

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

Joshua P M Newson^{1,¶}, Nichollas E Scott^{1,¶}, Ivy Yeuk Wah Chung², Tania Wong Fok Lung^{1,§}, Cristina Giogha³, Nancy Wang¹, Richard A Strugnell¹, Nat F Brown⁵, Mirosław Cygler², Jaclyn S Pearson³, and Elizabeth L Hartland^{1,3,4*}

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

27 ABSTRACT

28 Strains of *Salmonella* utilise two distinct type three secretion systems to deliver effector proteins
 29 directly into host cells. The *Salmonella* effectors SseK1 and SseK3 are arginine
 30 glycosyltransferases that modify mammalian death domain containing proteins with N-acetyl
 31 glucosamine (GlcNAc) when overexpressed ectopically or as recombinant protein fusions. Here, we
 32 combined Arg-GlcNAc glycopeptide immunoprecipitation and mass spectrometry to identify host
 33 proteins GlcNAcylated by endogenous levels of SseK1 and SseK3 during *Salmonella* infection. We
 34 observed that SseK1 modified the mammalian signaling protein TRADD, but not FADD as
 35 previously reported. Overexpression of SseK1 greatly broadened substrate specificity, while ectopic
 36 co-expression of SseK1 and TRADD increased the range of modified arginine residues within the
 37 death domain of TRADD. In contrast, endogenous levels of SseK3 resulted in modification of the
 38 death domains of receptors of the mammalian TNF superfamily, TNFR1 and TRAILR, at residues
 39 Arg³⁷⁶ and Arg²⁹³ respectively. Structural studies on SseK3 showed that the enzyme displays a
 40 classic GT-A glycosyltransferase fold and binds UDP-GlcNAc in a narrow and deep cleft with the
 41 GlcNAc facing the surface. Together our data suggests that Salmonellae carrying *sseK1* and *sseK3*
 42 employ the glycosyltransferase effectors to antagonise different components of death receptor
 43 signaling.

44

45

46 AUTHOR SUMMARY

47 Many Gram-negative pathogens employ type three secretion systems to translocate specialised
 48 bacterial effector proteins into the host cell. The activities of many *Salmonella* effectors are not well
 49 understood, and identifying the host targets of these effectors will lead to a better understanding of
 50 *Salmonella* pathogenesis. The overexpression of effectors *in vitro* can provide useful insights into
 51 their function, but these results may not represent the true biological role of these effectors during
 52 infection. In this study, we showed that endogenous levels of two *Salmonella* effectors target
 53 different host death domain containing signaling proteins that are required for inflammatory and
 54 cell death signaling. The study developed new tools to study cognate effector interactions and
 55 established the important of effector expression levels in defining natural and unnatural interactions.

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 glycosyltransferases [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.

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 (Asp²²³-Ala²²⁴-Asp²²⁵) and SseK3 (Asp²²⁶-Ala²²⁷-Asp²²⁸) abrogated their glycosyltransferase activity [9], consistent with findings for the homologous T3SS effector, NleB1, from EPEC [7, 8]. Mutation of a single glutamic acid (Glu²⁵³) 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 glycosylation 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 GlcNAcyate a given target, the approach is substrate specific and there requires knowledge of the potential targets. To identify the full range of host Arg-GlcNAcyated substrates, we developed a quantitative approach using mass spectrometry (MS) to enrich for arginine glycosylated peptides [13]. Using peptides derived from host cells infected with *S. Typhimurium ΔsseK123* over-expressing either SseK1 or inactive SseK1_{E255A},

label-free MS based quantification revealed a broad range of substrates that were only modified in the presence of active SseK1 (Fig. 2A, Supplementary Table 1, Supplementary 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 glycosylation [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 murine TRADD (mTRADD) and FADD (mFADD) (Fig. 2C, Supplementary Table 2, Supplementary 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 (Supplementary Table 3, Supplementary Fig. 2). This suggested TRADD was the preferred substrate when SseK1 was translocated at native levels during infection.

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

A previous report demonstrated that TRADD carrying the amino acid substitution Arg²³³ 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 Arg²⁴³ when SseK1 was overexpressed, but that Arg²³³ was the site of glycosylation when SseK1 was expressed at native levels during infection. The equivalent sites in human TRADD (hTRADD) are Arg²³⁵ and Arg²⁴⁵. To confirm these findings, we generated single and double Arg²³⁵ and Arg²⁴⁵ mutants of hTRADD and expressed these ectopically in HEK239T cells co-transfected with pEGFP-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, Supplementary Table 4). Under these conditions, we detected several further sites of modification at Arg²³⁹, Arg²⁷⁸, and Arg²²⁴, 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 (Supplementary Table 4). We detected different patterns of Arg-GlcNAcylation for each of these mutant Flag-TRADD variants, suggesting that deletion of a preferred glycosylation site caused a shift in site specificity despite comparable protein levels, as determined by MS analysis (Supplementary Table 5, Supplementary Fig 3). 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 Δ *sseK12* mutant or a triple Δ *sseK123* mutant, and we applied label free MS based quantification to screen for glycosylation events in a non-biased manner (Fig. 4A, Supplementary 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 Δ *sseK12* deletion mutant, while no Arg-GlcNAcylation was detected in cells infected with the Δ *sseK123* deletion mutant (Fig. 4B, Supplementary Table 6, Supplementary Fig. 4). Glycosylation of mTRAILR occurred at Arg²⁹³ while mTNFR1 was glycosylated at Arg³⁷⁶. 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). 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_{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_{DD} by GFP-SseK3 but not GFP-SseK3_{E258A} (Fig. 5A). Similarly, Flag-hTNFR1_{DD} was Arg-GlcNAcylated by GFP-SseK3 but not GFP-SseK3_{E258A} (Fig. 6A).

204 To validate the specific site of modification, recombinant GST-SseK3 was incubated with His-
 205 hTRAILR2_{DD} in the presence of the sugar donor UDP-GlcNAc, subjected to tryptic digestion and
 206 analysed by LC-MS (Supplementary Table 7 and 8). We detected hTRAILR2_{DD} modified at Arg³⁵⁹
 207 (Fig. 5B, 5C), equivalent to the modification of mTRAILR at Arg²⁹³ described above. In the
 208 absence of UDP-GlcNAc, the modification of hTRAILR2_{DD} was not observed (Fig 5D). For both
 209 experiments the unmodified peptide ³⁹⁶DASVHTLLDALETGER⁴¹² derived from mTRAILR was
 210 monitored as an internal control and showed comparable ion intensity within the sample input
 211 which was consistent with the total observed protein levels between samples (Supplementary Table
 212 8). Within these digests, we also observed evidence of self Arg-GlcNAcylation of SseK3, even
 213 without the addition of UDP-GlcNAc (Supplementary Table 7, Supplementary Fig 5 to 8).
 214 Immunoblots using anti-Arg-GlcNAc antibodies confirmed the glycosylation of hTRAILR2_{DD} in
 215 the presence of GST-SseK3 and UDP-GlcNAc as well as the apparent self-modification of SseK3
 216 (Fig. 5E).

217
 218 *In vitro* validation of hTNFR1 Arg-GlcNAcylation by GST-SseK3 was complicated by the
 219 insoluble nature of recombinant His-hTNFR1_{DD}. Instead, Flag-hTNFR1_{DD} was enriched from
 220 HEK293T cells by anti-Flag immunoprecipitation following co-transfection with pEGFP-SseK3
 221 (Fig. 6A). Protein samples were subjected to tryptic digestion ahead of analysis by LC-MS and
 222 Arg³⁷⁶ was confirmed as the preferred site of GlcNAcylation of Flag-hTNFR1_{DD} (Fig. 6B,
 223 Supplementary Table 9)

224
 225 To explore the strength of the protein-protein interactions between SseK3 and the novel substrates,
 226 TNFR1 and TRAILR, the yeast two-hybrid system was used to detect protein-protein interactions.
 227 Auxotrophic yeast strains were co-transformed to express SseK3 and the death domain of hTNFR1
 228 (hTNFR1_{DD}). Yeast expressing SseK3 and hTNFR1_{DD} grew when plated on selective media,
 229 indicating a stable interaction between SseK3 and the hTNFR1_{DD} (Fig 6C). In contrast, yeast co-

transformed to express SseK3 and the death domain of hTRAILR2 (hTRAILR2_{DD}) did not grow on selective media, suggesting the interaction between these two proteins is either weaker, transient or that the death domain alone was insufficient for binding (Fig 6C).

Further to understanding the biology of SseK3, we observed that HA tagged SseK3 showed distinct Golgi localisation following *Salmonella* infection, as reported previously for ectopic expression [9]. Interestingly, this localisation did not require the glycosyltransferase activity of SseK3, suggesting that SseK3 may contain a novel Golgi targeting motif (Fig. 7).

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 (Fig. 8A). There are two α -helices parallel to strands β 1- β 3 and two other helices parallel to the other three β -strands but on the opposite side of the β -sheet. An α -helical hairpin is inserted within the long connection between strands β 3 and β 4 of the central β -sheet that is somewhat separated from the rest of the protein and distant from the active site (Fig. 8A, 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

256 clearly visible in the crystal. However, an additional density was present nearby at the bottom of the
 257 long arm of L-groove and we have modelled GlcNAc into this density. In our structures, the UDP
 258 uracil O2 forms a hydrogen bond with the backbone amide of Phe53 while the uracil N3 atom
 259 forms a hydrogen bond with the backbone carbonyl oxygen of the same Phe53 (Fig. 8C). As well,
 260 the sidechains of Arg55 and Arg59 are close to the uracil O4. Moreover, the uracil ring forms a
 261 π -stacking interaction with Trp52 and Phe190. The ribose also forms hydrogen bonds to the protein,
 262 namely, C2 hydroxyl to the carbonyl of Gln51 and to the hydroxyl of Tyr224, while the C3
 263 hydroxyl is hydrogen bonded to the backbone amide of Ala227. In the presence of Mg^{2+} ion (in
 264 E258Q mutant), both phosphates ligand the ion together with Asp228 (from the Dx²²⁸ motif),
 265 Asp325 and Ser327 (Fig. 8D). In addition, the α -phosphate hydrogen bonds to OG of Ser333 and β -
 266 phosphate hydrogen bonds to OG of Ser333 and NE1 of Trp334. In the absence of the metal ion (in
 267 the native structure) UDP still binds to the enzyme and uridine makes the same contacts. The
 268 phosphates are shifted ~ 2 Å away from the Dx motif and make no close contacts with the protein.
 269 In this structure, Tyr334 is disordered.

270

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

275

276

277 **DISCUSSION**

278 Elucidating the biochemical activity of bacterial effector proteins, their preferred host substrates,
279 and contribution to virulence, remains a major priority of host-pathogen research to help elucidate
280 basic processes that are manipulated by highly evolved pathogens, and to provide for new
281 opportunities to disrupt disease causing processes. A number of studies have described putative
282 substrates of the *Salmonella* Typhimurium effectors SseK1 and SseK3 [8, 9, 11, 12] but these
283 studies have relied substantially on in vitro experiments using recombinant proteins or
284 overexpression of the effectors in mammalian cells. Here, we identified the substrates of SseK1 and
285 SseK3 when expressed at native levels during *S. Typhimurium* infection of RAW264.7 cells.

286
287 Although previous reports had suggested both TRADD and FADD were substrates of SseK1 [8, 9,
288 11], our data suggested that TRADD is the preferred substrate of SseK1, as we could not detect
289 modification of FADD during *S. Typhimurium* infection. TRADD plays a key role in the activation
290 of canonical NF- κ B signaling leading to pro-inflammatory cytokine secretion (reviewed in [17, 18])
291 and programmed cell death via TNF induced apoptosis or necroptosis [19]. Previous work has
292 suggested that SseK1 plays a role in inhibiting both NF- κ B activation and necroptotic cell death in
293 infected macrophages [9], and so it is likely that *S. Typhimurium* employs SseK1 to inhibit TNF
294 signaling as required. Despite this, single deletion mutants of *S. Typhimurium* have not clearly
295 demonstrated a requirement for SseK1 in mouse infection models [5, 20, 21]. We speculate that
296 SseK1 acts in concert with other effectors *in vivo* to achieve significant inhibition of NF- κ B
297 signaling during *S. Typhimurium* infection.

