Salmonella effectors SseK1 and SseK3 target death domain proteins in the TNF and TRAIL signaling pathways

Joshua P M Newson1,¶, Nichollas E Scott1,¶, Ivy Yeuk Wah Chung2, Tania Wong Fok Lung1,§, Cristina Giogha3, Jiyao Gan1,3, Nancy Wang1, Richard A Strugnell1, Nathaniel F Brown5, Miroslaw Cygler2, Jaclyn S Pearson3, and Elizabeth L Hartland1,3,4*

1 Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
2 Department of Biochemistry, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
3 Centre for Innate Immunity and Infectious Diseases, Hudson Institute of Medical Research, Clayton, Victoria, Australia
4 Department of Molecular and Translational Science, Monash University, Clayton, Victoria, Australia
5 Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada

¶ These authors contributed equally to this work
§ Present address: Department of Pediatrics, Columbia University at the Columbia University Medical Center, New York, New York, USA

Corresponding author: elizabeth.hartland@hudson.org.au Tel: (+61-3) 8572 2800

KEYWORDS

Salmonella, SseK, glycosyltransferase, death receptor signaling
ABSTRACT

Strains of *Salmonella* utilise two distinct type three secretion systems to deliver effector proteins directly into host cells. The *Salmonella* effectors SseK1 and SseK3 are arginine glycosyltransferases that modify mammalian death domain containing proteins with N-acetyl glucosamine (GlcNAc) when overexpressed ectopically or as recombinant protein fusions. Here, we combined Arg-GlcNAc glycopeptide immunoprecipitation and mass spectrometry to identify host proteins GlcNAcylated by endogenous levels of SseK1 and SseK3 during *Salmonella* infection. We observed that SseK1 modified the mammalian signaling protein TRADD, but not FADD as previously reported. Overexpression of SseK1 greatly broadened substrate specificity, while ectopic co-expression of SseK1 and TRADD increased the range of modified arginine residues within the death domain of TRADD. In contrast, endogenous levels of SseK3 resulted in modification of the death domains of receptors of the mammalian TNF superfamily, TNFR1 and TRAILR, at residues Arg\textsuperscript{376} and Arg\textsuperscript{293} respectively. Structural studies on SseK3 showed that the enzyme displays a classic GT-A glycosyltransferase fold and binds UDP-GlcNAc in a narrow and deep cleft with the GlcNAc facing the surface. Together our data suggest that salmonellae carrying *sseK1* and *sseK3* employ the glycosyltransferase effectors to antagonise different components of death receptor signaling.
INTRODUCTION

Pathogenic serovars of *Salmonella* utilise two type three secretion systems (T3SS), encoded by *Salmonella* pathogenicity island-1 and -2 (SPI-1 and SPI-2) to deliver distinct cohorts of effector proteins into host cells during infection (1, 2). These effector proteins subvert normal cellular processes and collectively enable the bacteria to invade and persist within host cells, partially through the manipulation of inflammatory cell signaling and programmed cell death (reviewed in (3, 4)). While the importance of effector translocation to pathogenesis is well established, the specific contribution of many effectors is still unclear. In particular, many effectors translocated by the SPI-2 encoded T3SS remain poorly characterised.

SseK1, SseK2, and SseK3 comprise a family of highly similar *Salmonella* effectors that are translocated by the SPI-2 T3SS during infection (5, 6). SseK family members show high sequence similarity to NleB1, a T3SS effector protein from enteropathogenic *Escherichia coli* (EPEC), which functions as an arginine glycosyltransferase and catalyses the addition of N-acetylglucosamine (GlcNAc) to arginine residues of the mammalian signaling adaptors FADD and TRADD, a modification termed Arg-GlcNAcylation (7, 8). A recent report provides evidence that SseK1 and SseK3, but not SseK2, also function as Arg-GlcNAc transferases (9). Mutation of a conserved DxD catalytic motif within SseK1 and SseK3 abrogates their glycosyltransferase activity (9), consistent with findings for NleB1 (7, 8). These studies describing catalytically important regions of the glycosyltransferases provide opportunities to better understand the function of these novel enzymes (10).

Although well recognised as glycosyltransferases, there are conflicting reports regarding the host substrates of the SseK effectors. One report suggested that recombinant SseK1 modifies recombinant TRADD *in vitro* (8), whereas a subsequent report suggested SseK1 modifies both TRADD and FADD when co-expressed ectopically in mammalian cell lines (9). Yet another report
using *in vitro* glycosylation assays suggested that recombinant SseK1 glycosylates GAPDH but not FADD (11). SseK3 on the other hand was reported to bind but not modify the E3-ubiquitin ligase TRIM32 (12), and was shown to weakly modify TRADD but not FADD (9). While these studies provide useful insights, the true role of these effectors is better interrogated through non-biased screens conducted under conditions that reflect endogenous levels of the effectors and host proteins (13).

Here, we explored the endogenous Arg-GlcNAc glycosyltransferase activity of SseK1 and SseK3 during *Salmonella* infection. Using a mass spectrometry-based approach to enrich arginine GlcNAcylated peptides from infected host cells (13), we found that SseK1 modified the signaling adaptor TRADD, while SseK3 modified the signaling receptors TNFR1 and TRAILR. In addition, we performed structural studies on SseK3 and showed that the enzyme displays a classic GT-A glycosyltransferase fold and binds UDP-GlcNAc in a narrow and deep cleft with the GlcNAc facing the surface. Together these studies suggest that *Salmonella* has evolved multiple means to manipulate death receptor signaling through the acquisition of arginine glycosyltransferases with differing substrate specificities. Furthermore, we have determined the crystal structure of SseK3 alone and with the UDP-GlcNAc substrate. The UDP is retained in the active site while GlcNAc is cleaved under crystallization conditions. However, in the structure of E258Q mutant the GlcNAc is retained within the binding site although cleaved of the UDP. A long, negatively charged depression indicates the entry site of the arginine sidechain for glycosylation.
RESULTS

Mutation of a conserved glutamic acid abrogates the catalytic activity of SseK1 and SseK3.

Previous reports have indicated that SseK1 and SseK3 function as arginine glycosyltransferases which modify a conserved arginine residue in the death domains of several mammalian immune signaling proteins (8, 9). Mutation of a conserved DxD catalytic motif within SseK1 (Asp^{223}-Ala^{224}-Asp^{225}) and SseK3 (Asp^{226}-Ala^{227}-Asp^{228}) abrogated their glycosyltransferase activity (9), consistent with findings for the homologous T3SS effector, NleB1, from EPEC (7, 8). Mutation of a single glutamic acid (Glu^{253}) also impairs the ability of NleB1 to inhibit NF-κB activation following TNF stimulation of mammalian cells (8), and renders NleB1 catalytically inactive (10). Here, we confirmed the importance of this conserved glutamic acid to the biochemical activity of SseK1 and SseK3 by infecting RAW264.7 cells with a *Salmonella* Typhimurium SL1344 triple ΔsseK123 deletion mutant (6) complemented with either native SseK1, -2 or -3, catalytic triad mutants, or mutants lacking the conserved glutamic acid (SseK1_{E255A}, SseK2_{E271A}, and SseK3_{E258A}). Cell lysates were assessed by immunoblot using an antibody specific for arginine glycosylation (Fig. 1A and B). Native SseK1 and SseK3, but not SseK2, catalysed Arg-GlcNAcylation, while all catalytic triad mutants and glutamic acid mutants showed no activity. Notably, overexpression of both SseK1 and SseK3 increased levels of arginine GlcNAcylation relative to wild-type SL1344, suggesting that overexpression may cause non-authentic Arg-GlcNAcylation of host substrates.

The host signaling adaptor TRADD is the preferred substrate of SseK1. A number of reports have used *in vitro* experiments to identify possible substrates of SseK1 (8, 9, 11). While these studies have demonstrated the ability of SseK1 to GlcNAcylate a given target, the approach is substrate specific and assumes prior knowledge of the potential targets. To identify the full range of host Arg-GlcNAcylated substrates, we developed a quantitative approach using mass spectrometry (MS) to enrich for arginine glycosylated peptides (13). Quantitative mass spectrometry was performed to compare the triple sseK123 mutant complemented with sseK1 or catalytically inactive sseK1_{E255A} (Fig. 2A), where levels of expression upon complementation were similar (Fig. 1A).
Using peptides derived from RAW264.7 cells infected with S. Typhimurium ΔsseK123 over-expressing either SseK1 or inactive SseK1E255A, label-free MS based quantification revealed a broad range of substrates that were only modified in the presence of active SseK1 (Fig. 2A, Supplemental Table 1, Supplemental Fig. 1A). Among these substrates, mouse TRADD (mTRADD) was Arg-GlcNAcylated at Arg243 (Fig. 2B), which represented a novel site of modification. Unexpectedly, we also detected Arg-GlcNAcylation of a large number of S. Typhimurium proteins, among them the two-component response regulators OmpR and ArcA, as well as the ribosome-associated proteins RpsA and RbfA, and transcription-associated proteins RpoD and NusA.

We have previously shown that over-expression of NleB1 leads to enhanced levels of cellular arginine GlcNAcylation (13). Given that the overexpression of SseK1 and SseK3 also appeared to increase the range of substrates relative to endogenous wild type levels of expression (Fig 1A and B), we explored the activity of endogenous SseK1 during S. Typhimurium infection of RAW264.7 cells using our strategy for enrichment of Arg-GlcNAcylated peptides (13). Here we compared Arg-GlcNAcylation during infection with a ΔsseK23 double deletion mutant versus a ΔsseK123 triple deletion mutant (6). In addition to data dependent MS acquisition to enhance the detection of Arg-GlcNAcylation events, we applied parallel reaction monitoring, a targeted high-resolution MS approach (14), to specifically observe the Arg-GlcNAcylated species of mTRADD and FADD (mFADD) (Fig. 2C, Supplemental Table 2, Supplemental Fig. 1B). Using this approach, we found that mTRADD was glycosylated at Arg233 (Fig. 2D). Arg233 is the equivalent of the previously reported Arg235 of human TRADD (hTRADD) which is GlcNAcylated by NleB1 from EPEC in vitro (8). There was no evidence of Arg-GlcNAcylated mFADD, despite the fact that mFADD was readily detectable within the input control samples (Supplemental Table 3, Supplemental Fig. 2). This suggested TRADD was the preferred substrate when SseK1 was translocated at native levels during infection. We confirmed that the same was true for SseK1 in human cells. HeLa229 and THP-1 were infected with ΔsseK23 and ΔsseK123 to compare GlcNAcylation by endogenous levels
of SseK1. One dominant band was observed with anti-Arg-GlcNAc antibodies that also reacted with anti-TRADD antibodies by immunoblot (Supplemental Fig. 3).

