances in computational chemistry have led to innovation in mass spectrometry-based glycan analysis [466]. As a result, the amount and specificity of structural information that can be obtained by mass spectrometry-based methods has significantly increased [466].

Alternatively, treatment with O-glycoproteases that cleave the peptide backbone based on the presence of specific O-linked glycans may be useful in obtaining more information about the exact glycan structure present on NleA. For example, OgpA, derived from *Akkermansia muciniphilia*, cleaves at the N-terminus of serine and threonine residues that are glycosidically linked to core 1 O-glycans [467]. Other bacteria-derived glycoproteases, including StcE, ZmpC, and CpaA, cleave glycosylated serine and threonine residues with unique specificities [468]. Treatment of modified NleA with various bacterial glycoproteases with unique specificities

followed by analysis by SDS-PAGE to observe cleaved protein sizes could provide information about the glycan motifs present in the structure of the host-mediated modification of NleA.

In addition to enzymes that cleave the peptide backbone, glycosidases can be used to release sugars from the glycan chain. Endoglycosidases remove oligosaccharides from the protein at the site of attachment or within the glycan, while exoglycosidases trim monosaccharides from the non-reducing termini of the glycan chain [469]. Although a universal endo-O-glycosidase capable of cleaving intact O-glycans from the protein has not been characterized, some glycosidases remove truncated O-glycan structures [467]. Commercially-available glycosidases from *Streptococcus pneumoniae* and *Enterococcus faecalis* release core 1 and some core 3 O-glycans [467]. Due to their innate specificity, exoglycosidase treatment is commonly used to simplify glycoproteomic analyses by trimming structures for targeted characterization of glycan epitopes [467]. In **Chapter 2**, we showed that treatment with a mixture of glycosidases that remove all N-linked and simple O-linked glycans did not reduce the apparent molecular weight of modified NleA. While we can conclude that N-glycans are not present on modified NleA, the limitation of the O-glycosidases present in the mixture to remove only truncated O-glycan structures is inconclusive for determining the structure of the modification. Thus, sequential treatment with exoglycosidases to remove terminal residues followed by endoglycosidase release of truncated structures may be more successful in reducing the apparent protein size of NleA by Western blot.

As we showed in **Chapter 2**, NleA transfected into host cells still gets modified, so this may be a potential method of increasing the amount of glycosylated NleA protein that can be collected from infected cell lysates. With transfection, NleA is translated in the cell rather than delivered fully translated like when it is translocated through the T3SS. We have not yet investigated whether there are any differences in protein folding or localization with this

discrepancy. Alternatively, an *E. coli* expression system that uses a dual-plasmid approach, one plasmid encoding the target protein and another the O-glycosylation machinery, can produce proteins modified by human-like O-glycosylation [470,471]. This technique was able to produce up to 100 mg/ml of target protein with 85% glycosylated [471]. However, as of now this method is limited to addition of Tn or T antigens. So, while this system may be useful for obtaining significant amounts of NleA modified by a single GalNAc, it would not be representative of the complete scale of NleA glycosylation in infected cells. Altogether, improvements in obtaining modified NleA and more sensitive mass spectrometry analysis are promising for the technique to be revisited in future studies.

Discovery of O-glycosylated proteins is far from exhausted, with O-glycoproteome analysis predicting that upwards of 83% of proteins entering the ER-Golgi secretory pathway are O-glycosylated [430]. In addition to the difficulties already mentioned, the analysis of O-glycosylation has added challenge due to the lack of consensus sequence. While prolines are often found adjacent to O-glycosylation sites, this is a mere statistical effect and is of little analytical use [466]. O-glycosylation positions within a protein are therefore considerably more challenging to identify than N-linkages [466]. In **Chapter 2**, we utilized NetOGlyc 4.0 software to identify 22 predicted modification sites in the EPEC NleA protein sequence. However, NetOGlyc 4.0 is incapable of predicting isoform-specific glycosylation sites or which GALNTs are responsible for their modification [472]. Although it is well accepted that most GALNT family members are isoenzymes with some degree of substrate redundancy, GALNTs also possess activity toward specific targets [473]. Recent findings demonstrate that GALNTs possess unique specificities for serine and threonine residues, peptide sequence, and prior remote or nearby glycosylation [473]. It was also shown that prior GalNAc glycosylation could direct subsequent glycosylation, with

flanking charged residues altering the rates of glycosylation [474–476]. These results suggest that nearby peptide charge is an additional factor regulating mucin-type O-glycosylation that is not accounted for by many existing glycosylation prediction software.

The ISOGlyP software was developed to address several of the aforementioned shortcomings, enabling glycosylation site prediction based on GALNT isoform preferences [472]. ISOGlyP predicts the propensity of a particular amino acid to be glycosylated by a given GALNT, where a residue with value greater than one is favoured and a value less than one is disfavoured. When the *C. rodentium* NleA protein sequence is input into ISOGlyP, several residues have values greater than one corresponding to many different GALNT isoforms. Of particular interest is threonine at position 177 in the amino acid sequence that has a predicted value of 14.79 and 15.47 for modification by GALNT1 and GALNT11, respectively. This threonine is contained within the serine- and threonine-rich region predicted to be highly modified by NetOGlyc 4.0 and absent from several clinical strains. In the catalytic domain of some GALNTs, a subset of residues form a hydrophobic "proline pocket" [477], resulting in a preference of some GALNTs for acceptor peptides containing a proline three residues to the C-terminus of the threonine or serine to which the GalNAc is added [478–480]. Interestingly, the threonine residue in the NleA sequence identified by ISOGlyP to be most highly favoured for O-GalNAc modification has a proline three residues towards the C-terminus. To determine if this threonine could be an initiator site from which further glycosylation follows, future experiments could perform single amino acid replacement, in which the threonine is mutated to an alanine, and observe whether NleA still gets modified.

Notably, ISOGlyP prediction of modification sites in the *C. rodentium*, EPEC, and EHEC NleA sequences differ in number and location, indicating that glycosylation of NleA proteins

between strains may be initiated at different sites by different GALNTs. Thus, although both the NleA protein and its modification are conserved among A/E pathogens, these differences suggest that the mechanism by which NleA modification is initiated and the GALNTs involved may reflect pathogen niche. Additionally, the modified serine- and threonine-rich region of NleA shown to be variable in **Chapter 2** was identified from EHEC clinical isolates. We have not yet investigated if similar contraction of suspected modified regions is observed among EPEC and *C. rodentium* strains.

5.3 The modification of NleA occurs in the host cell secretory pathway

Mucin-type protein O-glycosylation is initiated in the Golgi by the transfer of GalNAc to Ser and Thr residues of polypeptide acceptors by resident GALNTs. GALNTs have a common structure with membrane-spanning and luminal catalytic domains [481]. NleA was previously shown to localize to the Golgi following translocation [342], and, in **Chapter 2** we used inhibitors of the host cell secretory pathway to identify this as the site of modification of NleA. Treatment with brefeldin A and monensin limited, but did not fully abrogate, NleA glycosylation. This suggests that NleA must either span, cross, or disrupt the secretory pathway membranes. Therefore, we hypothesize that NleA is glycosylated in the Golgi, but the mechanism by which NleA accesses this compartment is unknown. NleA does not possess any known Golgi-targeting motifs but does contain a C-terminal class I PSD-95/Dlg1/ZO-1 (PDZ) binding motif conserved across A/E pathogens [426]. Deletion of the PDZ binding motif resulted in attenuated virulence and delayed trafficking of NleA to the Golgi [426]. While the PDZ binding motif may be involved in initial targeting of NleA to the secretory pathway, it does not seem to be required.

In canonical O-GalNAc glycosylation, the subcellular location of GALNTs is confined to the Golgi [455]. However, redistribution of GALNTs via COPI-dependent retrograde trafficking from the Golgi to the ER has been identified in response to certain stimuli and in cancer cells [482,483]. This mechanism results in increased levels of GALNT-mediated O-glycosylation and Tn antigen expression and has been associated with more aggressive cancer cell features such as increased migration and invasion [483,484]. In addition, this redistribution was determined to be specific to GALNT family enzymes, whereas multiple other glycosylation enzymes were not displaced from the Golgi [482]. Considering many of the immortalized cell lines utilized in the *in vitro* work of this thesis originate from naturally occurring cancers, it would be important to investigate whether GALNTs are subject to redistribution in these cells to determine if the pathway of O-glycosylation, specifically O-GalNAc initiation, is altered. Subsequent studies should therefore determine if O-glycosylation of NleA is initiated in the Golgi, in the typical pathway, or if NleA is partially or fully modified in the ER without requiring entry into the Golgi.

We demonstrated in **Chapter 3** that interaction of NleA with the Sec24 subunit of the COPII vesicle coat is necessary for NleA O-glycosylation. This data suggests that the NleA-Sec24 interaction precedes modification of NleA. The core COPII components include Sar1, Sec23, Sec24, Sec13, and Sec31 [485]. Sar1 is a small GTPase that recruits other coat proteins to the ER membrane [485]. Sec23 and Sec24 heterodimers form an inner coat complex while Sec13 and Sec31 heterotetramers form the outer coat complex [485]. Sec24 is the primary component responsible for cargo selection [486]. We have previously shown that NleA functions to inhibit protein trafficking primarily via its interaction with Sec24 [299]. One compelling question is whether NleA binds Sec24 for the purpose of getting modified and happens to prevent anterograde trafficking in the process or, does NleA bind Sec24 to inhibit protein trafficking and end up being

carried into the Golgi where it is modified. Our data suggests that glycosylation of NleA is not inherently beneficial to bacterial virulence or persistence, possibly supporting the idea of NleA modification as a host response or inadvertent occurrence. The interaction of NleA with other COPII proteins was not investigated in this work, as we have previously identified Sec23 and Sec24 as the only COP proteins that co-immunoprecipitated with NleA [299]. In addition, when co-affinity purification and liquid chromatography-tandem mass spectrometry was used to identify 19 intermediate or high confidence NleA-host protein interactions, Sec23 and Sec24 were the only COPII proteins detected [446]. For these reasons, we elected to focus on these interactions as a potential mechanism for NleA's entry into the host secretory system. Further investigation of NleA-host protein interactions is an interesting topic of future investigation.

The Golgi functions as the main sorting hub of the cell secretory pathway; proteins received from the ER are processed, sorted, and packaged for transport to their ultimate destinations: lysosomes, the plasma membrane, or secretion. Although we have localized the site of NleA modification to this pathway, we have not investigated what consequence O-glycosylation has on localization of NleA upon exit from the secretory pathway. Subcellular localization undoubtedly influences bacterial effector function by limiting access to molecular interaction partners or substrates. Therefore, determining the localization of NleA following translocation will be significant to characterizing the cellular function of the effector. Preliminary immunofluorescent visualization of WT NleA showed localization to the Golgi apparatus three hours post-infection, comparable to previous studies showing colocalization of NleA and Golgi marker signal [342]. As demonstrated in **Chapter 4**, unglycosylated NleA may be more stable in the host cell. However, based on the analysis performed, we cannot say if this is due to reduced secretion or degradation when NleA is not glycosylated. Therefore, we anticipate that *in vitro* infection of epithelial cells

followed by a time course of immunofluorescence imaging every 1-2 hours over a 24 hour time course will provide valuable insight into the localization of NleA, determine whether the effector's localization to specific compartments is necessary for the perturbations identified above, and investigate whether glycosylation of NleA affects its localization.

5.4 Absence of NleA modification does not significantly impact virulence but may increase protein stability

Bacterial effectors contribute to the success of bacterial infections by interfering with host pathways. Mechanisms of bacterial effector function include direct binding to host targets, functional mimicry, structural mimicry, and modification of host proteins [487]. Post-translational modification of effectors can further diversify and enable regulation of their function. In this thesis, we have presented compelling evidence for the modification of NleA by host factors, but two primary questions remain: what the cell biological consequences of NleA glycosylation are and how modification of NleA impacts its function. In **Chapter 2**, we determined that glycosylation of NleA was not required for virulence in an acute mortality mouse model. Moreover, bacteria expressing non-glycosylatable NleA outcompeted bacteria expressing WT NleA in competitive index experiments, resulting in a subtle persistence effect. As shown in **Chapter 4**, non-glycosylatable NleA was observed to be potentially more stable in the host cell. Together, these results indicate that the consequence of NleA modification may not be favourable to the pathogen, as initially expected, and instead may represent a host response to the injected protein.

We expect that if NleA glycosylation was a marker for secretion, we would find modified NleA protein in the culture supernatant, which we do not. However, if modified NleA was embedded in the cell membrane to serve as some kind of cell signaling molecule, our method of

probing intracellular FLAG signal, described in **Chapter 4**, would not be able to distinguish membrane-associated from cytoplasmic NleA. Structural modeling of NleA, shown in **Chapter 2**, predicted a central β-barrel structure with the N- and C-termini as unstructured loops. β-barrel proteins are ubiquitously found in the outer membranes of gram-negative bacteria where they serve essential functions in cargo transport, signaling, and act as porins, transporters, enzymes, and virulence factors [488]. In eukaryotes, β-barrels form one of the two main structural classes of integral membrane proteins [489]. Additionally, NleA was previously determined to be membrane-associated by subcellular fractionation and Western blotting [342]. NleA membrane-association resisted extraction with high salt and low pH, typically used for removal of proteins that are peripherally associated with membranes [342]. These findings support the notion of NleA integration into the host cell membrane following modification, however the function it performs there has not been determined.

Mucin-type O-glycans make up the majority of glycans in the human gut [490]. The stomach and duodenum epithelium express core 1 and core 2-based structures, while the rest of the small intestine contains core 3-based structures [491]. Glycans expressed by the colon epithelium are core 3 and core 4-based structures [491]. These core structures are further modified by Gal, GlcNAc, Fuc, sialic acid, and sulfate. Interestingly, gut epithelial glycan composition follows an increasing gradient of sialic acid from the ileum to the rectum that is associated with a reverse gradient of fucose [492]. In comparison to humans, the mouse intestinal glycome consists primarily of core 2-based glycans as well as some regional variation in the distribution of terminal epitopes [493,494]. In the gut, intestinal epithelial glycosylation has been implicated in inflammatory bowel disease (IBD) pathogenesis [490]. IBD has been associated with increased expression of truncated O-glycans and altered expression of terminal glycan structures [490].

Gut epithelial glycans are also involved in antigen uptake, immune tolerance, and inflammatory cell recruitment [490]. Intestinal epithelial O-glycans can directly regulate host-microbe interactions by serving as adhesins or providing nutrients for bacterial metabolism. If NleA were functioning in this way, we would have expected to see an increased benefit to bacterial survival in our mixed infection experiments, when we observed the opposite. Additionally, gut glycoproteins regulate leukocyte recruitment via receptor-ligand interactions [490]. For example, transepithelial migration of neutrophils from the blood to the intestinal lumen requires interaction with glycosylated receptors, including intercellular adhesion molecule 1 (ICAM-1) [490]. In IBD, increased neutrophil accumulation in the colon crypt bases form abscesses that are a hallmark of disease [495]. These results indicate a clear role of intestinal glycans in both the host immune response and inflammation. Using the system described in **Chapter 4**, we predict that we will identify a role for NleA glycosylation in one of these responses. Although we observed no effect of NleA modification on mortality, we did not further investigate whether the host-mediated glycosylation impacts host immunity and damage. We did, however, perform immunofluorescent visualization of bacterial burden in colon sections from these mice, where we observed no differences in crypt infiltration between bacteria expressing WT and non-glycosylable NleA but did see reduced infiltration of Δ*nleA* bacteria (data not shown). Due to the functions of glycans in immune signaling and cell recruitment, histopathological studies may have identified differences in GI pathology between mice infected due to host-mediated glycosylation of NleA that were not evident in our analyses.