298
299 Although SseK3 had been reported to interact with TRADD and TRIM32 [9, 12], here we found
300 that the preferred substrates of SseK3 were TNFR1 and TRAILR, both members of the mammalian
301 TNF receptor superfamily. TNFR1 responds to stimulation by extracellular TNF, and initiates a
302 signalling cascade culminating in either inflammatory cytokine production or programmed cell

303 death, as described above. Similarly, extracellular TRAIL binds the membrane-associated receptors
 304 TRAILR1 [22], TRAILR2 [23], TRAILR3 [24], and TRAILR4 [25]. However, only TRAILR1 and
 305 TRAILR2 contain full length death domains that stimulate a range of signaling events, including
 306 inflammatory cytokine production, extrinsic apoptotic cell death via caspases-8 and -3, necroptosis
 307 or even the promotion of cell survival via anti-apoptotic functions mediated by TRAF2
 308 ubiquitination of caspase-8 [26, 27] (reviewed in [16, 28, 29]). This diversity of signaling outcomes
 309 complicates the interrogation of bacterial manipulation of TRAIL signaling. However, TRAIL
 310 deficient mice showed no difference in susceptibility to infection with *Salmonella* [30], which we
 311 confirmed (data not shown) suggesting that TNFR1 may be a more important target for bacterial
 312 inhibition in mammalian hosts. Consistent with this, TNFR1 deficient mice are highly susceptible to
 313 *Salmonella* infection [31]. Interestingly, a number of studies have implicated *TRAIL* polymorphisms
 314 with increased susceptibility to *Salmonella* infection in chickens [32, 33], suggesting TRAIL may
 315 activate an important immune defence pathway to *Salmonella* in poultry that SseK3 activity may
 316 block.

317

318 In addition to the Arg-GlcNAcylation of host substrates, we also observed the apparent self-
 319 modification of SseK1 and SseK3 when overexpressed. This observation, in conjunction with
 320 increased levels of Arg-GlcNAcylation seen during overexpression, suggested that many substrates
 321 may be non-authentically glycosylated when the effectors are overexpressed. These findings are
 322 similar to a previous report that suggested overexpression of NleB1 also results in Arg-
 323 GlcNAcylation of a broader range of substrates [13]. We note in particular the numerous two-
 324 component response regulators that were Arg-GlcNAcylated during overexpression of SseK1 (Fig
 325 2A, Supplementary Table 1), perhaps suggesting a mechanism for effector-mediated regulation of
 326 two-component signaling outcomes. Surprisingly, these proteins also appeared to be Arg-
 327 GlcNAcyated by SseK1 and SseK3 during growth in LB (Supplementary Figure 9, Supplementary
 328 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 (Supplementary Figure 9). Overall this suggested that these events occur at low stoichiometry, and therefore the biological relevance of such glycosylation events is yet to be established.

Unlike SseK1, SseK3 shows a distinct subcellular localisation to the Golgi. As death receptors such as TNFR1 can be internalised and traffic with trans-Golgi vesicles [34], the localisation of SseK3 is consistent with modification of the receptors during infection. Interestingly, active glycosyltransferase activity was not required for Golgi localisation suggesting that the amino acid sequence of SseK3 may contain a novel Golgi targeting motif.

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 [35] (Fig. 8E). 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 α -1,3-galactosyltransferase [36]. The mobile C-terminus likely forms a gate that allows for an easy access of UDP-GlcNAc to access the site and helps to retain it there until the acceptor arginine binds nearby and the transfer reaction is

completed. To location and conformation of UDP observed in or 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, [37]) (fig. 8G. Moreover, the position of GlcNAc observed in the SseK3(E258Q) 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 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) [38]. 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. 8E). 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 [39]. Comparison of these structures and the recently deposited structure of NleB2 (PDB code 5H5Y) 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.

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

381 components of death receptor pathways, thereby subverting inflammatory and cell death responses
382 *in vivo*.
383

384 MATERIALS AND METHODS

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

388
389 **DNA cloning and purification.** The plasmids and primers used in this study are listed in Tables 2
390 and 3, respectively. DNA-modifying enzymes were used in accordance with the manufacturer's
391 instructions (New England BioLabs). Plasmids were extracted using the QIAGEN QIAprep Spin
392 Miniprep Kit. PCR products and restriction digests were performed using the Wizard SV Gel and
393 PCR Clean-Up System (Promega). pTrc99A-SseK1, pTrc99A-SseK2, and pTrc99A-SseK3 were
394 constructed by amplifying *sseK1*, *sseK2*, and *sseK3* from pEGFP-C2-SseK1, pEGFP-C2-SseK2,
395 and pEGFP-C2-SseK3 using the primer pairs SseK1_{F/R}, SseK2_{F/R}, and SseK3_{F/R}, respectively. The
396 PCR product was digested with *EcoRI* and *BamHI* and ligated into pTrc99A to produce a C-
397 terminal 2x hemagglutinin tag fusion to SseK1, SseK2, and SseK3. Constructs were transformed
398 into XL-1B cells, and verified by colony PCR and sequencing using the primer pair pTrc99A_{F/R}.
399 pEGFP-C2-SseK2 and pEGFP-C2-SseK3 were constructed by amplifying *sseK2* and *sseK3* from *S.*
400 Typhimurium SL1344 genomic DNA using the primer pairs GFPS2_{F/R} and GFPS3_{F/R} respectively
401 and AmpliTaq Gold DNA polymerase. The resultant PCR products were purified and ligated into
402 pGEM-T-Easy vector at an insert:vector molar ratio of 3:1. The ligation reactions were transformed
403 into XL-1 Blue cells and plated onto LA plates containing ampicillin and X-gal. Plasmids were
404 extracted and digested with *EcoRI* and *SalI* to release the bacterial genes, which were gel purified
405 and ligated into pre-digested pEGFP-C2. The ligation reactions were then transformed into XL-1
406 Blue cells and colony PCR was performed using primers pEGFP-C2_{F/R} to select positive clones.
407 The correct insert was confirmed by sequencing using the same primer pair. The primer pair
408 hTRAILR2DD_{F/R} was used to amplify the region encoding the death domain of TRAILR2 from
409 HeLa cDNA (Sigma Aldrich). The resulting amplicon was gel purified and digested with *EcoRI* and

410 *Bam*HI and ligated into pre-digested pGADT7. The ligation reactions were transformed into XL-1
411 Blue cells and plated on LA containing ampicillin. The correct insert was verified by colony PCR
412 and sequencing with the primer pair pGADT7-AD_{F/R}. pFlag-hTRAILR2_{DD} was constructed by
413 amplifying the death domain of human *TNFRSF10B* from pGADT7- hTRAILR2_{DD} using the
414 primer pair hTRAILR2_{DD-F}/ hTRAILR2_{DD-R}. The PCR product was digested with *Eco*RI and
415 *Bam*HI and ligated into p3xFlag-Myc-CMV-24 to produce an N-terminal 3xFlag fusion to
416 hTRAILR2_{DD}. pFlag-hTRAILR2_{DD} was transformed into XL-1B cells, and verified by colony PCR
417 and sequencing using the primer pair p3xFlag-Myc-CMV-24_{F/R}. pGEX-4T-1-SseK3 was
418 constructed by amplifying *sseK3* from pEGFP-C2-SseK3 using the primer pair GST-SseK3_{F/R}. The
419 PCR product was digested with *Eco*RI and *Sal*I and ligated into pre-digested pGEX-4T-1 to
420 produce an N-terminal GST fusion to SseK3. pGEX-4T-1-SseK3 was transformed into XL-1B
421 cells, and verified by colony PCR and sequencing using the primer pair pGEX-4T-1_{F/R}. The region
422 encoding the death domain of TNFR1 was amplified using the primer pair FLAGTNFR1DD_{F/R} and
423 pGADT7-TNFR1_{DD} as template. The resulting amplicon was gel purified and digested with *Bgl*II
424 and *Sal*I and ligated into pre-digested p3xFLAG-Myc-CMV-24 before transforming the reactions
425 into XL-1 Blue cells. Colony PCR and sequencing were performed with the primer pair p3xFlag-
426 Myc-CMV-24_{F/R} to ensure the correct insert has been ligated. pGBKT7-NleB1 was constructed by
427 digesting pGBT9-NleB1 [7], which carries *nleB1* flanked in between the restriction sites *Eco*RI and
428 *Bam*HI, and ligating into pGBKT7 digested with *Eco*RI and *Bam*HI. pGBKT7-SseK3 was
429 constructed by amplifying *sseK3* from pEGFP-C2-SseK3 using the primer pair SseK3_{F/R}. The PCR
430 product was ligated into the cloning vector pGEM-T-Easy and the ligation reaction was transformed
431 into XL-1 Blue cells. Transformants were then selected on LA containing ampicillin and X-gal.
432 Plasmids were extracted and digested with *Eco*RI and *Sal*I to release the bacterial gene, which was
433 then ligated into pre-digested pGBKT7 to create pGBKT7-SseK3. This plasmid was sequenced
434 using the primer pair pGBKT7_{F/R}. pGADT7-FADD_{DD}, pGADT7-TNFR1_{DD}, and pGADT7-
435 hTRAILR2_{DD} were constructed by amplifying the death domain regions of human FADD, TNFR1,

436 and TRAILR2 from HeLa cDNA using the primer pairs FADD_{DD-F/R}, TNFR1_{DD-F/R}, and
 437 hTRAILR2_{DD-F/R}, respectively. Amplified FADD_{DD} was digested with *EcoRI* and *BamHI*,
 438 TNFR1_{DD} was digested with *NdeI* and *BamHI*, and hTRAILR2_{DD} was digested with *EcoRI* and
 439 *BamHI*. Digested PCR products were ligated into pre-digested pGADT7 and transformed into XL-1
 440 Blue cells. Constructs were verified by colony PCR and sequencing using the primer pair
 441 pGADT7_{F/R}. For structural investigations, the segment corresponding to residues 25-335 of *sseK3*
 442 from *Salmonella Typhimurium* (strain SL1344) (Uniprot: A0A0H3NMP8) was cloned into vector
 443 pRL652, a derivative of vector pGEX-4T-1 (GE Healthcare) adapted for ligation-independent
 444 cloning. The construct contained a TEV-cleavable GST tag at the N-terminus.

445

446 **Site-directed mutagenesis.** Site-directed mutagenesis of plasmid constructs was performed using
 447 the QuikChange II Site-Directed Mutagenesis kit (Stratagene, California, USA), according to the
 448 manufacturer's instructions. pTrc99A-SseK1_{DxD(229-231)AAA}, pTrc99A-SseK2_{DxD(239-241)AAA}, and
 449 pTrc99A-SseK3_{DxD(226-228)AAA} were generated using pTrc99A-SseK1, pTrc99A-SseK2, or pTrc99A-
 450 SseK3 as template DNA and amplified by PCR using the primer pairs SseK1_{DxD-F/R}, SseK2_{DxD-F/R},
 451 or SseK3_{DxD-F/R}, respectively. pTrc99A-SseK1_{E255A}, pTrc99A-SseK2_{E271A}, and pTrc99A-SseK3_{E258A}
 452 were generated using pTrc99A-SseK1, pTrc99A-SseK2, or pTrc99A-SseK3 as template DNA and
 453 amplified by PCR using the primer pairs SseK1_{E255A-F/R}, SseK2_{E271A-F/R}, or SseK3_{E258A-F/R},
 454 respectively. pFlag-hTRADD_{R235A} and pFlag-hTRADD_{R245A} were generated using pFlag-hTRADD
 455 as template DNA and amplified by PCR using the primer pairs hTRADD_{R235A-F/R} and
 456 hTRADD_{R245A-F/R}, respectively. pFlag-hTRADD_{R235A/R245A} was generated using pFlag-
 457 hTRADD_{R235A} as template DNA and amplified by PCR using the primer pair hTRADD_{R245A-F/R}.

458 All PCR products were digested with *DpnI* at 37°C overnight before subsequent transformation into
 459 XL1-B cells. Plasmids were extracted and sequenced using the primer pairs pTrc99A_{F/R}, p3xFlag-
 460 Myc-CMV-24_{F/R}, or pEFGP-C2_{F/R}, as required.