**Overexpression of SseK1 alters the sites of glycosylation within the death domain of TRADD.**

A previous report demonstrated that mTRADD carrying the amino acid substitution Arg^{233} was Arg-GlcNAcylated at comparable levels to wild-type TRADD when co-expressed ectopically with GFP-SseK1 in mammalian cells (9). Our data indicated that mTRADD was glycosylated at Arg^{243} when SseK1 was overexpressed, but that Arg^{233} was the site of glycosylation when SseK1 was expressed at native levels during infection. The equivalent sites in human TRADD (hTRADD) are Arg^{235} and Arg^{245}. To determine the dominant modification site in hTRADD, we generated single and double Arg^{235} and Arg^{245} mutants and expressed these ectopically in HEK239T cells co-transfected with pEFGP-SseK1. Both Flag-hTRADD_{R235A} and Flag-hTRADD_{R245A} were Arg-GlcNAcylated at levels comparable to native Flag-hTRADD, while modification of the double mutant Flag-hTRADD_{R235A/R245A} was significantly reduced (Fig. 3A). To identify further possible sites of modification, Flag-hTRADD was enriched from cells co-transfected with pEGFP-SseK1 by anti-Flag immunoprecipitation, subjected to tryptic digestion and analysed by LC-MS (Fig. 3B, Supplemental Table 4). Under these conditions, we detected several further sites of modification at Arg^{239}, Arg^{278}, and Arg^{224}, suggesting that over-expression broadened the range of possible glycosylation sites within the single substrate. The anti-Flag enrichment and LC-MS approach was repeated with Flag-hTRADD_{R235A}, Flag-hTRADD_{R245A}, and Flag-hTRADD_{R235A/R245A} variants (Supplemental Table 4). We detected different patterns of Arg-GlcNAcylation for each of these mutant Flag-hTRADD variants, suggesting that deletion of a preferred glycosylation site caused a shift in site specificity despite comparable protein levels, as determined by MS analysis (Supplemental Table 5, Supplemental Fig 4). These data indicated that SseK1 was capable of modifying a range of arginine residues when expressed ectopically, and that over expression of SseK1 and related effectors may not replicate natural effector activity.
SseK3 glycosylates a conserved arginine residue in the mammalian death receptors TNFR1 and TRAILR. Similar to SseK1, we found that overexpression of SseK3 greatly increased arginine glycosylation activity relative to the levels generated by the bacterium during wild type *S. Typhimurium* infection (Fig. 1B). Hence, we applied our peptide enrichment strategy to identify Arg-GlcNAcylated substrates in the presence of native levels of SseK3. Arg-GlcNAcylated peptides were enriched from RAW264.7 cells infected with either a double ΔsseK12 mutant or a triple ΔsseK123 mutant, and we applied label free MS based quantification to screen for glycosylation events in a non-biased manner (Fig. 4A, Supplemental Table 6). Under these conditions, SseK3 modified specific arginine residues of mouse TNFR1 (mTNFR1) and TRAILR (mTRAILR), both death domain-containing receptors of the TNF superfamily (15). Arg-GlcNAcylation was observed in all three biological replicates of cells infected with the ΔsseK12 deletion mutant, while no Arg-GlcNAcylation was detected in cells infected with the ΔsseK123 deletion mutant (Fig. 4B, Supplemental Table 6). It should be noted all Arg-GlcNAcylation pulldowns demonstrated a high, >0.7, Pearson correlation, consistent with a high level of reproducibility between all biological replicates (Supplemental Fig. 5). Glycosylation of mTRAILR occurred at Arg^{293} while mTNFR1 was glycosylated at Arg^{376}. An alignment of protein sequences demonstrated that these sites correspond to a conserved arginine in the death domains of both proteins (Figs. 4C and D) and corresponds to the conserved arginine targeted by NleB1. No other Arg-GlcNAcylated peptides were detected under these conditions, suggesting that TRAILR and TNFR1 were the preferred substrates of SseK3.

To determine if SseK3 also modified human TRAILR, the Flag tagged death domain of hTRAILR2 (Flag-hTRAILR2\textsubscript{DD}) was enriched by anti-Flag immunoprecipitation from HEK293T cells co-transfected with pEGFP-SseK3. Of the four human isoforms of TRAILR, we focussed on hTRAILR2 as it shows the strongest sequence similarity to mTRAILR (16). Using the anti-Arg-GlcNAc antibody, we detected Arg-GlcNAcylation of Flag-hTRAILR2\textsubscript{DD} by GFP-SseK3 but not...
GFP-SseK3\textsubscript{E258A} (Fig. 5A). Similarly, Flag-hTNFR1\textsubscript{DD} was Arg-GlcNAcylated by GFP-SseK3 but not GFP-SseK3\textsubscript{E258A} (Fig. 6A).

To validate the specific site of modification, recombinant GST-SseK3 was incubated with His-hTRAILR\textsubscript{2DD} in the presence of the sugar donor UDP-GlcNAc, subjected to tryptic digestion and analysed by LC-MS (Supplemental Table 7 and 8). We detected hTRAILR\textsubscript{2DD} modified at Arg\textsuperscript{359} (Fig. 5B, 5C), equivalent to the modification of mTRAILR at Arg\textsuperscript{293} described above. In the absence of UDP-GlcNAc, the modification of hTRAILR\textsubscript{2DD} was not observed (Fig. 5D). For both experiments the unmodified peptide \textsuperscript{396}DASVHTLLDALETGER\textsuperscript{412} derived from mTRAILR was monitored as an internal control and showed comparable ion intensity within the sample input which was consistent with the total observed protein levels between samples (Supplemental Table 8). Within these digests, we also observed evidence of self Arg-GlcNAcylation of SseK3, even without the addition of UDP-GlcNAc (Supplemental Table 7, Supplemental Fig. 6-9) and modification of other endogenous \textit{Salmonella} proteins (Supplemental Fig. 10). Immunoblots using anti-Arg-GlcNAc antibodies confirmed the glycosylation of hTRAILR\textsubscript{2DD} in the presence of GST-SseK3 and UDP-GlcNAc as well as the apparent self-modification of SseK3 (Fig. 5E).

In vitro validation of hTNFR1 Arg-GlcNAcylation by GST-SseK3 was complicated by the insoluble nature of recombinant His-hTNFR1\textsubscript{DD}. In addition, since the overexpression of death domain containing proteins increases susceptibility to cell death, thereby affecting consistency of transfection and protein expression, Flag-hTNFR1\textsubscript{DD} was enriched from HEK293T cells by anti-Flag immunoprecipitation following co-transfection with pEGFP-SseK3 (Fig. 6A). This was also designed to overcome the high background we observe upon overexpression of SseK3 (Fig. 1B), here as EGFP-SseK3. Protein samples were subjected to tryptic digestion ahead of analysis by LC-MS and Arg\textsuperscript{376} was confirmed as the preferred site of GlcNAcylation of Flag-hTNFR1\textsubscript{DD} (Fig. 6B,
Supplemental Table 9). These experiments confirmed modification of the death domains of human TRAILR and TNFR1 by SseK3.

To explore interactions between SseK3 and the novel host proteins, TNFR1 and TRAILR, the yeast two-hybrid system was used to detect protein-protein interactions. Auxotrophic yeast strains were co-transformed to express SseK3 and the death domain of hTNFR1 (hTNFR1DD). Yeast expressing SseK3 and hTNFR1DD grew when plated on selective media, indicating a stable interaction between SseK3 and the hTNFR1DD in the yeast (Fig. 6C). In contrast, yeast co-transformed to express SseK3 and the death domain of hTRAILR2 (hTRAILR2DD) did not grow on selective media (Fig. 6C). Despite our inability to detect this interaction in the yeast, GlcNAcylation of hTRAILR2 by SseK3 requires an interaction between the target protein and the enzyme, even if weak and/or transient.

**Crystal structure of SseK3.** To get a better understanding of the catalytic mechanism of SseK3 and related Arg-GlcNAc glycosyltransferases and the importance of the E258 residue, we solved the structure of SseK3(25-335) co-crystallized with UDP-GlcNAc and the E258Q mutant with UDP-GlcNAc. The enzyme belongs to the α/β class and has a core GT-A glycosyltransferase fold, with a central six-stranded β-sheet (β5-β4-β3-β6-β11-β8) (Fig. 7A). There are two α-helices parallel to strands β3-β5 and two other helices parallel to the other three β-strands (β6-β11-β8) but on the opposite side of the β-sheet. The long connection between strands β5 and β6 of the central β-sheet contains several helices, two of which, α3-α4, form an α-helical hairpin that is somewhat separated from the rest of the protein and distant from the active site (Fig. 7A, B). Although the enzyme was co-crystallized with UDP-GlcNAc, only UDP is visible in the electron density of the native protein and there is no metal ion present (not added in the crystallization). UDP is located within a large, L-shaped groove with uridine filling the end of the short arm and the diphosphate extending to the middle corner of the L. The C-terminus of SseK3 is near the groove but the last five residues are disordered. In the SseK3 (E258Q) mutant, which crystallized in a different space group and in the
presence of Mg\(^{2+}\) and UDP-GlcNAc, the C-terminus is ordered and extends over the UDP binding site, partially covering it from the solvent. The SseK3 E258Q mutant retained some glycosyltransferase activity since UDP-GlcNAc was also hydrolysed and only UDP was clearly visible in the crystal. However, an additional density was present nearby at the bottom of the long arm of L-groove and we have modelled GlcNAc into this density. In our structures, the UDP uracil O2 forms a hydrogen bond with the backbone amide of Phe53 while the uracil N3 atom forms a hydrogen bond with the backbone carbonyl oxygen of the same Phe53 (Fig. 7C). As well, the sidechains of Arg55 and Arg59 are close to the uracil O4. Moreover, the uracil ring forms a π-stacking interaction with Trp52 and Phe190. The ribose also forms hydrogen bonds to the protein, namely, C2 hydroxyl to the carbonyl of Gln51 and to the hydroxyl of Tyr224, while the C3 hydroxyl is hydrogen bonded to the backbone amide of Ala227. In the presence of Mg\(^{2+}\) ion (in E258Q mutant), both phosphates ligand the ion together with Asp228 (from the DxD\(^{228}\) motif), Asp325 and Ser327 (Fig. 7D). In addition, the α-phosphate hydrogen bonds to OG of Ser333 and β-phosphate hydrogen bonds to OG of Ser333 and NE1 of Trp334. In the absence of the metal ion (in the native structure) UDP still binds to the enzyme and uridine makes the same contacts. The phosphates are shifted ~2 Å away from the DxD motif and make no close contacts with the protein. In this structure, Tyr334 is disordered.

Comparison of all independent molecules from crystal structures we determined showed that they superimpose with root-mean-squares deviation (rmsd) of 0.4-0.65 Å. The segments that deviated the most were the tip of the α-helical insertion mini-domain (aa 143-158) and the C-terminus, which is either well-ordered covering the bound UDP or completely disordered.
DISCUSSION

Interrogating the biochemical activity of bacterial effector proteins, their preferred host substrates, and contribution to virulence, remains a major priority of host-pathogen research to help elucidate basic processes that are manipulated by highly evolved pathogens, and to provide for new opportunities to disrupt disease causing processes. A number of studies have described putative substrates of the *Salmonella* Typhimurium effectors SseK1 and SseK3 (8, 9, 11, 12) but these studies have relied substantially on *in vitro* experiments using recombinant proteins or overexpression of the effectors in mammalian cells. Here, we identified the substrates of SseK1 and SseK3 when expressed at native levels during *S.* Typhimurium infection of RAW264.7 cells. Our initial approach was to examine endogenous levels of the effectors to detect the modification of host proteins under native infection conditions, in order to infer a hierarchy of host protein targeting and/or to sort true targets from artefacts of overexpression. This was followed by more detailed and targeted biochemistry and mutagenesis studies to confirm modification sites, which subsequently required the use of *in vitro* and mammalian cell overexpression systems.

Although previous reports had suggested both TRADD and FADD were substrates of SseK1 (8, 9, 11), our data suggested that TRADD is the preferred substrate of SseK1, as we could not detect modification of FADD during *S.* Typhimurium infection. TRADD plays a key role in the activation of canonical NF-κB signaling leading to pro-inflammatory cytokine secretion (reviewed in (17, 18)) and programmed cell death via TNF induced apoptosis or necroptosis (19). Previous work has suggested that SseK1 plays a role in inhibiting both NF-κB activation and necroptotic cell death in infected macrophages (9), and so it is likely that *S.* Typhimurium employs SseK1 to inhibit TNF signaling as required. Despite this, single deletion mutants of *S.* Typhimurium have not clearly demonstrated a requirement for SseK1 in mouse infection models (5, 20, 21). We speculate that SseK1 acts in concert with other effectors *in vivo* to achieve significant inhibition of NF-κB signaling during *S.* Typhimurium infection.
Although SseK3 had been reported to interact with TRADD and TRIM32 (9, 12), here we found that the preferred substrates of SseK3 were TNFR1 and TRAILR, both members of the mammalian TNF receptor superfamily. TNFR1 responds to stimulation by extracellular TNF, and initiates a signaling cascade culminating in either inflammatory cytokine production or programmed cell death, as described above. Similarly, extracellular TRAIL binds the membrane-associated receptors TRAILR1 (22), TRAILR2 (23), TRAILR3 (24), and TRAILR4 (25). However, only TRAILR1 and TRAILR2 contain full length death domains that stimulate a range of signaling events, including inflammatory cytokine production, extrinsic apoptotic cell death via caspases-8 and -3, necroptosis or even the promotion of cell survival via anti-apoptotic functions mediated by TRAF2 ubiquitination of caspase-8 (26, 27) (reviewed in (16, 28, 29)). This diversity of signaling outcomes complicates the interrogation of bacterial manipulation of TRAIL signaling. However, TRAIL-R deficient mice showed no difference in susceptibility to intraperitoneal infection with Salmonella compared to wild type mice (30), and we confirmed there was no difference in susceptibility of TRAIL-R deficient mice during systemic (Supplemental Fig. 11) or oral (data not shown) infection. This suggests that TNFR1 may be a more important target for bacterial inhibition in mammalian hosts. Consistent with this, TNFR1 deficient mice are highly susceptible to Salmonella infection (31). Interestingly, a number of studies have implicated TRAIL polymorphisms with increased susceptibility to Salmonella infection in chickens (32, 33), suggesting TRAIL may activate an important immune defence pathway to Salmonella in poultry that SseK3 activity may block.

