



Review

Type III Secretion Effectors with Arginine N-Glycosyltransferase Activity

Juan Luis Araujo-Garrido, Joaquín Bernal-Bayard and Francisco Ramos-Morales *

Departamento de Genética, Facultad de Biología, Universidad de Sevilla, 41012 Sevilla, Spain; jaraujo@us.es (J.L.A.-G.) jbbayard@us.es (J.B.-B.)

* Correspondence: f Ramos@us.es

Received: 12 February 2020; Accepted: 29 February 2020; Published: 2 March 2020

Abstract: Type III secretion systems are used by many Gram-negative bacterial pathogens to inject proteins, known as effectors, into the cytosol of host cells. These virulence factors interfere with a diverse array of host signal transduction pathways and cellular processes. Many effectors have catalytic activities to promote post-translational modifications of host proteins. This review focuses on a family of effectors with glycosyltransferase activity that catalyze addition of *N*-acetyl-D-glucosamine to specific arginine residues in target proteins, leading to reduced NF- κ B pathway activation and impaired host cell death. This family includes NleB from *Citrobacter rodentium*, NleB1 and NleB2 from enteropathogenic and enterohemorrhagic *Escherichia coli*, and SseK1, SseK2, and SseK3 from *Salmonella enterica*. First, we place these effectors in the general framework of the glycosyltransferase superfamily and in the particular context of the role of glycosylation in bacterial pathogenesis. Then, we provide detailed information about currently known members of this family, their role in virulence, and their targets.

Keywords: glycosyltransferases; type III secretion; effectors; NleB; SseK; *Citrobacter*; *Escherichia*; *Salmonella*; death domains

1. Introduction

Type III secretion systems (T3SSs) are fundamental tools for the interaction between many pathogenic and symbiotic Gram-negative bacteria and their hosts. T3SSs, also known as injectisomes, span the bacterial inner and outer membranes and form injection nanomachines that deliver proteins, known as effectors, into target eukaryotic cells to modulate a variety of cellular functions [1,2]. Effectors have diverse activities and targets, and typically have an N-terminal secretion domain and one or more C-terminal functional domains [3,4]. Effectors can be classified in four non-mutually exclusive categories according to their mechanisms of activity [5]: (i) effectors that mediate their activities through physical interactions with host targets; (ii) effectors that use functional or structural mimicry to alter host cell processes; (iii) effectors that promote post-translational modifications of host proteins; and (iv) effectors with protease activity. Examples of catalytic activities of effectors are ubiquitination, deubiquitination, phosphorylation, dephosphorylation, acetylation, deamidation, ADP ribosylation, cysteine methylation, and glycosylation. This review focuses on a specific family of T3SS effectors with glycosyltransferase activity that are found in *Salmonella enterica* and in attaching/effacing (A/E) pathogens.

Salmonellae are facultative intracellular Gram-negative bacterial pathogens belonging to the family Enterobacteriaceae. The genus *Salmonella* includes the species *S. bongori* and *S. enterica*, the latter divided into six subspecies and more than 2500 serovars [6]. *S. enterica* possesses two virulence-related T3SS [7–9], T3SS1 and T3SS2, that are encoded by *Salmonella* pathogenicity island 1 (SPI1) and *Salmonella* pathogenicity island 2 (SPI2) (Figure 1), respectively, whereas *S. bongori* lacks SPI2 [10].

The T3SS1 translocates effectors across the host cell plasma membrane that promote *Salmonella* invasion into epithelial cells in the intestine and stimulate localized inflammation [11]. The T3SS2 is expressed intracellularly in response to acidic pH and nutrient limitation found in the lumen of the *Salmonella*-containing vacuole (SCV), a modified phagolysosome which is the typical intracellular niche of this pathogen [12]. T3SS2 effectors are translocated through the SCV membrane, are important in intestinal and disseminated infections, and are required for growth within different host cell types [13]. Together, T3SS1 and T3SS2 secrete more than 40 effectors, some encoded by SPI1 or SPI2, but most encoded outside these islands, in other horizontally acquired regions of the genome [14–17].

Enterohemorrhagic *Escherichia coli* (EHEC) and enteropathogenic *E. coli* (EPEC) are extracellular human pathogens that attach to the intestinal epithelium and efface brush border microvilli forming A/E lesions [18]. EHEC causes hemorrhagic colitis and hemolytic uremic syndrome, whereas EPEC is a cause of infantile diarrhea. *Citrobacter rodentium* is another A/E pathogen which is the causative agent of murine transmissible colonic hyperplasia [19]. Lesions induced by *C. rodentium* in mice are similar to those caused by EHEC or EPEC in humans and these pathogens share a number of virulence determinants including the locus of enterocyte effacement (LEE) [20,21] (Figure 1). The LEE is a 34 kb pathogenicity island with 41 genes that encode all the structural components of a T3SS, seven of its effectors and their chaperones, proteins involved in adherence, and regulatory proteins [22]. Many effectors are encoded outside the LEE, and the total number of effectors is at least 29 in *C. rodentium*, 22 in EPEC, and 39 in EHEC [22]. A second gene cluster, potentially encoding a T3SS and known as ETT2, for *E. coli* T3SS 2, is present in many *E. coli* strains [23,24]. ETT2 resembles the *Salmonella* T3SS1 but in most cases it has been subjected to mutational attrition and there is no evidence for the secretion of effectors by this system [25]. However, this region is intact in some strains of *E. coli* and *Escherichia albertii*, and appears to be important for bacterial pathogenesis [25–27].

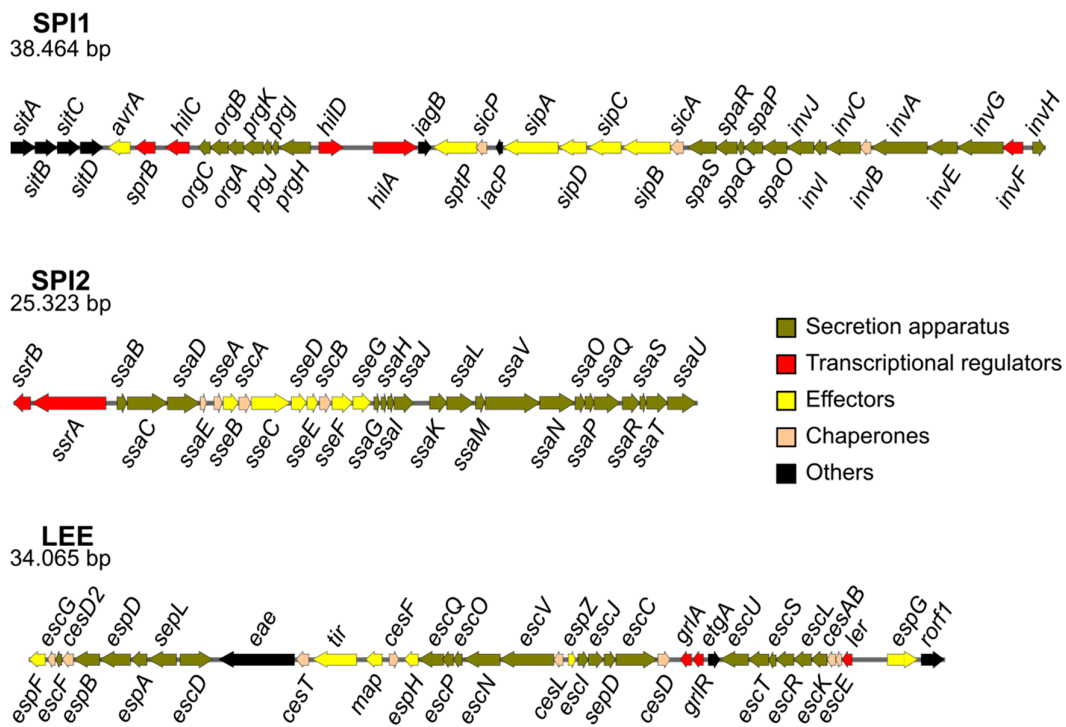


Figure 1. Genetic structure of *Salmonella* pathogenicity island 1 (SPI1) and *Salmonella* pathogenicity island 2 (SPI2) of *S. enterica* serovar Typhimurium strain 14028, and the locus of enterocyte effacement (LEE) island of *E. coli* O127:H6 strain E2348/69. Genes are colored according to their functional categories [22,28].

2. Glycosyltransferases

Glycosyltransferases (GTs) catalyze the formation of glycoside bonds using an activated donor sugar substrate that contains a nucleoside phosphate or a lipid phosphate leaving group [29]. Enzymes that use sugar mono- or diphosphonucleotides as donors are known as Leloir GTs [30]. Non-Leloir GTs use non-nucleotide donors, such as polyprenol pyrophosphates, polyprenol phosphates, sugar-1-phosphates, or sugar-1-pyrophosphates. Sugars can be transferred to a variety of biomolecules including glycans, lipids, proteins, peptides, and small molecules. The resulting glycosidic linkage can be an O-, S-, N- or C- covalent bond connecting a monosaccharide to another residue.

2.1. Classification

Traditionally, GTs were classified on the basis of their donor, acceptor, and product specificity, according to the recommendations of the International Union of Biochemistry and Molecular Biology (IUBMB), under the Enzyme Commission number EC.2.4. Given the limitations of this system, a new scheme that involves the classification of GTs into families on the basis of amino acid sequence similarities is now accepted [31,32]. Today, the carbohydrate active enzymes (CAZy) database (<http://www.cazy.org>) includes over 620,000 GT modules classified in more than 100 families and more than 13,000 nonclassified modules.

Depending on their mechanism of transfer, GTs can be classified as inverting or retaining, resulting in inversion ($\alpha \rightarrow \beta$) or retention ($\alpha \rightarrow \alpha$) of the stereochemistry of the donor's anomeric bond during the transfer, respectively [33]. Theoretical and experimental studies support a single-displacement S_N2 (substitution, nucleophilic, bimolecular)-like mechanism for inverting GTs [29]. This mechanism involves the formation of an oxocarbenium ion transition state (Figure 2). In contrast, the reaction mechanism of retaining glycosylation is controversial [34]. A double-displacement mechanism, in which the sugar is first transferred to the enzyme and then transferred to the acceptor, was suggested by analogy to retaining glycosidases [35]. Some experimental support was provided for this mechanism [36,37], particularly for the GT6 family members that have a carboxylate properly oriented to act as a catalytic nucleophile. But this is not the case for most retaining GTs, leading to the proposal of an alternative internal return S_Ni (substitution, nucleophilic, internal)-like mechanism, also termed front-face mechanism, in which the nucleophilic hydroxy group of the acceptor attacks the anomeric carbon atom from the same side from which the leaving group departs [38,39]. An orthogonal mechanism has also been proposed that differs from the S_Ni mechanism in the reaction profile and the time of bond formation and breakage [40] (Figure 2).

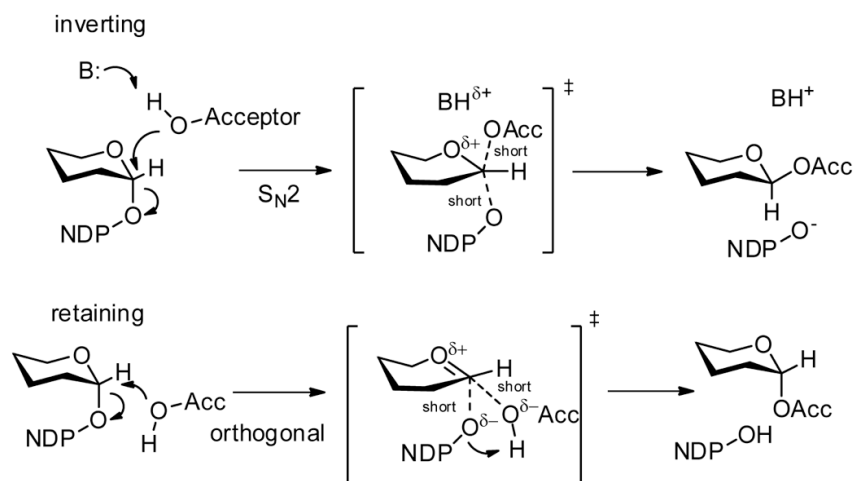


Figure 2. Glycosyltransferase mechanisms. An S_N2 (substitution, nucleophilic, bimolecular) process is accepted for inverting glycosyltransferases (GTs). Several mechanisms have been proposed for retaining enzymes. The orthogonal mechanism is depicted here. ‡: Transition state. Image by Dr.

Brock Schumann (Wikimedia Commons, CC BY-SA 3.0, <https://creativecommons.org/licenses/by-sa/3.0/deed.en>).