461

462 **Mammalian cell culture.** HEK293T cells (human embryonic kidney 293 cells expressing the SV40
463 large T-antigen, source: ATCC® CRL-3216) and RAW264.7 cells (murine leukemic monocyte-
464 macrophage cells source: Richard Strugnell, University of Melbourne) were maintained in DMEM,
465 low glucose with GlutaMAX™ supplement and pyruvate (DMEM (1X) + GlutaMAX(TM)-I)
466 (Gibco, Life Technologies, NY, USA). Tissue culture media was further supplemented with 10%
467 (v/v) heat-inactivated fetal bovine serum (FBS) (Thermo Fisher Scientific). Cells were maintained
468 in a 37 °C, 5% CO₂ incubator, and passaged to a maximum of 30 times. HEK293T cells were split
469 when cells reached 80 to 90% confluency with 1 ml 0.05% Trypsin-EDTA (1X) (Gibco,
470 LifeTechnologies) per 75 cm² of tissue culture, then resuspended with 10 volumes of DMEM
471 supplemented with FBS. RAW264.7 cells were physically detached with a cell scraper and further
472 diluted in fresh DMEM supplemented with FBS.

473

474 **Infection of RAW264.7 cells.** RAW264.7 cells were seeded to 24 well plates at a concentration of
475 3×10^5 cells per well one day before infection. 10 ml LB broths containing appropriate antibiotic
476 were inoculated with *Salmonella* strains and incubated at 37 °C overnight with shaking at 180 rpm.
477 On the day of infection, the OD₆₀₀ readings of the overnight culture were read and used to estimate
478 bacterial counts. Cells were then infected at a multiplicity of infection (MOI) of 10. 24 well plates
479 were centrifuged at 1500 rpm for 5 minutes at room temperature to promote and synchronise
480 infection. Infected cells were incubated at 37 °C, 5% CO₂ for 30 minutes. Culture media was
481 replaced with media containing 100 µg/ml gentamicin (Pharmacia, Washington, USA), and cells
482 were incubated at 37 °C, 5% CO₂ for a further 1 hour. Culture media was replaced with media
483 containing 10 µg/ml gentamicin, and where necessary, 1 mM IPTG, and cells were incubated at 37
484 °C, 5% CO₂ to the required time, post infection.

485

486 **Immunoblotting.** Cells were lysed in cold 1 x KalB lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM
487 NaCl, 1 mM EDTA, 1% (vol/vol) Triton X-100) supplemented with 2 mM Na₃VO₄, 10 mM NaF, 1

488 mM PMSF, and 1 x EDTA-free Complete protease inhibitor cocktail (Roche). Cell lysate was
 489 incubated for at least 30 minutes on ice, then cell debris was pelleted at 13000 rpm at 4 °C for 12
 490 minutes. The soluble protein fraction was mixed with 4 x Bolt® LDS sample buffer (Life
 491 Technologies) and DTT (Astral Scientific) to a final concentration of 50 mM. Proteins were boiled
 492 at 80 to 90 °C for 10 minutes, then loaded to Bolt® 4-12% Bis-Tris Plus gels (Life Technologies)
 493 alongside SeeBlue® pre-stained protein ladder (Life Technologies). Proteins were separated by
 494 electrophoresis using an XCell SureLock™ Mini-Cell system (Life Technologies) with 1 x Bolt®
 495 MES SDS or 1 x Bolt® MOPS SDS running buffer (Life Technologies), according to the
 496 manufacturer's instructions. Following electrophoresis, proteins were transferred onto nitrocellulose
 497 membranes using the iBlot2® gel transfer device (Life Technologies) and iBlot2® nitrocellulose
 498 transfer stacks (Life Technologies), according to the manufacturer's instructions. Membranes were
 499 blocked in 5% (w/v) skim milk in TBS (20 mM Tris, 50 mM NaCl, pH 8.0) with 0.1% (v/v) Tween
 500 20 at room temperature for at least 1 hour with shaking at 60 rpm. Membranes were rinsed and
 501 washed in TBS Tween, then probed one of the following primary antibodies as required at 4 °C
 502 overnight with shaking at 60 rpm: rabbit monoclonal anti-ArgGlcNAc (Abcam), mouse monoclonal
 503 anti-HA (BioLegend), mouse monoclonal anti-GFP (Roche), mouse monoclonal anti-Flag M2-HRP
 504 (Sigma), or mouse monoclonal anti-β-actin (Sigma). Membranes were again rinsed and washed in
 505 TBS Tween, then probed with anti-mouse or anti-rabbit IgG secondary antibodies conjugated to
 506 horseradish peroxidase (PerkinElmer) diluted in TBS with 5% BSA (Sigma) and 0.1% Tween
 507 (Sigma) at room temperature for one hour with shaking at 60 rpm. Membranes were rinsed and
 508 washed in TBS Tween at room temperature for at least 45 minutes with shaking at 60 rpm.
 509 Antibody binding was detected using chemiluminescent substrates for horseradish peroxidase
 510 (HRP) (ECL western blotting reagents (GE Healthcare) or ECL Prime western blotting reagent
 511 (Amersham, USA), according to the manufacturer's instructions, and visualised using an
 512 MFChemiBis imaging station.

513

514 **Transfection of HEK293T cells.** HEK293T cells were transfected using FuGENE®6 transfection
515 reagent (Promega), according to the manufacturer's instructions. Cells were transfected one day
516 after seeding to achieve 80 to 90% confluency. Transfection reagent was mixed with the reduced
517 serum medium Opti-MEM®I (1X) + GlutaMAX(TM)-I (Gibco, Life Technologies), and incubated
518 at room temperature for 5 minutes. Plasmid DNA was added at a transfection reagent:DNA ratio of
519 3:1, and incubated at room temperature for 25 minutes. The reaction was added to previously
520 seeded cells and incubated at 37 °C, 5% CO₂ for 16 to 24 hours.

521

522 **Immunoprecipitation of Flag-tagged fusion proteins.** At required time points post-transfection,
523 cells were lysed and the insoluble fraction removed as described above. Immunoprecipitation of
524 Flag-tagged proteins was performed using Anti-Flag® M2 Magnetic Beads (Sigma-Aldrich),
525 according to the manufacturer's instructions. Beads were first washed twice with lysis buffer, then
526 mixed with the remaining soluble protein fraction and incubated rotating at 4 °C overnight.
527 Following this, beads were washed three times with lysis buffer. Bound protein was eluted by
528 incubating beads in 60 µl of 150 µg/ml Flag peptide (Sigma-Aldrich), rotating at 4 °C for 30
529 minutes. Eluate was mixed with LDS and DTT, and boiled at 80 to 90 °C for 10 minutes. Input and
530 eluate samples were electrophoresed, then visualised by immunoblot as above.

531

532 **Yeast two hybrid assay.** Yeast strain *S. cerevisiae* Y2H Gold (Clontech, California, USA) was
533 transformed or cotransformed with plasmid DNA using the established lithium acetate method [40].
534 Transformants were plated to selective media as required to select for successful single or double
535 transformation. When validating interactions between two proteins, transformants were
536 subsequently plated to highly selective media. Briefly, *S. cerevisiae* Y2H Gold was streaked to
537 YPDA and incubated at 30 °C for 3 days. Healthy colonies were used to inoculate 10 ml YPDA
538 broth at a starting OD₆₀₀ of 0.2, and incubated at 30 °C with shaking at 200 rpm to an OD₆₀₀ of 0.6-
539 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 μ l 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/ μ l 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

592 rpm. Lysates were cooled and protein concentration determined using a BCA assay. 2 mg of protein
 593 from each sample was acetone precipitated by mixing 4 volumes of ice-cold acetone with one
 594 volume of sample. Samples were precipitated overnight at -20 °C and then spun down at 4000 G for
 595 10 minutes at 4 °C. The precipitated protein pellets were resuspended with 80% ice-cold acetone
 596 and precipitated for an additional 4 hours at -20 °C. Samples were spun down at 17000 G for 10
 597 minutes at 4 °C to collect precipitated protein, the supernatant was discarded and excess acetone
 598 driven off at 65 °C for 5 minutes.

599 Dried protein pellets were resuspended in 6 M urea, 2 M thiourea, 40 mM NH₄HCO₃ and reduced /
 600 alkylated prior to digestion with Lys-C (1/200 w/w) then trypsin (1/50 w/w) overnight as previously
 601 described [45]. Digested samples were acidified to a final concentration of 0.5% formic acid and
 602 desalted with 50 mg tC18 SEP-PAK (Waters corporation, Milford, USA) according to the
 603 manufacturer's instructions. Briefly, tC18 SEP-PAKs were conditioned with buffer B (80% ACN,
 604 0.1% formic acid), washed with 10 volumes of Buffer A* (0.1% TFA, 2% ACN), sample loaded,
 605 column washed with 10 volumes of Buffer A* and bound peptides eluted with buffer B then dried.

606 Peptide affinity purification was accomplished according to the protocol of Udeshi *et al.* [46],
 607 modified to allow for Arg-GlcNAc enrichment. Briefly, aliquots of 100 µl of Protein A/G plus
 608 Agarose beads (Santa Cruz, Santa Cruz CA) were washed three times with 1 ml of
 609 immunoprecipitation buffer (IAP, 10 mM Na₃PO₄, 50 mM NaCl, 50 mM MOPS, pH 7.2) and
 610 tumbled overnight with 10 µg of anti-Arg-GlcNAc antibody (ab195033, Abcam) at 4 °C. Coupled
 611 anti-Arg-GlcNAc beads were then washed three times with 1 ml of 100 mM sodium borate (pH 9)
 612 to remove non-bound proteins and cross-linked for 30 minutes rotating using 20 mM Dimethyl
 613 Pimelimidate (Thermo Fisher Scientific) in 100 mM HEPES, pH 8.0. Cross-linking was quenched
 614 by washing beads with 200 mM ethanolamine, pH 8.0, three times then rotating the beads in an
 615 additional 1 ml 200 mM ethanolamine, pH 8.0 for 2 hours at 4 °C. Beads were washed three times
 616 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-termini. 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 contaminants prior. For LFQ comparisons missing values were imputed using Perseus and Pearson correlations visualized using Matlab R2015a (<http://www.mathworks.com>).

Immunofluorescence microscopy. RAW264.7 cells were seeded onto coverslips at a density of 10⁵ cells per well 24 hours prior to infection. Cells were infected with *S. Typhimurium* SL1344 or

669 indicated mutant strains for 18 hours. Infected cells were fixed in 4% (w/v) paraformaldehyde
 670 (Sigma) in PBS for 12 minutes on ice, then washed 3 times with PBS, and incubated in NH₄Cl
 671 1:100 in PBS at room temperature for 20 minutes. Cells were washed twice with PBS, then
 672 incubated in 0.2% (v/v) Triton X-100 (Sigma) in PBS at room temperature for 3 minutes. Cells
 673 were washed three times with PBS, then blocked in 3% (w/v) BSA in PBS at room temperature for
 674 30 minutes. Coverslips were incubated in the following primary antibodies for 1 hour at room
 675 temperature: rabbit monoclonal anti-HA (Cell Signaling) and mouse monoclonal anti-Golgin-97
 676 (Invitrogen), used 1:200 in PBS with 3% BSA. Coverslips were then incubated in the following
 677 secondary antibodies for 30 minute at room temperature: anti-rabbit AlexaFluor 488 (Invitrogen)
 678 and anti-mouse AlexaFluor 568 (Invitrogen), used in 1:2000 in PBS with 3% BSA. Coverslips were
 679 subsequently incubated in Hoechst staining solution (Sigma) for 10 minutes, diluted 1:5000 in PBS.
 680 Coverslips were mounted onto microscope slides using Prolong Gold mounting medium (Life
 681 Technologies) and images were acquired using a Zeiss confocal laser scanning microscope with a
 682 100x EC Epiplan-Apochromat oil immersion objective.