In addition to the Arg-GlcNAcylation of host substrates, we also observed the apparent self-modification of SseK1 and SseK3 when overexpressed. This observation, in conjunction with increased levels of Arg-GlcNAcylation seen during overexpression, suggested that many substrates may be non-authentically glycosylated when the effectors are overexpressed. These findings are similar to a previous report that suggested overexpression of NleB1 also results in Arg-GlcNAcylation of a broader range of substrates (13). We note in particular the numerous two-
component response regulators that were Arg-GlcNAcylated during overexpression of SseK1 (Fig 2A, Supplemental Table 1), perhaps suggesting a mechanism for effector-mediated regulation of two-component signaling outcomes. Surprisingly, these proteins also appeared to be Arg-GlcNAcylated by SseK1 and SseK3 during growth in LB (Supplemental Fig. 10, Supplemental Table 10). Since these enrichment approaches did not provide information on the stoichiometry of these Arg-GlcNAcylation events, we attempted to interrogate this in Salmonella lysates without enrichment. However, this approach failed to detect any Arg-GlcNAcylation events, even for highly abundant proteins such as TufB and GlpK which we had found were modified following enrichment (Supplemental Fig. 10). Overall this suggested that these events occur at low stoichiometry, and therefore the biological relevance of such glycosylation events is yet to be established.

The crystal structure of SseK3 indicated that the enzyme has a GT-A glycosyltransferase fold and the closest structural homologs are enzymes from Carbohydrate-Active enZYmes Database (CAZy, http://www.cazy.org/) family GT44 and GT88, both containing bacterial retaining glucosyltransferases (34) (Fig. 7E). The structure showed the location of the substrate binding site as well as position of the active site residue, Glu258. The substrate binding groove has an L-shape and displays a negative electrostatic potential as would be expected to receive a positively charged arginine sidechain. The short arm of the groove is occupied by the UDP-GlcNAc and the long arm is most likely the binding site for the arginine sidechain. The E258Q mutant displayed some residual activity that prevented capturing the intact UDP-GlcNAc in the binding site but showed that in the absence of the acceptor, the GlcNAc is at least partially retained in the groove. Residual activity for the Glu-to-Gln mutation of a catalytic nucleophile was also observed in other retaining glucosyltransferases, for example, E317Q mutant of α1,3-galactosyltransferase (35). The mobile C-terminus likely forms a gate that allows for an easy access of UDP-GlcNAc to the site and helps to retain it there until the acceptor arginine binds nearby and the transfer reaction is completed. The location and conformation of UDP observed in our structure superimposes very well with UDP-Glc
from the crystal structure of one of the closest structural homologs, GT44 family *Clostridium difficile* toxin A (TcdA) glucosyltransferase domain (PDB code 3SRZ, (36)) (Fig. 7G). Moreover, the position of GlcNAc observed in the SseK3E258Q structure was in the vicinity of the Glc of UDP-Glc in toxin A. In this transferase, the UDP-Glc binding site is also covered by a loop that is in a different conformation in the absence of a UDP donor. The UDP conformation in SseK3 is also the same as UDP-GlcNAc in the structure of rabbit *N*-acetylglucosaminyltransferase I (PDB code 1FOA, (37)). In order to model a possible placement of the substrate arginine, we have superimposed UDP-GlcNAc from the latter structure on UDP. The GlcNAc fits snugly into the groove and leaves space for the arginine, which we placed in the long arm of the substrate binding groove (Fig. 7E). It appears that this position of arginine is plausible for the transferase reaction.

The role of the inserted mini-domain is unclear at the present time. Based on sequence alignment, this segment is present not only in SseK3 but also in SseK1/2 and NleB1/2 and we surmise that it plays a role in protein substrate recognition.

While this work was in progress, the structure of SseK3 crystallized under somewhat different conditions and in a different space group was published (38) and the structures of NleB2, SseK1 and SseK2 were published (39). Comparison of all these structures shows flexibility of the ~10 C-terminal residues, particularly in the absence of UDP. When ordered, this segment covers the active and substrate binding sites, strengthening the hypothesis that these residues participate in UDP-GlcNAc and substrate binding and release. The substrate binding to these proteins has yet to be established and the molecular basis for differences in their target specificities is not yet known. A comparison of the structures of SseK3 and SseK1 shows a stronger electrostatic potential around the substrate binding site in SseK1 (Supplemental Fig. S12A), together with a more open binding site (Supplemental Fig. 12B). This is caused by a shift of the 249-260 loop away from the substrate binding site in SseK1, while the cap (aa 130-177) is in nearly the same relative position. These differences likely affect the specificity for the substrate protein.
In summary, here we identified the endogenous host substrates of SseK1 and SseK3 during Salmonella infection. The dominant substrate of SseK1 was the signaling adaptor TRADD, while SseK3 Arg-GlcNAcylated the death domains of both TRAILR and TNFR1 at a conserved arginine residue. This suggests that Salmonella utilises the SseK effectors to antagonise multiple components of death receptor pathways, thereby subverting inflammatory and cell death responses in vivo.
MATERIALS AND METHODS

Strains and growth conditions. The strains used in this study are listed in Table 1. Bacteria were grown with shaking at 37°C in Luria-Bertani (LB) broth in the presence of ampicillin (100 µg/ml), streptomycin (50 µg/ml), or kanamycin (100 µg/ml) when required.

DNA cloning and purification. The plasmids and primers used in this study are listed in Tables 2 and 3, respectively. DNA-modifying enzymes were used in accordance with the manufacturer’s instructions (New England BioLabs). Plasmids were extracted using the QIAGEN QIAprep Spin Miniprep Kit. PCR products and restriction digests were performed using the Wizard SV Gel and PCR Clean-Up System (Promega). pTrc99A-SseK1, pTrc99A-SseK2, and pTrc99A-SseK3 were constructed by amplifying sseK1, sseK2, and sseK3 from pEGFP-C2-SseK1, pEGFP-C2-SseK2, and pEGFP-C2-SseK3 using the primer pairs SseK1F/R, SseK2F/R, and SseK3F/R, respectively. The PCR product was digested with EcoRI and BamHI and ligated into pTrc99A to produce a C-terminal 2x hemagglutinin tag fusion to SseK1, SseK2, and SseK3. Constructs were transformed into XL-1B cells, and verified by colony PCR and sequencing using the primer pair pTrc99A_F/R. pEGFP-C2-SseK2 and pEGFP-C2-SseK3 were constructed by amplifying sseK2 and sseK3 from S. Typhimurium SL1344 genomic DNA using the primer pairs GFPS2F/R and GFPS3F/R respectively and AmpliTaq Gold DNA polymerase. The resultant PCR products were purified and ligated into pGEM-T-Easy vector at an insert:vector molar ratio of 3:1. The ligation reactions were transformed into XL-1 Blue cells and plated onto LA plates containing ampicillin and X-gal. Plasmids were extracted and digested with EcoRI and SalI to release the bacterial genes, which were gel purified and ligated into pre-digested pEGFP-C2. The ligation reactions were then transformed into XL-1 Blue cells and colony PCR was performed using primers pEGFP-C2_F/R to select positive clones. The correct insert was confirmed by sequencing using the same primer pair. The primer pair hTRAILR2DD-F/R was used to amplify the region encoding the death domain of TRAILR2 from HeLa cDNA (Sigma Aldrich). The resulting amplicon was gel purified and digested with EcoRI and BamHI and ligated into pre-digested pGADT7. The ligation reactions were transformed into XL-1
Blue cells and plated on LA containing ampicillin. The correct insert was verified by colony PCR and sequencing with the primer pair pGADT7-AD
F/R. pFlag-hTRAILR2
DD was constructed by amplifying the death domain of human TNFRSF10B from pGADT7- hTRAILR2
DD using the primer pair hTRAILR2
DD-F/ hTRAILR2
DD-R. The PCR product was digested with EcoRI and BamHI and ligated into p3xFlag-Myc-CMV-24 to produce an N-terminal 3xFlag fusion to hTRAILR2
DD. pFlag-hTRAILR2
DD was transformed into XL-1B cells, and verified by colony PCR and sequencing using the primer pair p3xFlag-Myc-CMV-24
F/R. pGEX-4T-1-SseK3 was constructed by amplifying sseK3 from pEGFP-C2-SseK3 using the primer pair GST-SseK3
F/R. The PCR product was digested with EcoRI and SalI and ligated into pre-digested pGEX-4T-1 to produce an N-terminal GST fusion to SseK3. pGEX-4T-1-SseK3 was transformed into XL-1B cells, and verified by colony PCR and sequencing using the primer pair pGEX-4T-1
F/R. The region encoding the death domain of TNFR1 was amplified using the primer pair FLAGTNFR1DD
F/R and pGADT7-TNFR1
DD as template. The resulting amplicon was gel purified and digested with BglII and SalI and ligated into pre-digested p3xFLAG-Myc-CMV-24 before transforming the reactions into XL-1 Blue cells. Colony PCR and sequencing were performed with the primer pair p3xFlag-Myc-CMV-24
F/R to ensure the correct insert has been ligated. pGBK7-NleB1 was constructed by digesting pGBT9-NleB1 (7), which carries nleB1 flanked in between the restriction sites EcoRI and BamHI, and ligating into pGBK7 digested with EcoRI and BamHI. pGBK7-SseK3 was constructed by amplifying sseK3 from pEGFP-C2-SseK3 using the primer pair SseK3
F/R. The PCR product was ligated into the cloning vector pGEM-T-Easy and the ligation reaction was transformed into XL-1 Blue cells. Transformants were then selected on LA containing ampicillin and X-gal. Plasmids were extracted and digested with EcoRI and SalI to release the bacterial gene, which was then ligated into pre-digested pGBK7 to create pGBK7-SseK3. This plasmid was sequenced using the primer pair pGBK7
F/R. pGADT7-FADD
DD, pGADT7-TNFR1
DD, and pGADT7-hTRAILR2
DD were constructed by amplifying the death domain regions of human FADD, TNFR1, and TRAILR2 from HeLa cDNA using the primer pairs FADD
DD-F/R, TNFR1
DD-F/R, and
hTRAILR2DD-F/R, respectively. Amplified FADD was digested with EcoRI and BamHI, TNFR1DD was digested with NdeI and BamHI, and hTRAILR2DD was digested with EcoRI and BamHI. Digested PCR products were ligated into pre-digested pGADT7 and transformed into XL-1 Blue cells. Constructs were verified by colony PCR and sequencing using the primer pair pGADT7F/R. For structural investigations, the segment corresponding to residues 25-335 of sseK3 from Salmonella Typhimurium (strain SL1344) (Uniprot: A0A0H3NMP8) was cloned into vector pRL652, a derivative of vector pGEX-4T-1 (GE Healthcare) adapted for ligation-independent cloning. The construct contained a TEV-cleavable GST tag at the N-terminus.

Site-directed mutagenesis. Site-directed mutagenesis of plasmid constructs was performed using the QuikChange II Site-Directed Mutagenesis kit (Stratagene, California, USA), according to the manufacturer’s instructions. pTrc99A-SseK1DxD(229-231)AAA, pTrc99A-SseK2DxD(239-241)AAA, and pTrc99A-SseK3DxD(226-228)AAA were generated using pTrc99A-SseK1, pTrc99A-SseK2, or pTrc99A-SseK3 as template DNA and amplified by PCR using the primer pairs SseK1DxD-F/R, SseK2DxD-F/R, or SseK3DxD-F/R, respectively. pTrc99A-SseK1E255A, pTrc99A-SseK2E271A, and pTrc99A-SseK3E258A were generated using pTrc99A-SseK1, pTrc99A-SseK2, or pTrc99A-SseK3 as template DNA and amplified by PCR using the primer pairs SseK1E255A-F/R, SseK2E271A-F/R, or SseK3E258A-F/R, respectively. pFlag-hTRADD2 and pFlag-hTRADD2 were generated using pFlag-hTRADD as template DNA and amplified by PCR using the primer pairs hTRADD2-F/R and hTRADD2-F/R, respectively. pFlag-hTRADD2 was generated using pFlag-hTRADD2 as template DNA and amplified by PCR using the primer pair hTRADD2-F/R.

All PCR products were digested with DpnI at 37°C overnight before subsequent transformation into XL1-B cells. Plasmids were extracted and sequenced using the primer pairs pTrc99AF/R, p3xFlag-Myc-CMV-24F/R, or pEFGP-C2F/R, as required.