2.2. Structural Studies

GT structures (Figure 3) revealed that the majority of these enzymes fall into two general folds, called GT-A and GT-B, each of which includes members that are retaining and inverting GTs. This led to a classification of GTs in four clans: clan I for inverting enzymes with the GT-A fold, clan II for inverting enzymes with the GT-B fold, clan III for retaining enzymes with the GT-A fold, and clan IV for retaining enzymes with the GT-B fold [32]. Both fold types contain a Rossmann fold [41], a nucleotide-binding domain responsible for donor nucleotide recognition. A second domain is usually responsible for acceptor recognition and is more variable. Enzymes with the GT-B type fold contain two Rossmann-like folds separated by a deep wide crevice [42]. However, not all GTs fit into these clans. A third fold type, called GT-C fold, was predicted using computational methods [43–45], and the structure of several GTs with this fold, most of them belonging to family GT66, have been described [46]. These GTs usually contain an N-terminal transmembrane domain and a C-terminal globular domain with GT activity. In addition, the structure of the enzyme DUF1792 revealed a novel fold designated as GT-D type [47] and defined family GT101. Members of the GT51 family also have a distinct structure defined as lysozyme-type [48,49], and the recently published structure of TagA, a member of the GT26 family, suggests the definition of a novel fold termed GT-E [50]. A common feature found in many members of the GT-A and GT-C superfamilies is a DxD motif involved in binding to divalent metal ion and catalysis. However, the occurrence of this motif is not enough to predict a potential GT function since many GTs do not possess it and the DxD motif is found in more than a half of all described proteins [32]. Finally, given the modular nature of carbohydrate-active enzymes, a number of GTs contain noncatalytic domains potentially involved in substrate recognition or protein–protein interaction, and some polypeptides contain two GT modules belonging to the same or different families.

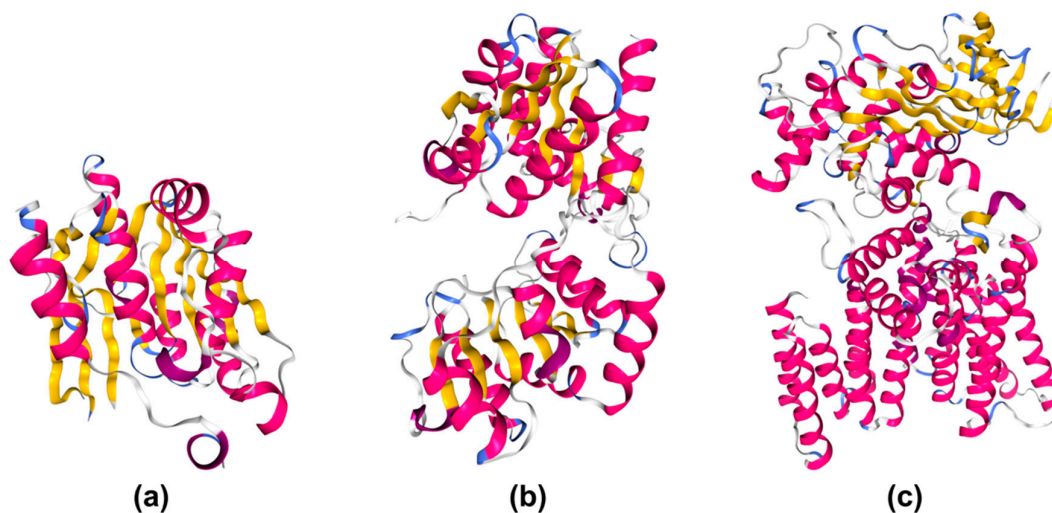


Figure 3. Glycosyltransferase folds. (a) Diagram of the GT-A fold protein SpsA from *Bacillus subtilis*, belonging to family GT2 (PDB ID: 1QG8) [51]; (b) diagram of the GT-B fold-type protein GtfB from *Amycolaptosis orientalis*, belonging to family GT1 (PDB ID: 1IIR) [52]; (c) diagram of the GT-C fold protein PglB from *Campylobacter lari*, belonging to family GT66 (PDB ID: 3RCE) [46]. Drawings created with NGL [53].

Table 1 shows a classification of GT families based on reaction mechanisms and fold types.

Table 1. Classification of glycosyltransferase families.

	Reaction Mechanism		
	Inverting	Retaining	Unknown
<i>GT-A</i>	2 ¹ , 7, 12, 13, 14, 16, 21, 29, 31, 42, 43, 82, 84, 109	6, 8, 15, 24, 27, 34, 44, 55, 62, 64, 78, 81, 88	
<i>GT-B</i>	1, 9, 10, 18, 19, 23, 28, 30, 33, 38, 41, 52, 63, 65, 68, 70, 80, 90, 104	3, 4, 5, 20, 35, 52, 72, 99, 107	
Fold type			
<i>GT-C</i>	22, 39, 50, 53, 57, 58, 59, 66, 83, 85, 87		
<i>Others</i>	26 (GT-E), 51 (lysozyme-type),		101 (GT-D)
<i>Unknown</i>	11, 17, 25, 37, 40, 47, 48, 49, 54, 56, 61, 67, 73, 74, 75, 76, 92, 94, 97, 98, 100, 102, 103, 105, 106, 108	32, 45, 60, 69, 71, 77, 79, 89, 93, 95, 96	91, 110

¹ Numbers correspond to the GT family numbers assigned in the carbohydrate active enzymes (CAZy) database.

3. Glycosylation in Bacterial Pathogenesis

As mentioned in the introduction, effectors from pathogenic bacteria often exhibit specific enzymatic activity against host cell targets as part of its infection strategy. Protein glycosylation, the most common post-translational modification, has been traditionally thought to be restricted to eukaryotic organisms. However, in the last years, many examples of bacterial virulence factors and toxins exhibiting GT activity have been characterized [54]. Bacterial glycosylation is broadly linked to bacterial pathogenesis [55]. It fulfills two main functions during infection: firstly, protein glycosylation enables bacterial adhesion to host cells and, secondly, leads to modification of specific host targets with the consequent manipulation of host cellular processes by the pathogen to their own benefit.

Bacterial GTs perform two types of protein modifications: N-linked glycosylation or O-linked glycosylation [29]. Whereas N-GTs conjugate glycans to side chain amide nitrogen of an asparagine in an Asn-X-Ser/Thr consensus sequence, O-glycosylation happens on the hydroxyl group of a serine or threonine residue. Both N-linked and O-linked glycosylation pathways may modify multiple proteins. Interestingly, there are examples of pathogenic bacteria, like *C. rodentium*, performing N- and O-glycosylation that are both critically required for colonization of the gastrointestinal tract in infected mice [54].

In this section we provide representative examples of well characterized bacterial GTs, attending to their specific biological role during infection and the type of modification they perform.

3.1. Glycosylation of Adhesins

One essential step to initiate infection is bacterial adhesion to the host epithelium. This first step depends sometimes on glycosylation of bacterial adhesins. Some examples are described below.

- A well-known example is the high-molecular-weight (HMW) system of *Haemophilus influenzae*. HMW1/HMW2 are two adhesins that are glycosylated at the cytoplasm by GTs HMW1C and HMW2C, respectively [56,57]. Then, they are transported to the bacterial surface and mediate adherence to respiratory epithelial cells [57]. Glucose and galactose are added to specific asparagine residues in the eukaryotic-like Asn-X-Ser/Thr consensus [58].
- Pgl (from Protein glycosylation) system in *Campylobacter jejuni* is another example of a glycosylated adhesin that has been considered a prototype for the relevance of bacterial

glycosylation in bacterial–host interaction [55]. It is an N-linked protein glycosylation system composed of several proteins, including various GTs that modify *C. jejuni* proteins leading to bacterial–host cell adhesion and invasion of host epithelium [59]. Disruption of the Pgl pathway results in reduction of human epithelial cells adhesion and decreases the ability to colonize mouse and chicken in vivo [59–61]. Modification catalyzed by the Pgl system consists of the addition of a heptasaccharide to asparagine residues within the eukaryotic-like Asn-X-Ser/Thr glycosylation consensus. The oligosaccharide is assembled on the cytoplasmic side of the inner membrane. Then, it is translocated by an ABC-transporter (PglK) [62]. Once in the periplasmic space, PglB, an oligo-saccharyltransferase, transfers the heptasaccharide to asparagine residues of target ligands [63,64].

- In addition to N-glycosylation-related mechanisms to improve bacterial adhesion, other systems based on O-glycosylation have been characterized. One example is the emerging bacterial autotransporter heptosyltransferase (BAHT) family found in several Gram-negative bacteria [65]. Autotransporters which belong to the T5SS family have a modular composition with an N-terminal passenger domain and a C-terminal β -barrel domain that drives the passenger across the outer membrane. The passenger domain may be heptosylated by a BAHT member on numerous serine residues. This modification occurs before translocation of the protein across the inner membrane and is required for bacterial adhesion to the host. Two well-characterized BAHT member are AAH (autotransporter adhesin heptosyltransferase), from diffusely adhering *E. coli* isolate 2787, and TibC from enterotoxigenic *E. coli* strain H10407 [65,66]. In addition to pathogenic *E. coli*, BAHT members have been identified in other bacteria, including *Citrobacter*, *Shigella*, *Salmonella*, and *Burkholderia* [65]. Further studies are necessary to fully characterize these interesting bacterial enzymes.
- Post-translational modification of bacterial proteins to enhance bacterial adhesion to the host is not restricted to Gram-negative bacteria. A well-known example of modified bacterial proteins in Gram-positive bacteria are the O-glycosylated serine-rich repeat proteins (SRRPs) [67]. These large glycoproteins, unique in Gram-positive bacteria, are crucial for biofilm formation and, more importantly, host cell adhesion [68]. The best-characterized SRRP is Fap1 from *Streptococcus parasanguinis*. Glycosylation of Fap1 relies on the activity of at least six different GTs, although the specific function and glycan transferred remains unclear [69]. This family of glycoproteins is highly conserved among streptococcal and staphylococcal species [67], thus, characterization of SRRPs GTs and understanding of their biological role are key to identify new therapeutic drug targets in order to tackle relevant Gram-positive pathogens.

3.2. Glycosylation of Flagellins

Glycosylation of flagellin is an important component of numerous flagellar systems in Archaea and Bacteria, playing an important role in assembly and virulence [70], since this organelle, in addition to conferring motility, can play a role in every step of the infection cycle [71]. Flagella O-glycosylation has been demonstrated in many polar flagellins from Gram-negative [72] and Gram-positive bacteria [73–80], and also in a limited number of lateral flagellins [72,81,82]. Bacteria can glycosylate flagellin at a varying number of acceptor sites, from a single site in *Burkholderia* spp. or *Listeria monocytogenes* to multiple sites in *C. jejuni* or *Selenomonas sputigena* [80,83–85]. Moreover, whereas some bacteria use a single monosaccharide, others show strain-to-strain glycan heterogeneity [72]. Selected examples are described below.

- In *Aeromonas caviae* Sch3N, the polar flagellins, FlaA and FlaB, are glycosylated with pseudoaminic acid glycans and this glycosylation is required for flagellar assembly. Genes involved in glycosylation are mapped in the O-antigen biosynthetic cluster [86,87]. In *Aeromonas hydrophila* AH-3, lateral flagellin is modified with a derivative of pseudaminic acid, whereas polar flagella are glycosylated with a heterogeneous glycan [88].
- *C. jejuni* FlaA is glycosylated at up to 19 sites that contribute to the flagellum serospecificity. FlaB is also glycosylated, but the number of sites remains to be established. The predominant O-glycans are derivatives of pseudoaminic acid and legionaminic acid. The genes encoding GTs

involved in flagellin glycosylation are located adjacent to the flagellin structural genes, *flaA* and *flaB* [85,89,90].

- In *Helicobacter pylori*, both the FlaA and FlaB structural proteins are glycosylated with pseudaminic acid at seven and ten sites, respectively. Biosynthesis of pseudaminic acid has been extensively studied in *H. pylori* and *C. jejuni* and involves six consecutive enzymatic steps before being transferred onto flagellins via O-linked glycosylation [91,92].
- The GT transferring pseudaminic acid onto flagellin has been studied in *Magnetospirillum magneticum* AMB-1 [93]. This GT, known as Maf, displays significant sequence identity with proteins from other Gram-negative bacteria, including Maf1 from *A. caviae* and PseE from *C. jejuni*, that have been shown to transfer pseudaminic acid onto flagellins in these organisms [94,95]. The structure of *M. magneticum* Maf displays a modified GT-A topology as found in CAZy families GT29 and GT42. In fact, the closest structural homologue is *C. jejuni* sialyltransferase Cst-II (PDB 1RO7), belonging to family GT42 [93].
- Glycosylation of *Clostridium difficile* flagellin is essential for motility. Two types of modification have been defined as type A and B. Type A modification involves an O-linked N-acetylglucosamine (GlcNAc) linked to a methylated threonine via a phosphodiester bond. Three GTs are involved in type B modification: GT1 and GT2 are responsible for the sequential addition of a GlcNAc and two rhamnoses, respectively, and GT3 is associated with the incorporation of a sulfonated peptidyl-amido sugar moiety [74].