683

684 **Protein expression and purification for structural studies.** The pRL652 plasmid containing
 685 GST-TEV-Ssek#(25-335) was transformed into BL21(DE3) strain. For protein expression a 15 mL
 686 overnight culture in LB was inoculated into one liter of terrific broth media supplemented with 100
 687 µg/ml of ampicillin. The inoculated cultures were grown at 37°C until the OD₆₀₀ reached 1.0. The
 688 cultures were transferred to 18°C, induced with 1mM of isopropyl β-D-1-thiogalactopyranoside
 689 (IPTG), left overnight to grow and harvested by centrifugation at 9,110 x g for 7 minutes.
 690 The cell pellet was resuspended in lysis buffer (50 mM Tris-HCl buffer pH 8.0, 10% glycerol, and
 691 0.1% Triton X) and the cells were lysed in a cell disruptor (Constant Systems Ltd., Northants,
 692 United Kingdom). Cell debris were removed by centrifugation at 28,965 x g for 30 minutes. The
 693 supernatant was loaded on 10 mL Glutathione-Superflow resin (Clontech) column equilibrated with
 694 standard buffer (20 mM Tris pH 8.0 and 150 mM NaCl). The column was washed with 5 column

695 volumes of standard buffer. The beads were then incubated with TEV protease (33 µg/mL) in 30
696 mL of standard buffer for overnight at room temperature. The flow-through containing cleaved
697 SseK3 was collected, concentrated to 30 mg/mL with the Millipore centrifugal filter with a
698 molecular weight cut-off of 10,000 for crystallization trials. The E258Q mutant was generated using
699 KOD Hot Start DNA Polymerase (Sigma-Aldrich) with mutagenic primers and WT plasmid as a
700 template according to the manufacturer's instructions. The E258Q mutant was purified using the
701 same protocol.

702

703 The seleno-methionine (SeMet) derivative of SseK3 was expressed in auxotrophic *E. coli* strain
704 B834(DE3). A 50 mL overnight culture in medium A (M9 medium, trace elements, glucose,
705 MgSO₄, CaCl₂, Biotin, 50 mg/mL methionine and thiamin) was used to inoculate 1 L of medium A.
706 All media used were supplemented with 100 µg/ml of Ampicillin. The cells were grown with
707 shaking at 37°C until OD₆₀₀ reached 1.0. The cells were pelleted at 4°C and then resuspended in 1 L
708 of media A without methionine. The culture was further incubated for 4 hours at 37°C and 50 mg of
709 SeMet was then added. After 30 minutes of incubation, the cultures were then induced with 1 mM
710 of IPTG and continued to grow for additional 10 h at 18°C. The cells were harvested by
711 centrifugation at 9,100 x g for 7 min. The SeMet-labelled protein was purified the same way as the
712 native protein.

713

714 **Protein Crystallization.** Initial crystals were obtained by screening using commercial and in-house
715 screens in a 96-well plate format. The crystallization was setup using Gryphon crystallization robot
716 (Art Robbins Instruments, Sunnyvale, CA). The best crystals were obtained using hanging-drop
717 vapor diffusion method at 20°C. 1 µL of protein solution supplemented with 3 mM of UDP-GlcNAc
718 was mixed with 1 µL of reservoir solution containing 0.5 M NaCl, 0.1 M Tris pH 8.5, 18% PEG
719 3350 and 5% MPD and suspended over 0.5 mL of reservoir. These crystals displayed P2₁ space
720 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 Mg^{2+} 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 $MgCl_2$. They displayed $P2_12_12_1$ 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, three molecules of TRIS, five molecules of methyl-pentanediol and 141 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.

747 **Table 1. Strains used in this study.**

Strain	Relevant characteristics	Reference
SL1344	Wild type <i>S. enterica</i> serovar Typhimurium strain SL1344	[6]
$\Delta sseK1/2/3$	SL1344 $\Delta sseK1\Delta sseK2\Delta sseK3$	[6]
$\Delta sseK2/3$	SL1344 $\Delta sseK2\Delta sseK3$	Nathaniel Brown
$\Delta sseK1/2$	SL1344 $\Delta sseK1\Delta sseK2$	[5]
<i>S. cerevisiae</i> Y2H Gold	<i>MATa</i> , <i>trp1-901</i> , <i>leu2-3, 112</i> , <i>ura3-52</i> , <i>his3-200</i> , <i>gal4Δ</i> , <i>gal80Δ</i> , <i>LYS2 :: GAL1_{UAS}-Gal1_{TATA}-His3</i> , <i>GAL2_{UAS}-Gal2_{TATA}-Ade2 URA3 :: MEL1_{UAS}-Mell_{TATA} AUR1-C MEL1</i>	Clontech

748

749

750 **Table 2. Plasmids used in this study.**

Plasmid	Relevant characteristics	Reference
pTrc99A	Low copy bacterial expression vector with inducible <i>lacI</i> promoter, Amp ^R	Pharmacia Biotech
pTrc99A-SseK1	<i>sseK1</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, Amp ^R	This study
pTrc99A-SseK1 _{DxD(229-231)AAA}	<i>sseK1</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, with DxD (229-231) catalytic motif mutated to AAA, Amp ^R	This study
pTrc99A-SseK1 _{E255A}	<i>sseK1</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, with Glu255 mutated to Ala, Amp ^R	This study
pTrc99A-SseK2	<i>sseK2</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, Amp ^R	This study
pTrc99A-SseK2 _{DxD(239-241)AAA}	<i>sseK2</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A with DxD(239-241) catalytic motif mutated to AAA, Amp ^R	This study
pTrc99A-SseK2 _{E271A}	<i>sseK2</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, with Glu271 mutated to Ala, Amp ^R	This study
pTrc99A-SseK3	<i>sseK3</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, Amp ^R	This study
pTrc99A-SseK3 _{DxD(226-228)AAA}	<i>sseK3</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A with DxD(226-228) catalytic motif mutated to AAA, Amp ^R	This study
pTrc99A-SseK3 _{E258A}	<i>sseK3</i> from <i>S. Typhimurium</i> SL1344 in pTrc99A, with Glu258 mutated to Ala, Amp ^R	This study
p3xFlag-Myc-CMV-24	Dual tagged N-terminal Met-3xFlag and C-terminal <i>c-myc</i> expression vector, Amp ^R	Sigma-Aldrich
pFlag-TRADD	Human TRADD in p3xFlag-Myc-CMV, Amp ^R	Jürg Tschopp
pFlag-TRADD _{R235A}	Human TRADD with Arg235 mutated to Ala, in p3xFlag-Myc-CMV, Amp ^R	This study

pFlag-TRADD _{R245A}	Human TRADD with Arg245 mutated to Ala, in p3xFlag- <i>Myc</i> -CMV, Amp ^R	This study
pFlag-TRADD _{R235A/R245A}	Human TRADD with Arg235 and Arg245 mutated to Ala, in p3xFlag- <i>Myc</i> -CMV, Amp ^R	This study
pFlag-hTRAILR2 _{DD}	Human TRAILR2 death domain in p3xFlag- <i>Myc</i> -CMV, Amp ^R	This study
pFlag-hTNFR1 _{DD}	Human TNFR1 death domain in p3xFlag- <i>Myc</i> -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	<i>sseK1</i> from <i>S. Typhimurium</i> SL1344 in pEGFP-C2, Kan ^R	This study
pEGFP-C2-SseK3	<i>sseK3</i> from <i>S. Typhimurium</i> SL1344 in pEGFP-C2, Kan ^R	This study
pEGFP-C2-SseK3 _{E258A}	<i>sseK3</i> from <i>S. Typhimurium</i> 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	<i>nleB1</i> from EPEC E2348/69 in pGBKT7, Kan ^R , Trp	This study
pGBKT7-SseK3	<i>sseK3</i> from <i>S. Typhimurium</i> 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	<i>sseK3</i> from <i>S. Typhimurium</i> SL1344 in pGEX-4T-1, Amp ^R	This study

751

752

753 **Table 3. Primers used in this study.**

Primer	Primer sequence 5'-3'
pTrc99A _F	CGGTTCTGGCAAATATTC
pTrc99A _R	GCAGTTCCCTACTCTCGC
p3xFlag-Myc-CMV-24 _F	AATGTCGTAATAACCCCGCCCCGTTGACGC
p3xFlag-Myc-CMV-24 _R	TATTAGGACAAGGCTGGTGGGCAC
pEGFP-C2 _F	AACACCCCCATCGGCG
pEGFP-C2 _R	GTAACCATTATAAGCTGC
pGBKT7-BD _F	AATACGACTCACTATAGG
pGBKT7-BD _R	CGTTTTTAAACCTAAGAGTC
pGADT7-AD _F	AATACGACTCACTATAGG
pGADT7-AD _R	GGTGCACGATGCACAG
pGEX-4T-1 _F	CGTATTGAAGCTATCCCACAA
pGEX-4T-1 _R	GGGAGCTGCATGTGTCAGAG
pET28a _F	AATACGACTCACTATAGG
pET28a _R	GCTAGTTATTGCTCAGCGG
SseK1 _F	CGGAATTCATGGAGCATTTAATTGTTATGATCCC
SseK1 _R	CGGGATCCCTACGCATAATCCGGGCACATCATAACGGATACG CATAATCCGGGCACATCATAACGGATACTGCACATGCCTCGC CC
SseK2 _F	CGGAATTCATGGGCACGTTTTTAATGCCG
SseK2 _R	CGGGATCCTTACGCATAATCCGGGCACATCATAACGGATACG CATAATCCGGGCACATCATAACGGATACTCCAAGAACTGGC AG
SseK3 _F	CGGAATTCATGTTTTCTCGAGTCAGAGG
SseK3 _R	CGGGATCCTTACGCATAATCCGGGCACATCATAACGGATACG CATAATCCGGGCACATCATAACGGATATCTCCAGGAGCTGAT AGTC
GST-SseK3 _F	CGCGAATTCATGTTTTCTCGAGTCAGAGGTTTTTC
GST-SseK3 _R	CGCGTCGACTTATCTCCAGGAGCTGATAGTCAAACCTGC
FADD _F	CGCGAATTCATGCTGTGTGCAGCATTTAACGTCATATG
FADD _R	CGCGGATCCTTACTGCTGAACCTCTTGTACCAGG
hTNFR1 _{DD-F}	CGCCATATGATGACGCTGTACGCCGTGGTGG
hTNFR1 _{DD-R}	CGCGGATCCTCACTCGATGTCCTCCAGGCAGC
hTRAILR2 _{DD-F}	CGCGAATTCATGGATCCCCTGAGACTCTGAGAC
hTRAILR2 _{DD-R}	CGCGGATCCTTAGAACTTTCCAGAGCTCAACAAGTG

SseK1 _{DXD-F}	GGTGTATATATCTTGCTGCTGCTATGATTATCACGGAAAA ACTGG
SseK1 _{DXD-R}	CCAGTTTTTCCGTGATAATCATAGCAGCAGCAAGATATAT ACACC
SseK2 _{DXD-F}	GTGGGTGCATATATCTTGCTGCAGCTATGTTACTTACTGAT AAAC
SseK2 _{DXD-R}	GTTTATCAGTAAGTAACATAGCTGCAGCAAGATATATGCA CCCAC
SseK3 _{DXD-F}	CTGGAGGTGGCTGCATATATCTTGCTGCTGCTATGTTACTT ACAG
SseK3 _{DXD-R}	CTGTAAGTAACATAGCAGCAGCAAGATATATGCAGCCAC CTCCAG
SseK1 _{E255A-F}	CGTGCTTCTATGGCAAACGGGATAATAGCT
SseK1 _{E255A-R}	AGCTATTATCCCGTTTGCCATAGAAGCACG
SseK2 _{E271A-F}	TGTTAGCCTTGCAAATGGGATTATTGCTGT
SseK2 _{E271A-R}	ACAGCAATAATCCCATTTGCAAGGCTAACA
SseK3 _{E258A-F}	GCATGAGTCTTGCAAATGGGATTATCGCCG
SseK3 _{E258A-R}	CGGCGATAATCCCATTTGCAAGACTCATGC
hTRADD _{R235A-F}	CGCAAGGTGGGGGCCTCACTGCAGCGAG
hTRADD _{R235A-R}	CTCGCTGCAGTGAGGCCCCCACCTTGCG
hTRADD _{R245A-F}	CGCCGGGTCCGCCAGCGCCCGG
hTRADD _{R245A-R}	CCGGGCGCTGGCGGACCCGGCG
GFPS1 _F	AAAGAATTCATGGAGCATTTAATTGTTATG
GFPS1 _R	AAAGGATCCCTACTGCACATGCCTCG
hTRAILR2 _{DD-F2}	CGCGAATTCATGGATCCCCTGAGACTCTGAGAC
hTRAILR2 _{DD-R2}	CCAAGCTTTTAGAACTTTCCAGAGCTCAACAAGTGG