In the context of highly dynamic host-pathogen interactions, it is difficult for us to postulate the cause and effect of bacterial and host responses. For example, we may speculate that modification is a host response to NleA hijacking protein trafficking, and that the bacteria responds

by eliminating the site of modification. But we cannot determine that the differences in the serine- and threonine-rich region of the NleA sequence observed in clinical isolates are not due to spontaneous diversification or when the sequence contraction occurred. Although NleA was originally thought to be specific to A/E pathogens, we recently identified *nleA* gene homologs in some strains of the intestinal pathogens *Salmonella* and *Shigella*, as well as in an environmental Oxalobacteraceae strain [496]. We performed limited analysis on the prevalence of *nleA* homologs in only one of these species, in which we identified presumptive *nleA* homologs in 16 of 10997 publicly available *Salmonella* genomes, fewer than 0.2% (unpublished). In addition, sequence alignment of homologs from all species showed multiple homolog sequences with similar contraction of the same serine- and threonine-rich region observed in EHEC clinical isolates (unpublished). The presence of *nleA* homologs in other bacteria may not be surprising, considering it is encoded on a mobile element. However, the presence of an NleA homolog in non-pathogenic environmental Oxalobacteraceae strain is confounding, both in origin and function. Whether this is truly a *nleA* homolog or an anomaly would require further investigation. We are mostly intrigued to find NleA homologs in strains of *Salmonella* and *Shigella*. *Salmonella* and *Shigella* are intracellular enteric pathogens that both possess a T3SS, thus it does not seem impossible that another enteric pathogen could pass *nleA* to one of these strains. Subsequent analyses should determine if effectors typically encoded on the same mobile genetic element as *nleA* are also present. This will provide insight on whether this effector is disseminated exclusively or as part of a larger mobile element. It will be important to analyze whether these NleA-like proteins have a similar post-translational modification and localization as NleA. To address the functional role of these homologs, our flow cytometry detection method outlined above can be used to determine

whether these effectors modulate the same or similar pathways as NleA, or whether their function is unique to the specific bacteria.

5.5 Conclusion

NleA has been shown to be required for bacterial virulence in *C. rodentium* infections, but the complete scope of the contribution of NleA during infection has not been well characterized. The work presented in this thesis demonstrates that NleA is modified by O-linked glycosylation following its translocation into the host cell, the first bacterial effector characterized to be modified in this way. Based on modified site and protein structure predictions, we identified a particular serine- and threonine-rich region of the NleA protein sequence that, when deleted, abrogates glycosylation of the protein. We identified the host secretory pathway as the site of modification and determined that NleA interaction with the COPII vesicle coat is required for its modification to occur. Furthermore, we determined that while glycosylation of NleA is not crucial for bacterial virulence, it resulted in a subtle persistence effect and may reduce NleA stability in the host cell. Additional work is required to explore the implications of NleA glycosylation on host response to infection. The cellular processes targeted by bacterial effector proteins are diverse and the mechanisms by which bacteria manipulate them are often unique. Understanding how effector proteins interact with and manipulate these cellular pathways has given rise to a number of important insights into fundamental cell biological processes. This thesis adds novel findings to our growing understanding of host-pathogen interactions and the crucial but complex role effectors play in modulating the responses involved.

REFERENCES

1. Havelaar AH, Kirk MD, Torgerson PR, Gibb HJ, Hald T, Lake RJ, et al. *World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010*. PLOS Med 2015;**12**.
2. Moon HW, Whipp SC, Argenzio RA, Levine MM, Giannella RA. *Attaching and Effacing Activities of Rabbit and Human Enteropathogenic Escherichia coli in Pig and Rabbit Intestines*. Infect Immun 1983;**41**:1340–51.
3. Nataro JP, Kaper JB. *Diarrheagenic Escherichia coli*. Clin Microbiol Rev 1998;**11**:142–201.
4. Goosney DL, Gruenheid S, Finlay BB. *Gut Feelings: Enteropathogenic E. coli (EPEC) Interactions with the Host*. Annu Rev Cell Dev Biol 2000;**16**:173–89.
5. Bray J. *Isolation of antigenically homogeneous strains of Bact. coli neapolitanum from summer diarrhœa of infants*. J Pathol Bacteriol 1945;**57**:239–47.
6. Kaper JB. *Defining EPEC*. Rev Microbiol Sao Paulo 1996;**27**:130–3.
7. Baldini MM, Kaper JB, Levine MM, Candy DCA, Moon HW. *Plasmid-mediated adhesion in enteropathogenic Escherichia coli*. J Pediatr Gastroenterol Nutr 1983;**2**:534–8.
8. Nataro JP, Maher KO, Mackie P, Kaper JB. *Characterization of Plasmids Encoding the Adherence Factor of Enteropathogenic Escherichia coli*. Infect Immun 1987;**55**:2370–7.
9. Girón JA, Ho ASY, Schoolnik GK. *An Inducible Bundle-Forming Pilus of Enteropathogenic Escherichia coli*. Science (80-) 1991;**254**:710–3.
10. Nataro JP, Scaletsky IC, Kaper JB, Levine MM, Trabulsi LR. *Plasmid-Mediated Factors Conferring Diffuse and Localized Adherence of Enteropathogenic Escherichia coli*. Infect Immun 1985;**48**:378–83.
11. Ingle DJ, Tauschek M, Edwards DJ, Hocking DM, Pickard DJ, Azzopardi KI, et al. *Evolution of atypical enteropathogenic E. coli by repeated acquisition of LEE pathogenicity island variants*. Nat Microbiol 2016;**1**.
12. Hazen TH, Donnenberg MS, Panchalingam S, Antonio M, Hossain A, Mandomando I, et al. *Genomic diversity of EPEC associated with clinical presentations of differing severity*. Nat Microbiol 2016;**1**.
13. Levine MM, Edelman R. *Enteropathogenic Escherichia coli of classic serotypes associated with infant diarrhea: epidemiology and pathogenesis.* Epidemiol Rev 1984;**6**:31–51.
14. Hill SM, Phillips AD, Walker-Smith JA, Phillips MAD. *Enteropathogenic Escherichia coli and life threatening chronic diarrhoea*. Gut 1991;**32**:154–8.
15. Bower JR, Congeni BL, Cleary TG, Stone RT, Wanger A, Murray BE, et al. *Escherichia coli O114:nonmotile as a Pathogen in an Outbreak of Severe Diarrhea Associated with a Day Care Center*. J Infect Dis 1989;**160**:243–7.
16. Wu S-X, Peng R-Q. *Studies on an outbreak of neonatal diarrhea caused by EPEC O127:H6 with plasmid analysis restriction analysis and outer membrane protein determination*. Acta Paediatr 1992;**81**:217–21.
17. Sekse C, Sunde M, Lindstedt BA, Hopp P, Bruheim T, Cudjoe KS, et al. *Potentially Human-Pathogenic Escherichia coli O26 in Norwegian Sheep Flocks*. Appl Environ Microbiol 2011;**77**:4949–58.
18. Moura RA, Sircili MP, Leomil L, Matté MH, Trabulsi LR, Elias WP, et al. *Clonal*

Relationship among Atypical Enteropathogenic Escherichia coli Strains Isolated from Different Animal Species and Humans. Appl Environ Microbiol 2009;**75**:7399–408.
19. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. *Recent Advances in Understanding Enteric Pathogenic Escherichia coli*. Clin Microbiol Rev 2013;**26**:822–80.
20. Vallance BA, Finlay BB. *Exploitation of host cells by enteropathogenic Escherichia coli*. Proc Natl Acad Sci U S A 2000;**97**:8799–806.
21. Donnenberg MS, Kaper JB. *Enteropathogenic Escherichia coli*. Infect Immun 1992;**60**:3953–61.
22. Levine MM, Nalin DR, Hornick RB, Bergquist EJ, Waterman DH, Young CR, et al. *Escherichia coli strains that cause diarrhea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive*. Lancet 1978;**1**:1119–22.
23. Dean P, Kenny B. *The effector repertoire of enteropathogenic E. coli: ganging up on the host cell*. Curr Opin Microbiol 2009;**12**:101–9.
24. Guttman JA, Samji FN, Li Y, Deng W, Lin A, Finlay BB. *Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens*. Cell Microbiol 2007;**9**:131–41.
25. Dean P, Maresca M, Schüller S, Phillips AD, Kenny B. *Potent diarrheagenic mechanism mediated by the cooperative action of three enteropathogenic Escherichia coli-injected effector proteins*. Proc Natl Acad Sci U S A 2006;**103**:1876–81.
26. Gill RK, Borthakur A, Hodges K, Turner JR, Clayburgh DR, Saksena S, et al. *Mechanism underlying inhibition of intestinal apical Cl-/OH- exchange following infection with enteropathogenic E. coli*. J Clin Invest 2007;**117**:428–37.
27. Viswanathan VK, Koutsouris A, Lukic S, Pilkinton M, Simonovic I, Simonovic M, et al. *Comparative Analysis of EspF from Enteropathogenic and Enterohemorrhagic Escherichia coli in Alteration of Epithelial Barrier Function*. Infect Immun 2004;**72**:3218–27.
28. Matsuzawa T, Kuwae A, Abe A. *Enteropathogenic Escherichia coli Type III Effectors EspG and EspG2 Alter Epithelial Paracellular Permeability*. Infect Immun 2005;**73**:6283–9.
29. Dean P, Kenny B. *Intestinal barrier dysfunction by enteropathogenic Escherichia coli is mediated by two effector molecules and a bacterial surface protein*. Mol Microbiol 2004;**54**:665–75.
30. Ochoa TJ, Contreras CA. *Enteropathogenic E. coli (EPEC) infection in children*. Curr Opin Infect Dis 2011;**24**:478–83.
31. Pawlowski SW, Alcantara C, Guerrant WR. *Diagnosis and Treatment of Acute or Persistent Diarrhea*. Gastroenterology 2009;**136**:1874–86.
32. Dupont HL. *Persistent Diarrhea A Clinical Review*. JAMA 2016;**315**:2712–23.
33. Kaper JB, Nataro JP, Mobley HLT. *Pathogenic Escherichia coli*. Nat Rev Microbiol 2004;**2**:123–40.
34. Trofa AF, Ueno-Olsen H, Oiwa R, Yoshikawa M. *Dr. Kiyoshi Shiga: Discoverer of the Dysentery Bacillus*. Clin Infect Dis 1999;**29**:1303–6.
35. Croxen MA, Finlay BB. *Molecular mechanisms of Escherichia coli pathogenicity*. Nat Rev Microbiol 2010;**8**:26–38.
36. *National Enteric Disease Surveillance: Shiga toxin-producing E. coli (STEC) Annual Report, 2017*. Centers Dis Control Prev 2021.

37. Elder JR, Bugarel M, den Bakker HC, Loneragan GH, Nightingale KK. *Interrogation of single nucleotide polymorphisms in gnd provides a novel method for molecular serogrouping of clinically important Shiga toxin producing Escherichia coli (STEC) targeted by regulation in the United States, including the "big six" non-O157* . J Microbiol Methods 2016;**129**:85–93.
38. Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, et al. *Hemorrhagic colitis associated with a rare Escherichia coli serotype*. N Engl J Med 1983;**308**:681–5.
39. Smith DGE, Hoey DEE, Sharp L, Currie C, Lingwood CA, Gally DL, et al. *Verotoxin 1 binding to intestinal crypt epithelial cells results in localization to lysosomes and abrogation of toxicity*. Cell Microbiol 2003;**5**:85–97.
40. Pruimboom-Brees IM, Morgan TW, Ackermann MR, Nystrom ED, Samuel JE, Cornick NA, et al. *Cattle lack vascular receptors for Escherichia coli O157:H7 Shiga toxins*. Proc Natl Acad Sci U S A 2000;**97**:10325–9.
41. Wells JG, Shipman LD, Greene KD, Sowers EG, Green JH, Cameron DN, et al. *Isolation of Escherichia coli Serotype O157:H7 and Other Shiga-Like-Toxin-Producing E. coli from Dairy Cattle*. J Clin Microbiol 1991;**29**:985–9.
42. Bielaszewska M, Schmidt H, Liesegang A, Prager R, Rabsch W, Tschape H, et al. *Cattle Can Be a Reservoir of Sorbitol-Fermenting Shiga Toxin-Producing Escherichia coli O157:H- Strains and a Source of Human Diseases*. J Clin Microbiol 2000;**38**:3470–3.
43. Hancock DD, Besser TE, Kinsel ML, Tarr PI, Rice DH, Paros MG. *The prevalence of Escherichia coli O157:H7 in dairy and beef cattle in Washington State*. Epidemiol Infect 1994;**113**:199–207.
44. Naylor SW, Gally DL, Christopher Low J. *Enterohaemorrhagic E. coli in veterinary medicine*. Int J Med Microbiol 2005;**295**:419–41.
45. Tuttle J, Gomez T, Doyle MP, Wells JG, Zhao T, Tauxe RM, et al. *Lessons from a large outbreak of Escherichia coli O157 : H7 infections : insights into the infectious dose and method of widespread contamination of hamburger patties*. Epidemiol Infect Infect 1999;**122**:185–92.
46. Nielsen EM, Scheutz F, Torpdahl M. *Continuous Surveillance of Shiga Toxin–Producing Escherichia coli Infections by Pulsed-Field Gel Electrophoresis Shows That Most Infections Are Sporadic*. Foodborne Pathog Dis 2006;**3**:81–7.
47. Elliott EJ, Robins-Browne RM, O'Loughlin E V., Bennett-Wood V, Bourke J, Henning P, et al. *Nationwide study of haemolytic uraemic syndrome: Clinical, microbiological, and epidemiological features*. Arch Dis Child 2001;**85**:125–31.
48. Grisaru S, Midgley JP, Bch BM, Hamiwka LA, Wade AW, Samuel SM. *Diarrhea-associated hemolytic uremic syndrome in southern Alberta: A long-term single-centre experience*. Paediatr Child Health 2011;**16**:337–40.
49. Thomas DE, Elliott EJ. *Interventions for preventing diarrhea-associated hemolytic uremic syndrome: systematic review*. BMC Public Health 2013;**13**.
50. Tarr PI, Gordon CA, Chandler WL. *Shiga-toxin-producing Escherichia coli and haemolytic uraemic syndrome*. Lancet 2005;**365**:1073–86.
51. Griffin PM, Tauxe R V. *The Epidemiology of Infections Caused by Escherichia coli O157:H7, Other Enterohemorrhagic E. coli, and the Associated Hemolytic Uremic Syndrome*. Epidemiol Rev 1991;**13**:60–98.
52. Dundas S, Todd WTA, Stewart AI, Murdoch PS, Chaudhuri AKR, Hutchinson SJ. *The*

Central Scotland Escherichia coli O157:H7 Outbreak: Risk Factors for the Hemolytic Uremic Syndrome and Death Among Hospitalized Patients. Clin Infect Dis 2001;**33**:923–31.
53. Carter AO, Borczyk AA, Carlson JAK, Harvey B, Hockin JC, Karmali MA, et al. *A severe outbreak of Escherichia coli O157:H7–associated hemorrhagic colitis in a nursing home*. N Engl J Med 1987;**317**:1496–500.
54. Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL. *Associations between Virulence Factors of Shiga Toxin-Producing Escherichia coli and Disease in Humans*. J Clin Microbiol 1999;**37**:497–503.
55. Byrne L, Vanstone GL, Perry NT, Launders N, Adak GK, Godbole G, et al. *Epidemiology and microbiology of Shiga toxin-producing Escherichia coli other than serogroup O157 in England, 2009-2013*. J Med Microbiol 2014;**63**:1181–8.
56. Gould LH, Mody RK, Ong KL, Clogher P, Cronquist AB, Garman KN, et al. *Increased Recognition of Non-O157 Shiga Toxin-Producing Escherichia coli Infections in the United States During 2000-2010: Epidemiologic Features and Comparison with E. coli O157 Infections*. Foodborne Pathog Dis 2013;**10**:453–60.
57. Hedican EB, Medus C, Besser JM, Juni BA, Koziol B, Taylor C, et al. *Characteristics of O157 versus Non-O157 Shiga Toxin-Producing Escherichia coli Infections in Minnesota, 2000-2006*. Clin Infect Dis 2009;**49**:358–64.
58. Hadler JL, Clogher P, Hurd S, Phan Q, Mandour M, Bemis K, et al. *Ten-Year Trends and Risk Factors for Non-O157 Shiga Toxin-Producing Escherichia coli Found Through Shiga Toxin Testing, Connecticut, 2000-2009*. Clin Infect Dis 2011;**53**:269–76.
59. Matussek A, Einemo IM, Jogenfors A, Löfdah S, Löfgren S. *Shiga Toxin-Producing Escherichia coli in Diarrheal Stool of Swedish Children: Evaluation of Polymerase Chain Reaction Screening and Duration of Shiga Toxin Shedding*. J Pediatric Infect Dis Soc 2016;**5**:147–51.
60. Ostroff SM, Tarr PI, Neill MA, Lewis JH, The S, Diseases I, et al. *Toxin Genotypes and Plasmid Profiles as Determinants of Systemic Sequelae in Escherichia coli O157:H7 Infections*. J Infect Dis 1989;**160**:994–8.
61. Strockbine NA, Marques LRM, Newland JW, Smith HW, Holmes RK, O'brien AD. *Two Toxin-Converting Phages from Escherichia coli O157:H7 Strain 933 Encode Antigenically Distinct Toxins with Similar Biologic Activities*. Infect Immun 1986;**53**:135–40.
62. O'Brien A, Tesh V, Donohue-Rolfe A, Jackson M, Olsnes S, Sandvig K, et al. *Shiga Toxin: Biochemistry, Genetics, Mode of Action, and Role in Pathogenesis*. Curr Top Microbiol Immunol 1992;**180**:65–94.
63. Unkmeir A, Schmidt H. *Structural Analysis of Phage-Borne stx Genes and Their Flanking Sequences in Shiga Toxin-Producing Escherichia coli and Shigella dysenteriae Type 1 Strains*. Infect Immun 2000;**68**:4856–64.
64. Kimmitt PT, Harwood CR, Barer MR. *Toxin Gene Expression by Shiga Toxin-Producing Escherichia coli: The Role of Antibiotics and the Bacterial SOS Response*. Emerg Infect Dis 2000;**6**:458–65.
65. Wagner PL, Livny J, Neely MN, Acheson DWK, Friedman DI, Waldor MK. *Bacteriophage control of Shiga toxin 1 production and release by Escherichia coli*. Mol Microbiol 2002;**44**:957–70.
66. Fraser ME, Chernaia MM, Kozlov Y V, James MNG. *Crystal structure of the holotoxin*