3.3. Modification of Host Cell Targets by Bacterial GTs

In addition to bacterial GTs that glycosylate their own proteins, we find another interesting group of bacterial proteins that modify host cell proteins, manipulating the cell physiology to counteract the host defense response. The large clostridial toxins TcdA/TcdB, the *Legionella* cytotoxic glycosyltransferase family (Lgt), and N-GTs from the NleB/SseK family are included in this group [54].

- Large clostridial glycosylating toxins: pathogenesis of *Clostridium* depends on the activity of two similar toxins, TcdA and TcdB, that exhibit GT activity in their N-terminal domains [96,97]. Both toxins perform O-glycosylation of small GTPases of the Rho family, inhibiting the regulatory functions of these proteins [98]. This modification occurs on the side chain hydroxyl group of a threonine residue, which is specific for each targeted Rho GTPase [99]. Another example of this family of toxins recently identified is PaTox from entomopathogenic *Photographus asymbiotica* [100].
- *Legionella pneumophila* Lgt family: *L. pneumophila* is the causative agent of legionellosis, a severe lung disease. This facultative intracellular pathogen uses a T4SS to inject more than 300 effectors into the host cell [101]. Among them, *L. pneumophila* secretes several GTs (Lgt1-3) that perform specific glycosylation of a serine residue (Ser⁵³) in the eukaryotic elongation factor 1A, a large GTPase and a component of the elongation complex in protein synthesis. This modification results in inhibition of protein synthesis and subsequently causes the death of infected cells [102,103]. Interestingly, Lgt1 share high structural homology with the clostridial toxin TcdB. Both GTs are members of the clan III of retaining enzymes with the GT-A fold [103–105].
- NleB/SseK family: we refer to this family of T3SS effectors in the next section.

4. Glycosyltransferases of the NleB Family

4.1. Members of the Family

T3SS effectors that catalyze the transfer of GlcNAc to arginines of host proteins have been studied in *C. rodentium*, EHEC, EPEC, and *S. enterica* (Table 2).

Table 2. NleB/SseK glycosyltransferases.

Organism.	Reference Strain	Effector Name	Protein Size (aa)	Genomic Context	References
<i>C. rodentium</i>	ICC168	NleB	329	Genomic island GI4	[106,107]
Enterohemorrhagic <i>E. coli</i> (EHEC) O157:H7	EDL933	NleB1	329	O-island 122/SpLE3	[106,108]
		NleB2	326	O-island 36/Phage CP-933K/Sp3	
Enteropathogenic <i>E. coli</i> (EPEC) O127:H6	E2348/69	NleB1	329	Integrative element IE6	[109,110]
		NleB2	326	Phage PP4	
<i>S. enterica</i> serovar Typhimurium	14028	SseK1	336		[111–113]
		SseK2	348		
		SseK3	335	Phage ST64B	

NleB was identified in *C. rodentium* together with other potential type III effectors encoded outside the LEE. They were designated NleA, NleB, NleC, NleD, NleE, NleF, and NleG for non-LEE-encoded effectors [106]. Two homologs of NleB are encoded in the genomes of EHEC and EPEC: NleB1 and NleB2 [106,108,109]. NleB1 is 89% identical to NleB from *C. rodentium*, and there is a difference in seven amino acids between the EHEC and EPEC versions. NleB2 is identical in both *E. coli* pathotypes but is only 60% identical to *C. rodentium* NleB, although an ortholog of *nleB2* is present in the genome of *C. rodentium* as a pseudogene [107]. Interestingly, *nleB2* is a non-neutral evolving gene, probably due to positive selection, suggesting that this gene is still in the process of optimizing its function [114]. *S. enterica* strains encode three homologs of NleB: SseK1, SseK2, and SseK3 [111,112]. Although the genes encoding these effectors are not located in the pathogenicity islands where the cognate T3SS are encoded, all of them have evidence of horizontal transfer and some are associated with other pathogenicity islands or prophages (Table 2). Interestingly, whereas *sseK1* and *sseK2* are present in most available *Salmonella* genome sequences, the phage-encoded SseK3 effector has a limited distribution [111]. A BLAST search on complete genomes available at the National Center for Biotechnology Information (NCBI) reveals that proteins identical or very similar to NleB1 or NleB2 are encoded in different strains of the emerging A/E pathogen *Escherichia albertii* [115], and significantly similar proteins are found in *Escherichia marmotae* [116] and *Yersinia hibernica* [117].

Percentages of identity between representative NleB/SseK proteins are shown in Table 3, and a phylogenetic analysis is shown in Figure 4.

Table 3. Percent identity data for NleB/SseK proteins provided by ClustalW.

	NleB <i>C. rodentium</i>	NleB1 EPEC	NleB1 EHEC	NleB <i>E. albertii</i>	NleB <i>E. marmotae</i>	NleB2	SseK1	SseK2	SseK3
NleB <i>C. rodentium</i>									
NleB1 EPEC	88.75								
NleB1 EHEC	89.06	97.87							
NleB1 <i>E. albertii</i>	88.45	97.57	99.09						
NleB <i>E. marmotae</i>	71.42	71.73	71.73	71.43					
NleB2 EPEC/EHEC	59.82	61.04	61.35	60.43	60.43				
SseK1 <i>S. enterica</i>	57.45	58.05	59.57	58.97	58.97	53.68			
SseK2 <i>S. enterica</i>	52.58	52.28	53.19	52.58	51.06	46.93	55.36		

SseK3 <i>S. enterica</i>	51.98	53.19	53.50	52.89	54.10	47.55	54.93	75.22	
NleB <i>Y. hibernica</i>	32.52	32.83	34.04	34.04	34.04	33.13	35.12	33.14	33.43

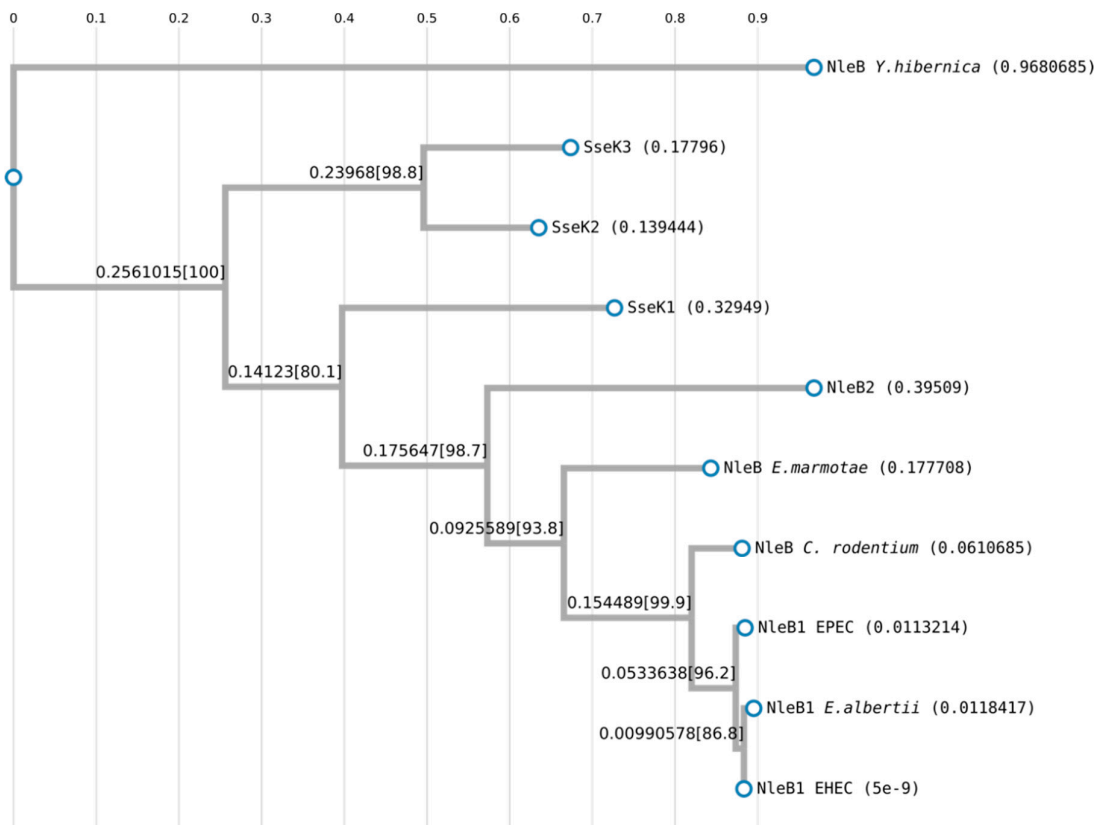


Figure 4. Phylogenetic analysis of NleB/SseK proteins. Alignment and phylogenetic reconstructions were performed using the function “build” of ETE3 v3.1.1 [118] as implemented on the GenomeNet (<https://www.genome.jp/tools/ete/>). The tree was constructed using FastTree v2.1.8 with default parameters [119]. Values at nodes are SH-like local support.

4.2. Expression, Translocation, and Subcellular Localization

Effectors of T3SSs are expected to be expressed in coordination with the synthesis of their cognate secretion apparatus. This is the case for effectors of the LEE, SPI1 or SPI2 systems that are encoded in the islands, but the issue is not so obvious for effectors encoded outside the islands, including the members of the NleB/SseK family.

In A/E organisms, Ler is the main transcriptional regulator of LEE. It is encoded in the island and its level is controlled by other transcriptional regulators that respond to different environmental stimuli, including host hormones, microbiota-liberated sugars and metabolites, molecular oxygen at the gut epithelium, and contact of bacteria with host epithelial cells [120]. Expression is undetectable when bacteria are grown in rich laboratory media, but rapid expression and secretion of proteins to the supernatant is observed upon transfer to MEM-HEPES medium, which is associated with the heterogeneous expression of the EspADB translocon apparatus [121]. Several *nle* genes, including *nleB*, are transcribed in EHEC under these secretion-permissive conditions, although induction of *nleB* expression was slight (twofold) and was not coordinated with the synthesis of EspA filaments at the single-cell level [122]. Secretion of NleB (NleB1 in EPEC) by the LEE-encoded T3SS has been demonstrated in *C. rodentium* and EPEC [106,110]. Transfection with GFP-fusions or translocation of TEM-1 β -lactamase fusions [123] from EPEC into HeLa cells reveals a diffuse cytosolic localization of NleB1 [110,124]. The regulatory region of EPEC *nleB2* contains a motif conserved among a subset of genes encoding effector proteins in A/E pathogens that was named NRIR (*nle* regulatory inverted

repeat). In contrast to other *nle* genes, like *nleA*, expression of *nleB2* is not regulated by Ler. However, it is repressed by another LEE-encoded regulator, GlrR, which is also a transcriptional repressor of the LEE, and by the global regulator H-NS [125]. An interesting observation is that similar to the LEE genes and *nleA*, *nleB2* is induced highly in DMEM (Dulbecco's modified Eagle's medium) cultures and repressed in LB (lysogeny broth) medium, suggesting that in spite of the different regulatory elements, expression of these genes is coregulated by environmental signals to ensure their coexpression with the T3SS [125]. Secretion of NleB2 was demonstrated by analyzing the secretome of EPEC using stable isotope labeling with amino acids in cell culture [126].