754

755

756 **Table 4. Data collection and structure refinement statistics.**

	SeMet SseK3	Native SseK3	SseK3 E258Q
Data collection			
Space group	P2 ₁	P2 ₁	P22 ₁ 2 ₁
<i>a,b,c</i> (Å)	96.2, 75.0, 97.3	74.5, 96.2, 90.7	74.9, 96.6, 107.2
□□□□□□□(°)	90, 108.9, 90	90, 93.3, 90	90, 90, 90
Wavelength (Å)	0.9794	0.9801	0.97949
Resolution (Å)	47.793-2.6 (2.66-2.6)	48.360-2.3 (2.35-2.3)	48.29-2.6 (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)
Refinement statistics			
R _{cryst} /R _{free}		0.2172 / 0.2502	
Wilson B (Å ²)		28.52	
Rmsd angles (°)		0.813	
Rmsd bonds (Å)		0.004	
Favoured (%)		97.68	
Allowed (%)		1.74	
PDB code		6CGI	6DUS

757

758

759 **ACKNOWLEDGMENTS**

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

766

767

768

REFERENCES

1. Galan JE, Curtiss R 3rd (1989) Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. Proc Natl Acad Sci USA 86: 6383–6387.
2. Shea JE, Hensel M, Gleeson C, Holden DW (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. Proc Natl Acad Sci USA 93: 2593–2597.
3. Jennings E, Thurston TLM, Holden DW (2017) *Salmonella* SPI-2 Type III Secretion System Effectors: Molecular Mechanisms And Physiological Consequences. Cell Host Microbe 22: 217–231.
4. LaRock DL, Chaudhary A, Miller SI (2015) Salmonellae interactions with host processes. Nat Rev Microbiol 13: 191–205.
5. Kujat Choy SL, Boyle EC, Gal-Mor O, Goode DL, Valdez Y, et al. (2004). SseK1 and SseK2 are novel translocated proteins of *Salmonella enterica* serovar typhimurium. Infect Immun 72:5115–5125.
6. Brown NF, Coombes BK, Bishop JL, Wickham ME, Lowden MJ, et al. (2011). *Salmonella* phage ST64B encodes a member of the SseK/NleB effector family. PLoS One 6:e17824.
7. Pearson JS, Giogha C, Ong SY, Kennedy CL, Kelly M, et al. (2013). A type III effector antagonizes death receptor signaling during bacterial gut infection. Nature 501: 247–251.
8. Li S, Zhang L, Yao Q, Li L, Dong N, et al. (2013). Pathogen blocks host death receptor signaling by arginine GlcNAcylation of death domains. Nature 501: 242–246.
9. Gunster RA, Matthews SA, Holden DW, Thurston TLM (2017). SseK1 and SseK3 Type III Secretion System Effectors Inhibit NF-κB Signaling and Necroptotic Cell Death in *Salmonella*-Infected Macrophages. Infect Immun 85: e00010–17.
10. Wong Fok Lung T, Giogha C, Creuzburg K, Ong SY, Pollock GL, et al. (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, et al. (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 TW, Giogha C, Hill MM, et al. (2015) SseK3 Is a *Salmonella* Effector That Binds TRIM32 and Modulates the Host's NF-κB Signaling Activity. PLoS One 10: e0138529.
13. Scott NE, Giogha C, Pollock GL, Kennedy CL, Webb AI, et al. (2017). The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). J Biol Chem 292: 17337–17350.
14. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ (2012) Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Mol Cell

Proteomics 11: 1475-1488.

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

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

17. Wertz IE, Dixit VM (2010) Regulation of death receptor signaling by the ubiquitin system. *Cell Death Differ* 17: 14-24.

18. Silke J (2011) The regulation of TNF signaling: what a tangled web we weave. *Curr Opin Immunol* 23: 620-626.

19. Vandenabeele P, Galluzzi L, Vanden Berghe T, Kroemer G (2010) Molecular mechanisms of necroptosis: an ordered cellular explosion. *Nat Rev Mol Cell Biol* 11: 700-714.

20. Buckner MMC, Croxen MA, Arena ET, Finlay BB (2011) 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 AS, Mushamiri I, Niemann GS, Brown RN, Adkins JN, et al. (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 AM, Gentz R, Ebner R, et al. (1997) The Receptor for the Cytotoxic Ligand TRAIL. *Science* 276: 111-113.

23. Walczak H, Degli-Esposti MA, Johnson RS, Smolak PJ, Waugh JY, et al. (1997) TRAIL-R2: a novel apoptosis-mediating receptor for TRAIL. *EMBO J* 16: 5386-5397.

24. Sheridan JP, Marsters SA, Pitti RM, Gurney A, Skubatch M, et al. (1997). Control of TRAIL-Induced Apoptosis by a Family of Signaling and Decoy Receptors. *Science* 277: 818-821.

25. Marsters SA, Sheridan JP, Pitti RM, Huang A, Skubatch M, et al. (1997) A novel receptor for Apo2L/TRAIL contains a truncated death domain. *Curr Biol* 7: 1003-1006.

26. Kischkel FC, Lawrence DA, Chuntharapai A, Schow P, Kim KJ, et al. (2000) Apo2L/TRAIL-dependent recruitment of endogenous FADD and caspase-8 to death receptors 4 and 5. *Immunity* 12: 611-620.

27. Sprick MR, Weigand MA, Rieser E, Rauch CT, Joo P, et al. (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.

28. LaFont E, Hartwig T, Walczak H (2018) Paving TRAIL's Path with Ubiquitin. *Trends Biochem Sci* 43: 44-60.

29. Jouan-Lanhout S, Arshad MI, Piquet-Pellorce C, Martin-Chouly C, Le Moigne-Muller G, et al. (2012) TRAIL induces necroptosis involving RIPK1/RIPK3-dependent PARP-1 activation. *Cell Death Differ* 19: 2003-2014.

30. Diehl GE, Yue HH, Hsieh K, Kuang AA, Ho M, et al. (2004) TRAIL-R as a negative regulator of innate immune cell responses. *Immunity* 21: 877-889.
31. Everest P, Roberts M, 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, Lamont SJ (2003). Association of INOS, TRAIL, TGF-beta2, TGF-beta3, and IgL genes with response to *Salmonella enteritidis* in poultry. *Genet Sel Evol* 35: S99-111.
33. Tohidi R, Idris IB, Malar Panandam J, Hair Bejo M (2013). The effects of polymorphisms in 7 candidate genes on resistance to *Salmonella Enteritidis* in native chickens. *Poult Sci* 92: 900-909.
34. Schneider-Brachert W, Tchikov V, Neumeyer J, Jakob M, Winoto-Morbach S, et al. (2004). Compartmentalization of TNF receptor 1 signaling: internalized TNF receptosomes as death signaling vesicles. *Immunity* 21: 415-428.
35. Campbell JA, Davies GJ, Bulone V, Henrissat B (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J* 329: 929-939.
36. Zhang Y, Swaminathan GJ, Deshpande A, Boix E, Natesh R, et al. (2003) Roles of individual enzyme-substrate interactions by alpha-1,3-galactosyltransferase in catalysis and specificity. *Biochemistry* 42: 13512-13521.
37. Pruitt RN, Chumbler NM, Rutherford SA, Farrow MA, Friedman DB, et al. (2012) Structural determinants of *Clostridium difficile* toxin A glucosyltransferase activity. *J Biol Chem* 287: 8013-8020.
38. Unligil UM, Zhou S, Yuwaraj S, Sarkar M, Schachter H, et al. (2000) X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* 19: 5269-5280.
39. Esposito D, Günster RA, Martino L, El Omari K, Wagner A, et al. (2018). Structural basis for the glycosyltransferase activity of the *Salmonella* effector SseK3. *J Biol Chem* 293: 5064-5078.
40. Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11: 355-360.
41. Kang D, Gho YS, Suh M, Kang C (2002) Highly sensitive and fast protein detection with Coomassie brilliant blue in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Bull Korean Chem Soc* 23: 1511-1512.
42. Ishihama Y, Rappsilber J, Mann M (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *J Proteome Res* 4: 988-994.
43. Rappsilber J, Mann M, 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 SJ, Azimifar SB, Mann M (2015) High-throughput phosphoproteomics reveals in

- vivo insulin signaling dynamics. Nat Biotechnol 33: 990-995.
45. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AV, et al. (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-MCP201.
46. Udeshi ND, Svinkina T, Mertins P, Kuhn E, Mani DR, et al. (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, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26: 1367-1372.
48. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, et al. (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, Sinitcyn P, Carlson A, Hein MY, et al. (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(Pt 2): 125–132.
51. Fodje M, Grochulski P, Janzen K, Labiuk S, Gorin J, et al. (2014) 08B1-1: an automated beamline for macromolecular crystallography experiments at the Canadian Light Source. J Synchrotron Radiat 21(Pt 3): 633-637.
52. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66(Pt 2): 213-221.
53. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66(Pt 4): 486-501.

SUPPORTING INFORMATION LEGENDS

Supplementary Table 1. Arginine-GlcNAcylated peptide pull downs from SseK1

overexpression during infection. A total of 133 unique arginine-GlcNAcylated peptides corresponding to 113 localized Arg-GlcNAc sites observed from triplicate biological infections with Δ sseK123 over-expressing either SseK1 or inactivated SseK1_{E255A}. The site of arginine-GlcNAcylation is provided with 113 sites localized, as defined by a localization score of >0.75, and an additional 35 sites unable to be confidently localized (localization score of <0.75). For assigned arginine glycopeptides and arginine glycosylation sites, the protein, gene, score, ion intensity, and biological replicate in which the identity was observed are provided.

Supplementary Table 2. Arginine-GlcNAcylated peptides pull downs from SSeK1 endogenous

levels infections. A total of 8 unique arginine-GlcNAcylated peptides corresponding to 11 localized Arg-GlcNAc sites observed from triplicate biological infections with Δ sseK123 or Δ sseK23. The site of arginine-GlcNAcylation is provided with 11 sites localized, as defined by a localization score of >0.75, and an additional 35 sites unable to be confidently localized (localization score of <0.75). For assigned arginine glycopeptides and arginine glycosylation sites, the protein, gene, score, ion intensity, and biological replicate in which the identity was observed are provided.

Supplementary Table 3: iBAQ based analysis of Δ sseK2/3 infections colon inputs for Arg-

GlcNAc enrichment. For the 5664 proteins observed from RAW264.7 infected with Δ sseK23, iBAQ intensity values generated using Maxquant are provided. For assigned proteins, the iBAQ values, score, summed ion intensity, number of MS/MS events, LFQ values and protein name gene, and biological replicate in which the identity was observed are provided.

Supplementary Table 4. Arginine-GlcNAcylated peptides and sites from co-transfection with

Flag-hTRADD variants and pEGFP-SseK1. A total of 15 unique arginine-GlcNAcylated

peptides are observed from hTRADD upon transfection of pEGFP-SseK1. Across the different Flag-hTRADD variants, different patterns of arginine-GlcNAcylation are observed with 13 sites of arginine-GlcNAcylation identified. For assigned arginine glycopeptides and sites, the protein, gene, score, ion intensity, and biological replicate in which the identity was observed are provided.

Supplementary Table 5. Proteins identified from co-transfection with Flag-hTRADD variants and pEGFP-SseK1. Protein analysis of Flag-hTRADD variants in co-transfected samples. For assigned proteins, the LFQ values, score, summed ion intensity, number of MS/MS events and protein name gene, and biological replicate in which the identity was observed are provided.

Supplementary Table 6. Arginine-GlcNAcylation peptides pull downs from SseK3 endogenous levels infections. A total of 14 unique arginine-GlcNAcylation peptides corresponding to 6 localized Arg-GlcNAc sites observed from triplicate biological infections with $\Delta sseK123$ or $\Delta sseK12$. The site of Arginine-GlcNAcylation provided with 6 sites localized, as defined by a localization score of >0.75 , and 1 additional sites unable to be confidently localized (localization score of <0.75). For assigned arginine glycopeptides and arginine glycosylation sites the protein, gene, score, ion intensity, and biological replicate in which the identity was observed are provided.