Mammalian cell culture. HEK293T cells (human embryonic kidney 293 cells expressing the SV40 large T-antigen, source: ATCC® CRL-3216), RAW264.7 cells (murine leukemic monocyte-macrophage cells, source: Richard Strugnell, University of Melbourne) and HeLa 229 cells (human
cervical carcinoma cells, source: Hayley J. Newton, University of Melbourne) were maintained in DMEM, low glucose with GlutaMAX™ supplement and pyruvate (DMEM (1X) + GlutaMAX(TM)-I) (Gibco, Life Technologies, NY, USA). THP-1 cells (human leukemic monocytes, source: Ashley Mansell, Hudson Institute) were maintained in RPMI medium 1640, with GlutaMAX™ supplement (RPMI (1X) + GlutaMAX(TM)-I) (Gibco, Life Technologies, NY, USA). Tissue culture media was further supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (Thermo Fisher Scientific). Cells were maintained in a 37 °C, 5% CO₂ incubator, and passaged to a maximum of 30 times. HEK293T cells and HeLa 229 cells were split when cells reached 80 to 90% confluency with 1 ml 0.05% Trypsin-EDTA (1X) (Gibco, Life Technologies) per 75 cm² of tissue culture, then resuspended with 10 volumes of DMEM supplemented with FBS. RAW264.7 cells were physically detached with a cell scraper and further diluted in fresh DMEM supplemented with FBS. THP-1 cells were split when the cell density reached 5 x 10⁵ to 8 x 10⁵ cells/ml per 75 cm² of tissue culture, then seeded at 1.5 x 10⁵ cells/ml with RPMI supplemented with FBS in 75 cm² of tissue culture flasks.

**Infection of mammalian cell lines.** RAW264.7 cells were seeded to 24 well plates at a concentration of 3 x 10⁵ cells per well one day before infection. HeLa 229 cells were seeded to 6 well plates at a concentration of 2.5 x 10⁶ per well the day before infection. THP-1 cells were seeded to 6 well plates at a concentration of 2.5 x 10⁶ per well three days before infection, 60 ng/ml of phorbol 12-myristate 13-acetate (Enzo Life Sciences) was supplemented in the RPMI medium. 10 mL LB broths containing appropriate antibiotic were inoculated with *Salmonella* strains and incubated at 37 °C overnight with shaking at 180 rpm. On the day of infection, the OD₆₀₀ readings of the overnight culture were read and used to estimate bacterial counts. Cells were then infected at a multiplicity of infection (MOI) of 10. 24 well plates were centrifuged at 1500 rpm for 5 minutes at room temperature to promote and synchronise infection. Infected cells were incubated at 37 °C, 5% CO₂ for 30 minutes. Culture media was replaced with media containing 100 µg/ml gentamicin (Pharmacia, Washington, USA), and cells were incubated at 37 °C, 5% CO₂ for a further 1 hour.
Culture media was replaced with media containing 10 µg/ml gentamicin, and where necessary, 1 mM IPTG, and cells were incubated at 37 °C, 5% CO₂ to the required time, post infection.

**Immunoblotting.** Cells were lysed in cold 1 x KalB lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% (vol/vol) Triton X-100) supplemented with 2 mM Na₃VO₄, 10 mM NaF, 1 mM PMSF, and 1 x EDTA-free Complete protease inhibitor cocktail (Roche). Cell lysate was incubated for at least 30 minutes on ice, then cell debris was pelleted at 13000 rpm at 4 °C for 12 minutes. The soluble protein fraction was mixed with 4 x Bolt® LDS sample buffer (Life Technologies) and DTT (Astral Scientific) to a final concentration of 50 mM. Proteins were boiled at 80 to 90 °C for 10 minutes, then loaded to Bolt® 4-12% Bis-Tris Plus gels (Life Technologies) alongside SeeBlue® pre-stained protein ladder (Life Technologies). Proteins were separated by electrophoresis using an XCell SureLock™ Mini-Cell system (Life Technologies) with 1 x Bolt® MES SDS or 1 x Bolt® MOPS SDS running buffer (Life Technologies), according to the manufacturer’s instructions. Following electrophoresis, proteins were transferred onto nitrocellulose membranes using the iBlot2® gel transfer device (Life Technologies) and iBlot2® nitrocellulose transfer stacks (Life Technologies), according to the manufacturer’s instructions. Membranes were blocked in 5% (w/v) skim milk in TBS (20 mM Tris, 50 mM NaCl, pH 8.0) with 0.1% (v/v) Tween 20 at room temperature for at least 1 hour with shaking at 60 rpm. Membranes were rinsed and washed in TBS Tween, then probed one of the following primary antibodies as required at 4 °C overnight with shaking at 60 rpm: rabbit monoclonal anti-ArgGlcNAc (Abcam), mouse monoclonal anti-HA (BioLegend), mouse monoclonal anti-GFP (Roche), mouse monoclonal anti-Flag M2-HRP (Sigma), mouse monoclonal anti-TRADD (BD Transduction Lab) or mouse monoclonal anti-β-actin (Sigma). Membranes were again rinsed and washed in TBS Tween, then probed with anti-mouse or anti-rabbit IgG secondary antibodies conjugated to horseradish peroxidase (PerkinElmer) diluted in TBS with 5% BSA (Sigma) and 0.1% Tween (Sigma) at room temperature for one hour with shaking at 60 rpm. Membranes were rinsed and washed in TBS Tween at room temperature for at least 45 minutes with shaking at 60 rpm. Antibody binding was detected using chemiluminescent
substrates for horseradish peroxidase (HRP) (ECL western blotting reagents (GE Healthcare) or ECL Prime western blotting reagent (Amersham, USA), according to the manufacturer’s instructions, and visualised using an MFChemiBis imaging station. For membranes required to be developed twice using two sets of primary antibodies, Restore™ PLUS Western Blot stripping buffer (Thermo Fisher Scientific) was applied on the blots with shaking at 60 rpm for 10 min following first time development. The stripped membranes were rinsed and washed in TBS Tween before being blocked in 5% (w/v) skim milk in TBS Tween with shaking at 60 rpm for 60 min at room temperature. Blocked membranes were rinsed and washed in TBS Tween, and then probed with the additional primary antibody.

**Transfection of HEK293T cells.** HEK293T cells were transfected using FuGENE® 6 transfection reagent (Promega), according to the manufacturer’s instructions. Cells were transfected one day after seeding to achieve 80 to 90% confluency. Transfection reagent was mixed with the reduced serum medium Opti-MEM®I (1X) + GlutaMAX(TM)-I (Gibco, Life Technologies), and incubated at room temperature for 5 minutes. Plasmid DNA was added at a transfection reagent:DNA ratio of 3:1, and incubated at room temperature for 25 minutes. The reaction was added to previously seeded cells and incubated at 37 °C, 5% CO₂ for 16 to 24 hours.

**Immunoprecipitation of Flag-tagged fusion proteins.** At required time points post-transfection, cells were lysed and the insoluble fraction removed as described above. Immunoprecipitation of Flag-tagged proteins was performed using Anti-Flag® M2 Magnetic Beads (Sigma-Aldrich), according to the manufacturer’s instructions. Beads were first washed twice with lysis buffer, then mixed with the remaining soluble protein fraction and incubated rotating at 4 °C overnight. Following this, beads were washed three times with lysis buffer. Bound protein was eluted by incubating beads in 60 µl of 150 µg/ml Flag peptide (Sigma-Aldrich), rotating at 4 °C for 30 minutes. Eluate was mixed with LDS and DTT, and boiled at 80 to 90 °C for 10 minutes. Input and eluate samples were electrophoresed, then visualised by immunoblot as above.
**Yeast two hybrid assay.** Yeast strain *S. cerevisiae* Y2H Gold (Clontech, California, USA) was transformed or cotransformed with plasmid DNA using the established lithium acetate method (40). Transformants were plated to selective media as required to select for successful single or double transformation. When validating interactions between two proteins, transformants were subsequently plated to highly selective media. Briefly, *S. cerevisiae* Y2H Gold was streaked to YPDA and incubated at 30 °C for 3 days. Healthy colonies were used to inoculate 10 ml YPDA broth at a starting OD₆₀₀ of 0.2, and incubated at 30 °C with shaking at 200 rpm to an OD₆₀₀ of 0.6-0.8. The yeast culture was centrifuged at 4000 rpm for 7 minutes, and the pellet was resuspended in sterile distilled water, and centrifuged again. The yeast culture was then resuspended in 100 mM lithium acetate, vortexed thoroughly, and centrifuged again. The lithium acetate supernatant was removed, and yeast were resuspended in 400 mM lithium acetate, vortexed thoroughly, and centrifuged again. The lithium acetate supernatant was removed, and yeast were resuspended in polyethylene glycol (PEG 3350, Sigma-Aldrich), 1 M lithium acetate, salmon sperm ssDNA at a final concentration of 2 mg/ml, and plasmid DNA as appropriate. This reaction was incubated at 30 °C for 30 minutes, then subjected to heat shock at 42 °C for 20 minutes. The reaction was briefly centrifuged, and the supernatant removed. The yeast pellet was resuspended in distilled water and plated to both SD/Trp-Leu and SD/Trp-Leu-Ade-His selective media, then incubated at 30 °C for 3 days.

**In vitro glycosylation assay.** His-tagged and GST-tagged fusion proteins were purified from bacterial cultures using Novagen His-Bind® purification kit or Novagen GST-Bind™ purification kit, respectively, according to the manufacturer’s instructions. Briefly, plasmids encoding either 6 x His-tagged or GST-tagged fusion proteins were transformed into BL21 C43 (DE3) *E. coli*. Overnight cultures grown in LB with appropriate antibiotics were used to inoculate a 200 mL LB subculture (1:100) which was grown for 3 hours at 37 °C with shaking at 180 rpm to an optical density of 0.6. Subcultures were induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG; AppliChem, Darmstadt, Germany), and grown for a further 3 hours. Cultures were then centrifuged.
at 10000 rpm at 4 °C for 15 minutes, then resuspended in the appropriate resuspension buffer. Resuspended bacteria were lysed using an EmulsiFlex-C3 High Pressure Homogenizer (Avestin), according to the manufacturer’s instructions. Lysates were centrifuged at 13000 rpm at room temperature for 30 minutes, and proteins were purified from the soluble fraction by either nickel- or glutathione-affinity chromatography, according to the manufacturer’s instructions.

Protein concentrations were determined using a bicinchoninic acid (BCA) kit (Thermo Fisher Scientific). Recombinant proteins (approximately 1 µg) were incubated alone or together, and in the presence of 1 mM UDP-GlcNAc (Sigma-Aldrich). Reactions were made to a total volume of 80 µl in TBS (50 mM Tris, 150 mM NaCl, pH 7.6) supplemented with 10 mM MgCl₂ and 10 mM MnCl₂. Reactions were incubated at 37 °C for 4-5 hours. To detect in vitro glycosylation, reactions were either electrophoresed and probed by Western blot as above, or processed for mass spectrometry analysis as below.

**Tryptic digest of gel-separated proteins.** Affinity purified proteins were separated using SDS-PAGE, fixed and visualized with Coomassie G-250 according to protocol of Kang et al. (41). Bands of interest were excised and destained in a 50:50 solution of 50 mM NH₄HCO₃ / 100% ethanol for 20 minutes at room temperature with shaking at 750 rpm. Destained samples were then washed with 100% ethanol, vacuum-dried for 20 minutes and rehydrated in 50 mM NH₄HCO₃ plus 10 mM DTT. Reduction was carried out for 60 minutes at 56 °C with shaking. The reducing buffer was then removed and the gel bands washed twice in 100% ethanol for 10 minutes to remove residual DTT. Reduced ethanol washed samples were sequentially alkylated with 55 mM Iodoacetamide in 50 mM NH₄HCO₃ in the dark for 45 minutes at room temperature. Alkylated samples were then washed with two rounds of 100% ethanol and vacuum-dried. Alkylated samples were then rehydrated with 12 ng/µl trypsin (Promega) in 40 mM NH₄HCO₃ at 4 °C for 1 hour. Excess trypsin was removed, gel pieces were covered in 40 mM NH₄HCO₃ and incubated overnight at 37 °C. Peptides were concentrated and desalted using C18 stage tips (42, 43) before analysis by LC-MS.
Enrichment of arginine-glycosylated peptides from infected cell lysate. Infected cells were washed three times in ice-cold PBS and lysed by scraping with ice-cold guanidinium chloride lysis buffer (6 M GdmCl, 100 mM Tris pH 8.5, 10 mM TCEP, 40 mM 2-Chloroacetamide) on a bed of ice according to the protocol of Humphrey et al. (44). Lysates were collected and boiled at 95 °C for 10 minutes with shaking at 2000 rpm to shear DNA and inactivate protease activity. Lysates were then cooled for 10 minutes on ice then boiled again at 95 °C for 10 minutes with shaking at 2000 rpm. Lysates were cooled and protein concentration determined using a BCA assay. 2 mg of protein from each sample was acetone precipitated by mixing 4 volumes of ice-cold acetone with one volume of sample. Samples were precipitated overnight at -20 °C and then spun down at 4000 G for 10 minutes at 4 °C. The precipitated protein pellets were resuspended with 80% ice-cold acetone and precipitated for an additional 4 hours at -20 °C. Samples were spun down at 17000 G for 10 minutes at 4 °C to collect precipitated protein, the supernatant was discarded and excess acetone driven off at 65 °C for 5 minutes.