In *Salmonella*, many different environmental stimuli and regulatory proteins affect SPI1 expression. HilA is a direct transcriptional activator, and the *hilA* gene is transcriptionally activated by HilD, HilC, and RtsA, three AraC-like regulatory proteins that create a complex feed-forward regulatory loop [127] that function partially by relieving repression mediated by H-NS [128]. The SPI1-encoded T3SS is deactivated within host cell vacuoles after invasion of epithelial cells or uptake into macrophages [129]. Two systems involved in this repression are the Rcs phosphorelay system [130] and the PhoQ/PhoP two-component system [131,132], which is critical for adaptation of *Salmonella* to the intravacuolar environment. The PhoQ/PhoP system is activated in the presence of low levels of divalent cations and low pH, that are found inside the SCV [133,134]. Interestingly, this regulatory system positively controls the expression of T3SS2 through transcriptional and post-transcriptional activation of the SsrA/SsrB two-component system, the main positive regulator of SPI2 [135]. Optimal expression of T3SS1 and T3SS2 can be achieved in vitro using appropriate media and culture conditions: LB medium with high osmolarity and microaerophilic conditions for the T3SS1, and LPM minimal medium with low pH, low Mg⁺⁺ concentration, and aeration conditions for the T3SS2. Significant amounts of SseK1 are produced under both sets of conditions [112], although transcription is higher under SPI2-inducing conditions [136]. This may be a consequence of the fact that transcription of the gene *sseK1* is not regulated by the main regulators of SPI1 or SPI2 (HilA, HilD, and SsrB), although it is directly activated by PhoP in a SsrB-independent manner. This pattern of expression is consistent with SseK1 being translocated through T3SS1 and T3SS2, with different patterns and kinetics depending on the specific host cell type (epithelial, macrophages, or fibroblasts) [136]. In contrast to SseK1, SseK2 and SseK3 were detected specifically under SPI2-inducing conditions and their expression was dependent on the SsrA/SsrB system, suggesting coordination with the expression of the T3SS2 [111,112]. T3SS2-dependent translocation of SseK2 into HeLa cells was shown using a CyaA fusion [112], whereas T3SS2-dependent secretion into the culture supernatant and translocation into HeLa cells of SseK3 was shown using an HA-tagged form of this effector [111]. Transfection of HeLa cells with a plasmid expressing a SseK1-GFP fusion indicated that it localized uniformly throughout the cytoplasm rather than associate with membrane components or other discrete structures within the host cell [112]. The same localization was found upon translocation of SseK1-HA into RAW264.7 macrophages [137]. In contrast, SseK2-HA and SseK3-HA were associated to the Golgi network when translocated into macrophages [137], a localization that was also observed after ectopic expression of GFP-SseK3 [138].

4.3. Enzymatic Activity and Structural Studies

Before the characterization of the activity of NleB/SseK effectors, *N*-glycosylation of arginine residues in proteins was rarely reported [139], and most studies on protein glycosylation had focused on the *O*-glycosylation of serine, threonine, and tyrosine residues or *N*-glycosylation of asparagine residues [55]. Therefore, the discovery that EPEC NleB1 has arginine GlcNAc transferase activity toward several proteins was surprising [140,141]. GlcNAcylation activity was previously reported for NleB from *C. rodentium*, although the specific target residues were not characterized [142]. More recently, arginine *N*-GlcNAcylation activity has been shown for *C. rodentium* NleB, EHEC NleB1, and *S. enterica* SseK1, whereas activity of NleB2 has not been detected and conflicting results exist about SseK2 and SseK3 [137,143].

The structures of several NleB/SseK family members were recently reported, indicating that these proteins possess the typical GT-A family fold and the metal-coordinating DxD motif essential

for ligand binding and enzymatic activity [144–147]. However, some controversy exists about the mechanism of transfer used by these glycosyltransferases. A study whose objective was the synthesis of an antibody that can specifically recognize arginine *N*-GlcNAcylated peptides suggested a β -anomeric configuration for the products of the reaction [148], implying that these enzymes are inverting GTs. However, the crystal structure of SseK3, which was reported in its free form and bound to UDP, GlcNAc, and manganese, suggested that it was a retaining arginine GT [144]. This conclusion was confirmed by another study that investigated, by NMR spectroscopy, the glycosidic bond configuration of a GlcNAc-GAPDH₁₈₇₋₂₀₃ glycopeptide formed enzymatically by incubation with SseK1 and UDP-GlcNAc/MnCl₂ [145,149]. A more recent study determined the crystal structure of EPEC NleB1 alone, NleB1 in complex with the death domain of FAS-associated death domain protein (FADD, one of its substrates, see below), and the death domain of TRADD (TNFR1-associated death domain protein) and RIPK1 (receptor-interacting serine/threonine-protein kinase 1) GlcNAcylated by NleB1 [146]. This study concluded that NleB1 is an inverting enzyme that converts the α -configuration in the UDP-GlcNAc donor to the β configuration in the arginine-GlcNAc product. The same study also cast doubts on the validity of previous conclusions since the study on SseK3 did not involve the analysis of the modified substrate and the SseK1 study forced the modification of a GAPDH peptide that may not be a physiological substrate. Finally, another study reporting the structure of SseK1 and SseK3 indicated that the closest structural homologs are enzymes from families GT44 and GT88, both containing retaining GTs (Table 1), but the configurations of the modified substrates were not studied [147]. Structures of SseK3 and NleB1 are shown in Figure 5.

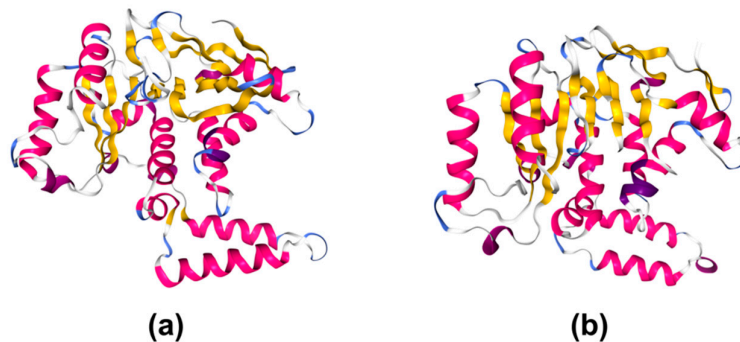


Figure 5. Structure of NleB/SseK family members. (a) Crystal structure of SseK3 from *S. enterica* serovar Typhimurium strain SL1344 (PDB ID: 6CGI) [147]; (b) crystal structure of NleB1 from *E. coli* O127:H6 (PDB ID: 6E66) [146]. Drawings created with NGL [53].

4.4. Role in Virulence

The role in virulence of particular T3SS effectors is often difficult to demonstrate due partially to the fact that many other effectors and virulence factors are involved, sometimes with opposite or potentially redundant effects. Apparent redundancy can also be a consequence of limitations of laboratory models of pathogenesis [150]. In the case of NleB/SseK effectors, involvement in virulence has been shown for some members of the family.

A signature-tagged mutagenesis screen in C57BL/6 mice identified NleB as an important virulence factor of *C. rodentium*: deletion of *nleB* resulted in reduced colonization of mice, with a competitive index (CI) in the colon of 0.069, and reduced colonic hyperplasia [110]. The GT activity of NleB is very relevant for this phenotype, since a mutant with an altered DxD catalytic site, NleB₂₂₁DAD-AAA₂₂₃, failed to complement the attenuation of the $\Delta nleB$ mutant. However, the fact that the catalytic mutant was able to provide partial complementation suggests that NleB may have functions that are independent of its GT activity [151]. Additional complementation experiments showed that EHEC NleB1 was able to restore the virulence of the *C. rodentium* $\Delta nleB$ mutant, whereas NleB2 was unable to complement the colonization defect. This study also identified a structural motif in NleB1, His₂₄₂-Glu₂₅₃-Asn₂₅₄, which was essential for activity and virulence [145]. The Glu₂₅₃ residue, together with Tyr₂₁₉ were identified in a previous study as essential for enzymatic activity and

virulence function of NleB1 [152]. Interestingly, the presence of *nleB1* in EHEC and EPEC strains is associated with colonization of sheep [153] and cattle [154], with human disease [155–159], and with an increase in disease severity [160–165]. In contrast, the role of NleB2 in virulence is currently unclear.

The first study describing *S. enterica* serovar Typhimurium effectors SseK1 and SseK2 failed to detect signs of attenuation for $\Delta sseK1$, $\Delta sseK2$, or $\Delta sseK1 \Delta sseK2$ mutants during infection of tissue culture cells or susceptible (Nramp1^s) BALB/c mice using a time-to-death assay after intraperitoneal infections [112]. Another study showed that the $\Delta sseK1 \Delta sseK2 \Delta sseK3$ triple mutant had significantly reduced growth levels in RAW264.7 macrophages but confirmed the lack of attenuation in mice [166]. A more sensitive assay using mixed infections of mice identified a mild but statistically significant attenuation for the *sseK1 sseK2* double mutant after oral infections (CI = 0.4–0.5 compared with the wild type). No difference was detected between the *sseK1 sseK2 sseK3* triple mutant and the *sseK1 sseK2* double mutant, suggesting that SseK3 has no role in virulence in this system [111]. Mild attenuation was also reported for an *sseK1* single mutant of *S. enterica* serovar Typhimurium after intraperitoneal and oral infections of BALB/c mice using the CI analysis [136]. In addition, an *sseK1* deletion mutant of *S. enterica* serovar Enteritidis showed decreased formation of biofilm, reduced intracellular survival in activated mouse peritoneal macrophages, and reduced pathogenicity in BALB/c mice [167]. The same group has recently shown that a $\Delta sseK2$ mutant of *S. enterica* serovar Typhimurium is also defective in biofilm formation and virulence [168]. Interestingly, a genomic screen for genes involved in long-term systemic infection in 129X1/SvJ (Nramp1^r) mice detected *sseK2* as one of the genes contributing to serovar Typhimurium infection by 28 days postinfection. The CI of the *sseK2* mutant in this system was 0.0041 in mesenteric lymph nodes, and 0.0026 in spleen [169]. This study underlines the importance of finding the right model to highlight the role of a given effector in virulence.

4.5. Targets of NleB/SseK Effectors

The most clearly demonstrated function of this family of T3SS effectors is the inhibition of NF- κ B pathway activation and of host cell death during infection. A first report suggested that NleB1 from EPEC contributed to NF- κ B inhibition by enhancing the activity of NleE, another effector that inhibits I κ B phosphorylation, leading to the stabilization of I κ B that retains NF- κ B in the cytoplasm [170]. Another study proposed a model where NleB1 acted at a different, upstream point in the NF- κ B signaling pathway that is activated in response to TNF α [124]. Involvement in the TNF receptor (TNFR) pathway, but not the TLR/ILR1 pathway, was confirmed using specific luciferase reporters [171]. This is a death receptor pathway that involves homo- or heterotypic interactions between intracellular death domains in receptors and death domains in the downstream adaptor proteins, including TRADD and FADD [172]. The mechanism of action of NleB1 was elucidated by two outstanding reports that used a yeast two-hybrid screen of a HeLa cDNA library and co-immunoprecipitation assays to identify host binding partners. NleB1 directly inactivated death domains in several proteins including TRADD, FADD, RIPK1, and TNFR1 [140,141]. This inhibition depended on the N-GlcNAc transferase activity of NleB1, which specifically modified Arg₂₃₅ and Arg₁₁₇ in the death domains of TRADD and FADD, respectively. These modifications blocked interactions between death domains, thereby disrupting inflammatory NF- κ B signaling, caspase 8-dependent apoptosis, and necroptosis [173]. NleB from *C. rodentium* showed the same activity, and the mouse model of infection with this A/E pathogen demonstrated that the GlcNAc transferase activity was crucial for virulence [140,141]. NleB2 from EPEC showed a lower enzymatic activity and only a fraction of TRADD was GlcNAcylated after transfection of HEK293T cells with a plasmid expressing NleB2 [140]. A more recent study combining an arginine-GlcNAc-specific enrichment method coupled with mass spectrometry concluded that the preferred target of EPEC NleB1 and *C. rodentium* NleB was FADD, and that overexpression of these effectors led to GlcNAcylation of additional proteins that may be nonphysiological targets [174].

At least some of the members of the family present in *S. enterica* serovar Typhimurium have an effect on the NF- κ B pathway. Similar to NleB1, SseK1 inhibited TNF- α -NF- κ B signaling and

catalyzed GlcNAcylation of TRADD after transfection of host cells [140]. Also, ectopic expression of SseK3 blocked TNF-stimulated NF- κ B activity and this effect depended on the conserved DxD catalytic motif [138]. Importantly, SseK1 and SseK3 inhibited *Salmonella*-induced NF- κ B activation and necroptotic cell death during infection in macrophages [137]. Although SseK2 was also able to inhibit NF- κ B activation after TNF- α stimulation, it is not clear if it is a bona fide inhibitor of this pathway since the effect was detected only after ectopic overexpression in mammalian cells but not during macrophage infections. Expression of SseK1 or SseK3 in a Δ sseK1/2/3 *Salmonella* mutant induced different arginine-GlcNAcylation patterns during infection of RAW264.7 macrophages, whereas expression of SseK2 did not result in detectable modifications. Furthermore, the death domain protein FADD was GlcNAcyated by SseK1, but not by SseK2 or SseK3, and the death domain protein TRADD was modified by SseK1 and, to a lesser extent, by SseK3 [137]. Another study confirmed that SseK1, SseK3, EHEC NleB1, EPEC NleB1, and *C. rodentium* NleB, but not SseK2 or NleB2, blocked TNF-mediated NF- κ B pathway activation. However, in conflict with the previous report, the same study reported glycosylation of FADD by SseK2, but not by SseK1 or SseK3 [143]. A recent study combined arginine-GlcNAc glycopeptide immunoprecipitation and mass spectrometry to identify host proteins GlcNAcyated by endogenous levels of SseK1 and SseK3 during *Salmonella* infection. This study confirmed modification of TRADD, but not FADD, by SseK1 and described the death domains of receptors TNFR1 and TRAILR as physiological SseK3 targets that were modified at residues Arg³⁷⁶ and Arg²⁹³, respectively [147].