Supplementary Table 7. Arginine-GlcNAcylation peptides and sites from in vitro assays of hTRAIL and SseK3. A total of 61 unique arginine-GlcNAcylation peptides corresponding to 15 Arg-GlcNAc sites within SseK3 and hTRAIL were observed within in-vitro Arg-GlcNAcylation assays. For assigned arginine glycopeptides and arginine glycosylation sites, the protein, gene, score, ion intensity, and in vitro condition in which the identity was observed are provided.

Supplementary Table 8. Proteins identified from in vitro assays of hTRAIL and SseK3. Protein analysis of in vitro assay samples. For assigned proteins, the LFQ values, score, summed ion

1014 intensity, number of MS/MS events and protein name gene, and biological replicate in which the
1015 identity was observed are provided.

1016

1017 **Supplementary Table 8. Proteins identified from *in vitro* assays of hTRAIL and SseK3.** Protein
1018 analysis of *in vitro* assay samples. For assigned proteins, the LFQ values, score, summed ion
1019 intensity, number of MS/MS events and protein name gene, and biological replicate in which the
1020 identity was observed are provided.

1021

1022 **Supplementary Table 9. Proteins identified from co-transfection of hTNFR1_{DD} and SseK3.**
1023 Protein analysis of co-transfection assay samples of hTNFR1_{DD}. For assigned proteins, the LFQ
1024 values, score, summed ion intensity, number of MS/MS events and protein name gene, and
1025 biological replicate in which the identity was observed are provided.

1026

1027 **Supplementary Table 10. Arginine-GlcNAcylated of Salmonella proteins.** A total of 40 unique
1028 arginine-GlcNAcylated peptides were observed from triplicate biological experiments of stationary
1029 phase Δ sseK123, Δ sseK12, Δ sseK23 and wild type SL1344. For assigned arginine glycopeptides
1030 and arginine glycosylation sites the protein, gene, score and biological replicate in which the
1031 identity was observed are provided.

1032

1033

1034 **Supplementary Figure 1. Pearson correlation of SseK1 related pulldowns.** To assess the
1035 reproducibility of arginine-GlcNAcylation, pull down heat maps of observed peptides are provided.
1036 A) Δ *sseK23* over-expressing either SseK1 or inactivated SseK1_{E255A}. The mean correlation
1037 between each replicate is 0.74922. B) Δ *sseK123* or *Salmonella* Δ *sseK23* infections. The mean
1038 correlation between each replicate is 0.70054.

1039
1040 **Supplementary Figure 2. Observed proteome of *Salmonella* Δ *sseK23* infection inputs.** During
1041 infection of RAW264.7 cells with Δ *sseK23*, 5664 proteins were observed across the input
1042 proteomes. Relative abundance of proteins based on the log₁₀ iBAQ intensity values generated
1043 using Maxquant are plotted based on ranked order. The position in the rank order of death domain
1044 containing proteins and TNFRSF proteins are shown. Within the input of all samples TRADD is
1045 higher in relative abundance then FADD.

1046
1047 **Supplementary Figure 3. FLAG-hTRADD levels observed in co-transfection studies.** The LFQ
1048 values for all co-transfection of FLAG-hTRADD variants are shown. Comparable levels of
1049 TRADD were observed within all samples.

1050
1051 **Supplementary Figure 4. Pearson correlation of SseK3 related pulldowns.** To assess the
1052 reproducibility of arginine-GlcNAcylation pull down, heat maps of observed peptides are provided
1053 for Δ *sseK12* or Δ *sseK123* infections. The mean correlation between each replicate is 0.737509.

1054
1055 **Supplementary Figure 5. The identification of Arg-GlcNAcylation sites within recombinant**
1056 **SseK3.** Manually curated EThcD spectra confirming the glycosylation of arginine-335 within the
1057 SseK3.

1058

1059 **Supplementary Figure 6. The identification of Arg-GlcNAcylation sites within recombinant**
1060 **SseK3.** Manually curated EThcD spectra showing glycosylation of arginine-305 within SseK3.
1061
1062 **Supplementary Figure 7. The identification of Arg-GlcNAcylation sites within recombinant**
1063 **SseK3.** Manually curated EThcD spectra showing glycosylation of arginine-153 within SseK3.
1064
1065 **Supplementary Figure 8. The identification of Arg-GlcNAcylation sites within recombinant**
1066 **SseK3.** Manually curated EThcD spectra showing glycosylation of arginine-184 within SseK3.
1067
1068 **Supplementary Figure 9. Arg-GlcNAcylation within stationary phase *Salmonella* grown in**
1069 **LB.** A) Ion Intensity heatmap of Arg-GlcNAcylation sites identified from biological triplicate
1070 pulldown from Δ sseK123, Δ sseK12, Δ sseK23 and wild type SL1344 during stationary phase growth
1071 in LB. All arginine-GlcNAcylation peptides were dependent on SseK1 and SseK3 B) Pearson
1072 correlation of Arg-GlcNAc pulldowns within LB grown strains of *Salmonella*. To assess the
1073 reproducibility of arginine-GlcNAcylation pull down heat maps of observed peptides are provided.
1074
1075
1076

FIGURE LEGENDS

Figure 1. Immunoblot of RAW264.7 cells infected with derivatives of *S. Typhimurium*

SL1344. Wild type *S. Typhimurium*, a triple Δ *sseK123* 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 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* Δ *sseK123* complemented with either SseK1-HA or SseK1E_{255A}-HA. Arg-GlcNAcylated peptides are presented as a volcano plot depicting mean ion intensity peptide ratios of SseK1-HA versus SseK1E_{255A}-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* Δ *sseK123* or *S. Typhimurium* Δ *sseK23*. Arginine-glycosylated peptides are presented as a volcano plot depicting mean log₂ ion intensity peptide ratios of Δ *sseK123* versus Δ *sseK23* 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 Arg²³³ 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* Δ *sseK123* or Δ *sseK12*. Arginine-glycosylated peptides are presented as a volcano plot depicting mean log₂ ion intensity peptide ratios of Δ *sseK123* versus Δ *sseK12* plotted against logarithmic *t* test *p* values from biological triplicate experiments. Arg-GlcNAcyated peptides are annotated by gene name and shaded blue. (B) Heat map showing observed ion intensity of Arg-GlcNAcyated peptides between biological triplicates. (C) Partial sequence alignment showing observed GlcNAcyated 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_{DD} and pEGFP-SseK3 or pEGFP-SseK3_{E258A}. Proteins

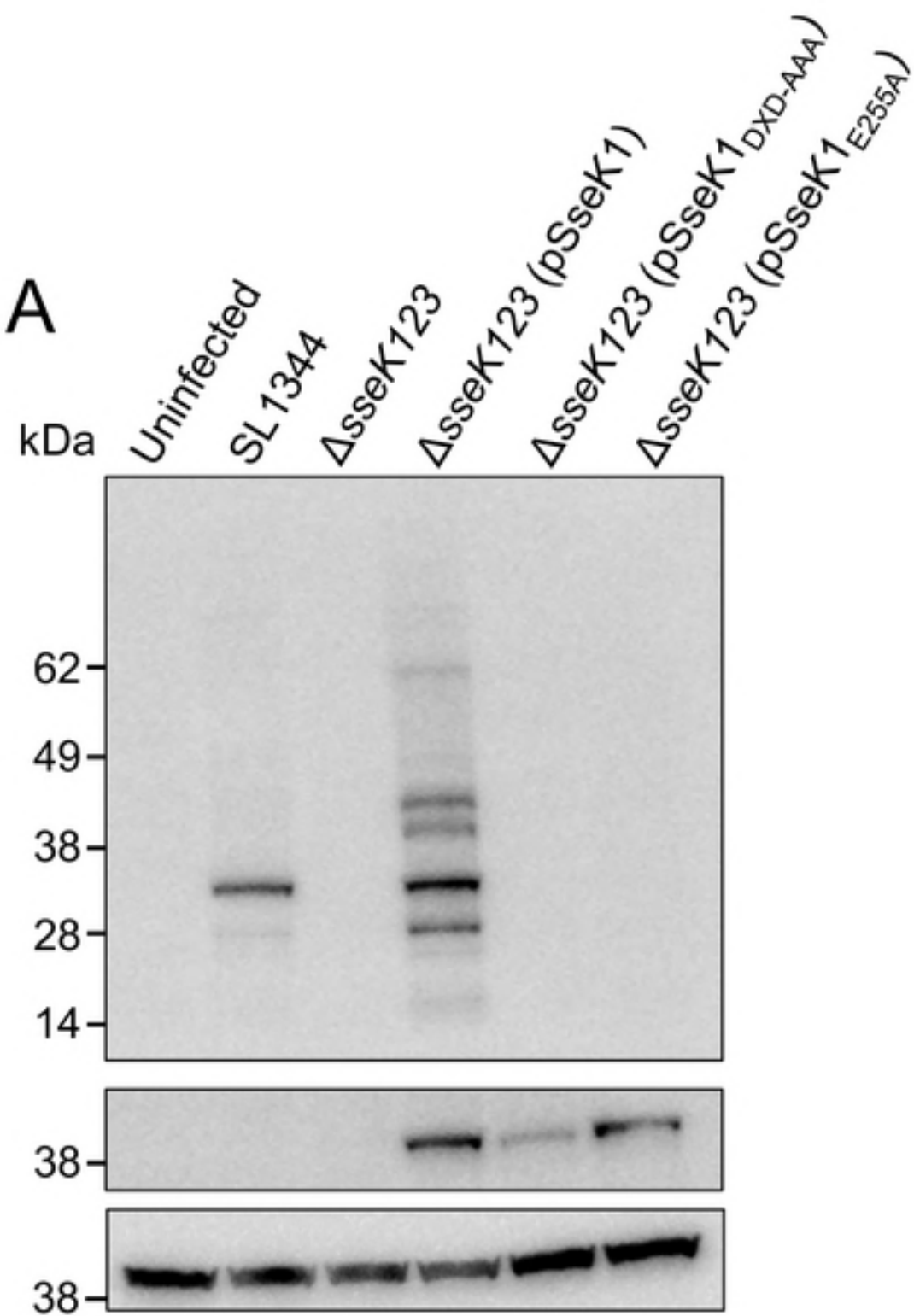
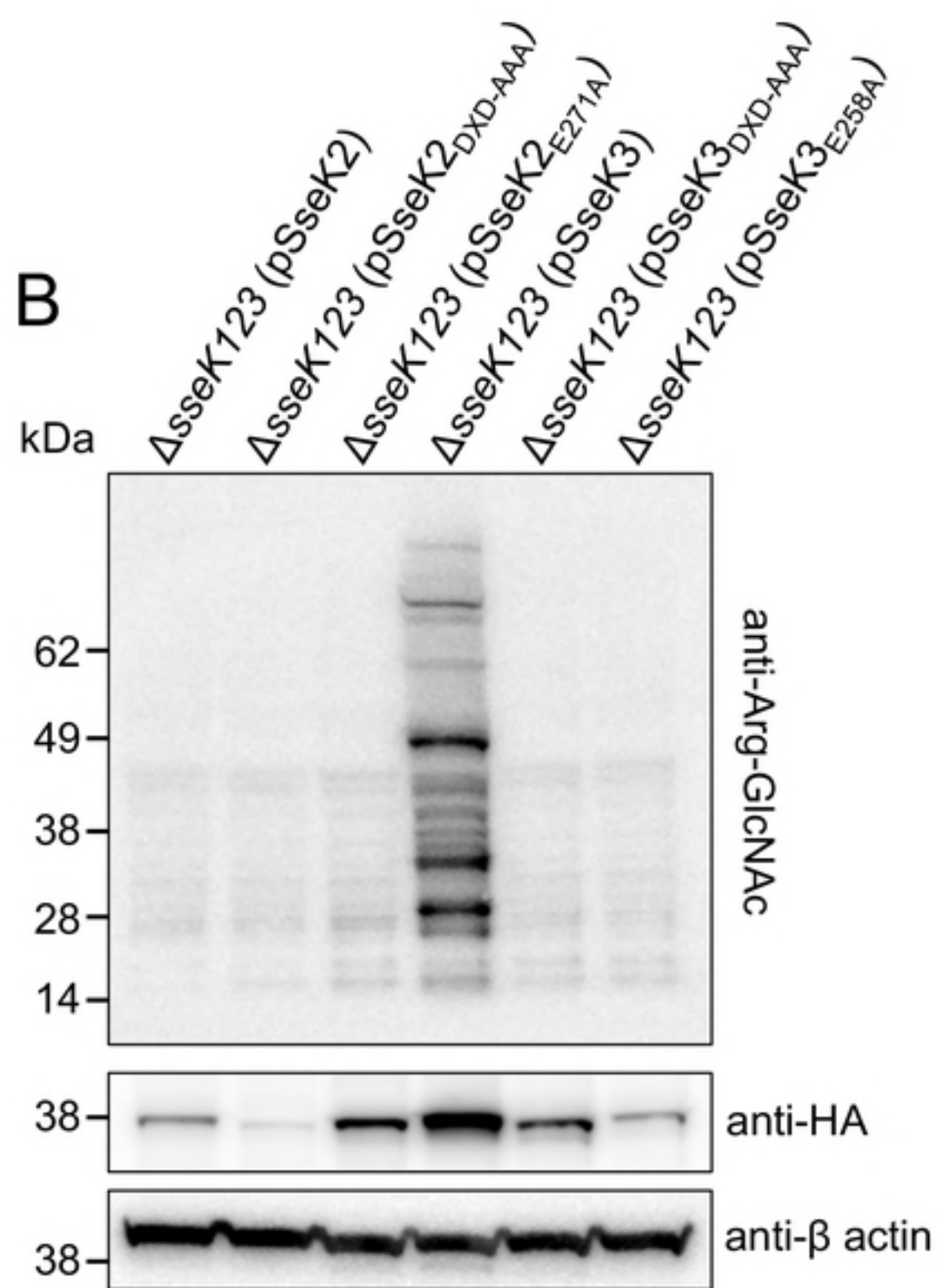
were detected with anti-Arg-GlcNAc, anti-GFP and anti-Flag as indicated. Antibodies to β -actin were used as a loading control. Representative immunoblot of at least three independent experiments. (B) LC-MS analysis of tryptic digest derived from co-incubation of recombinant His-hTRAILR2_{DD} and GST-SseK3 in the presence of UDP-GlcNAc. (C) LC-MS analysis of tryptic digest fractions derived from co-incubation of recombinant His-hTRAILR2_{DD} and GST-SseK3 with no sugar donor. (D) HCD fragmentation of recombinant His-hTRAILR2_{DD} incubated with GST-SseK3 and UDP-GlcNAc, Andromeda score 167.28. (E) Immunoblot of recombinant His-hTRAILR2_{DD} 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_{DD}. 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_{DD} and pEGFP-SseK3 or pEGFP-SseK3_{E258A}. Proteins were detected with anti-Arg-GlcNAc, anti-GFP and anti-Flag 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_{DD} 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 pGBKT7-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 pGBKT7-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 pGBKT7-SseK3 or co-transformation with pGBKT7 and pGADT7-hTNFR1 DD or pGADT7-hTRAILR2 DD.