from S. dysenteriae at 2.5A resolution. Struct Biol 1994;**1**:59–64.
67. Stein PE, Boodhoo A, Tyrrell GJ, Brunton JL, Read RJ. *Crystal structure of the cell-binding B oligomer of verotoxin-1 from E. coli*. Nature 1992;**355**:748–50.
68. Jacewicz M, Clausen H, Nudelman E, Donohue-rolfe A, Keusch GT. *Isolation of a Shigella Toxin-Binding Glycolipid from Rabbit Jejumum and HeLa Cells and Its Identification as Globotriaosylceramide*. J Exp Med 1986;**163**:1391–404.
69. Lindberg AA, Brown JE, Strömberg N, Westling-Ryd M, Schultz JE, Karlsson KA. *Identification of the Carbohydrate Eeceptor for Shiga Toxin Produced by Shigella dysenteriae Type 1*. J Biol Chem 1987;**262**:1779–85.
70. Waddell T, Cohen A, Lingwood CA. *Induction of verotoxin sensitivity in receptor-deficient cell lines using the receptor glycolipid globotriosylceramide*. Proc Natl Acad Sci U S A 1990;**87**:7898–901.
71. Karmali MA. *Infection by shiga toxin-producing Escherichia coli: An overview*. Appl Biochem Biotechnol - Part B Mol Biotechnol 2004;**26**:117–22.
72. Te Loo DMWM, Van Hinsbergh VWM, Van Den Heuvel LPWJ, Monnens LAH. *Detection of Verocytotoxin Bound to Circulating Polymorphonuclear Leukocytes of Patients with Hemolytic Uremic Syndrome*. J Am Soc Nephrol 2001;**12**:800–6.
73. Boyd B, Lingwood C. *Verotoxin receptor glycolipid in human renal tissue*. Nephron 1989;**51**:207–10.
74. Louise CB, Obrig TG. *Specific Interaction of Escherichia coli O157 : H7-Derived Shiga-like Toxin II with Human Renal Endothelial Cells*. J Infect Dis 1995;**172**:1397–401.
75. Gerber A, Karch H, Allerberger F, Verweyen HM, Zimmerhackl LB. *Clinical course and the role of Shiga toxin-producing Escherichia coli infection in the hemolytic-uremic syndrome in pediatric patients, 1997-2000, in Germany and Austria: A prospective study*. J Infect Dis 2002;**186**:493–500.
76. Braune SA, Wichmann D, Von Heinz MC, Nierhaus A, Becker H, Meyer TN, et al. *Clinical features of critically ill patients with shiga toxin-induced hemolytic uremic syndrome*. Crit Care Med 2013;**41**:1702–10.
77. Rigamonti D, Simonetti GD. *Direct cardiac involvement in childhood hemolytic-uremic syndrome: case report and review of the literature*. Eur J Pediatr 2016;**175**:1927–31.
78. Delmas Y, Vendrely B, Clouzeau B, Bachir H, Bui HN, Lacraz A, et al. *Outbreak of Escherichia coli O104:H4 haemolytic uraemic syndrome in France: Outcome with eculizumab*. Nephrol Dial Transplant 2014;**29**:565–72.
79. Johannes L, Römer W. *Shiga toxins - from cell biology to biomedical applications*. Nat Rev Microbiol 2010;**8**:105–16.
80. Panos GZ, Betsi GI, Falagas ME. *Systematic review: are antibiotics detrimental or beneficial for the treatment of patients with Escherichia coli O157:H7 infection?*. Aliment Pharmacol Ther 2006;**24**:731–42.
81. Agger M, Scheutz F, Villumsen S, Mølbak K, Petersen AM. *Antibiotic treatment of verocytotoxin-producing Escherichia coli (VTEC) infection: A systematic review and a proposal*. J Antimicrob Chemother 2015;**70**:2440–6.
82. Freedman SB, Xie J, Neufeld MS, Hamilton WL, Hartling L, Tarr PI. *Shiga Toxin-Producing Escherichia coli Infection, Antibiotics, and Risk of Developing Hemolytic Uremic Syndrome: A Meta-analysis*. Clin Infect Dis 2016;**62**:1251–8.
83. Safdar N, Said A, Gangnon RE. *Risk of Hemolytic Uremic Syndrome After Antibiotic Treatment of Escherichia coli O157:H7 Enteritis*. JAMA 2009;**288**:996–1001.

84. Albert MJ, Alam K, Islam M, Montanaro J, Rahman ASMH, Haider K, et al. *Hafnia alvei, a Probable Cause of Diarrhea in Humans*. Infect Immun 1991;**59**:1507–13.
85. Muchaamba F, Barmettler K, Treier A, Houf K, Stephan R. *Microbiology and Epidemiology of Escherichia albertii—An Emerging Elusive Foodborne Pathogen*. Microorganisms 2022;**10**:1–21.
86. Asoshima N, Matsuda M, Shigemura K, Honda M, Yoshida H, Hiwaki H, et al. *Identification of Escherichia albertii as a Causative Agent of a Food-Borne Outbreak Occurred in 2003*. Jpn J Infect Dis 2014;**67**:139–40.
87. Ooka T, Ogura Y, Katsura K, Seto K, Kobayashi H, Kawano K, et al. *Defining the Genome Features of Escherichia albertii, an Emerging Enteropathogen Closely Related to Escherichia coli*. Genome Biol Evol 2015;**7**:3170–9.
88. Ooka T, Tokuoka E, Furukawa M, Nagamura T, Ogura Y, Arisawa K, et al. *Human Gastroenteritis Outbreak Associated with Escherichia albertii, Japan*. Emerg Infect Dis 2013;**19**:144–6.
89. Konno T, Yatsuyanagi J, Takahashi S, Kumagai Y, Wada E, Chiba M, et al. *Isolation and Identification of Escherichia albertii from a Patient in an Outbreak of Gastroenteritis*. Jpn J Infect Dis 2012;**65**:203–7.
90. Baba A, Ebuchi S, Uryu K, Hiwaki H, Ogata K, Washimi E, et al. *An Outbreak of Water-Borne Gastroenteritis Caused by Diarrheagenic Escherichia coli Possessing eae Gene*. Jpn J Infect Dis 2006;**59**:59–60.
91. Inglis TJJ, Merritt AJ, Bzdyl N, Lansley S, Urosevic MN. *First bacteraemic human infection with Escherichia albertii*. New Microbes New Infect 2015;**8**:171–3.
92. Brandal LT, Tunsjø HS, Ranheim TE, Løbersli I, Lange H, Wester AL. *Shiga Toxin 2a in Escherichia albertii*. J Clin Microbiol 2015;**53**:1454–5.
93. Oaks JL, Besser TE, Walk ST, Gordon DM, Beckmen KB, Burek KA, et al. *Escherichia albertii in Wild and Domestic Birds*. Emerg Infect Dis 2010;**16**:638–46.
94. Maeda E, Murakami K, Sera N, Ito K, Fujimoto S. *Detection of Escherichia albertii from chicken meat and giblets*. J Vet Med Sci 2015;**77**:871–3.
95. Asoshima N, Matsuda M, Shigemura K, Honda M, Yoshida H, Oda T, et al. *Isolation of Escherichia albertii from Raw Chicken Liver in Fukuoka City, Japan*. Jpn J Infect Dis 2015;**68**:248–50.
96. Wang H, Li Q, Bai X, Xu Y, Zhao A, Sun H, et al. *Prevalence of eae-positive, lactose non-fermenting Escherichia albertii from retail raw meat in China*. Epidemiol Infect 2016;**144**:45–52.
97. Maheux AF, Boudreau DK, Bergeron MG, Rodriguez MJ. *Characterization of Escherichia fergusonii and Escherichia albertii isolated from water*. J Appl Microbiol 2014;**117**:597–609.
98. Ooka T, Seto K, Kawano K, Kobayashi H, Etoh Y, Ichihara S, et al. *Clinical Significance of Escherichia albertii*. Emerg Infect Dis 2012;**18**:488–92.
99. Barthold SW, Coleman GL, Jacoby RO, Livstone EM, Jonas AM. *Transmissible Murine Colonic Hyperplasia*. Vet Pathol 1978;**15**:223–36.
100. Itoh K, Matsui T, Tsuji K, Mitsuoka T, Ueda K. *Genetic control in the susceptibility of germfree inbred mice to infection by Escherichia coli O115a,c:K(B)*. Infect Immun 1988;**56**:930–5.
101. Petty NK, Feltwell T, Pickard D, Clare S, Toribio AL, Fookes M, et al. *Citrobacter rodentium is an unstable pathogen showing evidence of significant genomic flux*. PLoS

Pathog 2011;**7**.
102. Petty NK, Bulgin R, Crepin VF, Cerdeño-Tárraga AM, Schroeder GN, Quail MA, et al. *The Citrobacter rodentium genome sequence reveals convergent evolution with human pathogenic Escherichia coli*. J Bacteriol 2010;**192**:525–38.
103. Schauer DB, Falkow S. *The eae Gene of Citrobacter freundii Biotype 4280 is Necessary for Colonization in Transmissible Murine Colonic Hyperplasia*. Infect Immun 1993;**61**:4654–61.
104. Schauer DB, Falkow S. *Attaching and effacing locus of a Citrobacter freundii biotype that causes transmissible murine colonic hyperplasia*. Infect Immun 1993;**61**:2486–92.
105. Wiles S, Clare S, Harker J, Huett A, Young D, Dougan G, et al. *Organ specificity, colonization and clearance dynamics in vivo following oral challenges with the murine pathogen Citrobacter rodentium*. Cell Microbiol 2004;**6**:963–72.
106. Mundy R, Macdonald TT, Dougan G, Frankel G, Wiles S. *Citrobacter rodentium of mice and man*. Cell Microbiol 2005;**7**:1697–706.
107. Wiles S, Dougan G, Frankel G. *Emergence of a "hyperinfectious" bacterial state after passage of Citrobacter rodentium through the host gastrointestinal tract*. Cell Microbiol 2005;**7**:1163–72.
108. Ritchie JM. *Animal Models of Enterohemorrhagic Escherichia coli Infection* . Microbiol Spectr 2014;**2**.
109. Crepin VF, Collins JW, Habibzay M, Frankel G. *Citrobacter rodentium mouse model of bacterial infection*. Nat Protoc 2016;**11**:1851–76.
110. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, et al. *Host-Mediated Inflammation Disrupts the Intestinal Microbiota and Promotes the Overgrowth of Enterobacteriaceae (DOI:10.1016/j.chom.2007.06.010)*. Cell Host Microbe 2007;**2**:204.
111. Hoffmann C, Hill DA, Minkah N, Kirn T, Troy A, Artis D, et al. *Community-wide response of the gut microbiota to enteropathogenic Citrobacter rodentium infection revealed by deep sequencing*. Infect Immun 2009;**77**:4668–78.
112. Clare S, John V, Walker AW, Hill JL, Abreu-Goodger C, Hale C, et al. *Enhanced susceptibility to citrobacter rodentium infection in microRNA-155-Deficient mice*. Infect Immun 2013;**81**:723–32.
113. Diez E, Zhu L, Teatero SA, Paquet M, Roy MF, Loredo-Osti JC, et al. *Identification and characterization of Cri1, a locus controlling mortality during Citrobacter rodentium infection in mice*. Genes Immun 2011;**12**:280–90.
114. Teatero SA, Thomassin JL, Zhu L, Diez E, Malo D, Gruenheid S. *The Cri1 locus is the common genetic cause of susceptibility to Citrobacter rodentium infection in C3H and FVB mouse strains*. Gut Microbes 2011;**2**:173–7.
115. Papapietro O, Teatero S, Thanabalasuriar A, Yuki KE, Diez E, Zhu L, et al. *R-Spondin 2 signalling mediates susceptibility to fatal infectious diarrhoea*. Nat Commun 2013;**4**.
116. Jin Y-R, Yoon JK. *The R-spondin family of proteins: emerging regulators of WNT signaling*. Int J Biochem Cell Biol 2012;**44**:2278–87.
117. Kang E, Yousefi M, Gruenheid S. *R-spondins are expressed by the intestinal stroma and are differentially regulated during Citrobacter rodentium- and dssinduced colitis in mice*. PLoS One 2016;**11**:1–13.
118. Kang E, Zhou G, Yousefi M, Cayrol R, Xia J, Gruenheid S. *Loss of disease tolerance during Citrobacter rodentium infection is associated with impaired epithelial*

differentiation and hyperactivation of T cell responses. Sci Rep 2018;**8**:1–14.
119. Borenshtein D, Fry RC, Groff EB, Nambiar PR, Carey VJ, Fox JG, et al. *Diarrhea as a cause of mortality in a mouse model of infectious colitis.* Genome Biol 2008;**9**.
120. Ghosh S, Dai C, Brown K, Rajendiran E, Makarenko S, Baker J, et al. *Colonic microbiota alters host susceptibility to infectious colitis by modulating inflammation, redox status, and ion transporter gene expression*. Am J Physiol - Gastrointest Liver Physiol 2011;**301**:39–49.
121. Zheng D, Liwinski T, Elinav E. *Interaction between microbiota and immunity in health and disease*. Cell Res 2020;**30**:492–506.
122. Freter R. *The Fatal Enteric Cholera Infection in the Guinea Pig, Achieved by Inhibition of Normal Enteric Flora*. J Infect Dis 1955;**97**:57–65.
123. Bohnhoff M, Drake BL, Miller CP. *Effect of streptomycin on susceptibility of intestinal tract to experimental Salmonella infection*. Proc Soc Exp Biol Med 1954;**86**:132–7.
124. Kamada N, Kim Y-G, Sham HP, Vallance BA, Puente JL, Martens EC, et al. *Regulated Virulence Controls the Ability of a Pathogen to Compete with the Gut Microbiota*. Science (80-) 2012;**336**:1325–9.
125. Wlodarska M, Willing B, Keeney KM, Menendez A, Bergstrom KS, Gill N, et al. *Antibiotic treatment alters the colonic mucus layer and predisposes the host to exacerbated Citrobacter rodentium-induced colitis*. Infect Immun 2011;**79**:1536–45.
126. Rodrigues DM, Sousa AJ, Johnson-Henry KC, Sherman PM, Gareau MG. *Probiotics are effective for the prevention and treatment of citrobacter rodentium-induced colitis in mice*. J Infect Dis 2012;**206**:99–109.
127. Willing BP, Vacharaksa A, Croxen M, Thanachayanont T, Finlay BB. *Altering host resistance to infections through microbial transplantation*. PLoS One 2011;**6**:2–10.
128. Bergstrom K, Xia L. *The barrier and beyond: Roles of intestinal mucus and mucin-type O-glycosylation in resistance and tolerance defense strategies guiding host-microbe symbiosis*. Gut Microbes 2022;**14**.
129. Atuma C, Strugala V, Allen A, Holm L. *The adherent gastrointestinal mucus gel layer: Thickness and physical state in vivo*. Am J Physiol - Gastrointest Liver Physiol 2001;**280**:922–9.
130. Ermund A, Schütte A, Johansson MEV, Gustafsson JK, Hansson GC. *Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches*. Am J Physiol - Gastrointest Liver Physiol 2013;**305**:341–7.
131. Hansson GC. *Role of mucus layers in gut infection and inflammation*. Curr Opin Microbiol 2012;**15**:57–62.
132. Bergstrom K, Shan X, Casero D, Batushansky A, Lagishetty V, Jacobs JP, et al. *Proximal colon-derived O-glycosylated mucus encapsulates and modulates the microbiota*. Science (80-) 2020;**370**:467–72.
133. Johansson MEV, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. *The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria*. Proc Natl Acad Sci U S A 2008;**105**:15064–9.
134. Johansson MEV, Holmén Larsson JM, Hansson GC. *The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions*. Proc Natl Acad Sci U S A 2011;**108**:4659–65.
135. Li H, Limenitakis JP, Fuhrer T, Geuking MB, Lawson MA, Wyss M, et al. *The outer*