Some proteins that do not contain death domains have also been described as targets of NleB/SseK effectors. An early paper identified glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as a binding partner and substrate of the glycosyltransferase activity of *C. rodentium* NleB [151]. The same study described the interaction of GAPDH with the TNFR-associated factor 2 (TRAF2), a protein required for TNF- α -mediated NF- κ B activation. Data presented in this report suggested a model in which GlcNAcylation of GAPDH by NleB would disrupt the TRAF2-GAPDH interaction to suppress TRAF2 polyubiquitination and NF- κ B activation. Another study from the same authors suggested that this modification of GAPDH also inhibited the interaction with TRAF3 and reduced Lys⁶³-linked TRAF3 ubiquitination, leading to inhibition of type I IFN signaling [142]. Although several researchers were unable to detect NleB-dependent GlcNAcylation of GAPDH [140,146,174], this modification was confirmed by other studies [143,145,175], and mass spectrometry and site-directed mutagenesis demonstrated that EHEC NleB1 glycosylates GAPDH on Arg¹⁹⁷ and Arg²⁰⁰ [143]. In addition, GAPDH was also shown to be a target for SseK1 [143].

A new role has been recently suggested for EPEC NleB1 through its interaction and modification of HIF-1 α , a master regulator of cellular O₂ homeostasis. NleB1 catalyzes GlcNAcylation at Arg¹⁸ at the N-terminus of HIF-1 α , thereby enhancing its transcriptional activity and inducing downstream gene expression to alter host glucose metabolism [176].

There are also host binding partners described for NleB/SseK proteins whose GlcNAcylation has not been demonstrated. This is the case of TRIM32, the first binding partner identified for SseK3 [138]. TRIM32 is an E3 ubiquitin ligase that is not required for SseK3 to inhibit NF- κ B signaling during infection of macrophages [137]. DRG2 (developmentally-regulated GTP-binding protein 2), LRR18 (leucine-rich repeat-containing protein 18), and POLR2E (DNA-directed RNA polymerases I, II, and III subunit RPABC1), identified in a yeast two-hybrid screen [177], and MAP7 (ensconsin), identified by a quantitative proteomics approach and confirmed by coimmunoprecipitation [178], are novel binding partners of EHEC NleB1, although the GlcNAcylation of these proteins and the physiological role of these interactions have not been shown. A summary of host targets of NleB/SseK effectors is depicted in Figure 6.

Interestingly, a recent report has shown that NleB1 can also glycosylate a bacterial protein, GshB (glutathione synthetase) on Arg²⁵⁶. In fact, this was the most abundant arginine-GlcNAcyated target enriched from samples of HEK293T cells infected with EHEC [179]. GshB was also glycosylated in vitro and in vivo by *C. rodentium* NleB and in vitro by *Salmonella* SseK1. In vitro glycosylation of GshB enhances GshB-mediated production of glutathione, which provides resistance to oxidative stress. Consistent with that, both the *nleB* and *gshB* mutants of *C. rodentium* had a significant growth defect

Author Contributions: F.R.-M. designed the structure of the manuscript. J.L.A.-G., J.B.-B. and F.R.-M. wrote and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The work in the laboratory is funded by the Spanish Ministerio de Economía, Industria y Competitividad—Agencia Estatal de Investigación, and the European Regional Development Fund, grant number SAF2016-75365-R; and by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 842629.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the work.

References

1. Wagner, S.; Grin, I.; Malmshaimer, S.; Singh, N.; Torres-Vargas, C.E.; Westerhausen, S. Bacterial type III secretion systems: A complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiol. Lett.* **2018**, *365*, fny201.
2. Galán, J.E.; Lara-Tejero, M.; Marlovits, T.C.; Wagner, S. Bacterial type III secretion systems: Specialized nanomachines for protein delivery into target cells. *Annu. Rev. Microbiol.* **2014**, *68*, 415–438.
3. Galán, J.E. Common themes in the design and function of bacterial effectors. *Cell Host Microbe* **2009**, *5*, 571–579.
4. Ma, W.; Guttman, D.S. Evolution of prokaryotic and eukaryotic virulence effectors. *Curr. Opin. Plant Biol.* **2008**, *11*, 412–419.
5. Scott, N.E.; Hartland, E.L. Post-translational mechanisms of host subversion by bacterial effectors. *Trends Mol. Med.* **2017**, *23*, 1088–1102.
6. Brenner, F.W.; Villar, R.G.; Angulo, F.J.; Tauxe, R.; Swaminathan, B. *Salmonella* nomenclature. *J. Clin. Microbiol.* **2000**, *38*, 2465–2467.
7. Ochman, H.; Soncini, F.C.; Solomon, F.; Groisman, E.A. Identification of a pathogenicity island required for *Salmonella* survival in host cells. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7800–7804.
8. Shea, J.E.; Hensel, M.; Gleeson, C.; Holden, D.W. Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2593–2597.
9. Galán, J.E.; Curtiss, R. Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6383–6387.
10. Fookes, M.; Schroeder, G.N.; Langridge, G.C.; Blondel, C.J.; Mammina, C.; Connor, T.R.; Seth-Smith, H.; Vernikos, G.S.; Robinson, K.S.; Sanders, M.; et al. *Salmonella bongori* provides insights into the evolution of the salmonellae. *PLoS Pathog.* **2011**, *7*, e1002191.
11. Lou, L.; Zhang, P.; Piao, R.; Wang, Y. *Salmonella* pathogenicity island 1 (SPI-1) and its complex regulatory network. *Front. Cell. Infect. Microbiol.* **2019**, *9*, 270.
12. Tuli, A.; Sharma, M. How to do business with lysosomes: *Salmonella* leads the way. *Curr. Opin. Microbiol.* **2019**, *47*, 1–7.
13. Jennings, E.; Thurston, T.L.M.; Holden, D.W. *Salmonella* SPI-2 type III secretion system effectors: Molecular mechanisms and physiological consequences. *Cell Host Microbe* **2017**, *22*, 217–231.
14. Habyarimana, F.; Sabag-Daigle, A.; Ahmer, B.M.M. The SdiA-regulated gene *srcE* encodes a type III secreted effector. *J. Bacteriol.* **2014**, *196*, 2301–2312.
15. Sun, H.; Kamanova, J.; Lara-Tejero, M.; Galán, J.E. A family of *Salmonella* type III secretion effector proteins selectively targets the NF- κ B signaling pathway to preserve host homeostasis. *PLoS Pathog.* **2016**, *12*, e1005484.
16. Jaslow, S.L.; Gibbs, K.D.; Fricke, W.F.; Wang, L.; Pittman, K.J.; Mammel, M.K.; Thaden, J.T.; Fowler, V.G.; Hammer, G.E.; Elfenbein, J.R.; et al. *Salmonella* activation of STAT3 signaling by SarA effector promotes intracellular replication and production of IL-10. *Cell Rep.* **2018**, *23*, 3525–3536.
17. Ramos-Morales, F. Impact of *Salmonella enterica* type III secretion system effectors on the eukaryotic host cell. *ISRN Cell Biol.* **2012**, *2012*, 1–36.
18. Frankel, G.; Phillips, A.D.; Rosenshine, I.; Dougan, G.; Kaper, J.B.; Knutton, S. Enteropathogenic and enterohaemorrhagic *Escherichia coli*: More subversive elements. *Mol. Microbiol.* **1998**, *30*, 911–921.
19. Barthold, S.W. The microbiology of transmissible murine colonic hyperplasia. *Lab. Anim. Sci.* **1980**, *30*, 167–173.
20. Deng, W.; Li, Y.; Vallance, B.A.; Finlay, B.B. Locus of enterocyte effacement from *Citrobacter rodentium*:

- Sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. *Infect. Immun.* **2001**, *69*, 6323–6335.
21. Jarvis, K.G.; Girón, J.A.; Jerse, A.E.; Mcdaniel, T.K.; Donnenberg, M.S.; Kaper, J.B. 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. USA* **1995**, *92*, 7996–8000.
 22. Gaytán, M.O.; Martínez-Santos, V.I.; Soto, E.; González-Pedrajo, B. Type three secretion system in attaching and effacing pathogens. *Front. Cell. Infect. Microbiol.* **2016**, *6*, 129.
 23. Zhou, M.; Guo, Z.; Duan, Q.; Hardwidge, P.R.; Zhu, G. *Escherichia coli* type III secretion system 2: A new kind of T3SS? *Vet. Res.* **2014**, *45*, 32.
 24. Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; et al. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* **2001**, *8*, 11–22.
 25. Shulman, A.; Yair, Y.; Biran, D.; Sura, T.; Otto, A.; Gophna, U.; Becher, D.; Hecker, M.; Ron, E.Z. The *Escherichia coli* type III secretion system 2 has a global effect on cell surface. *MBio* **2018**, *9*, e01070-18.
 26. Ooka, T.; Ogura, Y.; Katsura, K.; Seto, K.; Kobayashi, H.; Kawano, K.; Tokuoka, E.; Furukawa, M.; Harada, S.; Yoshino, S.; et al. Defining the genome features of *Escherichia albertii*, an emerging enteropathogen closely related to *Escherichia coli*. *Genome Biol. Evol.* **2015**, *7*, 3170–3179.
 27. Ideses, D.; Gophna, U.; Paitan, Y.; Chaudhuri, R.R.; Pallen, M.J.; Ron, E.Z. A degenerate type III secretion system from septicemic *Escherichia coli* contributes to pathogenesis. *J. Bacteriol.* **2005**, *187*, 8164–8171.
 28. Hurley, D.; McCusker, M.P.; Fanning, S.; Martins, M. *Salmonella*-host interactions—Modulation of the host innate immune system. *Front. Immunol.* **2014**, *5*, 481.
 29. Lairson, L.L.; Henrissat, B.; Davies, G.J.; Withers, S.G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555.
 30. Mestrom, L.; Przypis, M.; Kowalczykiewicz, D.; Pollender, A.; Kumpf, A.; Marsden, S.R.; Bento, I.; Jarzębski, A.B.; Szymańska, K.; Chruściel, A.; et al. Leloir glycosyltransferases in applied biocatalysis: A multidisciplinary approach. *Int. J. Mol. Sci.* **2019**, *20*, 5263.
 31. Campbell, J.A.; Davies, G.J.; Bulone, V.; Henrissat, B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **1997**, *326*, 929–939.
 32. Coutinho, P.M.; Deleury, E.; Davies, G.J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **2003**, *328*, 307–317.
 33. Sinnott, M.L. Catalytic mechanism of enzymic glycosyl transfer. *Chem. Rev.* **1990**, *90*, 1171–1202.
 34. Liang, D.M.; Liu, J.H.; Wu, H.; Wang, B. Bin; Zhu, H.J.; Qiao, J.J. Glycosyltransferases: Mechanisms and applications in natural product development. *Chem. Soc. Rev.* **2015**, *44*, 8350–8374.
 35. Davies, G.J. Sweet secrets of synthesis. *Nat. Struct. Biol.* **2001**, *8*, 98–100.
 36. Vocadlo, D.J.; Davies, G.J.; Laine, R.; Withers, S.G. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **2001**, *412*, 835–838.
 37. Soya, N.; Fang, Y.; Palcic, M.M.; Klassen, J.S. Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology* **2011**, *21*, 547–552.
 38. Lira-Navarrete, E.; Iglesias-Fernández, J.; Zandberg, W.F.; Compañón, I.; Kong, Y.; Corzana, F.; Pinto, B.M.; Clausen, H.; Peregrina, J.M.; Vocadlo, D.J.; et al. Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 8206–8210.
 39. Ardèvol, A.; Rovira, C. The molecular mechanism of enzymatic glycosyl transfer with retention of configuration: Evidence for a short-lived oxocarbenium-like species. *Angew. Chem. Int. Ed.* **2011**, *50*, 10897–10901.
 40. Schuman, B.; Evans, S. V; Fyles, T.M. Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. *PLoS ONE* **2013**, *8*, e71077.
 41. Dodson, E.; Harding, M.M.; Hodgkin, D.C.; Rossmann, M.G. The crystal structure of insulin: III. Evidence for a 2-fold axis in rhombohedral zinc insulin. *J. Mol. Biol.* **1966**, *16*, 227–241.
 42. Schuman, B.; Alfaro, J.A.; Evans, S.V. Glycosyltransferase Structure and Function. In *Bioactive Conformation I*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 217–257.
 43. Kikuchi, N.; Kwon, Y.D.; Gotoh, M.; Narimatsu, H. Comparison of glycosyltransferase families using the profile hidden Markov model. *Biochem. Biophys. Res. Commun.* **2003**, *310*, 574–579.
 44. Liu, J.; Mushegian, A. Three monophyletic superfamilies account for the majority of the known