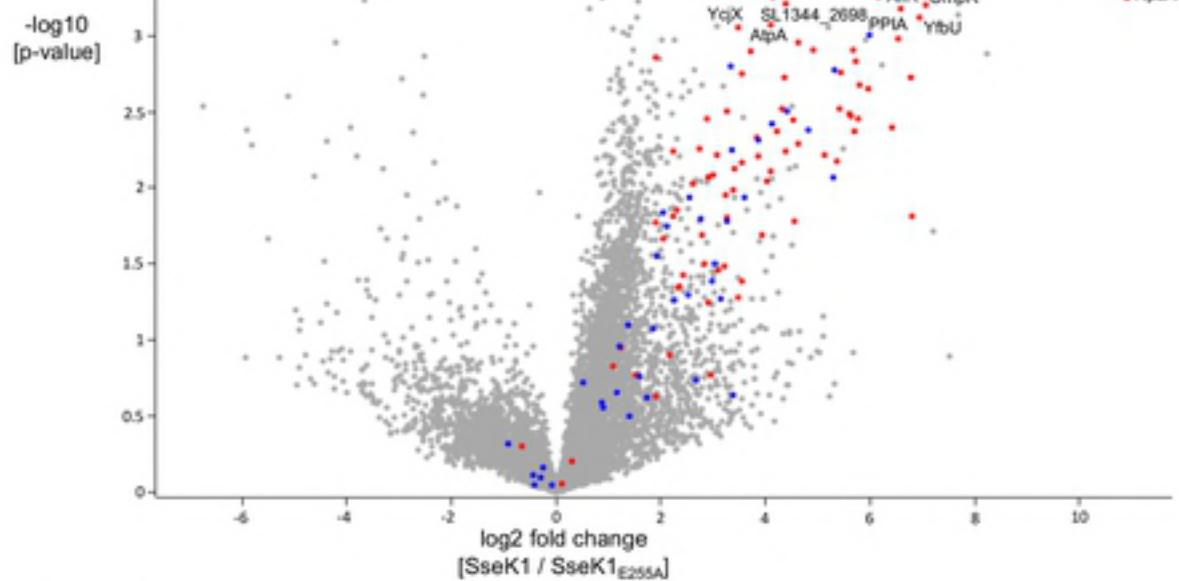
Figure 7. Golgi localisation of SseK3. Representative immunofluorescence fields of RAW264.7 cells infected with derivatives of *S. Typhimurium* SL1344 as indicated. The Δ *sseK123* triple mutant was complemented with a plasmid encoding HA-tagged SseK3 or catalytically-inactive HA-tagged SseK3_{E258A}. Overexpression of effectors was induced during host cell infection by the addition of 1 mM IPTG. Cells were fixed and permeabilized 20 h after infection. The Golgi apparatus was visualised using anti-Golgin-97 antibodies (red). HA-tagged effector proteins were visualised using anti-HA antibodies (green). Cell nuclei were visualised with Hoechst stain (blue).

Figure 8. SseK3 structure. (A) Cartoon representation of the SseK3-UDP complex. The structure of SseK3 is shown with the bound UDP drawn as sticks. SseK3 displays a GT-A fold and contains an α -helical insertion (marked by a dashed line). (B) Topology diagram of SseK3 with helices shown as red cylinders and strands as pink arrows. (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 Mg²⁺ ion and the hydrogen bonds to the diphosphate in the E258Q active site. The Mg²⁺ 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

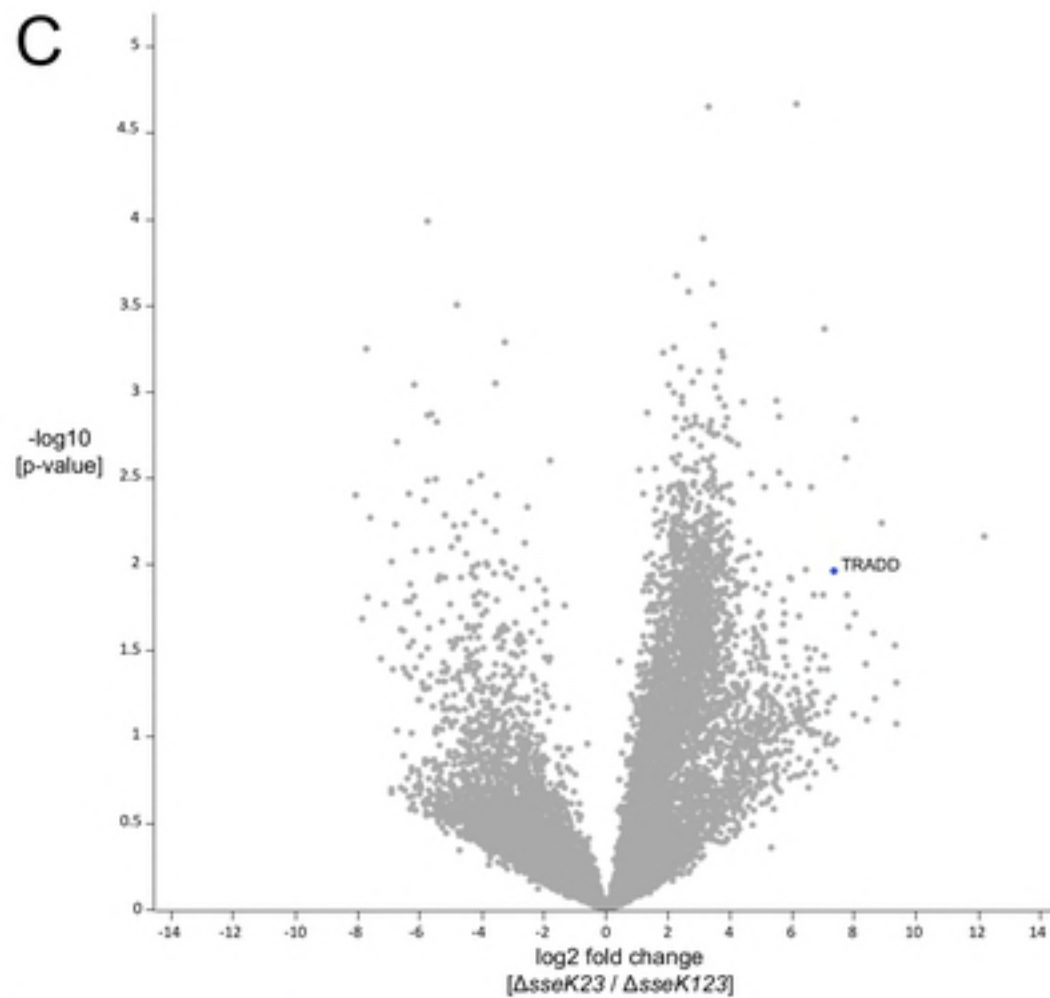
1181 UDP and GlcNAc from SseK3(E258Q) (white carbons) with UDP-Glc from toxin A
 1182 glucosyltransferase domain (pink carbons).
 1183
 1184
 1185