Dried protein pellets were resuspended in 6 M urea, 2 M thiourea, 40 mM NH₄HCO₃ and reduced / alkylated prior to digestion with Lys-C (1/200 w/w) then trypsin (1/50 w/w) overnight as previously described (45). Digested samples were acidified to a final concentration of 0.5% formic acid and desalted with 50 mg tC18 SEP-PAK (Waters corporation, Milford, USA) according to the manufacturer’s instructions. Briefly, tC18 SEP-PAKs were conditioned with buffer B (80% ACN, 0.1% formic acid), washed with 10 volumes of Buffer A* (0.1% TFA, 2% ACN), sample loaded, column washed with 10 volumes of Buffer A* and bound peptides eluted with buffer B then dried. Peptide affinity purification was accomplished according to the protocol of Udeshi et al. (46), modified to allow for Arg-GlcNAc enrichment. Briefly, aliquots of 100 µl of Protein A/G plus Agarose beads (Santa Cruz, Santa Cruz CA) were washed three times with 1 ml of immunoprecipitation buffer (IAP, 10 mM Na₃PO₄, 50 mm NaCl, 50 mM MOPS, pH 7.2) and tumbled overnight with 10 µg of anti-Arg-GlcNAc antibody (ab195033, Abcam) at 4 °C. Coupled anti-Arg-GlcNAc beads were then washed three times with 1 ml of 100 mM sodium borate (pH 9)
to remove non-bound proteins and cross-linked for 30 minutes rotating using 20 mM Dimethyl Pimelimidate (Thermo Fisher Scientific) in 100 mM HEPES, pH 8.0. Cross-linking was quenched by washing beads with 200 mM ethanolamine, pH 8.0, three times then rotating the beads in an additional 1 ml 200 mM ethanolamine, pH 8.0 for 2 hours at 4 °C. Beads were washed three times with IAP buffer and used immediately.

Purified peptides were resuspended in 1 ml IAP buffer and the pH checked to ensure compatibility with affinity conditions. Peptide lysates were then added to the prepared cross-linked anti-Arg-GlcNAc antibody beads and rotated for 3 hours at 4 °C. Upon completion antibody beads were spun down at 3000 G for 2 minutes at 4 °C and the unbound peptide lysates collected. Antibody beads were then washed six times with 1 ml of ice-cold IAP buffer and Arg-GlcNAc peptides eluted using two rounds of acid elution. For each elution round, 100 µl of 0.2% TFA was added and antibody beads allowed to stand at room temperature with gentle shaking every minute for 10 minutes. Peptide supernatants were collected and desalted using C18 stage tips (42, 43) before analysis by LC-MS.

Identification of arginine-glycosylated affinity enriched peptides and Flag-tagged proteins using reversed phase LC-MS. Purified peptides prepared were re-suspend in Buffer A* and separated using a two-column chromatography set up composed of a PepMap100 C18 20 mm x 75 µm trap and a PepMap C18 500 mm x 75 µm analytical column (Thermo Fisher Scientific). Samples were concentrated onto the trap column at 5 µL/min for 5 minutes and infused into an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific) at 300 nl/minute via the analytical column using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific). 125 minutes gradients were run altering the buffer composition from 1% buffer B to 28% B over 90 minutes, then from 28% B to 40% B over 10 minutes, then from 40% B to 100% B over 2 minutes, the composition was held at 100% B for 3 minutes, and then dropped to 3% B over 5 minutes and held at 3% B for another 15 minutes. The Lumos™ Mass Spectrometer was operated in a data-dependent mode automatically switching between the acquisition of a single Orbitrap MS
scan (120,000 resolution) every 3 seconds and Orbitrap EThcD for each selected precursor (maximum fill time 100 ms, AGC 5*104 with a resolution of 30000 for Orbitrap MS-MS scans). For parallel reaction monitoring (PRM) experiments the known tryptic Arg-modified sites of TRADD (8) and FADD (7) (Uniprot accession: B2RRZ7 and Q3U0V2 respectively) were monitored using the predicted m/z for the +2 and +3 charge states. Data-independent acquisition was performed by switching between the acquisition of a single Orbitrap MS scan (120000 resolution, m/z 300-1500) every 3 seconds and Orbitrap EThcD for each PRM precursor (maximum fill time 100 ms, AGC 5*104 with a resolution of 60000 for Orbitrap MS-MS scans).

**Mass spectrometry data analysis.** Identification of proteins and Arg-glycosylated peptides was accomplished using MaxQuant (v1.5.3.1) (47). Searches were performed against the Mouse (Uniprot proteome id UP000000589 - Mus musculus, downloaded 18-05-2016, 50306 entries), *Salmonella Typhimurium SL1344* (Uniprot proteome id UP000008962- *Salmonella Typhimurium* SL1344, downloaded 18-05-2016, 4,657 entries) or human (Uniprot proteome id UP000005640-*Homo sapiens*, downloaded 24/10/2013, 84,843 entries) proteomes depending on the samples with carbamidomethylation of cysteine set as a fixed modification. Searches were performed with trypsin cleavage specificity allowing 2 miscleavage events and the variable modifications of oxidation of methionine, N-Acetylhexosamine addition to arginine (Arg-GlcNAc) and acetylation of protein N-terminals. The precursor mass tolerance was set to 20 parts-per-million (ppm) for the first search and 10 ppm for the main search, with a maximum false discovery rate (FDR) of 1.0% set for protein and peptide identifications. To enhance the identification of peptides between samples the Match Between Runs option was enabled with a precursor match window set to 2 minutes and an alignment window of 10 minutes. For label-free quantitation, the MaxLFQ option within Maxquant (48) was enabled in addition to the re-quantification module. The resulting protein group output was processed within the Perseus (v1.4.0.6) (49) analysis environment to remove reverse matches and common protein contaminates prior. For LFQ comparisons missing values were imputed using Perseus and Pearson correlations visualized using Matlab R2015a (http://www.mathworks.com).
The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD010769. Data can be accessed using the username: reviewer55228@ebi.ac.uk  Password: hkVeAKlw.

Mass spectrometry experimental design and statistical rationale. Both quantitative (identification of Arg-GlcNAcylated substrates) and qualitative (assessing site of localizations) were undertaken in this study with a full list of replicate numbers, number of experimental groups provided in Supplemental Experimental Design Table. For quantitative studies, six samples (condition of interest vs negative control) corresponding to three biological replicates of each condition were generated in parallel and subjected to identical sample handling. For Arg-GlcNAcylated pulldowns, negative control data was derived from three biological replicates of cells infected with ΔsseK123. For overexpression-based studies, negative control data was derived from cells infected with AsseK123 expressing inactivated SseK1E255A. Prepared samples were run as a single randomized batch with blanks between samples to minimise carryover. Reproducibility between biological repeats was assessed using Pearson correlations and these are provided as Supplemental Material. All MS/MS assignments for identified peptides within this work can be accessed using MS-viewer (http://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) via the unique identifier provided in Supplemental Table 11. Missing values were imputed using Perseus with the imputed population randomly assigned a LFQ values with a downshifted LFQ mean of 2.5 standard derivations and width of 0.3 width standard derivations compared to the observed values. Statistical assessment of imputed pulldowns was accomplished in Perseus using student t-tests with a Benjamini–Hochberg FDR set to 0.05. Identified Arg-GlcNAcylated sites of interested were manually inspected and annotated to ensure the correctness of the assignment and localization of the modification.

Protein expression and purification for structural studies. The pRL652 plasmid containing GST-TEV-SseK3(25-335) was transformed into BL21(DE3) strain. For protein expression, a 15 mL overnight culture in LB was inoculated into one liter of terrific broth media supplemented with 100
μg/ml of ampicillin. The inoculated cultures were grown at 37°C until the OD_{600} reached 1.0. The cultures were transferred to 18°C, induced with 1mM of isopropyl β-D-1-thiogalactopyranoside (IPTG), left overnight to grow and harvested by centrifugation at 9,110 x g for 7 minutes.

The cell pellet was resuspended in lysis buffer (50 mM Tris-HCl buffer pH 8.0, 10% glycerol, and 0.1% Triton X) and the cells were lysed in a cell disruptor (Constant Systems Ltd., Northants, United Kingdom). Cell debris were removed by centrifugation at 28,965 x g for 30 minutes. The supernatant was loaded on 10 mL Glutathione-Superflow resin (Clontech) column equilibrated with standard buffer (20 mM Tris pH 8.0 and 150 mM NaCl). The column was washed with 5 column volumes of standard buffer. The beads were then incubated with TEV protease (33 μg/mL) in 30 mL of standard buffer for overnight at room temperature. The flow-through containing cleaved SseK3 was collected, concentrated to 30 mg/mL with the Millipore centrifugal filter with a molecular weight cut-off of 10,000 for crystallization trials. The E258Q mutant was generated using KOD Hot Start DNA Polymerase (Sigma-Aldrich) with mutagenic primers and WT plasmid as a template according to the manufacturer’s instructions. The E258Q mutant was purified using the same protocol.

The seleno-methionine (SeMet) derivative of SseK3 was expressed in auxotrophic E. coli strain B834(DE3). A 50 mL overnight culture in medium A (M9 medium, trace elements, glucose, MgSO_{4}, CaCl_{2}, Biotin, 50 mg/mL methionine and thiamin) was used to inoculate 1 L of medium A. All media used were supplemented with 100 μg/ml of Ampicillin. The cells were grown with shaking at 37°C until OD_{600} reached 1.0. The cells were pelleted at 4°C and then resuspended in 1 L of media A without methionine. The culture was further incubated for 4 hours at 37°C and 50 mg of SeMet was then added. After 30 minutes of incubation, the cultures were then induced with 1 mM of IPTG and continued to grow for additional 10 h at 18°C. The cells were harvested by centrifugation at 9,100 x g for 7 min. The SeMet-labelled protein was purified the same way as the native protein.
**Protein Crystallization.** Initial crystals were obtained by screening using commercial and in-house screens in a 96-well plate format. The crystallization was setup using Gryphon crystallization robot (Art Robbins Instruments, Sunnyvale, CA). The best crystals were obtained using hanging-drop vapor diffusion method at 20 °C. 1 µL of protein solution supplemented with 3 mM of UDP-GlcNAc was mixed with 1 µL of reservoir solution containing 0.5 M NaCl, 0.1 M Tris pH 8.5, 18% PEG 3350 and 5% MPD and suspended over 0.5 mL of reservoir. These crystals displayed P2₁ space group symmetry. The SeMet-containing SseK3 crystallized at slightly different conditions and these crystals had the same space group symmetry as the native crystals but they differed in cell dimensions (Table 4). The best crystals of the SseK3(E258Q) mutant in complex with UDP-GlcNAc and Mg²⁺ were obtained at somewhat different conditions, 0.5 M NaCl, 0.1 M Tris pH 8.5, 12% PEG 3350 and 5% MPD supplemented with 6 mM UDP-GlcNAc and 6 mM MgCl₂. They displayed P2₁2₁2₁ space group symmetry (Table 4).