mucus layer hosts a distinct intestinal microbial niche. Nat Commun 2015;**6**.
136. Van Klinken BJW, Einerhand AWC, Duits LA, Makkink MK, Tytgat KMAJ, Renes IB, et al. *Gastrointestinal expression and partial cDNA cloning of murine Muc2*. Am J Physiol - Gastrointest Liver Physiol 1999;**276**:115–24.
137. Allen A, Hutton DA, Pearson JP. *The MUC2 gene product: A human intestinal mucin*. Int J Biochem Cell Biol 1998;**30**:797–801.
138. Gustafsson JK, Johansson MEV. *The role of goblet cells and mucus in intestinal homeostasis*. Nat Rev Gastroenterol Hepatol 2022;**19**:785–803.
139. Marcobal A, Southwick AM, Earle KA, Sonnenburg JL. *A refined palate: Bacterial consumption of host glycans in the gut*. Glycobiology 2013;**23**:1038–46.
140. Bell A, Juge N. *Mucosal glycan degradation of the host by the gut microbiota*. Glycobiology 2021;**31**:691–6.
141. Bergstrom KSB, Kissoon-Singh V, Gibson DL, Ma C, Montero M, Sham HP, et al. *Muc2 protects against lethal infectious colitis by disassociating pathogenic and commensal bacteria from the colonic mucosa*. PLoS Pathog 2010;**6**.
142. Luo Q, Kumar P, Vickers TJ, Sheikh A, Lewis WG, Rasko DA, et al. *Enterotoxigenic escherichia coli secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells*. Infect Immun 2014;**82**:509–21.
143. Nesta B, Valeri M, Spagnuolo A, Rosini R, Mora M, Donato P, et al. *SslE Elicits Functional Antibodies That Impair In Vitro Mucinase Activity and In Vivo Colonization by Both Intestinal and Extraintestinal Escherichia coli Strains*. PLoS Pathog 2014;**10**.
144. Hews CL, Tran SL, Wegmann U, Brett B, Walsham ADS, Kavanaugh D, et al. *The StcE metalloprotease of enterohaemorrhagic Escherichia coli reduces the inner mucus layer and promotes adherence to human colonic epithelium ex vivo*. Cell Microbiol 2017;**19**:1–10.
145. In J, Foulke-Abel J, Zachos NC, Hansen AM, Kaper JB, Bernstein HD, et al. *Enterohemorrhagic Escherichia coli Reduces Mucus and Intermicrovillar Bridges in Human Stem Cell-Derived Colonoids*. Cmgh 2016;**2**:48-62.e3.
146. Bhullar K, Zarepour M, Yu H, Yang H, Croxen M, Stahl M, et al. *The serine protease autotransporter Pic modulates Citrobacter rodentium pathogenesis and its innate recognition by the host*. Infect Immun 2015;**83**:2636–50.
147. Erdem AL, Avelino F, Xicohtencatl-Cortes J, Girón JA. *Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing Escherichia coli*. J Bacteriol 2007;**189**:7426–35.
148. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, Fleckenstein JM. *Eata, an immunogenic protective antigen of enterotoxigenic escherichia coli, degrades intestinal mucin*. Infect Immun 2014;**82**:500–8.
149. Grys TE, Walters LL, Welch RA. *Characterization of the StcE protease activity of Escherichia coli O157:H7*. J Bacteriol 2006;**188**:4646–53.
150. Loonen LMP, Stolte EH, Jaklofsky MTJ, Meijerink M, Dekker J, Van Baarlen P, et al. *REG3γ-deficient mice have altered mucus distribution and increased mucosal inflammatory responses to the microbiota and enteric pathogens in the ileum*. Mucosal Immunol 2014;**7**:939–47.
151. Johansson MEV, Hansson GC. *Immunological aspects of intestinal mucus and mucins*. Nat Rev Immunol 2016;**16**:639–49.
152. Gibson DL, Ma C, Rosenberger CM, Bergstrom KSB, Valdez Y, Huang JT, et al. *Toll-like*

receptor 2 plays a critical role in maintaining mucosal integrity during Citrobacter rodentium-induced colitis. Cell Microbiol 2008;**10**:388–403.
153. Gibson DL, Ma C, Bergstrom KSB, Huang JT, Man C, Vallance BA. *MyD88 signalling plays a critical role in host defence by controlling pathogen burden and promoting epithelial cell homeostasis during Citrobacterrodentium-induced colitis*. Cell Microbiol 2008;**10**:618–31.
154. Lebeis SL, Bommarius B, Parkos CA, Sherman MA, Kalman D. *TLR Signaling Mediated by MyD88 Is Required for a Protective Innate Immune Response by Neutrophils to Citrobacter rodentium* . J Immunol 2007;**179**:566–77.
155. Khan MA, Ma C, Knodler LA, Valdez Y, Rosenberger CM, Deng W, et al. *Toll-like receptor 4 contributes to colitis development but not to host defense during Citrobacter rodentium infection in mice*. Infect Immun 2006;**74**:2522–36.
156. Ciaston I, Dobosz E, Potempa J, Koziel J. *The subversion of toll-like receptor signaling by bacterial and viral proteases during the development of infectious diseases*. Mol Aspects Med 2022;**88**:1–47.
157. Bergstrom KS, Sham HP, Zarepour M, Vallance BA. *Innate host responses to enteric bacterial pathogens: A balancing act between resistance and tolerance*. Cell Microbiol 2012;**14**:475–84.
158. Mullineaux-Sanders C, Sanchez-Garrido J, Hopkins EGD, Shenoy AR, Barry R, Frankel G. *Citrobacter rodentium–host–microbiota interactions: immunity, bioenergetics and metabolism*. Nat Rev Microbiol 2019;**17**:701–15.
159. Geddes K, Rubino SJ, Magalhaes JG, Streutker C, Le Bourhis L, Cho JH, et al. *Identification of an innate T helper type 17 response to intestinal bacterial pathogens*. Nat Med 2011;**17**:837–44.
160. Kim YG, Kamada N, Shaw MH, Warner N, Chen GY, Franchi L, et al. *The Nod2 Sensor Promotes Intestinal Pathogen Eradication via the Chemokine CCL2-Dependent Recruitment of Inflammatory Monocytes*. Immunity 2011;**34**:769–80.
161. Zheng Y, Valdez PA, Danilenko DM, Hu Y, Sa SM, Gong Q, et al. *Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens*. Nat Med 2008;**14**:282–9.
162. Seo SU, Kuffa P, Kitamoto S, Nagao-Kitamoto H, Rousseau J, Kim YG, et al. *Intestinal macrophages arising from CCR2+ monocytes control pathogen infection by activating innate lymphoid cells*. Nat Commun 2015;**6**.
163. Ishigame H, Kakuta S, Nagai T, Kadoki M, Nambu A, Komiyama Y, et al. *Differential Roles of Interleukin-17A and -17F in Host Defense against Mucoepithelial Bacterial Infection and Allergic Responses*. Immunity 2009;**30**:108–19.
164. Kinnebrew MA, Buffie CG, Diehl GE, Zenewicz LA, Leiner I, Hohl TM, et al. *Interleukin 23 Production by Intestinal CD103 +CD11b + Dendritic Cells in Response to Bacterial Flagellin Enhances Mucosal Innate Immune Defense*. Immunity 2012;**36**:276–87.
165. Rathinam VAK, Vanaja SK, Fitzgerald KA. *Regulation of inflammasome signaling*. Nat Immunol 2012;**13**:333–42.
166. Rathinam VAK, Vanaja SK, Waggoner L, Sokolovska A, Becker C, Stuart LM, et al. *TRIF licenses caspase-11-dependent NLRP3 inflammasome activation by gram-negative bacteria*. Cell 2012;**150**:606–19.
167. Kayagaki N, Warming S, Lamkanfi M, Walle L Vande, Louie S, Dong J, et al. *Non-canonical inflammasome activation targets caspase-11*. Nature 2011;**479**:117–21.

168. Liu Z, Zaki MH, Vogel P, Gurung P, Finlay BB, Deng W, et al. *Role of inflammasomes in host defense against citrobacter rodentium infection*. J Biol Chem 2012;**287**:16955–64.
169. Simmons CP, Clare S, Ghaem-Maghami M, Uren TK, Rankin J, Huett A, et al. *Central role for B lymphocytes and CD4+ T cells in immunity to infection by the attaching and effacing pathogen Citrobacter rodentium*. Infect Immun 2003;**71**:5077–86.
170. Vallance BA, Deng W, Knodler LA, Finlay BB. *Mice lacking T and B lymphocytes develop transient colitis and crypt hyperplasia yet suffer impaired bacterial clearance during Citrobacter rodentium infection*. Infect Immun 2002;**70**:2070–81.
171. Backert I, Koralov SB, Wirtz S, Kitowski V, Billmeier U, Martini E, et al. *STAT3 Activation in Th17 and Th22 Cells Controls IL-22–Mediated Epithelial Host Defense during Infectious Colitis*. J Immunol 2014;**193**:3779–91.
172. Higgins LM, Frankel G, Connerton I, Gonçalves NS, Dougan G, MacDonald TT. *Role of Bacterial Intimin in Colonic Hyperplasia and Inflammation*. Science (80-) 1999;**285**:588–91.
173. Wang X, Ota N, Manzanillo P, Kates L, Zavala-Solorio J, Eidenschenk C, et al. *Interleukin-22 alleviates metabolic disorders and restores mucosal immunity in diabetes*. Nature 2014;**514**:237–41.
174. Maaser C, Housley MP, Iimura M, Smith JR, Vallance BA, Finlay BB, et al. *Clearance of Citrobacter rodentium requires B cells but not secretory immunoglobulin A (IgA) or IgM antibodies*. Infect Immun 2004;**72**:3315–24.
175. Bry L, Brenner MB. *Critical Role of T Cell-Dependent Serum Antibody, but Not the Gut-Associated Lymphoid Tissue, for Surviving Acute Mucosal Infection with Citrobacter rodentium, an Attaching and Effacing Pathogen*. J Immunol 2004;**172**:433–41.
176. Kamada N, Sakamoto K, Seo S-U, Zeng MY, Kim Y-G, Cascalho M, et al. *Humoral Immunity in the Gut Selectively Targets Phenotypically Virulent Attaching-and-Effacing Bacteria for Intraluminal Elimination*. Cell Host Microbe 2015;**17**:617–27.
177. Drexler M. Secret Agents: The Menace of Emerging Infections. The National Academies Press; 2002.
178. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey H V., Domazet-Lošo T, Douglas AE, et al. *Animals in a bacterial world, a new imperative for the life sciences*. Proc Natl Acad Sci U S A 2013;**110**:3229–36.
179. Diard M, Hardt WD. *Evolution of bacterial virulence*. FEMS Microbiol Rev 2017;**41**:679–97.
180. Che D, Hasan MS, Chen B. *Identifying pathogenicity islands in bacterial pathogenomics using computational approaches*. Pathogens 2014;**3**:36–56.
181. Hacker J, Bender L, Ott M, Wingender J, Lund B, Marre R, et al. *Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal Escherichia coli isolates*. Microb Pathog 1990;**8**:213–25.
182. Almagro-Moreno S. *How Bacterial Pathogens Emerge*. Am Sci 2022;**110**:162–84.
183. Perna NT, Iii GP, Burland V, Mau B, Glasner JD, Rose DJ, et al. *Genome sequence of enterohaemorrhagic Escherichia coli O157:H7*. Nature 2001;**409**:529–33.
184. Donnenberg MS, Giron JA, Nataro JP. *A plasmid-encoded type IV fimbrial gene of enteropathogenic Escheriehia coli associated with localized adherence*. Mol Microbiol 1992;**6**:3427–37.
185. Tobe T, Sasakawa C. *Role of bundle-forming pilus of enteropathogenic Escherichia coli in host cell adherence and in microcolony development*. Cell Microbiol 2001;**3**:579–85.

186. Tobe T, Sasakawa C. *Species-specific cell adhesion of enteropathogenic Escherichia coli is mediated by type IV bundle-forming pili*. Cell Microbiol 2002;**4**:29–42.
187. Levine MM, Nataro JP, Karch H, Baldini MM, Kaper JB, Black RE, et al. *The Diarrheal Response of Humans to Some Classic Serotypes of Enteropathogenic Escherichia coli is Dependent on a Plasmid Encoding an Enteroadhesiveness Factor*. J Infect Dis 1985;**152**:550–9.
188. Mundy R, Pickard D, Wilson RK, Simmons CP, Dougan G, Frankel G. *Identification of a novel type IV pilus gene cluster required for gastrointestinal colonization of Citrobacter rodentium*. Mol Microbiol 2003;**48**:795–809.
189. Xicohtencatl-Cortes J, Monteiro-Neto V, Ledesma MA, Jordan DM, Francetic O, Kaper JB, et al. *Intestinal adherence associated with type IV pili of enterohemorrhagic Escherichia coli O157:H7*. J Clin Invest 2007;**117**:3519–29.
190. McDaniel TK, Jarvis KG, Donnenberg MS, Kaper JB. *A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens*. Proc Natl Acad Sci U S A 1995;**92**:1664–8.
191. Elliott SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng YK, Lai LC, et al. *The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic Escherichia coli E2348/69*. Mol Microbiol 1998;**28**:1–4.
192. Wong ARC, Pearson JS, Bright MD, Munera D, Robinson KS, Lee SF, et al. *Enteropathogenic and enterohaemorrhagic Escherichia coli: Even more subversive elements*. Mol Microbiol 2011;**80**:1420–38.
193. Jarvis KG, Girón JA, Jerse AE, Mcdaniel TK, Donnenberg MS, Kaper JB. *Enteropathogenic Escherichia coli contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation*. Proc Natl Acad Sci U S A 1995;**92**:7996–8000.
194. Deng W, Li Y, Vallance BA, Finlay BB. *Locus of Enterocyte Effacement from Citrobacter rodentium: Sequence Analysis and Evidence for Horizontal Transfer among Attaching and Effacing Pathogens*. Infect Immun 2001;**69**:6323–35.
195. Perna NT, Mayhew GF, Pósfai G, Elliott S, Donnenberg MS, Kaper JB, et al. *Molecular evolution of a pathogenicity island from enterohemorrhagic Escherichia coli O157:H7*. Infect Immun 1998;**66**:3810–7.
196. Zhu C, Agin TS, Elliott SJ, Johnson LA, Thate TE, Kaper JB, et al. *Complete nucleotide sequence and analysis of the locus of enterocyte effacement from rabbit diarrheagenic Escherichia coli RDEC-1*. Infect Immun 2001;**69**:2107–15.
197. Hyma KE, Lacher DW, Nelson AM, Bumbaugh AC, Janda JM, Strockbine NA, et al. *Evolutionary genetics of a new pathogenic Escherichia species: Escherichia albertii and related Shigella boydii strains*. J Bacteriol 2005;**187**:619–28.
198. McDaniel TK, Kaper JB. *A cloned pathogenicity island from enteropathogenic Escherichia coli confers the attaching and effacing phenotype on E. coli K-12*. Mol Microbiol 1997;**23**:399–407.
199. Elliott SJ, Yu J, Kaper JB. *The cloned locus of enterocyte effacement from enterohemorrhagic Escherichia coli O157:H7 is unable to confer the attaching and effacing phenotype upon E. coli K-12*. Infect Immun 1999;**67**:4260–3.
200. Hacker J, Kaper JB. *Pathogenicity Islands and the Evolution of Microbes*. Annu Rev Microbiol 2000;**54**:641–79.
201. Furniss RCD, Clements A. *Regulation of the Locus of Enterocyte Effacement in Attaching*