- glycosyltransferases. *Protein Sci.* **2003**, *12*, 1418–1431.
45. Rosén, M.L.; Edman, M.; Sjöström, M.; Wieslander, A. Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *J. Biol. Chem.* **2004**, *279*, 38683–38692.
 46. Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K.P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **2011**, *474*, 350–356.
 47. Zhang, H.; Zhu, F.; Yang, T.; Ding, L.; Zhou, M.; Li, J.; Haslam, S.M.; Dell, A.; Erlandsen, H.; Wu, H. The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat. Commun.* **2014**, *5*, 4339.
 48. Lovering, A.L.; De Castro, L.H.; Lim, D.; Strynadka, N.C.J. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science (80-)* **2007**, *315*, 1402–1405.
 49. Yuan, Y.; Barrett, D.; Zhang, Y.; Kahne, D.; Sliz, P.; Walker, S. Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 5348–5353.
 50. Kattke, M.D.; Gosschalk, J.E.; Martinez, O.E.; Kumar, G.; Gale, R.T.; Cascio, D.; Sawaya, M.R.; Philips, M.; Brown, E.D.; Clubb, R.T. Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *PLoS Pathog.* **2019**, *15*, e1007723.
 51. Charnock, S.J.; Davies, G.J. Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **1999**, *38*, 6380–6385.
 52. Mulichak, A.M.; Losey, H.C.; Walsh, C.T.; Garavito, R.M. Structure of the UDP-glucosyltransferase GtFB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **2001**, *9*, 547–557.
 53. Rose, A.S.; Bradley, A.R.; Valasatava, Y.; Duarte, J.M.; Prlić, A.; Rose, P.W. NGL viewer: Web-based molecular graphics for large complexes. *Bioinformatics* **2018**, *34*, 3755–3758.
 54. Lu, Q.; Li, S.; Shao, F. Sweet talk: Protein glycosylation in bacterial interaction with the host. *Trends Microbiol.* **2015**, *23*, 630–641.
 55. Nothaft, H.; Szymanski, C.M. Protein glycosylation in bacteria: Sweeter than ever. *Nat. Rev. Microbiol.* **2010**, *8*, 765–778.
 56. Grass, S.; Buscher, A.Z.; Swords, W.E.; Apicella, M.A.; Barenkamp, S.J.; Ozchlewski, N.; St Geme, J.W. 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*, 737–751.
 57. St Geme, J.W.; Grass, S. Secretion of the *Haemophilus influenzae* HMW1 and HMW2 adhesins involves a periplasmic intermediate and requires the HMWB and HMWC proteins. *Mol. Microbiol.* **1998**, *27*, 617–630.
 58. Gross, J.; Grass, S.; Davis, A.E.; Gilmore-Erdmann, P.; Townsend, R.R.; St. Geme, J.W. The *Haemophilus influenzae* HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. *J. Biol. Chem.* **2008**, *283*, 26010–26015.
 59. Szymanski, C.M.; Burr, D.H.; Guerry, P. *Campylobacter* protein glycosylation affects host cell interactions. *Infect. Immun.* **2002**, *70*, 2242–2244.
 60. Karlyshev, A. V.; Everest, P.; Linton, D.; Cawthraw, S.; Newell, D.G.; Wren, B.W. The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **2004**, *150*, 1957–1964.
 61. Hendrixson, D.R.; DiRita, V.J. Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol. Microbiol.* **2004**, *52*, 471–484.
 62. Perez, C.; Gerber, S.; Boilevin, J.; Bucher, M.; Darbre, T.; Aebi, M.; Reymond, J.-L.; Locher, K.P. Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* **2015**, *524*, 433–438.
 63. Wacker, M.; Linton, D.; Hitchen, P.G.; Nita-Lazar, M.; Haslam, S.M.; North, S.J.; Panico, M.; Morris, H.R.; Dell, A.; Wren, B.W.; et al. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **2002**, *298*, 1790–1793.
 64. Bokhari, H.; Maryam, A.; Shahid, R.; Siddiqi, A.R. Oligosaccharyltransferase PglB of *Campylobacter jejuni* is a glycoprotein. *World J. Microbiol. Biotechnol.* **2020**, *36*, 9.
 65. Lu, Q.; Yao, Q.; Xu, Y.; Li, L.; Li, S.; Liu, Y.; Gao, W.; Niu, M.; Sharon, M.; Ben-Nissan, G.; et al. An iron-containing dodecameric heptosyltransferase family modifies bacterial autotransporters in pathogenesis. *Cell Host Microbe* **2014**, *16*, 351–363.
 66. Benz, I.; Schmidt, M.A. Glycosylation with heptose residues mediated by the aah gene product is essential

- for adherence of the AIDA-I adhesin. *Mol. Microbiol.* **2001**, *40*, 1403–1413.
67. Zhou, M.; Wu, H. Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology* **2009**, *155*, 317–327.
 68. Lizcano, A.; Sanchez, C.J.; Orihuela, C.J. A role for glycosylated serine-rich repeat proteins in gram-positive bacterial pathogenesis. *Mol. Oral Microbiol.* **2012**, *27*, 257–269.
 69. Zhu, F.; Zhang, H.; Yang, T.; Haslam, S.M.; Dell, A.; Wu, H. Engineering and dissecting the glycosylation pathway of a streptococcal serine-rich repeat adhesin. *J. Biol. Chem.* **2016**, *291*, 27354–27363.
 70. Logan, S.M. Flagellar glycosylation—A new component of the motility repertoire? *Microbiology* **2006**, *152*, 1249–1262.
 71. Duan, Q.; Zhou, M.; Zhu, L.; Zhu, G. Flagella and bacterial pathogenicity. *J. Basic Microbiol.* **2013**, *53*, 1–8.
 72. Merino, S.; Tomás, J.M. Gram-negative flagella glycosylation. *Int. J. Mol. Sci.* **2014**, *15*, 2840–2857.
 73. Delvaux, N.A.; Thoden, J.B.; Holden, H.M. Molecular architectures of Pen and Pal: Key enzymes required for CMP-pseudaminic acid biosynthesis in *Bacillus thuringiensis*. *Protein Sci.* **2018**, *27*, 738–749.
 74. Valiente, E.; Bouché, L.; Hitchen, P.; Faulds-Pain, A.; Songane, M.; Dawson, L.F.; Donahue, E.; Stabler, R.A.; Panico, M.; Morris, H.R.; et al. Role of glycosyltransferases modifying type B flagellin of emerging hypervirulent *Clostridium difficile* lineages and their impact on motility and biofilm formation. *J. Biol. Chem.* **2016**, *291*, 25450–25461.
 75. Bouché, L.; Panico, M.; Hitchen, P.; Binet, D.; Sastre, F.; Faulds-Pain, A.; Valiente, E.; Vinogradov, E.; Aubry, A.; Fulton, K.; et al. The type B flagellin of hypervirulent *Clostridium difficile* is modified with novel sulfonated peptidylamido-glycans. *J. Biol. Chem.* **2016**, *291*, 25439–25449.
 76. De Maayer, P.; Cowan, D.A. Comparative genomic analysis of the flagellin glycosylation island of the Gram-positive thermophile *Geobacillus*. *BMC Genom.* **2016**, *17*, 913.
 77. Janesch, B.; Schirmeister, F.; Maresch, D.; Altmann, F.; Messner, P.; Kolarich, D.; Schäffer, C. Flagellin glycosylation in *Paenibacillus alvei* CCM 2051T. *Glycobiology* **2016**, *26*, 74–87.
 78. Twine, S.M.; Reid, C.W.; Aubry, A.; McMullin, D.R.; Fulton, K.M.; Austin, J.; Logan, S.M. Motility and flagellar glycosylation in *Clostridium difficile*. *J. Bacteriol.* **2009**, *191*, 7050–7062.
 79. Twine, S.M.; Paul, C.J.; Vinogradov, E.; McNally, D.J.; Brisson, J.R.; Mullen, J.A.; McMullin, D.R.; Jarrell, H.C.; Austin, J.W.; Kelly, J.F.; et al. Flagellar glycosylation in *Clostridium botulinum*. *FEBS J.* **2008**, *275*, 4428–4444.
 80. Schirm, M.; Kalmokoff, M.; Aubry, A.; Thibault, P.; Sandoz, M.; Logan, S.M. Flagellin from *Listeria monocytogenes* is glycosylated with β -O-linked N-acetylglucosamine. *J. Bacteriol.* **2004**, *186*, 6721–6727.
 81. Haya, S.; Tokumaru, Y.; Abe, N.; Kaneko, J.; Aizawa, S.I. Characterization of lateral flagella of *Selenomonas ruminantium*. *Appl. Environ. Microbiol.* **2011**, *77*, 2799–2802.
 82. Merino, S.; Aquilini, E.; Fulton, K.M.; Twine, S.M.; Tomás, J.M. The polar and lateral flagella from *Plesiomonas shigelloides* are glycosylated with legionaminic acid. *Front. Microbiol.* **2015**, *6*, 649.
 83. Scott, A.E.; Twine, S.M.; Fulton, K.M.; Titball, R.W.; Essex-Lopresti, A.E.; Atkins, T.P.; Prior, J.L. Flagellar glycosylation in *Burkholderia pseudomallei* and *Burkholderia thailandensis*. *J. Bacteriol.* **2011**, *193*, 3577–3587.
 84. Rath, C.B.; Schirmeister, F.; Figl, R.; Seeberger, P.H.; Schäffer, C.; Kolarich, D. Flagellin glycoproteomics of the periodontitis associated pathogen *Selenomonas sputigena* reveals previously not described O-glycans and rhamnose fragment rearrangement occurring on the glycopeptides. *Mol. Cell. Proteom.* **2018**, *17*, 721–736.
 85. Thibault, P.; Logan, S.M.; Kelly, J.F.; Brisson, J.R.; Ewing, C.P.; Trust, T.J.; Guerry, P. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **2001**, *276*, 34862–34870.
 86. Parker, J.L.; Day-Williams, M.J.; Tomas, J.M.; Stafford, G.P.; Shaw, J.G. Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *Microbiologyopen* **2012**, *1*, 149–160.
 87. Rabaan, A.A.; Gryllos, I.; Tomás, J.M.; Shaw, J.G. Motility and the polar flagellum are required for *Aeromonas caviae* adherence to HEP-2 cells. *Infect. Immun.* **2001**, *69*, 4257–4267.
 88. Wilhelms, M.; Fulton, K.M.; Twine, S.M.; Tomás, J.M.; Merino, S. Differential glycosylation of polar and lateral flagellins in *Aeromonas hydrophila* AH-3. *J. Biol. Chem.* **2012**, *287*, 27851–27862.
 89. Dorrell, N.; Mangan, J.A.; Laing, K.G.; Hinds, J.; Linton, D.; Al-Ghusein, H.; Barrell, B.G.; Parkhill, J.; Stoker, N.G.; Karlyshev, A. V.; et al. Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Res.* **2001**, *11*, 1706–1715.