**Data collection and structure determination.** For data collection, the crystals were soaked briefly in a cryo-protecting solution containing 30% MPD and 70% reservoir solution and flash-cooled in liquid nitrogen. Diffraction data were collected at the Canadian Light Source (CLS) beam line 08B1-1 for the native crystal and 08ID-1 for the SeMe-labelled crystal. Data were processed using XDS program (50) with the AutoProcess script (51). The structure was solved by single-wavelength anomalous dispersion (SAD) method using program AutoSol in Phenix program suite (52). The initial model was refined using Phenix software interspaced with manual rebuilding using COOT (53). The initial model of the SseK3 molecule was placed in the context of the native dataset by molecular replacement. There are four molecules in the asymmetric unit. The refinement continued with the Phenix software until convergence was reached with R_work=0.2172 and R_free=0.2502. The final model contains residues 27-330 for molecule A, 28-329 for molecule B, 26-329 for molecule C and D. Each molecule contains bound UDP. There are 292 solvent molecules. The structure of the E258Q mutant was solved by molecular replacement and contains two molecules in the asymmetric unit. The refinement converged with R_work=0.177 and R_free=0.225. The final model
contains residues 26-335 in molecule A and 28-335 in molecule B, UDP, Mg$^{2+}$ and GlcNAc bound to each molecule, two molecules of TRIS, two molecules of methyl-pentanediol and 186 water molecules. Data collection and refinement statistics are shown in Table 4. The coordinates and structure factors were deposited in the PDB data bank with accession numbers 6CGI and 6DUS.

**Mouse infection studies.** All experiments involving mice were approved by the Animal Ethics Committee at The University of Melbourne, under project number 1613898. All experiments were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, 8th edition, 2013. C57BL/6 and Trail-R$^{-/-}$ mice were obtained from John Silke at the Walter and Elisa Hall Institute, Melbourne, Australia, and were acclimatised at the Peter Doherty Institute for Infection and Immunity for 1 week prior to infection. For intravenous infection, *S. Typhimurium* strain BRD509 (54, 55) was sub-cultured from a fresh broth culture, grown shaking at 180 rpm until the mid-logarithmic phase, and frozen down with 10% glycerol at -80°C until use. Mice were intravenously injected via the tail vein with $2 \times 10^5$ CFU in 200 µl sterile PBS. At designated time points, mice were euthanised by CO$_2$ asphyxiation and spleen and liver were resected for enumeration of bacterial load and quantification of cytokine levels. Serum was also collected for cytokine analysis.

**Cytokine analysis.** The BD™ Cytometric Bead Array (CBA) Mouse Inflammation Kit (BD Biosciences) was used to quantify cytokines in tissue homogenates and serum as per manufacturer’s instructions. Prepared CBA assay samples were analysed for the presence of six mouse cytokine markers of inflammation (IL-6, IL-10, MCP-1, IFN$\gamma$, TNF and IL-12p70) using a BD™ FACS Canto II. The detection limit (dotted line) for cytokine concentration was determined as the lowest data point on the standard curve that had >80% actual recovery compared to expected concentration.
| Strain            | Relevant characteristics                                                                 | Reference   |
|-------------------|-------------------------------------------------------------------------------------------|-------------|
| SL1344            | Wild type *S. enterica* serovar Typhimurium strain SL1344                                  | Nathaniel Brown |
| ΔsseK1/2/3        | SL1344 ΔsseK1ΔsseK2ΔsseK3                                                                 | (6)         |
| ΔsseK2/3          | SL1344 ΔsseK2ΔsseK3                                                                     | (6)         |
| ΔsseK1/2          | SL1344 ΔsseK1ΔsseK2                                                                     | (5)         |
| BRD509            | SL1344, *aroA aroD*                                                                      | (54, 55)    |
| *S. cerevisiae* Y2H Gold | *MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1_{UAS}–Gal1_{TATA}–His3, GAL2_{UAS}–Gal2_{TATA}–Ade2 URA3 : : MEL1_{UAS}–Mel1_{TATA} AUR1-C MEL1* | Clontech    |
Table 2. Plasmids used in this study.

| Plasmid | Relevant characteristics | Reference |
|---------|--------------------------|-----------|
| pTrc99A | Low copy bacterial expression vector with inducible lacI promoter, Amp<sup>R</sup> | Pharmacia Biotech |
| pTrc99A-SseK1 | sseK1 from S. Typhimurium SL1344 in pTrc99A, Amp<sup>R</sup> | This study |
| pTrc99A-SseK1<sub>DxD(229-231)AAA</sub> | sseK1 from S. Typhimurium SL1344 in pTrc99A, with DxD (229-231) catalytic motif mutated to AAA, Amp<sup>R</sup> | This study |
| pTrc99A-SseK1<sub>E255A</sub> | sseK1 from S. Typhimurium SL1344 in pTrc99A, with Glu255 mutated to Ala, Amp<sup>R</sup> | This study |
| pTrc99A-SseK2 | sseK2 from S. Typhimurium SL1344 in pTrc99A, Amp<sup>R</sup> | This study |
| pTrc99A-SseK2<sub>DxD(239-241)AAA</sub> | sseK2 from S. Typhimurium SL1344 in pTrc99A, with DxD(239-241) catalytic motif mutated to AAA, Amp<sup>R</sup> | This study |
| pTrc99A-SseK2<sub>E271A</sub> | sseK2 from S. Typhimurium SL1344 in pTrc99A, with Glu271 mutated to Ala, Amp<sup>R</sup> | This study |
| pTrc99A-SseK3 | sseK3 from S. Typhimurium SL1344 in pTrc99A, Amp<sup>R</sup> | This study |
| pTrc99A-SseK3<sub>DxD(226-228)AAA</sub> | sseK3 from S. Typhimurium SL1344 in pTrc99A, with DxD(226-228) catalytic motif mutated to AAA, Amp<sup>R</sup> | This study |
| pTrc99A-SseK3<sub>E258A</sub> | sseK3 from S. Typhimurium SL1344 in pTrc99A, with Glu258 mutated to Ala, Amp<sup>R</sup> | This study |
| p3xFlag-<i>Myc</i>-CMV-24 | Dual tagged N-terminal Met-3xFlag and C-terminal c-myc expression vector, Amp<sup>R</sup> | Sigma-Aldrich |
| pFlag-TRADD | Human TRADD in p3xFlag-Myc-CMV, Amp<sup>R</sup> | Jürg Tschopp |
| pFlag-TRADD<sub>R235A</sub> | Human TRADD with Arg235 mutated to Ala, in p3xFlag-Myc-CMV, Amp<sup>R</sup> | This study |
| pFlag-TRADD<sub>R245A</sub> | Human TRADD with Arg245 mutated to Ala, in p3xFlag-Myc-CMV, Amp<sup>R</sup> | This study |
| pFlag-TRADD<sub>R235A/R245A</sub> | Human TRADD with Arg235 and Arg245 mutated to Ala, in p3xFlag-Myc-CMV, Amp<sup>R</sup> | This study |
| Vector Name | Description | Source |
|-------------|-------------|--------|
| pFlag-hTRAILR2_{DD} | Human TRAILR2 death domain in p3xFlag-Myc-CMV, Amp\(^R\) | This study |
| pFlag-hTNFR1_{DD} | Human TNFR1 death domain in p3xFlag-Myc-CMV, Amp\(^R\) | This study |
| pEGFP-C2 | Expression vector carrying EGFP fused to the N terminus of the partner protein, Kan\(^R\) | Clontech |
| pEGFP-C2-SseK1 | sseK1 from S. Typhimurium SL1344 in pEGFP-C2, Kan\(^R\) | This study |
| pEGFP-C2-SseK3 | sseK3 from S. Typhimurium SL1344 in pEGFP-C2, Kan\(^R\) | This study |
| pEGFP-C2-SseK3_{E258A} | sseK3 from S. Typhimurium SL1344 in pEGFP-C2, with Glu258 mutated to Ala, Kan\(^R\) | This study |
| pET28a | Bacterial expression vector with T7lac promoter including N-terminal 6 x Histidine tag, Kan\(^R\) | Novagen |
| pET28a-hTRAILR2_{DD} | Human TRAILR2 death domain in pET28a, Kan\(^R\) | This study |
| pGBKT7 | High copy number yeast expression vector carrying a GAL4 DNA binding domain, Kan\(^R\) (bacterial selection), Trp (selectable marker in yeast) | Clontech |
| pGBKT7-NleB1 | nleB1 from EPEC E2348/69 in pGBKT7, Kan\(^R\), Trp | This study |
| pGBKT7-SseK3 | sseK3 from S. Typhimurium SL1344 in pGBKT7, Kan\(^R\), Trp | This study |
| pGADT7-AD | High copy number yeast expression vector carrying a GAL4 activation domain, Amp\(^R\) (bacterial selection), Leu (selectable marker in yeast) | Clontech |
| pGADT7-AD-FADD_{DD} | Death domain of human FADD in pGADT7-AD, Amp\(^R\), Leu | This study |
| pGADT7-AD-TNFR1_{DD} | Death domain of human TNFR1 in pGADT7-AD, Amp\(^R\), Leu | This study |
| pGADT7-AD-hTRAILR2_{DD} | Human TRAILR2 death domain in pGADT7-AD, Amp\(^R\), Leu | This study |
| pGEX-4T-1 | Low copy number N-terminal glutathione-S-transferase fusion vector, Amp\(^R\) | GE Healthcare |
| pGEX-SseK3 | sseK3 from S. Typhimurium SL1344 in pGEX-4T-1, Amp\(^R\) | This study |
### Table 3. Primers used in this study.

| Primer                  | Primer sequence 5’-3’                          |
|-------------------------|-----------------------------------------------|
| pTrc99AF                | CGGTTCTGGCAAAATATTC                            |
| pTrc99AR                | GCAGTTCCCTACTCTCGC                            |
| p3xFlag-Myo-CMV-24F     | AATGTCGTAATAACCCCCCGCCCGTTGACGC               |
| p3xFlag-Myo-CMV-24R     | TATTAGGACAAGGCTGTTGCGAC                      |
| pEGFP-C2F               | AACACccccCATCGCG                             |
| pEGFP-C2R               | GTAACcATTATAAGCtGC                            |
| pGBK7T7-BDF            | AATACGACTC ACTATAGG                           |
| pGBK7T7-BDR            | GGTGcACGATGCACAG                             |
| pGADT7-ADF             | AATACGACTDATAGG                               |
| pGADT7-ADR             | GGTGCACGATGCACAG                             |
| pGEX-4T-1F             | CGTATTGAAAGCTATCCACAA                       |
| pGEX-4T-1R             | GGGAGctGCATGTGCAGAG                         |
| pET28aF                | AATACGACTCATATAGG                             |
| pET28aR                | GCTAGTTATTGCTCAGCGG                         |
| SseK1F                 | CGGAATTCATGGGACATTTAATTTGTTATGATCCC         |
| SseK1R                 | CGGGATCCCTACGCAATAATCCGGCACATCATCAGGATACG    |
| SseK2F                 | CGGAATTCATGTGGTGGTCTCGAGGCAGG              |
| SseK2R                 | CGGGATCCCTACGCAATAATCCGGCACATCATCAGGATACG    |
| SseK3F                 | CGGAATTCATGTGGTGGTCTCGAGGCAGG              |
| SseK3R                 | CGGGATCCCTACGCAATAATCCGGCACATCATCAGGATACG    |
| GST-SseK3F             | CGCGAATTCATGTTTCTCGAGTCAGAGTTTTC           |
| GST- SseK3R            | CGCGTCACTTTCATCTCGAGGAGCTGAGTCTAAACTGC      |
| FADD                   | CGCGAATTCATGCTGAGCAGCATTTAACGTATAG            |
| FADD                    | CGCGGATCTTCTACGTAACCTTGTCAAGG               |
| hTNFR1DD-F             | CGCCCATATGATGACGCTGACGCGAGTCTGACGAACCTGAGAGA|
| hTNFR1DD-R             | CGCCCATATGATGACGCTGACGCGAGTCTGACGAACCTGAGAGA|
| hTRAILR2DD-F           | CGCGAATTCATGGATCCACTGAGACTGACGAGGCAGAGGCCAG|
| hTRAILR2DD-R           | CGCGAATTCATGGATCCACTGAGACTGACGAGGCCAGAGGCCAG|
| SseK1DXD-F             | GGTGTATATATTCCTGGTCTGGTCTATGATTATCAGGAAAAACTGG |
SseK1DXD-R  CCAGTTTTTCCGTGATAATCATAGCAGCAGCAAGATATATACACC
SseK2DXD-F  GTGGGTGCTATATCTTTGCTGACGCTATGTTACTTACTGATAAAC
SseK2DXD-R  GTTTATCAGTAAGTAACATAGCTGCAGCAAGATATATGCCAAC
SseK3DXD-F  CTGGAGGTGGCTGATATATCTTGCTGCTGCTATGTTACTTACAG
SseK3DXD-R  CTGTAAGTAACATAGCAGCAGCAAGATATATGCAGCCACCTCCAG
SseK3E255A-F  CGTGCTTTCTATGGCAAACGGGATAATAGCT
SseK3E255A-R  AGCTATTATCCCGTTTGCCATAGAAGCACG
SseK3E271A-F  TGTTAGCCTTGCAAAATGGGATTATTGCTGT
SseK3E271A-R  ACAGCAATAATCCCCATTGCAAGGCTAACA
SseK3E258A-F  GCATGAGTCTTGCAAAATGGGATTATCGCG
SseK3E258A-R  CGGCGATAATCCCCATTTGGCAAGACTCATGC
hTRADDR235A-F  CGCAAGGTGGGGGCCTCAGCGACGCAG
hTRADDR235A-R  CTCGCTGCACTGAGGCCCCCACCCTTSGG
hTRADDR245A-F  CGCCGCGTCCGCCACGCGGCACGG
hTRADDR245A-R  CGGCGACGCTGGGACCCGGGCG
GFPS1F  AAAGAATTCATGGAGCATTTAATTGTTAG
GFPS1R  AAAGGATCCCTACTGCACATGCCTCG
hTRAILR2DD-F2  CGCGAATTCATGGATCCCACTGAGACTCTGAGAC
hTRAILR2DD-R2  CCAAGCTTTTAGAACCCTTCCAGCCTCAACAAATGAG
Table 4. Data collection and structure refinement statistics.