and Effacing Pathogens. J Bacteriol 2017;**200**.
202. Mellies JL, Elliott SJ, Sperandio V, Donnenberg MS, Kaper JB. *The Per regulon of enteropathogenic Escherichia coli: Identification of a regulatory cascade and a novel transcriptional activator, the locus of enterocyte effacement (LEE)-encoded regulator (Ler)*. Mol Microbiol 1999;**33**:296–306.
203. Bustamante VH, Santana FJ, Calva E, Puente JL. *Transcriptional regulation of type III secretion genes in enteropathogenic Escherichia coli: Ler antagonizes H-NS-dependent repression*. Mol Microbiol 2001;**39**:664–78.
204. Gomez-Duarte OG, Kaper JB. *A plasmid-encoded regulatory region activates chromosomal eaeA expression in enteropathogenic Escherichia coli*. Infect Immun 1995;**63**:1767–76.
205. Deng W, Puente JL, Gruenheid S, Li Y, Vallance BA, Vázquez A, et al. *Dissecting virulence: Systematic and functional analyses of a pathogenicity island*. Proc Natl Acad Sci U S A 2004;**101**:3597–602.
206. Islam MS, Bingle LEH, Pallen MJ, Busby SJW. *Organization of the LEE1 operon regulatory region of enterohaemorrhagic Escherichia coli O157:H7 and activation by GrlA*. Mol Microbiol 2011;**79**:468–83.
207. Jiménez R, Cruz-Migoni SB, Huerta-Saquero A, Bustamante VH, Puente JL. *Molecular characterization of GrlA, a specific positive regulator of ler expression in enteropathogenic Escherichia coli*. J Bacteriol 2010;**192**:4627–42.
208. Padavannil A, Jobichen C, Mills E, Velazquez-Campoy A, Li M, Leung KY, et al. *Structure of GrlR-GrlA complex that prevents GrlA activation of virulence genes*. Nat Commun 2013;**4**:2–11.
209. Islam MS, Krachler AM. *Mechanosensing regulates virulence in Escherichia coli O157:H7*. Gut Microbes 2016;**7**:63–7.
210. Alsharif G, Ahmad S, Islam MS, Shah R, Busby SJ, Krachler AM. *Host attachment and fluid shear are integrated into a mechanical signal regulating virulence in Escherichia coli O157:H7*. Proc Natl Acad Sci U S A 2015;**112**:5503–8.
211. Hueck CJ. *Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants*. Microbiol Mol Biol Rev 1998;**62**:379–433.
212. Cornelis GR. *The type III secretion injectisome*. Nat Rev Microbiol 2006;**4**:811–25.
213. Slater SL, Sågfors AM, Pollard DJ, Ruano-Gallego D, Frankel G. *The Type III Secretion System of Pathogenic Escherichia coli*. Curr Top Microbiol Immunol 2018;**461**:51–72.
214. Deng W, Li Y, Hardwidge PR, Frey EA, Pfuetzner RA, Lee S, et al. *Regulation of type III secretion hierarchy of translocators and effectors in attaching and effacing bacterial pathogens*. Infect Immun 2005;**73**:2135–46.
215. Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, et al. *Assembly, structure, function and regulation of type III secretion systems*. Nat Rev Microbiol 2017;**15**:323–37.
216. Gauthier A, Puente JL, Finlay BB. *Secretin of the enteropathogenic Escherichia coli type III secretion system requires components of the type III apparatus for assembly and localization*. Infect Immun 2003;**71**:3310–9.
217. Knutton S, Rosenshine M, Pallen MJ, Nisan I, Neves BC, Bain C, et al. *A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells*. EMBO J 1998;**17**:2166–76.
218. Sekiya K, Ohishi M, Ogino T, Tamano K, Sasakawa C, Abe A. *Supermolecular structure*

of the enteropathogenic Escherichia coli type III secretion system and its direct interaction with the EspA-sheath-like structure. Proc Natl Acad Sci U S A 2001;**98**:11638–43.
219. Gaytán MO, Martínez-Santos VI, Soto E, González-Pedrajo B. *Type three secretion system in attaching and effacing pathogens*. Front Cell Infect Microbiol 2016;**6**:1–25.
220. Ide T, Laarmann S, Greune L, Schillers H, Oberleithner H, Schmidt MA. *Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic Escherichia coli*. Cell Microbiol 2001;**3**:669–79.
221. Akeda Y, Galán JE. *Chaperone release and unfolding of substrates in type III secretion*. Nature 2005;**437**:911–5.
222. Gauthier A, Finlay BB. *Translocated Intimin Receptor and Its Chaperone Interact with ATPase of the Type III Secretion Apparatus of Enteropathogenic Escherichia coli*. J Bacteriol 2003;**185**:6747–55.
223. Chen L, Ai X, Portaliou AG, Minetti CA, Remeta DP, Economou A, et al. *Substrate-Activated Conformational Switch on Chaperones Encodes aTargeting Signal in Type III Secretion*. Cell Rep 2013;**3**.
224. Sorg JA, Blaylock B, Schneewind O. *Secretion signal recognition by YscN, the Yersinia type III secretion ATPase*. Proc Natl Acad Sci U S A 2006;**103**:16490–5.
225. Andrade A, Pardo JP, Espinosa N, Pérez-Hernández G, González-Pedrajo B. *Enzymatic characterization of the enteropathogenic Escherichia coli type III secretion ATPase EscN*. Arch Biochem Biophys 2007;**468**:121–7.
226. Thomas NA, Deng W, Baker N, Puente J, Finlay BB. *Hierarchical delivery of an essential host colonization factor in enteropathogenic Escherichia coli*. J Biol Chem 2007;**282**:29634–45.
227. Mills E, Baruch K, Charpentier X, Kobi S, Rosenshine I. *Real-Time Analysis of Effector Translocation by the Type III Secretion System of Enteropathogenic Escherichia coli*. Cell Host Microbe 2008;**3**:104–13.
228. Ghosh P. *Process of Protein Transport by the Type III Secretion System*. Microbiol Mol Biol Rev 2004;**68**:771–95.
229. Crawford JA, Kaper JB. *The N-terminus of enteropathogenic Escherichia coli (EPEC) Tir mediates transport across bacterial and eukaryotic cell membranes*. Mol Microbiol 2002;**46**:855–68.
230. Munera D, Crepin VF, Marches O, Frankel G. *N-terminal type III secretion signal of enteropathogenic Escherichia coli translocator proteins*. J Bacteriol 2010;**192**:3534–9.
231. Deng W, Yu HB, Li Y, Finlay BB. *SepD/SepL-dependent secretion signals of the type III secretion system translocator proteins in enteropathogenic Escherichia coli*. J Bacteriol 2015;**197**:1263–75.
232. Blocker A, Komoriya K, Aizawa SI. *Type III secretion systems and bacterial flagella: Insights into their function from structural similarities*. Proc Natl Acad Sci U S A 2003;**100**:3027–30.
233. Van Gijsegem F, Cornelis GR. *Assembly and Function of Type Iii Secretory Systems*. Annu Rev Microbiol 2000;**54**:735–74.
234. Page AL, Parsot C. *Chaperones of the type III secretion pathway: Jacks of all trades*. Mol Microbiol 2002;**46**:1–11.
235. Parsot C, Hamiaux C, Page AL. *The various and varying roles of specific chaperones in type III secretion systems*. Curr Opin Microbiol 2003;**6**:7–14.

236. Wainwright LA, Kaper JB. *EspB and EspD require a specific chaperone for proper secretion from enteropathogenic Escherichia coil*. Mol Microbiol 1998;**27**:1247–60.
237. Abe A, De Grado M, Pfuetzner RA, Sánchez-SanMartín C, DeVinney R, Puente JL, et al. *Enteropathogenic Escherichia coli translocated intimin receptor, Tir, requires a specific chaperone for stable secretion*. Mol Microbiol 1999;**33**:1162–75.
238. Neves BC, Mundy R, Petrovska L, Dougan G, Knutton S, Frankel G. *CesD2 of enteropathogenic Escherichia coli is a second chaperone for the type III secretion translocator protein EspD*. Infect Immun 2003;**71**:2130–41.
239. Creasey EA, Delahay RM, Bishop AA, Shaw RK, Kenny B, Knutton S, et al. *CesT is a bivalent enteropathogenic Escherichia coli chaperone required for translocation of both Tir and Map*. Mol Microbiol 2003;**47**:209–21.
240. Creasey EA, Friedberg D, Shaw RK, Umanski T, Knutton S, Rosenshine I, et al. *CesAB is an enteropathogenic Escherichia coli chaperone for the type-III translocator proteins EspA and EspB*. Microbiology 2003;**149**:3639–47.
241. Thomas NA, Deng W, Puente JL, Frey EA, Yip CK, Strynadka NCJ, et al. *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*. Mol Microbiol 2005;**57**:1762–79.
242. Su MSW, Kao HC, Lin CN, Syu WJ. *Gene l0017 encodes a second chaperone for EspA of Enterohaemorrhagic Escherichia coli O157 : H7*. Microbiology 2008;**154**:1094–103.
243. Sal-Man N, Setiaputra D, Scholz R, Deng W, Yu ACY, Strynadka NCJ, et al. *EscE and EscG are cochaperones for the type III needle protein EscF of enteropathogenic escherichia coli*. J Bacteriol 2013;**195**:2481–9.
244. Allison SE, Tuinema BR, Everson ES, Sugiman-Marangos S, Zhang K, Junop MS, et al. *Identification of the docking site between a type III secretion system ATPase and a chaperone for effector cargo*. J Biol Chem 2014;**289**:23734–44.
245. Elliott SJ, Hutcheson SW, Dubois MS, Mellies JL, Wainwright LA, Batchelor M, et al. *Identification of CesT, a chaperone for the type III secretion of Tir in enteropathogenic Escherichia coli*. Mol Microbiol 1999;**33**:1176–89.
246. Garmendia, Junkal, Frankel Gad CVF. *Enteropathogenic and Enterohemorrhagic Escherichia coli Infections: Translocation, Translocation, Translocation*. Infect Immun 2005;**73**:2573–85.
247. Guttman JA, Finlay BB. *Subcellular alterations that lead to diarrhea during bacterial pathogenesis*. Trends Microbiol 2008;**16**:535–42.
248. Kenny B, Devinney R, Stein M, Reinscheid DJ, Frey EA, Finlay BB. *Enteropathogenic E . coli (EPEC) Transfers Its Receptor for Intimate Adherence into Mammalian Cells*. Cell 1997;**91**:511–20.
249. Kalman D, Weiner OD, Goosney DL, Sedat JW, Finlay BB, Abo A, et al. *Enteropathogenic E. coli acts through WASP and Arp2/3 complex to form actin pedestals*. Nat Cell Biol 1999;**1**:389–91.
250. Gruenheid S, Devinney R, Bladt F, Goosney D, Gelkop S, Gish GD, et al. *Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells*. Nat Cell Biol 2001;**3**.
251. Xue Y, Du M, Sheng H, Hovde CJ, Zhu MJ. *Escherichia coli O157:H7 suppresses host autophagy and promotes epithelial adhesion via Tir-mediated and cAMP-independent activation of protein kinase A*. Cell Death Discov 2017;**3**.

252. Yan D, Wang X, Luo L, Cao X, Ge B. *Inhibition of TLR signaling by a bacterial protein containing immunoreceptor tyrosine-based inhibitory motifs*. Nat Immunol 2012;**13**:1063–71.
253. Goddard PJ, Sanchez-Garrido J, Slater SL, Kalyan M, Ruano-Gallego D, Marchès O, et al. *Enteropathogenic Escherichia coli Stimulates Effector-Driven Rapid Caspase-4 Activation in Human Macrophages*. Cell Rep 2019;**27**:1008-1017.e6.
254. Zhong Q, Roumeliotis TI, Kozik Z, Cepeda-Molero M, Fernández LÁ, Shenoy AR, et al. Clustering of Tir during enteropathogenic E. coli infection triggers calcium influx–dependent pyroptosis in intestinal epithelial cells. vol. 18. 2020.
255. Yan D, Quan H, Wang L, Liu F, Liu H, Chen J, et al. *Enteropathogenic Escherichia coli Tir recruits cellular SHP-2 through ITIM motifs to suppress host immune response*. Cell Signal 2013;**25**:1887–94.
256. Zhou R, Chen Z, Hao D, Wang Y, Zhang Y, Yi X, et al. *Enterohemorrhagic Escherichia coli Tir inhibits TAK1 activation and mediates immune evasion*. Emerg Microbes Infect 2019;**8**:734–48.
257. Kenny B, Ellis S, Leard AD, Warawa J, Mellor H, Jepson MA. *Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic Escherichia coli effector molecules*. Mol Microbiol 2002;**44**:1095–107.
258. Alto NM, Shao F, Lazar CS, Brost RL, Chua G, Mattoo S, et al. *Identification of a bacterial type III effector family with G protein mimicry functions*. Cell 2006;**124**:133–45.
259. Berger CN, Crepin VF, Jepson MA, Arbeloa A, Frankel G. *The mechanisms used by enteropathogenic Escherichia coli to control filopodia dynamics*. Cell Microbiol 2009;**11**:309–22.
260. Iizumi Y, Sagara H, Kabe Y, Azuma M, Kume K, Ogawa M, et al. *The Enteropathogenic E. coli Effector EspB Facilitates Microvillus Effacing and Antiphagocytosis by Inhibiting Myosin Function*. Cell Host Microbe 2007;**2**:383–92.
261. Hamaguchi M, Hamada D, Suzuki KN, Sakata I, Yanagihara I. *Molecular basis of actin reorganization promoted by binding of enterohaemorrhagic Escherichia coli EspB to α-catenin*. FEBS J 2008;**275**:6260–7.
262. Taylor KA, Luther PW, Donnenberg MS. *Expression of the EspB Protein of Enteropathogenic Escherichia coil within HeLa Cells Affects Stress Fibers and Cellular Morphology*. Infect Immun 1999;**67**:120–5.
263. Luo W, Donnenberg MS. *Analysis of the Function of Enteropathogenic Escherichia coli EspB by Random Mutagenesis*. Infect Immun 2006;**74**:810–20.
264. Nougayrède JP, Donnenberg MS. *Enteropathogenic Escherichia coli EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway*. Cell Microbiol 2004;**6**:1097–111.
265. Marchès O, Batchelor M, Shaw RK, Patel A, Cummings N, Nagai T, et al. *EspF of enteropathogenic Escherichia coli binds sorting nexin 9*. J Bacteriol 2006;**188**:3110–5.
266. Hodges K, Alto NM, Ramaswamy K, Dudeja PK, Hecht G. *The enteropathogenic Escherichia coli effector protein EspF decreases sodium hydrogen exchanger 3 activity*. Cell Microbiol 2008;**10**:1735–45.
267. Alto NM, Weflen AW, Rardin MJ, Yarar D, Lazar CS, Tonikian R, et al. *The type III effector EspF coordinates membrane trafficking by the spatiotemporal activation of two eukaryotic signaling pathways*. J Cell Biol 2007;**178**:1265–78.
268. Dong N, Zhu Y, Lu Q, Hu L, Zheng Y, Shao F. *Structurally distinct bacterial TBC-like*

GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. Cell 2012;**150**:1029–41.
269. Furniss RCD, Slater S, Frankel G, Clements A. *Enterohaemorrhagic E. coli modulates an ARF6:Rab35 signaling axis to prevent recycling endosome maturation during infection*. J Mol Biol 2016;**428**:3399–407.
270. Hardwidge PR, Deng W, Vallance BA, Rodriguez-Escudero I, Cid VJ, Molina M, et al. *Modulation of host cytoskeleton function by the enteropathogenic Escherichia coli and Citrobacter rodentium effector protein EspG*. Infect Immun 2005;**73**:2586–94.
271. Selyunin AS, Sutton SE, Weigele BA, Reddick LE, Orchard RC, Bresson SM, et al. *The assembly of a GTPase–kinase signalling complex by a bacterial catalytic scaffold*. Nature 2011;**469**:107–11.
272. Glotfelty LG, Zahs A, Hodges K, Shan K, Alto NM, Hecht GA. *Enteropathogenic E. coli Effectors EspG1/G2 Disrupt Microtubules, Contribute to Tight Junction Perturbation and Inhibit Restoration*. Cell Microbiol 2014;**16**:1767–83.
273. Tu X, Nisan I, Yona C, Hanski E, Rosenshine I. *EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic Escherichia coli*. Mol Microbiol 2003;**47**:595–606.
274. Dong N, Liu L, Shao F. *A bacterial effector targets host DH-PH domain RhoGEFs and antagonizes macrophage phagocytosis*. EMBO J 2010;**29**:1363–76.
275. N. Berger C, F. Crepin V, Baruch K, Mousnier A, Rosenshine I, Frankel G. *EspZ of Enteropathogenic and Enterohemorrhagic Escherichia coli*. MBio 2012;**3**:1–12.
276. Shames SR, Croxen MA, Deng W, Finlay BB. *The type III system-secreted effector EspZ localizes to host mitochondria and interacts with the translocase of inner mitochondrial membrane 17b*. Infect Immun 2011;**79**:4784–90.
277. Simpson N, Shaw R, Crepin VF, Mundy R, FitzGerald AJ, Cummings N, et al. *The enteropathogenic Escherichia coli type III secretion system effector Map binds EBP50/NHERF1: Implication for cell signalling and diarrhoea*. Mol Microbiol 2006;**60**:349–63.
278. Zihni C, Mills C, Matter K, Balda MS. *Tight junctions: From simple barriers to multifunctional molecular gates*. Nat Rev Mol Cell Biol 2016;**17**:564–80.
279. Deibel C, Krämer S, Chakraborty T, Ebel F. *EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90 kDa protein*. Mol Microbiol 1998;**28**:463–74.
280. De Grado M, Abe A, Gauthier A, Steele-Mortimer O, DeVinney R, Finlay BB. *Identification of the intimin-binding domain of Tir of enteropathogenic Escherichia coli*. Cell Microbiol 1999;**1**:7–17.
281. Hartland EL, Batchelor M, Delahay RM, Hale C, Matthews S, Dougan G, et al. *Binding of intimin from enteropathogenic Escherichia coli to Tir and to host cells*. Mol Microbiol 1999;**32**:151–8.
282. Kenny B. *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*. Mol Microbiol 1999;**31**:1229–41.
283. Campellone KG, Leong JM. *Tails of two Tirs: actin pedestal formation by enteropathogenic E. coli and enterohemorrhagic E. coli O157:H7*. Curr Opin Microbiol 2003;**7**:82–90.
284. Goosney DL, Vinney RDE. *Recruitment of Cytoskeletal and Signaling Proteins to*

Enteropathogenic and Enterohemorrhagic Escherichia coli Pedestals. Infect Immun 2001;**69**:3315–22.
285. Campellone KG, Robbins D, Leong JM. *EspF U Is a Translocated EHEC Effector that Interacts with Tir and N-WASP and Promotes Nck-Independent Actin Assembly*. Dev Cell 2004;**7**:217–28.
286. Brady MJ, Campellone KG, Ghildiyal M, Leong JM. *Enterohaemorrhagic and enteropathogenic Escherichiacoli Tir proteins trigger a common Nck-independent actin assembly pathway*. Cell Microbiol 2007;**9**:2242–53.
287. Garmendia J, Phillips AD, Carlier MF, Chong Y, Schüller S, Marches O, et al. *TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton*. Cell Microbiol 2004;**6**:1167–83.
288. Matsuzawa T, Kuwae A, Yoshida S, Sasakawa C, Abe A. *Enteropathogenic Escherichia coli activates the RhoA signaling pathway via the stimulation of GEF-H1*. EMBO J 2004;**23**:3570–82.
289. Shaw RK, Smollett K, Cleary J, Garmendia J, Straatman-Iwanowska A, Frankel G, et al. *Enteropathogenic Escherichia coli type III effectors EspG and EspG2 disrupt the microtubule network of intestinal epithelial cells*. Infect Immun 2005;**73**:4385–90.
290. Mundy R, Petrovska L, Smollett K, Simpson N, Wilson RK, Yu J, et al. *Identification of a Novel Citrobacter rodentium Type III Secreted Protein, EspI, and Roles of This and Other Secreted Proteins in Infection*. Infect Immun 2004;**72**:2288–302.
291. Nagai T, Abe A, Sasakawa C. *Targeting of enteropathogenic Escherichia coli EspF to host mitochondria is essential for bacterial pathogenesis: Critical role of the 16th leucine residue in EspF*. J Biol Chem 2005;**280**:2998–3011.
292. Crane JK, McNamara BP, Donnenberg MS. *Role of EspF in host cell death induced by enteropathogenic Escherichia coli*. Cell Microbiol 2001;**3**:197–211.
293. Kanack KJ, Crawford JA, Tatsuno I, Karmali MA, Kaper JB. *SepZ/EspZ is secreted and translocated into HeLa cells by the enteropathogenic Escherichia coli type III secretion system*. Infect Immun 2005;**73**:4327–37.
294. Mills E, Baruch K, Aviv G, Nitzan M, Rosenshine I. *Dynamics of the type III secretion system activity of enteropathogenic Escherichia coli*. MBio 2013;**4**.
295. Serapio-Palacios A, Finlay BB. *Dynamics of expression, secretion and translocation of type III effectors during enteropathogenic Escherichia coli infection*. Curr Opin Microbiol 2020;**54**:67–76.
296. Tobe T, Beatson SA, Taniguchi H, Abe H, Bailey CM, Fivian A, et al. *An extensive repetoire of type III secretion effectors in Escherichia coli O157 and the role of lambdoid phages in their dissemination*. Proc Natl Acad Sci U S A 2006;**103**:14941–6.
297. Iguchi A, Thomson NR, Ogura Y, Saunders D, Ooka T, Henderson IR, et al. *Complete genome sequence and comparative genome analysis of enteropathogenic Escherichia coli O127:H6 strain E2348/69*. J Bacteriol 2009;**91**:347–54.
298. Yen H, Sugimoto N, Tobe T. *Enteropathogenic Escherichia coli Uses NleA to Inhibit NLRP3 Inflammasome Activation*. PLoS Pathog 2015;**11**:1–23.
299. Kim J, Thanabalasuriar A, Chaworth-Musters T, Fromme JC, Frey EA, Lario PII, et al. *The Bacterial Virulence Factor NleA Inhibits Cellular Protein Secretion by Disrupting Mammalian COPII Function*. Cell Host Microbe 2007;**2**:160–71.
300. Kelly M, Hart E, Mundy R, Marchès O, Wiles S, Badea L, et al. *Essential Role of the Type III Secretion System Effector NleB in Colonization of Mice by Citrobacter*

rodentium. Infect Immun 2006;**74**:2328–37.
301. Wickham ME, Lupp C, Mascarenhas M, Vázquez A, Coombes BK, Brown NF, et al. *Bacterial genetic determinants of non-O157 STEC outbreaks and hemolytic-uremic syndrome after infection*. J Infect Dis 2006;**194**:819–27.
302. Li S, Zhang L, Yao Q, Li L, Dong N, Rong J, et al. *Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains*. Nature 2013;**501**:242–6.
303. Pearson JS, Giogha C, Ong SY, Kennedy CL, Kelly M, Robinson KS, et al. *A type III effector antagonizes death receptor signalling during bacterial gut infection*. Nature 2013;**501**:247–51.
304. Baruch K, Gur-Arie L, Nadler C, Koby S, Yerushalmi G, Ben-Neriah Y, et al. *Metalloprotease type III effectors that specifically cleave JNK and NF-κB*. EMBO J 2011;**30**:221–31.
305. Pearson JS, Riedmaier P, Marchès O, Frankel G, Hartland EL. *A type III effector protease NleC from enteropathogenic Escherichia coli targets NF-κB for degradation*. Mol Microbiol 2011;**80**:219–30.
306. Yen H, Ooka T, Iguchi A, Hayashi T, Sugimoto N, Tobe T. *Nlec, a type III secretion protease, compromises NF-kB activation by targeting p65/rela*. PLoS Pathog 2010;**6**.
307. Wickham ME, Lupp C, Vázquez A, Mascarenhas M, Coburn B, Coombes BK, et al. *Citrobacter rodentium virulence in mice associates with bacterial load and the type III effector NleE*. Microbes Infect 2007;**9**:400–7.
308. Zhang L, Ding X, Cui J, Xu H, Chen J, Gong YN, et al. *Cysteine methylation disrupts ubiquitin-chain sensing in NF-κB activation*. Nature 2012;**481**:204–10.
309. Yao Q, Zhang L, Wan X, Chen J, Hu L, Ding X, et al. *Structure and Specificity of the Bacterial Cysteine Methyltransferase Effector NleE Suggests a Novel Substrate in Human DNA Repair Pathway*. PLoS Pathog 2014;**10**.
310. Echtenkamp F, Deng W, Wickham ME, Vazquez A, Puente JL, Thanabalasuriar A, et al. *Characterization of the NleF effector protein from attaching and effacing bacterial pathogens*. FEMS Microbiol Lett 2008;**281**:98–107.
311. Blasche S, Mortl M, Steuber H, Siszler G, Nisa S, Schwarz F, et al. *The E. coli Effector Protein NleF Is a Caspase Inhibitor*. PLoS One 2013;**8**.
312. Wu B, Skarina T, Yee A, Jobin MC, Dileo R, Semesi A, et al. *NleG type 3 effectors from enterohaemorrhagic Escherichia coli are U-Box E3 ubiquitin ligases*. PLoS Pathog 2010;**6**.
313. Valleau D, Little DJ, Borek D, Skarina T, Quaile AT, Leo R Di, et al. *Functional diversification of the NleG effector family in enterohemorrhagic Escherichia coli*. Proc Natl Acad Sci U S A 2018;**115**:10004–9.
314. García-Angulo VA, Deng W, Thomas NA, Finlay BB, Puente JL. *Regulation of expression and secretion of NleH, a new non-locus of enterocyte effacement-encoded effector in Citrobacter rodentium*. J Bacteriol 2008;**190**:2388–99.
315. Hemrajani C, Marches O, Wiles S, Girard F, Dennis A, Dziva F, et al. *Role of NleH, a type III secreted effector from attaching and effacing pathogens, in colonization of the bovine, ovine, and murine gut*. Infect Immun 2008;**76**:4804–13.
316. Hemrajani C, Berger CN, Robinson KS, Marchès O, Mousnier A, Frankel G. *NleH effectors interact with Bax inhibitor-1 to block apoptosis during enteropathogenic Escherichia coli infection*. Proc Natl Acad Sci U S A 2010;**107**:3129–34.
317. Gao X, Wan F, Mateo K, Callegari E, Wang D, Deng W, et al. *Bacterial effector binding*

to ribosomal protein S3 subverts NF-κB function. PLoS Pathog 2009;**5**.
318. Marchès O, Ledger TN, Boury M, Ohara M, Tu X, Goffaux F, et al. *Enteropathogenic and enterohaemorrhagic Escherichia coli deliver a novel effector called Cif, which blocks cell cycle G2/M transition*. Mol Microbiol 2003;**50**:1553–67.
319. Ritchie JM, Brady MJ, Riley KN, Ho TD, Campellone KG, Herman IM, et al. *EspF U, a type III-translocated effector of actin assembly, fosters epithelial association and late-stage intestinal colonization by E. coli O157:H7*. Cell Microbiol 2008;**10**:836–47.
320. Sallee NA, Rivera GM, Dueber JE, Vasilescu D, Mullins RD, Mayer BJ, et al. *The pathogen protein EspFU hijacks actin polymerization using mimicry and multivalency*. Nature 2008;**454**:1005–8.
321. Smith K, Humphreys D, Hume PJ, Koronakis V. *Enteropathogenic Escherichia coli Recruits the Cellular Inositol Phosphatase SHIP2 to Regulate Actin-Pedestal Formation*. Cell Host Microbe 2010;**7**:13–24.
322. Weiss SM, Ladwein M, Schmidt D, Ehinger J, Lommel S, Städing K, et al. *IRSp53 Links the Enterohemorrhagic E. coli Effectors Tir and EspFU for Actin Pedestal Formation*. Cell Host Microbe 2009;**5**:244–58.
323. Dahan S, Wiles S, La Ragione RM, Best A, Woodward MJ, Stevens MP, et al. *EspJ is a prophage-carried type III effector protein of attaching and effacing pathogens that modulates infection dynamics*. Infect Immun 2005;**73**:679–86.
324. Marchès O, Covarelli V, Dahan S, Cougoule C, Bhatta P, Frankel G, et al. *EspJ of enteropathogenic and enterohaemorrhagic Escherichia coli inhibits opsono-phagocytosis*. Cell Microbiol 2008;**10**:1104–15.
325. Pollard DJ, Berger CN, So EC, Yu L, Hadavizadeh K, Jennings P, et al. *Broad-spectrum regulation of nonreceptor tyrosine kinases by the bacterial ADP-ribosyltransferase Espj*. MBio 2018;**9**.
326. Young JC, Clements A, Lang AE, Garnett JA, Munera D, Arbeloa A, et al. *The Escherichia coli effector EspJ blocks Src kinase activity via amidation and ADP ribosylation*. Nat Commun 2014;**5**:5887.
327. Pearson JS, Giogha C, Mühlen S, Nachbur U, Pham CLL, Zhang Y, et al. *EspL is a bacterial cysteine protease effector that cleaves RHIM proteins to block necroptosis and inflammation*. Nat Microbiol 2017;**2**.
328. Deng W, De Hoog CL, Yu HB, Li Y, Croxen MA, Thomas NA, et al. *A comprehensive proteomic analysis of the type III secretome of Citrobacter rodentium*. J Biol Chem 2010;**285**:6790–800.
329. Arbeloa A, Bulgin RR, Mackenzie G, Shaw RK, Pallen MJ, Crepin VF, et al. *Subversion of actin dynamics by EspM effectors of attaching and effacing bacterial pathogens*. Cell Microbiol 2008;**10**:1429–41.
330. Simovitch M, Sason H, Cohen S, Zahavi EE, Melamed-Book N, Weiss A, et al. *EspM inhibits pedestal formation by enterohaemorrhagic Escherichia coli and enteropathogenic E. coli and disrupts the architecture of a polarized epithelial monolayer*. Cell Microbiol 2010;**12**:489–505.
331. Dennis A, Kudo T, Kruidenier L, Girard F, Crepin VF, MacDonald TT, et al. *The p50 subunit of NF-κB is critical for in vivo clearance of the noninvasive enteric pathogen Citrobacter rodentium*. Infect Immun 2008;**76**:4978–88.
332. Nadler C, Baruch K, Kobi S, Mills E, Haviv G, Farago M, et al. *The type III secretion effector NleE inhibits NF-κB activation*. PLoS Pathog 2010;**6**.