90. Alm, R.A.; Guerry, P.; Power, M.E.; Trust, T.J. Variation in antigenicity and molecular weight of *Campylobacter coli* VC167 flagellin in different genetic backgrounds. *J. Bacteriol.* **1992**, *174*, 4230–4238.
91. Schoenhofen, I.C.; McNally, D.J.; Brisson, J.-R.; Logan, S.M. Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: Synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* **2006**, *16*, 8C–14C.
92. Schirm, M.; Soo, E.C.; Aubry, A.J.; Austin, J.; Thibault, P.; Logan, S.M. Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol. Microbiol.* **2003**, *48*, 1579–1592.
93. Sulzenbacher, G.; Roig-Zamboni, V.; Lebrun, R.; Guérardel, Y.; Murat, D.; Mansuelle, P.; Yamakawa, N.; Qian, X.X.; Vincentelli, R.; Bourne, Y.; et al. Glycosylate and move! The glycosyltransferase Maf is involved in bacterial flagella formation. *Environ. Microbiol.* **2018**, *20*, 228–240.
94. McNally, D.J.; Hui, J.P.M.; Aubry, A.J.; Mui, K.K.K.; Guerry, P.; Brisson, J.R.; Logan, S.M.; Soo, E.C. Functional characterization of the flagellar glycosylation locus in *Campylobacter jejuni* 81–176 using a focused metabolomics approach. *J. Biol. Chem.* **2006**, *281*, 18489–18498.
95. Parker, J.L.; Lowry, R.C.; Couto, N.A.S.; Wright, P.C.; Stafford, G.P.; Shaw, J.G. Maf-dependent bacterial flagellin glycosylation occurs before chaperone binding and flagellar T3SS export. *Mol. Microbiol.* **2014**, *92*, 258–272.
96. Voth, D.E.; Ballard, J.D. *Clostridium difficile* toxins: Mechanism of action and role in disease. *Clin. Microbiol. Rev.* **2005**, *18*, 247–263.
97. Kuehne, S.A.; Cartman, S.T.; Heap, J.T.; Kelly, M.L.; Cockayne, A.; Minton, N.P. The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* **2010**, *467*, 711–713.
98. Just, I.; Selzer, J.; Wilm, M.; von Eichel-Streiber, C.; Mann, M.; Aktories, K. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. *Nature* **1995**, *375*, 500–503.
99. Genth, H.; Aktories, K.; Just, I. Monoglucosylation of RhoA at threonine 37 blocks cytosol-membrane cycling. *J. Biol. Chem.* **1999**, *274*, 29050–29056.
100. Jank, T.; Bogdanović, X.; Wirth, C.; Haaf, E.; Spoerner, M.; Böhmer, K.E.; Steinemann, M.; Orth, J.H.C.; Kalbitzer, H.R.; Warscheid, B.; et al. A bacterial toxin catalyzing tyrosine glycosylation of Rho and deamidation of Gq and Gi proteins. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1273–1280.
101. Qiu, J.; Luo, Z.-Q. *Legionella* and *Coxiella* effectors: Strength in diversity and activity. *Nat. Rev. Microbiol.* **2017**, *15*, 591–605.
102. Belyi, Y.; Tabakova, I.; Stahl, M.; Aktories, K. Lgt: A family of cytotoxic glucosyltransferases produced by *Legionella pneumophila*. *J. Bacteriol.* **2008**, *190*, 3026–3035.
103. Lü, W.; Du, J.; Stahl, M.; Tzivelekidis, T.; Belyi, Y.; Gerhardt, S.; Aktories, K.; Einsle, O. Structural basis of the action of glucosyltransferase Lgt1 from *Legionella pneumophila*. *J. Mol. Biol.* **2010**, *396*, 321–331.
104. Reinert, D.J.; Jank, T.; Aktories, K.; Schulz, G.E. Structural basis for the function of *Clostridium difficile* toxin B. *J. Mol. Biol.* **2005**, *351*, 973–981.
105. Belyi, Y.; Stahl, M.; Sovkova, I.; Kaden, P.; Luy, B.; Aktories, K. Region of elongation factor 1A1 involved in substrate recognition by *Legionella pneumophila* glucosyltransferase Lgt1. Identification of Lgt1 as a retaining glucosyltransferase. *J. Biol. Chem.* **2009**, *284*, 20167–20174.
106. Deng, W.; Puente, J.L.; Gruenheid, S.; Li, Y.; Vallance, B.A.; Vázquez, A.; Barba, J.; Ibarra, J.A.; O'Donnell, P.; Metalnikov, P.; et al. Dissecting virulence: Systematic and functional analyses of a pathogenicity island. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3597–3602.
107. Petty, N.K.; Bulgin, R.; Crepin, V.F.; Cerdeño-Tárraga, A.M.; Schroeder, G.N.; Quail, M.A.; Lennard, N.; Corton, C.; Barron, A.; Clark, L.; et al. The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. *J. Bacteriol.* **2010**, *192*, 525–538.
108. Perna, N.T.; Plunkett, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; et al. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **2001**, *409*, 529–533.
109. Iguchi, A.; Thomson, N.R.; Ogura, Y.; Saunders, D.; Ooka, T.; Henderson, I.R.; Harris, D.; Asadulghani, M.; Kurokawa, K.; Dean, P.; et al. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *J. Bacteriol.* **2009**, *91*, 347–354.
110. Kelly, M.; Hart, E.; Mundy, R.; Marchès, O.; Wiles, S.; Badea, L.; Luck, S.; Tauschek, M.; Frankel, G.; Robins-Browne, R.M.; et al. Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. *Infect. Immun.* **2006**, *74*, 2328–2337.

111. Brown, N.F.; Coombes, B.K.; Bishop, J.L.; Wickham, M.E.; Lowden, M.J.; Gal-Mor, O.; Goode, D.L.; Boyle, E.C.; Sanderson, K.L.; Finlay, B.B.; et al. *Salmonella* phage ST64B encodes a member of the SseK/NleB effector family. *PLoS ONE* **2011**, *6*, e17824.
112. Kujat Choy, S.L.; Boyle, E.C.; Gal-Mor, O.; Goode, D.L.; Valdez, Y.; Vallance, B.A.; Finlay, B.B. SseK1 and SseK2 are novel translocated proteins of *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* **2004**, *72*, 5115–5125.
113. Jarvik, T.; Smillie, C.; Groisman, E.A.; Ochman, H. Short-term signatures of evolutionary change in the *Salmonella enterica* serovar Typhimurium 14028 genome. *J. Bacteriol.* **2010**, *192*, 560–567.
114. Eswarappa, S.M.; Janice, J.; Balasundaram, S. V.; Chakravorty, D. Non-neutral evolution in non-LEE-encoded type III effectors of attaching and effacing *Escherichia coli*. *Microbes Infect.* **2013**, *15*, 147–151.
115. Bhatt, S.; Egan, M.; Critelli, B.; Kouse, A.; Kalman, D.; Upreti, C. The evasive enemy: Insights into the virulence and epidemiology of the emerging attaching and effacing pathogen *Escherichia albertii*. *Infect. Immun.* **2019**, *87*, e00254-18.
116. Liu, S.; Feng, J.; Pu, J.; Xu, X.; Lu, S.; Yang, J.; Wang, Y.; Jin, D.; Du, X.; Meng, X.; et al. Genomic and molecular characterisation of *Escherichia marmotae* from wild rodents in Qinghai-Tibet plateau as a potential pathogen. *Sci. Rep.* **2019**, *9*, 10619.
117. Nguyen, S. V.; Muthappa, D.M.; Hurley, D.; Donoghue, O.; McCabe, E.; Anes, J.; Schaffer, K.; Murphy, B.P.; Buckley, J.F.; Fanning, S. *Yersinia hibernica* sp. nov., isolated from pig-production environments. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 2023–2027.
118. Huerta-Cepas, J.; Serra, F.; Bork, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **2016**, *33*, 1635–1638.
119. Price, M.N.; Dehal, P.S.; Arkin, A.P. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650.
120. Hughes, D.T.; Sperandio, V. Inter-kingdom signalling: Communication between bacteria and their hosts. *Nat. Rev. Microbiol.* **2008**, *6*, 111–120.
121. Roe, A.J.; Yull, H.; Naylor, S.W.; Woodward, M.J.; Smith, D.G.E.; Gally, D.L. Heterogeneous surface expression of EspA translocon filaments by *Escherichia coli* O157:H7 is controlled at the posttranscriptional level. *Infect. Immun.* **2003**, *71*, 5900–5909.
122. Roe, A.J.; Tysall, L.; Dransfield, T.; Wang, D.; Fraser-Pitt, D.; Mahajan, A.; Constandinou, C.; Inglis, N.; Downing, A.; Talbot, R.; et al. Analysis of the expression, regulation and export of NleA-E in *Escherichia coli* O157:H7. *Microbiology* **2007**, *153*, 1350–1360.
123. Charpentier, X.; Oswald, E. Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J. Bacteriol.* **2004**, *186*, 5486–5495.
124. Newton, H.J.; Pearson, J.S.; Badea, L.; Kelly, M.; Lucas, M.; Holloway, G.; Wagstaff, K.M.; Dunstone, M.A.; Sloan, J.; Whisstock, J.C.; et al. The type III effectors NleE and NleB from enteropathogenic *E. coli* and OspZ from *Shigella* block nuclear translocation of NF-kappaB p65. *PLoS Pathog.* **2010**, *6*, e1000898.
125. García-Angulo, V.A.; Martínez-Santos, V.I.; Villaseñor, T.; Santana, F.J.; Huerta-Saquero, A.; Martínez, L.C.; Jiménez, R.; Lara-Ochoa, C.; Téllez-Sosa, J.; Bustamante, V.H.; et al. A distinct regulatory sequence is essential for the expression of a subset of *nle* genes in attaching and effacing *Escherichia coli*. *J. Bacteriol.* **2012**, *194*, 5589–5603.
126. Deng, W.; Yu, H.B.; De Hoog, C.L.; Stoyanov, N.; Li, Y.; Foster, L.J.; Finlay, B.B. Quantitative proteomic analysis of type III secretome of enteropathogenic *Escherichia coli* reveals an expanded effector repertoire for attaching/effacing bacterial pathogens. *Mol. Cell. Proteom.* **2012**, *11*, 692–709.
127. Ellermeier, C.D.; Ellermeier, J.R.; Schlauch, J.M. HilD, HilC and RtsA constitute a feed forward loop that controls expression of the SPI1 type three secretion system regulator *hilA* in *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* **2005**, *57*, 691–705.
128. Schechter, L.M.; Jain, S.; Akbar, S.; Lee, C.A. The small nucleoid-binding proteins H-NS, HU, and Fis affect *hilA* expression in *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* **2003**, *71*, 5432–5435.
129. Laughlin, R.C.; Knodler, L.A.; Barhoumi, R.; Ross Payne, H.; Wu, J.; Gomez, G.; Pugh, R.; Lawhon, S.D.; Bäuml, A.J.; Steele-Mortimer, O.; et al. Spatial segregation of virulence gene expression during acute enteric infection with *Salmonella enterica* serovar Typhimurium. *MBio* **2014**, *5*, e00946-13.
130. Mouslim, C.; Delgado, M.; Groisman, E.A. Activation of the RcsC/YojN/RcsB phosphorelay system attenuates *Salmonella* virulence. *Mol. Microbiol.* **2004**, *54*, 386–395.