|                       | SeMet SseK3 | Native SseK3 | SseK3 E258Q |
|-----------------------|-------------|--------------|--------------|
|                       | PDB: 6CGI   | PDB: 6DUS    |              |
| **Data Collection Statistics** |            |              |              |
| Space group           | P2₁         | P2₁          | P2₂₂₂₂      |
| a, b, c (Å)           | 96.2 75.0 97.3 | 74.5 96.2 90.7 | 96.5 107.1 74.7 |
| α, β, γ (°)           | 90.0 108.9 90.0 | 90.0 93.3 90.0 | 90.0 90.0 90.0 |
| Wavelength (Å)        | 0.9794      | 0.9801       | 0.9795       |
| Resolution (Å)        | 47.793-2.6 (2.66-2.60) | 48.360 – 2.3 (2.35-2.3) | 74.74 – 2.6 (2.67-2.6) |
| Total Reflections     | 536851 (39870) | 227638 (16899) | 162097 (12225) |
| Unique reflections    | 78948 (5859) | 56770 (4200) | 24555 (1799) |
| R_meas                | 0.128 (0.539) | 0.11 (0.599) | 0.077 (0.510) |
| CC1/2                 | 99.7 (96.5) | 99.7 (87.6) | 99.9 (94.5) |
| Completeness (%)      | 99.8 (99.9) | 99.8 (99.8) | 100 (100) |
| Redundancy            | 6.8 (13.21) | 4.0 (4.02) | 6.6 (6.77) |
| (I)/σ (I)             | 12.62 (5.64) | 11.54 (2.98) | 18.75 (3.8) |
| Wilson B (Å²)         | 28.52       | 34.83        |              |
| **Refinement statistics** |            |              |              |
| R_work/R_free         | 0.217 / 0.250 | 0.177/0.225 |              |
| Rmsd on angles (°)    | 0.813       | 0.685        |              |
| Rmsd on bonds (Å)     | 0.004       | 0.003        |              |
| Clash score           | 3.48        | 6.34         |              |
| Favored (%)           | 97.68       | 97.04        |              |
| Allowed (%)           | 1.74        | 2.8          |              |
| Outliers (%)          | 0.58        | 0.16         |              |
| Isotropic B-factor (Å²)| 35.58       | 39.82        |              |
| Protein (atoms)       | 9651        | 4855         |              |
| Ligand (atoms)        | 100         | 114          |              |
| Solvent (atoms)       | 292         | 186          |              |
ACKNOWLEDGMENTS

Graphical abstract was created with BioRender. The authors are indebted to Jürg Tschopp for the gift of pFlag-TRADD. Research described in this paper was performed using beamline 08B1-1 and 08ID-1 at the Canadian Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research.
REFERENCES

1. Galan, J. E., and Curtiss, R., 3rd (1989) Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *Proc Natl Acad Sci U S A* 86, 6383-6387

2. Shea, J. E., Hensel, M., Gleeson, C., and Holden, D. W. (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 93, 2593-2597

3. Jennings, E., Thurston, T. L. M., and Holden, D. W. (2017) *Salmonella* SPI-2 Type III Secretion System Effectors: Molecular Mechanisms And Physiological Consequences. *Cell Host Microbe* 22, 217-231

4. LaRock, D. L., Chaudhary, A., and Miller, S. I. (2015) *Salmonellae* interactions with host processes. *Nat Rev Microbiol* 13, 191-205

5. Kujat Choy, S. L., Boyle, E. C., Gal-Mor, O., Goode, D. L., Valdez, Y., Vallance, B. A., and Finlay, B. B. (2004) SseK1 and SseK2 are novel translocated proteins of *Salmonella enterica* serovar *typhimurium*. *Infect Immun* 72, 5115-5125

6. 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., and Finlay, B. B. (2011) *Salmonella* phage ST64B encodes a member of the SseK/NleB effector family. *PLoS One* 6, e17824

7. Pearson, J. S., Giogha, C., Ong, S. Y., Kennedy, C. L., Kelly, M., Robinson, K. S., Lung, T. W., Mansell, A., Riedmaier, P., Oates, C. V., Zaid, A., Muhlen, S., Crepin, V. F., Marches, O., Ang, C. S., Williamson, N. A., O’Reilly, L. A., Bankovacki, A., Nachbur, U., Infusini, G., Webb, A. I., Silke, J., Strasser, A., Frankel, G., and Hartland, E. L. (2013) A type III effector antagonizes death receptor signalling during bacterial gut infection. *Nature* 501, 247-251

8. Li, S., Zhang, L., Yao, Q., Li, L., Dong, N., Rong, J., Gao, W., Ding, X., Sun, L., Chen, X., Chen, S., and Shao, F. (2013) Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains. *Nature* 501, 242-246

9. Gunster, R. A., Matthews, S. A., Holden, D. W., and Thurston, T. L. (2017) SseK1 and SseK3 Type III Secretion System Effectors Inhibit NF-kappaB Signaling and Necroptotic Cell Death in *Salmonella*-Infected Macrophages. *Infect Immun* 85

10. Wong Fok Lung, T., Giogha, C., Creuzburg, K., Ong, S. Y., Pollock, G. L., Zhang, Y., Fung, K. Y., Pearson, J. S., and Hartland, E. L. (2016) Mutagenesis and Functional Analysis of the Bacterial Arginine Glycosyltransferase Effector NleB1 from Enteropathogenic *Escherichia coli*. *Infect Immun* 84, 1346-1360

11. El Qaidi, S., Chen, K., Halim, A., Siukstaite, L., Rueter, C., Hurtado-Guerrero, R., Clausen, H., and Hardwidge, P. R. (2017) NleB/SseK effectors from *Citrobacter rodentium*, *Escherichia coli*, and *Salmonella enterica* display distinct differences in host substrate specificity. *J Biol Chem* 292, 11423-11430

12. Yang, Z., Soderholm, A., Lung, T. W., Giogha, C., Hill, M. M., Brown, N. F., Hartland, E., and Teasdale, R. D. (2015) SseK3 Is a *Salmonella* Effector That Binds TRIM32 and Modulates the Host's NF-kappaB Signalling Activity. *PLoS One* 10, e0138529

13. Scott, N. E., Giogha, C., Pollock, G. L., Kennedy, C. L., Webb, A. I., Williamson, N. A., Pearson, J. S., and Hartland, E. L. (2017) The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). *J Biol Chem* 292, 17337-17350

14. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., and Coon, J. J. (2012) Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* 11, 1475-1488

15. Aggarwal, B. B. (2003) Signalling pathways of the TNF superfamily: a double-edged sword. *Nat Rev Immunol* 3, 745-756

16. von Karstedt, S., Montinario, A., and Walczak, H. (2017) Exploring the TRAILs less travelled: TRAIL in cancer biology and therapy. *Nat Rev Cancer* 17, 352-366

17. Wertz, I. E., and Dixit, V. M. (2010) Regulation of death receptor signaling by the ubiquitin
system. Cell Death Differ 17, 14-24
18. Silke, J. (2011) The regulation of TNF signalling: what a tangled web we weave. Curr Opin Immunol 23, 620-626
19. Vandenabeele, P., Galluzzi, L., Vanden Berghe, T., and Kroemer, G. (2010) Molecular mechanisms of necroptosis: an ordered cellular explosion. Nat Rev Mol Cell Biol 11, 700-714
20. Buckner, M. M. C., Croxen, M., Arena, E. T., and Finlay, B. B. (2014) 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 2, 208-216
21. Kidwai, A. S., Mushamiri, I., Niemann, G. S., Brown, R. N., Adkins, J. N., and Heffron, F. (2013) Diverse secreted effectors are required for Salmonella persistence in a mouse infection model. PLoS One 8, e70753
22. Pan, G., O'Rourke, K., Chinnaiyan, A. M., Gentz, R., Ebner, R., Ni, J., and Dixit, V. M. (1997) The receptor for the cytotoxic ligand TRAIL. Science 276, 111-113
23. Walczak, H., Degli-Esposti, M. A., Johnson, R. S., Smolak, P. J., Waugh, J. Y., Boiani, N., Timour, M. S., Gerhart, M. J., Schooley, K. A., Smith, C. A., Goodwin, R. G., and Rauch, C. T. (1997) TRAIL-R2: a novel apoptosis-mediating receptor for TRAIL. Embo j 16, 5386-5397
24. Sheridan, J. P., Marsters, S. A., Rieser, E., Rauch, C. T., Joo, P., Blenis, J., Krammer, P. H., and Walczak, H. (2000) FADD/MORT1 and caspase-8 are recruited to TRAIL receptors 1 and 2 and are essential for apoptosis mediated by TRAIL receptor 2. Immunity 12, 599-609
25. Lafont, E., Hartwig, T., and Walczak, H. (2018) Paving TRAIL’s Path with Ubiquitin. Trends Biochem Sci 43, 44-60
26. Jouan-Lanhouet, S., Arshad, M. I., Piquet-Pellorce, C., Martin-Chouly, C., Le Moigne-Muller, G., Van Herreweghe, F., Takahashi, N., Sergent, O., Lagadic-Gossman, D., Vandenabeele, P., Samson, M., and Dimanche-Boitrel, M. T. (2012) TRAIL induces necroptosis involving RIPK1/RIPK3-dependent PARP-1 activation. Cell Death Differ 19, 2003-2014
30. Diehl, G. E., Yue, H. H., Hsieh, K., Kuang, A. A., Ho, M., Morici, L. A., Lenz, L. L., Cado, D., Riley, L. W., and Winoto, A. (2004) TRAIL-R as a negative regulator of innate immune cell responses. Immunity 21, 877-889
31. Everest, P., Roberts, M., and Dougan, G. (1998) Susceptibility to Salmonella typhimurium infection and effectiveness of vaccination in mice deficient in the tumor necrosis factor alpha p55 receptor. Infect Immun 66, 3355-3364
32. Malek, M., and Lamont, S. J. (2003) Association of INOS, TRAIL, TGF-beta2, TGF-beta3, and IgL genes with response to Salmonella enteritidis in poultry. Genetics, selection, evolution : GSE 35 Suppl 1, S99-111
33. Tohidi, R., Idris, I. B., Malar Panandam, J., and Hair Bejo, M. (2013) The effects of polymorphisms in 7 candidate genes on resistance to Salmonella enteritidis in native chickens. Poultry Science 92, 900-909
34. Campbell, J. A., Davies, G. J., Bulone, V., and Henriassat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem J 326 ( Pt 3), 929-939
35. Zhang, Y., Swaminathan, G. J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K. R.,
and Brew, K. (2003) Roles of individual enzyme-substrate interactions by alpha-1,3-galactosyltransferase in catalysis and specificity. *Biochemistry* 42, 13512-13521

36. Pruitt, R. N., Chumbler, N. M., Rutherford, S. A., Farrow, M. A., Friedman, D. B., Spiller, B., and Lacy, D. B. (2012) Structural determinants of *Clostridium difficile* toxin A glucosyltransferase activity. *J Biol Chem* 287, 8013-8020

37. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* 19, 5269-5280

38. Esposito, D., Gunster, R. A., Martino, L., El Omari, K., Wagner, A., Thurston, T. L. M., and Rittinger, K. (2018) Structural basis for the glycosyltransferase activity of the *Salmonella* effector SseK3. *J Biol Chem* 293, 5064-5078