333. Mühlen S, Ruchaud-Sparagano MH, Kenny B. *Proteasome-independent degradation of canonical NFκB complex components by the NleC protein of pathogenic Escherichia coli*. J Biol Chem 2011;**286**:5100–7.
334. Sham HP, Shames SR, Croxen MA, Ma C, Chan JM, Khan MA, et al. *Attaching and effacing bacterial effector NleC suppresses epithelial inflammatory responses by inhibiting NF-κB and p38 mitogen-activated protein kinase activation*. Infect Immun 2011;**79**:3552–62.
335. Shames SR, Bhavsar AP, Croxen MA, Law RJ, Mak SHC, Deng W, et al. *The pathogenic Escherichia coli type III secreted protease NleC degrades the host acetyltransferase p300*. Cell Microbiol 2011;**13**:1542–57.
336. Gao X, Wang X, Pham TH, Feuerbacher LA, Lubos ML, Huang M, et al. *NleB, a Bacterial Effector with Glycosyltransferase Activity, Targets GAPDH Function to Inhibit NF-κB Activation*. Cell Host Microbe 2013;**13**:87–99.
337. Edelblum KL, Yan F, Yamaoka T, Polk DB. *Regulation of apoptosis during homeostasis and disease in the intestinal epithelium*. Inflamm Bowel Dis 2006;**12**:413–24.
338. Allen RT, Hunter WJ, Agrawal DK. *Morphological and biochemical characterization and analysis of apoptosis*. J Pharmacol Toxicol Methods 1997;**37**:215–28.
339. Crane JK, Majumdar S, Pickhardt DF. *Host cell death due to enteropathogenic Escherichia coli has features of apoptosis*. Infect Immun 1999;**67**:2575–84.
340. Pham TH, Gao X, Tsai K, Olsen R, Wan F, Hardwidge PR. *Functional differences and interactions between the escherichia coli type III secretion system effectors NleH1 and NleH2*. Infect Immun 2012;**80**:2133–40.
341. Samba-louaka A, Nougayrède JP, Watrin C, Jubelin G, Oswald E, Taieb F. *Bacterial cyclomodulin Cif blocks the host cell cycle by stabilizing the cyclin-dependent kinase inhibitors p21waf1 and p27kip1*. Cell Microbiol 2008;**10**:2496–508.
342. Gruenheid S, Sekirov I, Thomas NA, Deng W, O'Donnell P, Goode D, et al. *Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic Escherichia coli O157:H7*. Mol Microbiol 2004;**51**:1233–49.
343. Ruano-Gallego D, Sanchez-Garrido J, Kozik Z, Núñez-Berrueco E, Cepeda-Molero M, Mullineaux-Sanders C, et al. *Type III secretion system effectors form robust and flexible intracellular virulence networks*. Science (80-) 2021;**371**.
344. Lee MCS, Miller EA, Goldberg J, Orci L, Schekman R. *Bi-Directional Protein Transport Between the ER and Golgi*. Annu Rev Cell Dev Biol 2004;**20**:87–123.
345. Thanabalasuriar A, Bergeron J, Gillingham A, Mimee M, Thomassin JL, Strynadka N, et al. *Sec24 interaction is essential for localization and virulence-associated function of the bacterial effector protein NleA*. Cell Microbiol 2012;**14**:1206–18.
346. Thanabalasuriar A, Kim J, Gruenheid S. *The inhibition of COPII trafficking is important for intestinal epithelial tight junction disruption during enteropathogenic Escherichia coli and Citrobacter rodentium infection*. Microbes Infect 2013;**15**:738–44.
347. Thanabalasuriar A, Koutsouris A, Weflen A, Mimee M, Hecht G, Gruenheid S. *The bacterial virulence factor NleA is required for the disruption of intestinal tight junctions by enteropathogenic Escherichia coli*. Cell Microbiol 2010;**12**:31–41.
348. McNamara BP, Koutsouris A, O'Connell CB, Nougayréde JP, Donnenberg MS, Hecht G. *Translocated EspF protein from enteropathogenic Escherichia coli disrupts host intestinal barrier function*. J Clin Invest 2001;**107**:621–9.
349. Cepeda-Molero M, Schüller S, Frankel G, Fernández LÁ. Systematic Deletion of Type III

Secretion System Effectors in Enteropathogenic E. coli Unveils the Role of Non-LEE Effectors in A/E Lesion Formation. In: Rodrigo L, editor. E. Coli Infect. - Importance Early Diagnosis Effic. Treat., IntechOpen; 2020.
350. Tahir J, Rashid M, Afzal AJ. *Post-translational modifications in effectors and plant proteins involved in host–pathogen conflicts*. Plant Pathol 2019;**68**:628–44.
351. Popa CM, Tabuchi M, Valls M. *Modification of Bacterial Effector Proteins Inside Eukaryotic Host Cells*. Front Cell Infect Microbiol 2016;**6**:1–13.
352. Khoury GA, Baliban RC, Floudas CA. *Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database*. Sci Rep 2011;**1**:1–5.
353. Olsen J V., Mann M. *Status of large-scale analysis of posttranslational modifications by mass spectrometry*. Mol Cell Proteomics 2013;**12**:3444–52.
354. Tegtmeyer N, Backert S. *Role of Abl and Src family kinases in actin-cytoskeletal rearrangements induced by the Helicobacter pylori CagA protein*. Eur J Cell Biol 2011;**90**:880–90.
355. Backert S, Tegtmeyer N, Selbach M. *The versatility of Helicobacter pylori CagA effector protein functions: The master key hypothesis (Helicobacter 15 (163-176))*. Helicobacter 2010;**15**:483.
356. Sougleri IS, Papadakos KS, Zadik MP, Mavri-Vavagianni M, Mentis AF, Sgouras DN. *Helicobacter pylori CagA protein induces factors involved in the epithelial to mesenchymal transition (EMT) in infected gastric epithelial cells in an EPIYA-phosphorylation-dependent manner*. FEBS J 2016;**283**:206–20.
357. Warawa J, Kenny B. *Phosphoserine modification of the enteropathogenic Escherichia coli Tir molecule is required to trigger conformational changes in Tir and efficient pedestal elongation*. Mol Microbiol 2001;**42**:1269–80.
358. Campellone KG, Leong JM. *Nck-independent actin assembly is mediated by two phosphorylated tyrosines within enteropathogenic Escherichia coli Tir*. Mol Microbiol 2005;**56**:416–32.
359. Ismaili A, Philpott DJ, Dytoc MT, Sherman PM. *Signal transduction responses following adhesion of verocytotoxin- producing Escherichia coli*. Infect Immun 1995;**63**:3316–26.
360. DeVinney R, Stein M, Reinscheid D, Abe A, Ruschkowski S, Finlay BB. *Enterohemorrhagic Escherichia coli O157:H7 Produces Tir, Which Is Translocated to the Host Cell Membrane but Is Not Tyrosine Phosphorylated*. Infect Immun 1999;**67**:2389–98.
361. DeVinney R, Puente JL, Gauthier A, Goosney D, Brett Finlay B, Finlay BB. *Enterohaemorrhagic and enteropathogenic Escherichia coli use a different Tir-based mechanism for pedestal formation*. Mol Microbiol 2001;**41**:1445–58.
362. Kenny B. *The enterohaemorrhagic Escherichia coli (serotype O157:H7) Tir molecule is not functionally interchangeable for its enteropathogenic E. coli (serotype O127:H6) homologue*. Cell Microbiol 2001;**3**.
363. Campellone KG, Giese A, Donald J, Leong JM. *A tyrosine-phosphorylated 12-amino-acid sequence of enteropathogenic Escherichia coli Tir binds the host adaptor protein Nck and is required for Nck localization to actin pedestals*. Mol Microbiol 2002;**43**:1227–41.
364. Brandt S, Kenny B, Rohde M, Martinez-Quiles N, Backert S. *Dual infection system identifies a crucial role for PKA-mediated serine phosphorylation of the EPEC-Tir-injected effector protein in regulating Rac1 function*. Cell Microbiol 2009;**11**:1254–71.

365. Backert S, Kenny B, Gerhard R, Tegtmeyer N, Bran S. *PKA-mediated phosphorylation of EPEC-Tir at serine residues 434 and 463: A novel pathway in regulating Rac1 GTPase function*. Gut Microbes 2010;**1**:94–9.
366. Behrends C, Harper JW. *Constructing and decoding unconventional ubiquitin chains*. Nat Struct Mol Biol 2011;**18**:520–8.
367. Ramanathan HN, Ye Y. *Cellular strategies for making monoubiquitin signals*. Crit Rev Biochem Mol Biol 2012;**47**:17–28.
368. Angot A, Vergunst A, Genin S, Peeters N. *Exploitation of eukaryotic ubiquitin signaling pathways by effectors translocated by bacterial type III and type IV secretion systems*. PLoS Pathog 2007;**3**:0001–13.
369. Zhou Y, Zhu Y. *Diversity of bacterial manipulation of the host ubiquitin pathways*. Cell Microbiol 2015;**17**:26–34.
370. Narayanan LA, Edelmann MJ. *Ubiquitination as an efficient molecular strategy employed in Salmonella infection*. Front Immunol 2014;**5**:1–7.
371. Kubori T, Galán JE. *Temporal regulation of Salmonella virulence effector function by proteasome-dependent protein degradation*. Cell 2003;**115**:333–42.
372. Thomas M, Holden DW. *Ubiquitination-A Bacterial Effector's Ticket to Ride*. Cell Host Microbe 2009;**5**:309–11.
373. Patel JC, Hueffer K, Lam TT, Galán JE. *Diversification of a Salmonella Virulence Protein Function by Ubiquitin-Dependent Differential Localization*. Cell 2009;**137**:283–94.
374. Knodler LA, Winfree S, Drecktrah D, Ireland R, Steele-Mortimer O. *Ubiquitination of the bacterial inositol phosphatase, SopB, regulates its biological activity at the plasma membrane*. Cell Microbiol 2009;**11**:1652–70.
375. Zhang Y, Higashide W, Dai S, Sherman DM, Zhou D. *Recognition and ubiquitination of Salmonella type III effector SopA by a ubiquitin E3 ligase, HsRMA1*. J Biol Chem 2005;**280**:38682–8.
376. Zhang Y, Higashide WM, McCormick BA, Chen J, Zhou D. *The inflammation-associated Salmonella SopA is a HECT-like E3 ubiquitin ligase*. Mol Microbiol 2006;**62**:786–93.
377. Guo C, Henley JM. *Wrestling with stress: Roles of protein SUMOylation and deSUMOylation in cell stress response*. IUBMB Life 2014;**66**:71–7.
378. Wilson VG. *Sumoylation at the host-pathogen interface*. Biomolecules 2012;**2**:203–27.
379. Verma S, Mohapatra G, Ahmad SM, Rana S, Jain S, Khalsa JK, et al. *Salmonella Engages Host MicroRNAs To Modulate SUMOylation: a New Arsenal for Intracellular Survival* . Mol Cell Biol 2015;**35**:2932–46.
380. Beyer AR, Truchan HK, May LJ, Walker NJ, Borjesson DL, Carlyon JA. *The Anaplasma phagocytophilum effector AmpA hijacks host cell SUMOylation*. Cell Microbiol 2015;**17**:504–19.
381. Chen B, Sun Y, Niu J, Jarugumilli GK, Wu X. *Protein Lipidation in Cell Signaling and Diseases: Function, Regulation, and Therapeutic Opportunities*. Cell Chem Biol 2018;**25**:817–31.
382. Reinicke AT, Hutchinson JL, Magee AI, Mastroeni P, Trowsdale J, Kelly AP. *A Salmonella typhimurium effector Protein SifA is modified by host cell prenylation and S-acylation machinery*. J Biol Chem 2005;**280**:14620–7.
383. Al-quadan T, Price CT, London N, Schueler-furman O, Abukwaik Y. *Anchoring of bacterial effectors to host membranes through host-mediated lipidation by prenylation: A common paradigm*. Trends Microbiol 2011;**19**:573–9.

384. Varki A. *Biological roles of glycans*. Glycobiology 2017;**27**:3–49.
385. Reily C, Stewart TJ, Renfrow MB, Novak J. *Glycosylation in health and disease*. Nat Rev Nephrol 2019;**15**:346–66.
386. Laine RA. *A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 × 10 structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems*. Glycobiology 1994;**4**:759–67.
387. Nyathi Y, Wilkinson BM, Pool MR. *Co-translational targeting and translocation of proteins to the endoplasmic reticulum*. Biochim Biophys Acta - Mol Cell Res 2013;**1833**:2392–402.
388. Johnson N, Powis K, High S. *Post-translational translocation into the endoplasmic reticulum*. Biochim Biophys Acta 2013;**1833**:2403–9.
389. Spiro RG. *Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds*. Glycobiology 2002;**12**.
390. Yue J, Huang R, Lan Z, Xiao B, Luo Z. *Abnormal glycosylation in glioma: related changes in biology, biomarkers and targeted therapy*. Biomark Res 2023;**11**:1–21.
391. Stanley P, Moremen KW, Lewis NE, Taniguchi N, Aebi M. N-Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials Glycobiol. 4th ed., Cold Spring Harbor Laboratory Press; 2022.
392. Aebi M. *N-linked protein glycosylation in the ER*. Biochim Biophys Acta - Mol Cell Res 2013;**1833**:2430–7.
393. Brockhausen I, Wandall HH, Hagen KG Ten, Stanley P. O-GalNAc Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials Glycobiol. 4th ed., Cold Spring Harbor Laboratory Press; 2022.
394. Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, Reis CA, et al. *Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:Polypeptide N-acetylgalactosaminyltransferases in Drosophila, Caenorhabditis elegans, and mammals*. J Biol Chem 2002;**277**:22623–38.
395. Ten Hagen KG, Tran DT, Gerken TA, Stein DS, Zhang Z. *Functional characterization and expression analysis of members of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family from Drosophila melanogaster*. J Biol Chem 2003;**278**:35039–48.
396. Jensen PH, Kolarich D, Packer NH. *Mucin-type O-glycosylation - Putting the pieces together*. FEBS J 2010;**277**:81–94.
397. Fu C, Zhao H, Wang Y, Cai H, Xiao Y, Zeng Y, et al. *Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen*. HLA 2016;**88**:275–86.
398. Zachara NE, Akimoto Y, Boyce M, Hart GW. The O-GlcNAc Modification. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials Glycobiol. 4th ed., Cold Spring Harbor Laboratory Press; 2022.
399. Wolfert MA, Boons G-J. *Adaptive immune activation: glycosylation does matter*. Nat Chem Biol 2013;**9**:776–84.
400. Clark MC, Baum LG. *T cells modulate glycans on CD43 and CD45 during development and activation, signal regulation, and survival*. Ann N Y Acad Sci 2012;**1253**:58–67.
401. Quast I, Peschke B, Lünemann JD. *Regulation of antibody effector functions through IgG Fc N-glycosylation*. Cell Mol Life Sci 2017;**74**:837–47.
402. Kaneko Y, Nimmerjahn F, Ravetch J V. *Anti-inflammatory activity of immunoglobulin G*

resulting from Fc sialylation. Science (80-) 2006;**313**:670–3.
403. Nothaft H, Szymanski CM. *Protein glycosylation in bacteria: sweeter than ever*. Nat Rev Microbiol 2010;**8**:765–78.
404. Brimer CD, Montie TC. *Cloning and comparison of fliC genes and identification of glycosylation in the flagellin of Pseudomonas aeruginosa a-type strains*. J Bacteriol 1998;**180**:3209–17.
405. Castric P. *PilO, a gene required for glycosylation of Pseudomonas aeruginosa 1244 pilin*. Microbiology 1995;**141**:1247–54.
406. Stimson E, Virji M, Makepeace K, Dell A, Morris HR, Payne G, et al. *Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose*. Mol Microbiol 1995;**17**:1201–14.
407. Grass S, Buscher AZ, Swords WE, Apicella MA, Barenkamp SJ, Ozchlewski N, et al. *The Haemophilus influenzae HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis*. Mol Microbiol 2003;**48**:737–51.
408. Szymanski CM, Ruijin Y, Ewing CP, Trust TJ, Guerry P. *Evidence for a system of general protein glycosylation in Campylobacter jejuni*. Mol Microbiol 1999;**32**:1022–30.
409. Dobos KM, Khoo K-HK-H, Swiderek KM, Brennan PJ, Belisle JT, Dobos (K, et al. *Definition of the Full Extent of Glycosylation of the 45-Kilodalton Glycoprotein of Mycobacterium tuberculosis Chemical evidence for the true glycosylation of mycobacterial proteins was recently provided in the context of the 45-kDa MPT 32 secreted protei*. J Bacteriol 1996;**178**:2498–506.
410. Wu H, Mintz KP, Ladha M, Fives-Taylor PM. *Isolation and characterization of Fap1, a fimbriae-associated adhesin of Streptococcus parasanguis FW213*. Mol Microbiol 1998;**28**:487–500.
411. Benz I, Schmidt MA. *Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin*. Mol Microbiol 2001;**40**:1403–13.
412. Lindenthal C, Elsinghorst EA. *Identification of a glycoprotein produced by enterotoxigenic Escherichia coli*. Infect Immun 1999;**67**:4084–91.
413. Forrest S, Welch M. *Arming the troops: Post-translational modification of extracellular bacterial proteins*. Sci Prog 2020;**103**:1–22.
414. Merino S, Tomás JM. *Gram-negative flagella glycosylation*. Int J Mol Sci 2014;**15**:2840–57.
415. Frankel G, Phillips AD, Rosenshine I, Dougan G, Kaper JB, Knutton S. *Enteropathogenic and enterohaemorrhagic Escherichia coli: More subversive elements*. Mol Microbiol 1998;**30**:911–21.
416. Knutton S, Lloyd DR, McNeish AS. *Adhesion of enteropathogenic Escherichia coli to human intestinal enterocytes and cultured human intestinal mucosa*. Infect Immun 1987;**55**:69–77.
417. Deng W, Vallance BA, Li Y, Puente JL, Finlay BB. *Citrobacter rodentium translocated intimin receptor (Tir) is an essential virulence factor needed for actin condensation, intestinal colonization and colonic hyperplasia in mice*. Mol Microbiol 2003;**48**:95–115.
418. Vallance BA, Khan MA, Deng W, Gruenheid S, Finlay BB. *Modeling enteropathogenic and enterohemorrhagic E. coli infections and disease*. Drug Discov Today Dis Model 2004;**1**:73–9.