131. Golubeva, Y.A.; Sadik, A.Y.; Ellermeier, J.R.; Slauch, J.M. Integrating global regulatory input into the *Salmonella* pathogenicity Island 1 type III secretion system. *Genetics* **2012**, *190*, 79–90.
132. Palmer, A.D.; Kim, K.; Slauch, J.M. PhoP-mediated repression of the SPI1 T3SS in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **2019**, *201*, e00264-19.
133. Prost, L.R.; Miller, S.I. The Salmonellae PhoQ sensor: Mechanisms of detection of phagosome signals. *Cell. Microbiol.* **2008**, *10*, 576–582.
134. Kato, A.; Groisman, E.A.; Howard Hughes Medical Institute The PhoQ/PhoP regulatory network of *Salmonella enterica*. *Adv. Exp. Med. Biol.* **2008**, *631*, 7–21.
135. Bijlsma, J.J.E.; Groisman, E.A. The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica*. *Mol. Microbiol.* **2005**, *57*, 85–96.
136. Baisón-Olmo, F.; Galindo-Moreno, M.; Ramos-Morales, F. Host cell type-dependent translocation and PhoP-mediated positive regulation of the effector SseK1 of *Salmonella enterica*. *Front. Microbiol.* **2015**, *6*, 396.
137. Günster, R.A.; Matthews, S.A.; Holden, D.W.; Thurston, T.L.M. SseK1 and SseK3 T3SS effectors inhibit NF- κ B signalling and necroptotic cell death in *Salmonella*-infected macrophages. *Infect. Immun.* **2017**, doi:10.1128/IAI.00010-17.
138. Yang, Z.; Soderholm, A.; Lung, T.W.F.; Giogha, C.; Hill, M.M.; Brown, N.F.; Hartland, E.; Teasdale, R.D. SseK3 Is a *Salmonella* effector that binds TRIM32 and modulates the host's NF- κ B signalling activity. *PLoS ONE* **2015**, *10*, e0138529.
139. Singh, D.G.; Lomako, J.; Lomako, W.M.; Whelan, W.J.; Meyer, H.E.; Serwe, M.; Metzger, J.W. beta-Glucosylarginine: a new glucose-protein bond in a self-glucosylating protein from sweet corn. *FEBS Lett.* **1995**, *376*, 61–64.
140. Li, S.; Zhang, L.; Yao, Q.; Li, L.; Dong, N.; Rong, J.; Gao, W.; Ding, X.; Sun, L.; Chen, X.; et al. Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains. *Nature* **2013**, *501*, 242–246.
141. Pearson, J.S.; Giogha, C.; Ong, S.Y.; Kennedy, C.L.; Kelly, M.; Robinson, K.S.; Lung, T.W.F.; Mansell, A.; Riedmaier, P.; Oates, C.V.L.; et al. A type III effector antagonizes death receptor signalling during bacterial gut infection. *Nature* **2013**, *501*, 247–251.
142. Gao, X.; Pham, T.H.; Feuerbacher, L.A.; Chen, K.; Hays, M.P.; Singh, G.; Rueter, C.; Hurtado-Guerrero, R.; Hardwidge, P.R. *Citrobacter rodentium* NleB protein inhibits Tumor Necrosis Factor (TNF) receptor-associated factor 3 (TRAF3) ubiquitination to reduce host type I interferon production. *J. Biol. Chem.* **2016**, *291*, 18232–18238.
143. El Qaidi, S.; Chen, K.; Halim, A.; Siukstaite, L.; Rueter, C.; Hurtado-Guerrero, R.; Clausen, H.; Hardwidge, P.R. NleB/SseK effectors from *Citrobacter rodentium*, *Escherichia coli*, and *Salmonella enterica* display distinct differences in host substrate specificity. *J. Biol. Chem.* **2017**, *292*, 11423–11430.
144. Esposito, D.; Günster, R.A.; Martino, L.; El Omari, K.; Wagner, A.; Thurston, T.L.M.; Rittinger, K. Structural basis for the glycosyltransferase activity of the *Salmonella* effector SseK3. *J. Biol. Chem.* **2018**, *293*, 5064–5078.
145. Park, J.B.; Kim, Y.H.; Yoo, Y.; Kim, J.; Jun, S.-H.; Cho, J.W.; El Qaidi, S.; Walpole, S.; Monaco, S.; García-García, A.A.; et al. Structural basis for arginine glycosylation of host substrates by bacterial effector proteins. *Nat. Commun.* **2018**, *9*, 4283.
146. Ding, J.; Pan, X.; Du, L.; Yao, Q.; Xue, J.; Yao, H.; Wang, D.-C.; Li, S.; Shao, F. Structural and functional insights into host death domains inactivation by the bacterial arginine GlcNAcyltransferase effector. *Mol. Cell* **2019**, *74*, 922–935.e6.
147. Newson, J.P.; Scott, N.E.; Yeuk Wah Chung, I.; Wong Fok Lung, T.; Giogha, C.; Gan, J.; Wang, N.; Strugnelli, R.; Brown, N.F.; Cygler, M.; et al. *Salmonella* effectors SseK1 and SseK3 target death domain proteins in the TNF and TRAIL signaling pathways. *Mol. Cell. Proteom.* **2019**, doi:10.1074/mcp.RA118.001093.
148. Pan, M.; Li, S.; Li, X.; Shao, F.; Liu, L.; Hu, H.-G. Synthesis of and specific antibody generation for glycopeptides with arginine N-GlcNAcylation. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 14517–14521.
149. Park, J.B.; Yoo, Y.; Cho, H.-S. Structural insights showing how arginine is able to be glycosylated by pathogenic effector proteins. *BMB Rep.* **2018**, *51*, 609–610.
150. Ghosh, S.; O'Connor, T.J. Beyond paralogs: The multiple layers of redundancy in bacterial pathogenesis. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 467.
151. Gao, X.; Wang, X.; Pham, T.H.; Feuerbacher, L.A.; Lubos, M.-L.; Huang, M.; Olsen, R.; Mushegian, A.; Slawson, C.; Hardwidge, P.R. NleB, a bacterial effector with glycosyltransferase activity, targets GAPDH function to inhibit NF- κ B activation. *Cell Host Microbe* **2013**, *13*, 87–99.

152. Wong Fok Lung, T.; Giogha, C.; Creuzburg, K.; Ong, S.Y.; Pollock, G.L.; Zhang, Y.; Fung, K.Y.; Pearson, J.S.; Hartland, E.L. Mutagenesis and functional analysis of the bacterial arginine glycosyltransferase effector NleB1 from enteropathogenic *Escherichia coli*. *Infect. Immun.* **2016**, *84*, 1346–1360.
153. Cornick, N.A.; Pitzer, J.; Helgersson, A.F.; Madsen, M.L.; Kurth, K.T.; Xiao, Q.; Minion, F.C. Use of signature-tagged mutagenesis to identify genes associated with colonization of sheep by *E. coli* O157:H7. *Vet. Microbiol.* **2017**, *201*, 177–182.
154. Misyurina, O.; Asper, D.J.; Deng, W.; Finlay, B.B.; Rogan, D.; Potter, A.A. The role of Tir, EspA, and NleB in the colonization of cattle by Shiga toxin producing *Escherichia coli* O26:H11. *Can. J. Microbiol.* **2010**, *56*, 739–747.
155. Afset, J.E.; Bruant, G.; Brousseau, R.; Harel, J.; Anderssen, E.; Bevanger, L.; Bergh, K. Identification of virulence genes linked with diarrhea due to atypical enteropathogenic *Escherichia coli* by DNA microarray analysis and PCR. *J. Clin. Microbiol.* **2006**, *44*, 3703–3711.
156. Bugarel, M.; Martin, A.; Fach, P.; Beutin, L. Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) *Escherichia coli* strains: A basis for molecular risk assessment of typical and atypical EPEC strains. *BMC Microbiol.* **2011**, *11*, 142.
157. Baines, D.; Sumarah, M.; Kuldau, G.; Juba, J.; Mazza, A.; Masson, L. Aflatoxin, fumonisin and Shiga toxin-producing *Escherichia coli* infections in calves and the effectiveness of celmanax®/dairyman's choice™ applications to eliminate morbidity and mortality losses. *Toxins (Basel)* **2013**, *5*, 1872–1895.
158. Bugarel, M.; Beutin, L.; Fach, P. Low-density macroarray targeting non-locus of enterocyte effacement effectors (*nle* genes) and major virulence factors of Shiga toxin-producing *Escherichia coli* (STEC): A new approach for molecular risk assessment of STEC isolates. *Appl. Environ. Microbiol.* **2010**, *76*, 203–211.
159. Wickham, M.E.; Lupp, C.; Mascarenhas, M.; Vázquez, A.; Coombes, B.K.; Brown, N.F.; Coburn, B.A.; Deng, W.; Puente, J.L.; Karmali, M.A.; et al. Bacterial genetic determinants of non-O157 STEC outbreaks and hemolytic-uremic syndrome after infection. *J. Infect. Dis.* **2006**, *194*, 819–827.
160. Buvens, G.; Piérard, D. Virulence profiling and disease association of verocytotoxin-producing *Escherichia coli* O157 and non-O157 isolates in Belgium. *Foodborne Pathog. Dis.* **2012**, *9*, 530–535.
161. de Boer, R.F.; Ferdous, M.; Ott, A.; Scheper, H.R.; Wisselink, G.J.; Heck, M.E.; Rossen, J.W.; Kooistra-Smid, A.M.D. Assessing the public health risk of Shiga toxin-producing *Escherichia coli* by use of a rapid diagnostic screening algorithm. *J. Clin. Microbiol.* **2015**, *53*, 1588–1598.
162. Ferdous, M.; Friedrich, A.W.; Grundmann, H.; de Boer, R.F.; Croughs, P.D.; Islam, M.A.; Kluytmans-van den Bergh, M.F.Q.; Kooistra-Smid, A.M.D.; Rossen, J.W.A. Molecular characterization and phylogeny of Shiga toxin-producing *Escherichia coli* isolates obtained from two Dutch regions using whole genome sequencing. *Clin. Microbiol. Infect.* **2016**, *22*, 642.e1–642.e9.
163. Matussek, A.; Jernberg, C.; Einemo, I.-M.; Monecke, S.; Ehrlich, R.; Engelmann, I.; Löfgren, S.; Mernelius, S. Genetic makeup of Shiga toxin-producing *Escherichia coli* in relation to clinical symptoms and duration of shedding: a microarray analysis of isolates from Swedish children. *Eur. J. Clin. Microbiol. Infect. Dis.* **2017**, *36*, 1433–1441.
164. Vieira, M.A.; Dos Santos, L.F.; Dias, R.C.B.; Camargo, C.H.; Pinheiro, S.R.S.; Gomes, T.A.T.; Hernandez, R.T. Atypical enteropathogenic *Escherichia coli* as aetiological agents of sporadic and outbreak-associated diarrhoea in Brazil. *J. Med. Microbiol.* **2016**, *65*, 998–1006.
165. Xu, Y.; Bai, X.; Jin, Y.; Hu, B.; Wang, H.; Sun, H.; Fan, R.; Fu, S.; Xiong, Y. High prevalence of virulence genes in specific genotypes of atypical enteropathogenic *Escherichia coli*. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 109.
166. Buckner, M.M.C.; Croxen, M. a; Arena, E.T.; Finlay, B.B. A comprehensive study of the contribution of *Salmonella enterica* serovar Typhimurium SPI2 effectors to bacterial colonization, survival, and replication in typhoid fever, macrophage, and epithelial cell infection models. *Virulence* **2011**, *2*, 208–216.
167. Yang, Y.; Yu, C.; Ding, K.; Zhang, C.; Liao, C.; Jia, Y.; Li, J.; Cheng, X. Role of the *sseK1* gene in the pathogenicity of *Salmonella enterica* serovar Enteritidis in vitro and in vivo. *Microb. Pathog.* **2018**, *117*, 270–275.
168. Zhang, X.; He, L.; Zhang, C.; Yu, C.; Yang, Y.; Jia, Y.; Cheng, X.; Li, Y.; Liao, C.; Li, J.; et al. The impact of *sseK2* deletion on *Salmonella enterica* serovar Typhimurium virulence in vivo and in vitro. *BMC Microbiol.* **2019**, *19*, 182.
169. Lawley, T.D.; Chan, K.; Thompson, L.J.; Kim, C.C.; Govoni, G.R.; Monack, D.M. Genome-wide screen for *Salmonella* genes required for long-term systemic infection of the mouse. *PLoS Pathog.* **2006**, *2*, e11.

170. Nadler, C.; Baruch, K.; Kobi, S.; Mills, E.; Haviv, G.; Farago, M.; Alkalay, I.; Bartfeld, S.; Meyer, T.F.; Ben-Neriah, Y.; et al. The type III secretion effector NleE inhibits NF- κ B activation. *PLoS Pathog.* **2010**, *6*, e1000743.
171. Ruchaud-Sparagano, M.-H.; Mühlen, S.; Dean, P.; Kenny, B. The enteropathogenic *E. coli* (EPEC) Tir effector inhibits NF- κ B activity by targeting TNF α receptor-associated factors. *PLoS Pathog.* **2011**, *7*, e1002414.
172. Molloy, S. Bacterial pathogenesis: A sweet interaction with death for EPEC. *Nat. Rev. Microbiol.* **2013**, *11*, 659.
173. Pearson, J.S.; Hartland, E.L. A surprising sweetener from enteropathogenic *Escherichia coli*. *Gut Microbes* **2014**, *5*, 766–769.
174. Scott, N.E.; Giogha, C.; Pollock, G.L.; Kennedy, C.L.; Webb, A.I.; Williamson, N.A.; Pearson, J.S.; Hartland, E.L. The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). *J. Biol. Chem.* **2017**, *292*, 17337–17350.
175. El Qaidi, S.; Zhu, C.; McDonald, P.; Roy, A.; Maity, P.K.; Rane, D.; Perera, C.; Hardwidge, P.R. High-throughput screening for bacterial glycosyltransferase inhibitors. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 435.
176. Xu, C.; Liu, X.; Zha, H.; Fan, S.; Zhang, D.; Li, S.; Xiao, W. A pathogen-derived effector modulates host glucose metabolism by arginine GlcNAcylation of HIF-1 α protein. *PLoS Pathog.* **2018**, *14*, e1007259.
177. Blasche, S.; Arens, S.; Ceol, A.; Siszler, G.; Schmid, M.A.; Häuser, R.; Schwarz, F.; Wuchty, S.; Aloy, P.; Uetz, P.; et al. The EHEC-host interactome reveals novel targets for the translocated intimin receptor. *Sci. Rep.* **2014**, *4*, 7531.
178. Law, R.J.; Law, H.T.; Scurll, J.M.; Scholz, R.; Santos, A.S.; Shames, S.R.; Deng, W.; Croxen, M.A.; Li, Y.; de Hoog, C.L.; et al. Quantitative mass spectrometry identifies novel host binding partners for pathogenic *Escherichia coli* type III secretion system effectors. *J. Proteome Res.* **2016**, *15*, 1613–1622.
179. El Qaidi, S.; Scott, N.E.; Hays, M.P.; Geisbrecht, B. V.; Watkins, S.; Hardwidge, P.R. An intra-bacterial activity for a T3SS effector. *Sci. Rep.* **2020**, *10*, 1073.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).