39. Park, J. B., Kim, Y. H., Yoo, Y., Kim, J., Jun, S. H., Cho, J. W., El Qaidi, S., Walpole, S., Monaco, S., Garcia-Garcia, A. A., Wu, M., Hays, M. P., Hurtado-Guerrero, R., Angulo, J., Hardwidge, P. R., Shin, J. S., and Cho, H. S. (2018) Structural basis for arginine glycosylation of host substrates by bacterial effector proteins. *Nat Commun* 9, 4283

40. Gietz, R. D., Schiestl, R. H., Willems, A. R., and Woods, R. A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast (Chichester, England)* 11, 355-360

41. Kang, D.-H., Gho, Y.-S., Suh, M.-K., and Kang, C.-H. (2002) Highly sensitive and fast protein detection with coomassie brilliant blue in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Bulletin of the Korean Chemical Society* 23, 1511-1512

42. Ishihama, Y., Rappsilber, J., and Mann, M. (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *J Proteome Res* 5, 988-994

43. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2, 1896-1906

44. Humphrey, S. J., Azimifar, S. B., and Mann, M. (2015) High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nature Biotechnology* 33, 990-995

45. Scott, N. E., Parker, B. L., Connolly, A. M., Paulech, J., Edwards, A. V., Crossett, B., Falconer, L., Kolarich, D., Djordjevic, S. P., Hojrup, P., Packer, N. H., Larsen, M. R., and Cordwell, S. J. (2011) Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol Cell Proteomics* 10, M000031-mcp000201

46. Udeshi, N. D., Svinikina, T., Mertins, P., Kuhn, E., Mani, D. R., Qiao, J. W., and Carr, S. A. (2013) Refined preparation and use of anti-diglycine remnant (K-epsilon-GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics* 12, 825-831

47. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26, 1367-1372

48. Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 13, 2513-2526

49. Tyanova, S., Temu, T., Siniteyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13, 731-740

50. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66, 125-132

51. Fodje, M., Grochulski, P., Janzen, K., Labiuk, S., Gorin, J., and Berg, R. (2014) 08B1-1: an automated beamline for macromolecular crystallography experiments at the Canadian Light Source. *Journal of synchrotron radiation* 21, 633-637
52. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213-221

53. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501

54. Hoiseth, S. K., and Stocker, B. A. D. (1981) Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature* 291, 238-239

55. Strugnell, R., Dougan, G., Chatfield, S., Charles, I., Fairweather, N., Tite, J., Li, J. L., Beesley, J., and Roberts, M. (1992) Characterization of a *Salmonella typhimurium aro* vaccine strain expressing the P.69 antigen of *Bordetella pertussis*. *Infect Immun* 60, 3994-4002
FIGURE LEGENDS

Figure 1. Immunoblot of RAW264.7 cells infected with derivatives of S. Typhimurium SL1344. Wild type S. Typhimurium, a triple Δssek123 mutant and triple mutant complemented with plasmids encoding one of HA-tagged SseK1, SseK2 or SseK3, or with catalytically-inactive effector derivatives were used to infect RAW264.7 cells for 20 h, as indicated (A and B). Overexpression of the effectors was induced during host cell infection by the addition of 1 mM IPTG. RAW264.7 cells were lysed and proteins detected by immunoblot with anti-ArgGlcNAc, anti-HA, and anti-β-actin antibodies as indicated. Representative immunoblot of at least three independent experiments.

Figure 2. Enrichment of peptides Arg-GlcNAcylated by SseK1 derived from Salmonella-infected RAW264.7 cells. (A) Label-free quantification of Arg-GlcNAc peptides immunoprecipitated from RAW264.7 cells infected with S. Typhimurium ΔsseK123 complemented with either SseK1-HA or SseK1E255A-HA. Arg-GlcNAcylated peptides are presented as a volcano plot depicting mean ion intensity peptide ratios of SseK1-HA versus SseK1E255A-HA plotted against logarithmic t test p values from biological triplicate experiments. Arg-GlcNAcylated peptides with corresponding t test p values below 0.001 are annotated by protein name, with human peptides shaded blue and bacterial peptides shaded red. (B) Manually curated EThcD spectra showing glycosylation of Arg243 within the death domain of mouse TRADD, Andromeda score 249.32. Within MS/MS spectra NL denote neutral loss associated ions. (C) Parallel reaction monitoring of ArgGlcNAc peptide immunoprecipitated from RAW264.7 cells infected with S. Typhimurium ΔsseK123 or S. Typhimurium ΔsseK23. Arginine-glycosylated peptides are presented as a volcano plot depicting mean log2 ion intensity peptide ratios of ΔsseK123 versus ΔsseK23 plotted against logarithmic t test p values from biological triplicate experiments. Arg-GlcNAcylated peptides are annotated by protein name and shaded blue. (D) Manually curated EThcD spectra
showing glycosylation of Arg^{233} within the death domain of mouse TRADD, Andromeda score 71.45. Within MS/MS spectra NL denote neutral loss associated ions.

**Figure 3. Mutagenesis of putative SseK1 glycosylation sites of TRADD.** (A) Immunoblot showing Arg-GlcNAcylation of ectopically expressed Flag-hTRADD or Flag-hTRADD mutants in HEK293T cells co-transfected with pEGFP-SseK1. Cells were harvested for immunoblotting and detected with anti-ArgGlcNAc, anti-GFP, and anti-Flag antibodies. Antibodies to β-actin were used as a loading control. Representative immunoblot of at least three independent experiments. (B) Manually curated EThcD spectra of Arg-GlcNAcylated Flag-hTRADD enriched by anti-Flag immunoprecipitation following ectopic expression in HEK293T cells and co-transfection with pEGFP-SseK1. Various observed sites of Arg-GlcNAcylation are highlighted in red, and presented alongside corresponding M/Z values and observed Andromeda scores. Within MS/MS spectra NL denote neutral loss associated ions.

**Figure 4. Identifying substrates of SseK3 by Arg-GlcNAc peptide enrichment.** (A) Label-free quantification of Arg-GlcNAc peptide immunoprecipitated from RAW264.7 cells infected with *S. Typhimurium ΔsseK123* or ΔsseK12. Arginine-glycosylated peptides are presented as a volcano plot depicting mean log2 ion intensity peptide ratios of ΔsseK123 versus ΔsseK12 plotted against logarithmic t test *p* values from biological triplicate experiments. Arg-GlcNAcylated peptides are annotated by gene name and shaded blue. (B) Heat map showing observed ion intensity of Arg-GlcNAcylated peptides between biological triplicates. (C) Partial sequence alignment showing observed GlcNAcylated arginine residue is conserved between identified substrates. (D) Manually curated HCD spectra of arginine glycosylated TNFRSF10B/TNFR1 (upper) and TNFRSF1A/TRAILR (lower). Observed sites of Arg-GlcNAcylation are highlighted in red, and presented alongside corresponding M/Z values and observed Andromeda scores.
Figure 5. *In vitro* validation of host substrate modifications by SseK3. (A) Immunoblot of inputs and immunoprecipitates (IP) of anti-Flag immunoprecipitations performed on lysates of HEK293T cells co-transfected with pFlag-hTRAILR2<sub>DD</sub> and pEGFP-SseK3 or pEGFP-SseK3<sub>E258A</sub>. Proteins were detected with anti-Arg-GlcNAc, anti-GFP and anti-Flag antibodies as indicated. Antibodies to β-actin were used as a loading control. Representative immunoblot of at least three independent experiments. (B) LC-MS analysis of tryptic digest derived from co-incubation of recombinant His-hTRAILR2<sub>DD</sub> and GST-SseK3 in the presence of UDP-GlcNAc. (C) LC-MS analysis of tryptic digest fractions derived from co-incubation of recombinant His-hTRAILR2<sub>DD</sub> and GST-SseK3 with no sugar donor. (D) HCD fragmentation of recombinant His-hTRAILR2<sub>DD</sub> incubated with GST-SseK3 and UDP-GlcNAc, Andromeda score 167.28. (E) Immunoblot of recombinant His-hTRAILR2<sub>DD</sub> and GST-SseK3 following co-incubation at 37°C for 5 hours. Proteins were detected with anti-Arg-GlcNAc, anti-GST, and anti-His antibodies as indicated. Arrow indicates Arg-GlcNAcylated His-hTRAILR2<sub>DD</sub>. Representative immunoblot of at least three independent experiments.

Figure 6. *In vitro* binding studies of SseK3 and Arg-GlcNAcylation of human TNFR1. (A) Immunoblot of input and immunoprecipitate (IP) of anti-Flag immunoprecipitations performed on lysates of HEK293T cells co-transfected with pFlag-hTNFR1<sub>DD</sub> and pEGFP-SseK3 or pEGFP-SseK3<sub>E258A</sub>. Proteins were detected with anti-Arg-GlcNAc, anti-GFP and anti-Flag antibodies as indicated. β-actin detection was used as a loading control. Representative immunoblot of at least three independent experiments. (B) EThcD fragmentation of Flag-hTNFR1<sub>DD</sub> enriched from HEK293T cells by anti-Flag immunoprecipitation following co-transfection with pEGFP-SseK3, peptide confirmed by manual annotation. (C) *S. cerevisiae* Y2HGold co-transformed with pGBK7-SseK3 and pGADT7-hTRAILR2 DD or pGADT7-hTNFR1 DD and plated onto selective media to select for plasmid carriage (DDO) or to select for protein-protein interactions (QDO). *S. cerevisiae* Y2HGold co-transformed with pGBK7-NleB1 and pGADT7-FADD DD was used as a
positive control for protein-protein interactions. Self-activation by the bait or prey fusion proteins was discounted by co-transformation of *S. cerevisiae* Y2HGold with pGADT7 and pGBK7-SseK3 or co-transformation with pGBK7 and pGADT7-hTNFR1 DD or pGADT7-hTRAILR2 DD.

**Figure 7. SseK3 structure.** (A) Cartoon representation of the SseK3-UDP complex. The structure of SseK3 is shown with the bound UDP drawn as sticks. Secondary structures are numbered. SseK3 displays a GT-A fold and contains an α-helical insertion (marked by a dashed line) The secondary structure elements are rainbow coloured, from blue at the N-terminus to red at the C-terminus. (B) Topology diagram of SseK3 with colours of secondary structures matching those in panel A. (C) Coordination of UDP in the E258Q active site. Residues in SseK3 are shown in green and the UDP moiety is shown in cyan. Hydrogen bonds are shown as black dashed lines. Arg59 is directed toward the uracil while Arg55 is mobile and assumes different conformations in various structures. (D) Coordination of the Mg$^{2+}$ ion and the hydrogen bonds to the diphosphate in the E258Q active site. The Mg$^{2+}$ ion is shown as a cyan sphere. (E) Superposition of SseK3 (red) and GT44 family *Clostridium difficile* toxin A (TcdA) glucosyltransferase domain (PDB code 3SRZ, blue). Toxin A domain is larger than SseK3 and the segments that have no correspondence in SseK3 are painted gray. Two segments of SseK3 without correspondence in toxin A are painted pink. (F) The UDP-GlcNAc and arginine are modelled into the active site of SseK3. SseK3 is represented as a solvent accessible surface coloured by the electrostatic potential (red – negative, blue – positive). UDP_GlcNAc was taken from structure 3SRZ and placed in SseK3 based on the position of UDP in the SseK3 crystal structure. Arginine was positioned in the long arm of the groove and can be easily accommodated. In this position, the arginine could accept GlcNAc from UDP-GlcNAc. The UDP moiety occupies the short arm of the L-shaped groove and an acceptor arginine is modelled into the long arm. (G) The superposition of UDP and GlcNAc from SseK3(E258Q) (white carbons) with UDP-Glc from toxin A glucosyltransferase domain (pink carbons). The figures were prepared with PyMOL software (Schrodinger Inc, Cambridge, MA).
Figure 6

A

Flag-TNFR1co + GFP
Flag + GFP-TNFR1co + GAD7-SseK3
Flag + GFP + GAD7-SseK3
Flag + GFP + GAD7-SseK3co

B

anti-AIP-GlcNAc
anti-Flag
anti-GFP
anti-β-actin

1% input

C

Tumor necrosis factor receptor superfamily member 1A

1 – pGBK7-NleB1 + pGADT7-FADD DD
2 – pGBK7-SseK3 + pGAD7 DD
3 – pGBK7 + pGADT7-TNFR1 DD
4 – pGBK7-SseK3 + pGAD7-TNFR1 DD

1 – pGBK7-NleB1 + pGADT7-FADD DD
2 – pGBK7-SseK3 + pGAD7 DD
3 – pGBK7 + pGADT7-hTRAILR2 DD
4 – pGBK7-SseK3 + pGAD7-hTRAILR2 DD