419. Phillips N, Hayward RD, Koronakis V. *Phosphorylation of the enteropathogenic E. coli receptor by the Src-family kinase c-Fyn triggers actin pedestal formation*. Nat Cell Biol 2004;**6**:618–25.
420. Marcus SL, Knodler LA, Finlay BB. *Salmonella enterica serovar Typhimurium effector SigD/SopB is membrane-associated and ubiquitinated inside host cells*. Cell Microbiol 2002;**4**:435–46.
421. Mukhopadhyay D, Riezman H. *Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling*. Science (80-) 2007;**315**:201–5.
422. Coombes BK, Wickham ME, Mascarenhas M, Gruenheid S, Brett Finlay B, Karmali MA. *Molecular Analysis as an Aid To Assess the Public Health Risk of Non-O157 Shiga Toxin-Producing Escherichia coli Strains*. Appl Environ Microbiol 2008;**74**:2153–60.
423. Mundy R, Jenkins C, Yu J, Smith H, Frankel G. *Distribution of espI among clinical enterohaemorrhagic and enteropathogenic Escherichia coli isolates*. J Med Microbiol 2004;**53**:1145–9.
424. Wickham MEMEME, Brown NFNFNF, Boyle ECECEC, Coombes BKBKBK, Finlay BBB. *Virulence Is Positively Selected by Transmission Success between Mammalian Hosts*. Curr Biol 2007;**17**:783–8.
425. Creuzburg K, Recktenwald J, Kuhle V, Herold S, Hensel M, Schmidt H. *The Shiga Toxin 1-Converting Bacteriophage BP-4795 Encodes an NleA-Like Type III Effector Protein*. J Bacteriol 2005;**187**:8494–8.
426. Lee SF, Kelly M, Mcalister A, Luck SN, Garcia EL, Hall RA, et al. *A C-terminal class I PDZ binding motif of EspI/NleA modulates the virulence of attaching and effacing Escherichia coli and Citrobacter rodentium*. Cell Microbiol 2008;**10**:499–513.
427. Struwe WB, Robinson C V. *Relating glycoprotein structural heterogeneity to function – insights from native mass spectrometry*. Curr Opin Struct Biol 2019;**58**:241–8.
428. Krieger M, Reddy P, Kozarsky K, Kingsley D, Hobbie L, Penman M. Analysis of the Synthesis, Intracellular Sorting, and Function of Glycoproteins Using a Mammalian Cell Mutant with Reversible Glycosylation Defects. Methods Cell Biol., vol. 32, Academic Press; 1989, p. 57–84.
429. Donnenberg MS, Kaper JB. *Construction of an eae Deletion Mutant of Enteropathogenic Escherichia coli by Using a Positive-Selection Suicide Vector*. Infect Immun 1991;**59**:4310–7.
430. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTBG, et al. *Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology*. EMBO J 2013;**32**:1478–88.
431. Doms RW, Russ G, Yewdell JW. *Brefeldin A Redistributes Resident and Itinerant Golgi Proteins to the Endoplasmic Reticulum*. J Cell Biol 1989;**109**:61–72.
432. Fujiwara T, Oda K, Yokota S, Takatsuki A, Ikehara Y. *Brefeldin A Causes Disassembly of the Golgi Complex and Accumulation of Secretory Proteins in the Endoplasmic Reticulum*. J Biol Chem 1988;**263**:18545–52.
433. Lippincott-Schwartz J, Yuan LC, Bonifacino JS, Klausner RD. *Rapid Redistribution of Golgi Proteins into the ER in Cells Treated with Brefeldin A: Evidence for Membrane Cycling from Golgi to ER*. Cell 1989;**56**:801–13.
434. Lippincott-Schwartz J, Donaldson JG, Schweizer A, Berger EG, Hauri HP, Yuan LC, et al. *Microtubule-Dependent Retrograde Transport of Proteins into the ER in the Presence of Brefeldin A Suggests an ER Recycling Pathway*. Cell 1990;**60**:821–36.

435. Misumi Y, Misumis Y, Miki K, Takatsukig A, Tamurag G, Ikeharab Y. *Novel Blockade by Brefeldin A of Intracellular Transport of Secretory Proteins in Cultured Rat Hepatocytes*. J Biol Chem 1986;**261**:11398–403.
436. Kubo RT, Pigeon ML. *Expression of membrane IGM by a human B lymphoblastoid cell line in the presence of monensin*. Mol Immunol 1983;**20**:345–8.
437. Mollenhauer HH, James Morré D, Rowe LD. *Alteration of intracellular traffic by monensin; mechanism, specificity and relationship to toxicity*. Biochim Biophys Acta 1990;**1031**:225–46.
438. Tartakoff AM. *Perturbation of Vesicular Traffic with the Carboxylic Ionophore Monensin*. Cell 1983;**32**:1026–8.
439. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega*. Mol Syst Biol 2011;**7**.
440. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. *Accurate prediction of protein structures and interactions using a three-track neural network*. Science (80-) 2021;**373**:871–6.
441. Mitra N, Sinha S, Ramya TNC, Surolia A. *N-linked oligosaccharides as outfitters for glycoprotein folding, form and function*. Trends Biochem Sci 2006;**31**:156–63.
442. Parodi AJ. *Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation.* Biochem J 2000;**348**:1–13.
443. Dwek RA. *Glycobiology: Toward Understanding the Function of Sugars*. Chem Rev 1996;**96**:683–720.
444. Lis H, Sharon N. *Protein glycosylation: Structural and functional aspects*. Eur J Biochem 1993;**218**:1–27.
445. Schüller S, Chong Y, Lewin J, Kenny B, Frankel G, Phillips AD. *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*. Cell Microbiol 2007;**9**:1352–64.
446. Sontag RL, Nakayasu ES, Brown RN, Niemann GS, Sydor MA, Sanchez O, et al. *Identification of Novel Host Interactors of Effectors Secreted by Salmonella and Citrobacter*. MSystems 2016;**1**.
447. Chantret I, Rodolosse A, Barbat A, Dussaulx E, Brot-Laroche E, Zweibaum A, et al. *Differential expression of sucrase-isomaltase in clones isolated from early and late passages of the cell line Caco-2: evidence for glucose-dependent negative regulation*. J Cell Sci 1994;**107**:213–25.
448. Ullman-Culleré MH, Foltz CJ. *Body Condition Scoring: A Rapid and Accurate Method for Assessing Health Status in Mice*. Lab Anim Sci 1999;**49**:319–23.
449. Rong Fu Wang, Kushner SR. *Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli*. Gene 1991;**100**:195–9.
450. Edwards RA, Keller LH, Schifferli DM. *Improved allelic exchange vectors and their use to analyze 987P fimbria gene expression*. Gene 1998;**207**:149–57.
451. Jarvis KG, Kaper JB. *Secretion of Extracellular Proteins by Enterohemorrhagic Escherichia coli via a Putative Type III Secretion System*. Infect Immun 1996;**64**:4826–9.
452. Burns L, Giannakopoulou N, Zhu L, Xu YZ, Khan RH, Bekal S, et al. *The bacterial virulence factor NleA undergoes host-mediated O-linked glycosylation*. Mol Microbiol 2023;**119**:161–73.

453. Esko JD, Wandall HH, Stanley P, Varki A, Cummings RD, Esko JD, et al. Glycosylation Mutants of Cultured Mammalian Cells. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials Glycobiol. 4th ed., Cold Spring Harbor Laboratory Press; 2022.
454. Tollefsen SE, Kornfeld R. *Isolation and characterization of lectins from Vicia villosa*. J Biol Chem 1983;**258**:5165–71.
455. Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. *Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family*. Glycobiology 2012;**22**:736–56.
456. Sun X, Zhan M, Sun X, Liu W, Meng X. *C1GALT1 in health and disease*. Oncol Lett 2021;**22**.
457. Wang Y, Ju T, Ding X, Xia B, Wang W, Xia L, et al. *Cosmc is an essential chaperone for correct protein O-glycosylation*. Proc Natl Acad Sci U S A 2010;**107**:9228–33.
458. Sanchez-Garrido J, Ruano-Gallego D, Choudhary JS, Frankel G. *The type III secretion system effector network hypothesis*. Trends Microbiol 2022;**30**:524–33.
459. Burns L, Le Mauff F, Gruenheid S. *Direct evidence of host-mediated glycosylation of NleA and its dependence on interaction with the COPII complex*. Gut Microbes 2024;**16**:2305477.
460. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. *Fiji: An open-source platform for biological-image analysis*. Nat Methods 2012;**9**:676–82.
461. Shental-Bechor D, Levy Y. *Effect of glycosylation on protein folding: A close look at thermodynamic stabilization*. Proc Natl Acad Sci U S A 2008;**105**:8256–61.
462. Kuan SF, Byrd JC, Basbaum C, Kim YS. *Inhibition of mucin glycosylation by aryl-N-acetyl-α-galactosaminides in human colon cancer cells*. J Biol Chem 1989;**264**:19271–7.
463. Tran DT, Ten Hagen KG. *Mucin-type O-glycosylation during development*. J Biol Chem 2013;**288**:6921–9.
464. Puri KD, Gopalakrishnan B, Surolia A. *Carbohydrate binding specificity of the Tn-antigen binding lectin from Vicia villosa seeds (VVLB4)*. FEBS Lett 1992;**312**:208–12.
465. Bechtella L, Chunsheng J, Fentker K, Ertürk GR, Safferthal M, Polewski Ł, et al. *Ion mobility-tandem mass spectrometry of mucin-type O-glycans*. Nat Commun 2024;**15**.
466. Grabarics M, Lettow M, Kirschbaum C, Greis K, Manz C, Pagel K. *Mass Spectrometry-Based Techniques to Elucidate the Sugar Code*. Chem Rev 2022;**122**:7840–908.
467. Bagdonaite I, Malaker SA, Polasky DA, Riley NM, Schjoldager K, Vakhrushev SY, et al. *Glycoproteomics*. Nat Rev Methods Prim 2022;**2**.
468. Vainauskas S, Guntz H, McLeod E, McClung C, Ruse C, Shi X, et al. *A Broad-Specificity O-Glycoprotease That Enables Improved Analysis of Glycoproteins and Glycopeptides Containing Intact Complex O-Glycans*. Anal Chem 2022;**94**:1060–9.
469. Kobata A. *Exo- and endoglycosidases revisited*. Proc Japan Acad Ser B Phys Biol Sci 2013;**89**:97–117.
470. Sim L, Thompson N, Geissner A, Withers SG, Wakarchuk WW. *Mammalian sialyltransferases allow efficient Escherichia coli-based production of mucin-type O-glycoproteins but can also transfer Kdo*. Glycobiology 2022;**32**:429–40.
471. Du T, Buenbrazo N, Kell L, Rahmani S, Sim L, Withers SG, et al. *A Bacterial Expression Platform for Production of Therapeutic Proteins Containing Human-like O-Linked Glycans*. Cell Chem Biol 2019;**26**:203-212.e5.
472. Mohl JE, Gerken TA, Leung MY. *ISOGlyP: De novo prediction of isoform-specific*

mucin-type O-glycosylation. Glycobiology 2021;**31**:168–72.
473. de las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. *Polypeptide GalNAc-Ts: from redundancy to specificity*. Curr Opin Struct Biol 2019;**56**:87–96.
474. May C, Ji S, Syed ZA, Revoredo L, Daniel EJP, Gerken TA, et al. *Differential splicing of the lectin domain of an O-glycosyltransferase modulates both peptide and glycopeptide preferences*. J Biol Chem 2020;**295**:12525–36.
475. de las Rivas M, Paul Daniel EJ, Narimatsu Y, Compañón I, Kato K, Hermosilla P, et al. *Molecular basis for fibroblast growth factor 23 O-glycosylation by GalNAc-T3*. Nat Chem Biol 2020;**16**:351–60.
476. Ballard CJ, Paserba MR, Paul Daniel EJ, Hurtado-Guerrero R, Gerken TA. *Polypeptide N-acetylgalactosaminyltransferase (GalNAc-T) isozyme surface charge governs charge substrate preferences to modulate mucin type O-glycosylation*. Glycobiology 2023;**33**:817–36.
477. Fritz TA, Raman J, Tabak LA. *Dynamic association between the catalytic and lectin domains of human UDP-GalNAc:polypeptide α-N-acetylgalactosaminyltransferase-2*. J Biol Chem 2006;**281**:8613–9.
478. O'Connell B, Tabak LA, Ramasubbu N. *The influence of flanking sequences on O-glycosylation*. Biochem Biophys Res Commun 1991;**180**:1024–30.
479. Wilson IBH, Gavel Y, Von Heijne G. *Amino acid distributions around O-linked glycosylation sites*. Biochem J 1991;**275**:529–34.
480. Gerken TA, Jamison O, Perrine CL, Collette JC, Moinova H, Ravi L, et al. *Emerging paradigms for the initiation of mucin-type protein O-glycosylation by the polypeptide GalNAc transferase family of glycosyltransferases*. J Biol Chem 2011;**286**:14493–507.
481. Paulson JC, Colley KJ. *Glycosyltransferases*. J Biol Chem 1989;**264**:17615–8.
482. Gill DJ, Chia J, Senewiratne J, Bard F. *Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes*. J Cell Biol 2010;**189**:843–58.
483. Nguyen AT, Chia J, Ros M, Hui KM, Saltel F, Bard F. *Organelle Specific O-Glycosylation Drives MMP14 Activation, Tumor Growth, and Metastasis*. Cancer Cell 2017;**32**:639-653.e6.
484. Gill DJ, Tham KM, Chia J, Wang SC, Steentoft C, Clausen H, et al. *Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness*. Proc Natl Acad Sci U S A 2013;**110**.
485. Tang VT, Ginsburg D. *Cargo selection in endoplasmic reticulum–to–Golgi transport and relevant diseases*. J Clin Invest 2023;**133**:1–10.
486. Miller E, Antonny B, Hamamoto S, Schekman R. *Cargo selection into COPII vesicles is driven by the Sec24p subunit*. EMBO J 2002;**21**:6105–13.
487. Scott NE, Hartland EL. *Post-translational Mechanisms of Host Subversion by Bacterial Effectors*. Trends Mol Med 2017;**23**:1088–102.
488. Fairman JW, Noinaj N, Buchanan SK. *The structural biology of β-barrel membrane proteins: A summary of recent reports*. Curr Opin Struct Biol 2011;**21**:523–31.
489. Roumia AF, Theodoropoulou MC, Tsirigos KD, Nielsen H, Bagos PG. *Landscape of Eukaryotic Transmembrane Beta Barrel Proteins*. J Proteome Res 2020;**19**:1209–21.
490. Kudelka MR, Stowell SR, Cummings RD, Neish AS. *Intestinal epithelial glycosylation in homeostasis and gut microbiota interactions in IBD*. Nat Rev Gastroenterol Hepatol 2020;**17**:597–617.
491. Tailford LE, Crost EH, Kavanaugh D, Juge N. *Mucin glycan foraging in the human gut*

microbiome. Front Genet 2015;**5**.
492. Robbe C, Capon C, Maes E, Rousset M, Zweibaum A, Zanetta JP, et al. *Evidence of regio-specific glycosylation in human intestinal mucins: Presence of an acidic gradient along the intestinal tract*. J Biol Chem 2003;**278**:46337–48.
493. Arike L, Holmén-Larsson J, Hansson GC. *Intestinal Muc2 mucin O-glycosylation is affected by microbiota and regulated by differential expression of glycosyltranferases*. Glycobiology 2017;**27**:318–28.
494. Larsson JMH, Thomsson KA, Rodríguez-Piñeiro AM, Karlsson H, Hansson GC. *Studies of mucus in mouse stomach , small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution*. Am J Physiol Gastrointest Liver Physiol 2013;**305**:357–63.
495. Brazil JC, Parkos CA. *Pathobiology of neutrophil – epithelial interactions* 2016;**273**:94–111.
496. Coelho C, Veríssimo A, Tiago I. *High-Quality Draft Genome Sequences of Crenobacter cavernae Strain CAVE-375 and Oxalobacteriaceae sp. Strain CAVE-383, Two Bacteria Isolated from Dripping Water in a Karstic Cave in Portugal*. Microbiol Resour Announc 2019;**8**.