A**B**

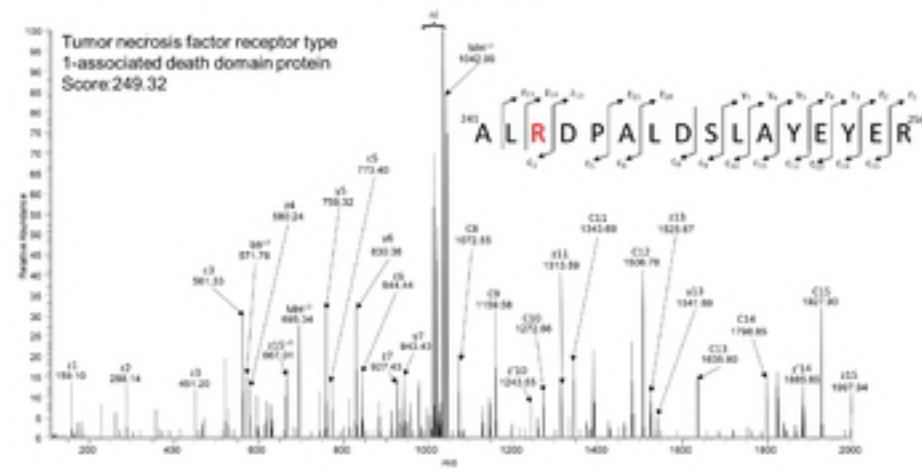
A



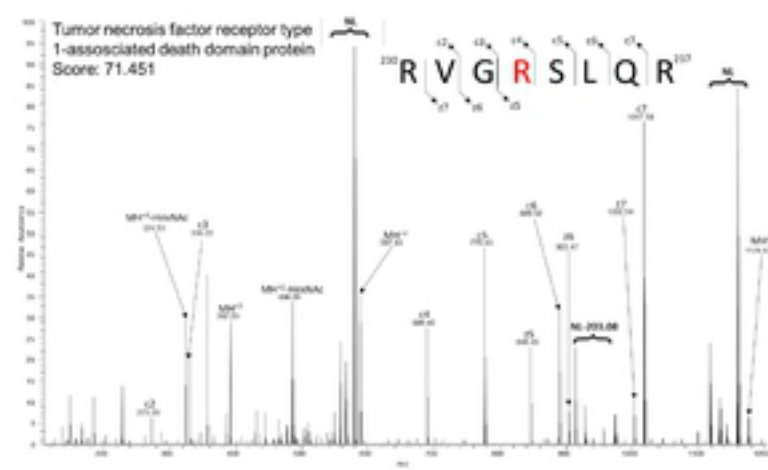
C

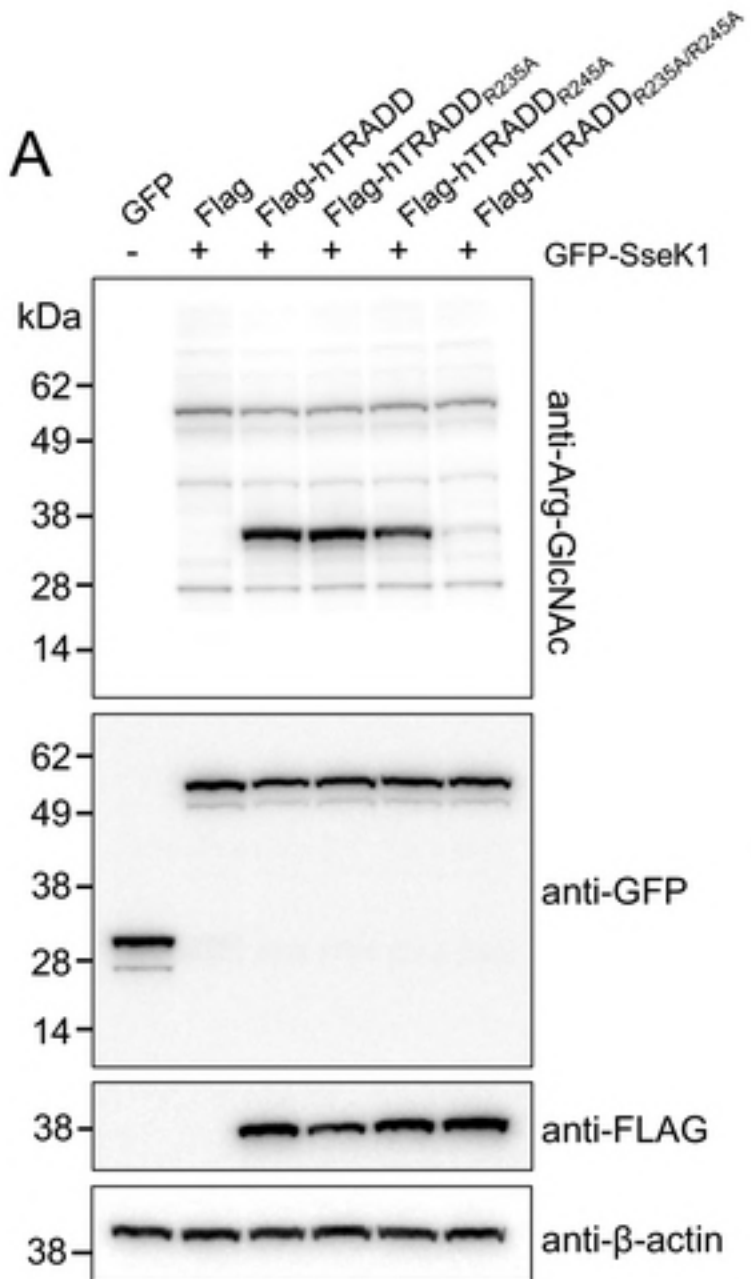
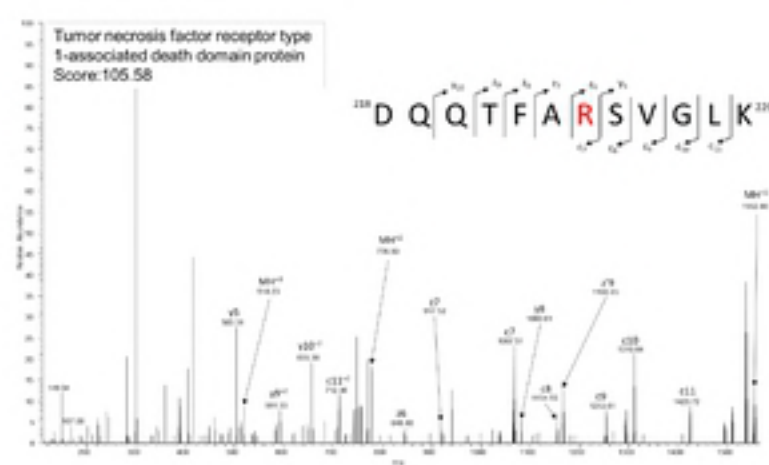
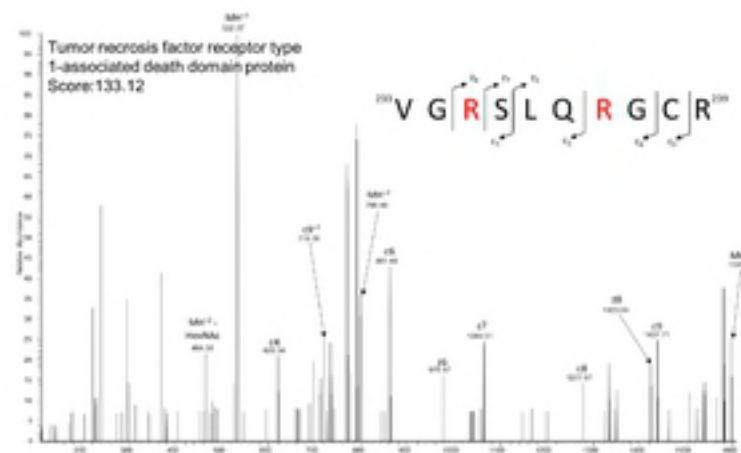
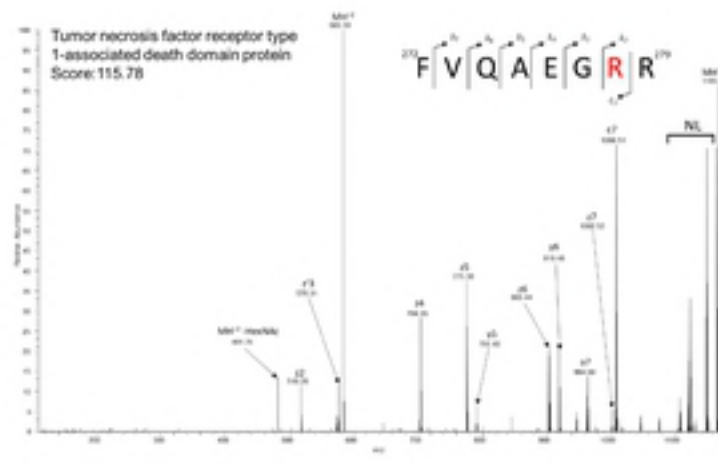
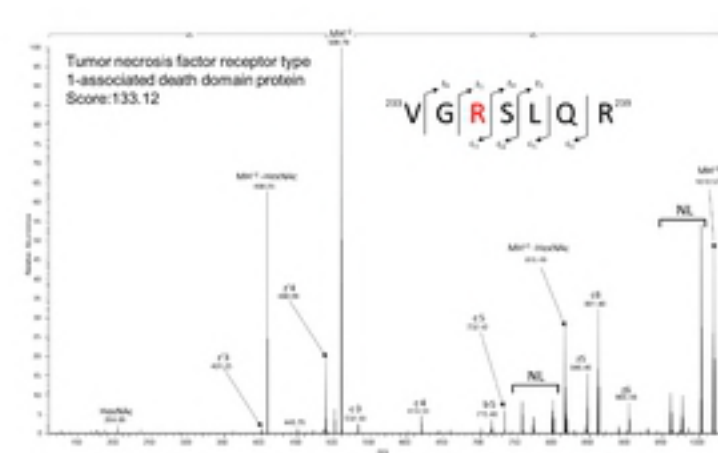


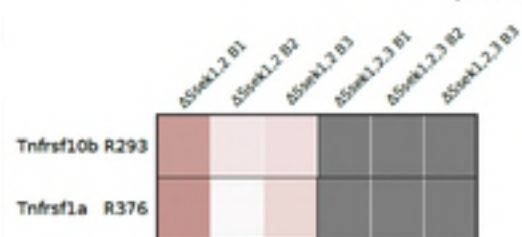
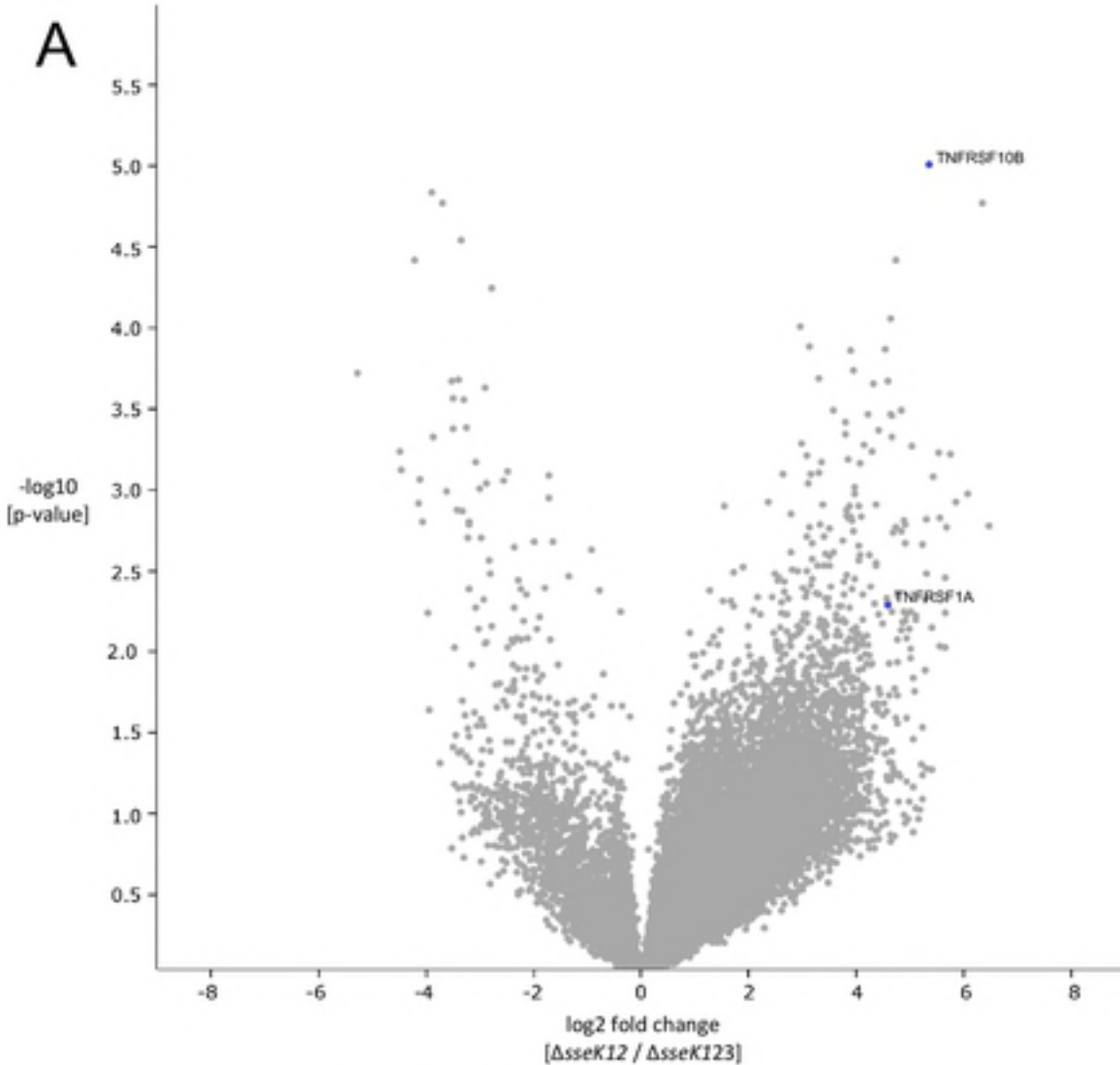
B



D



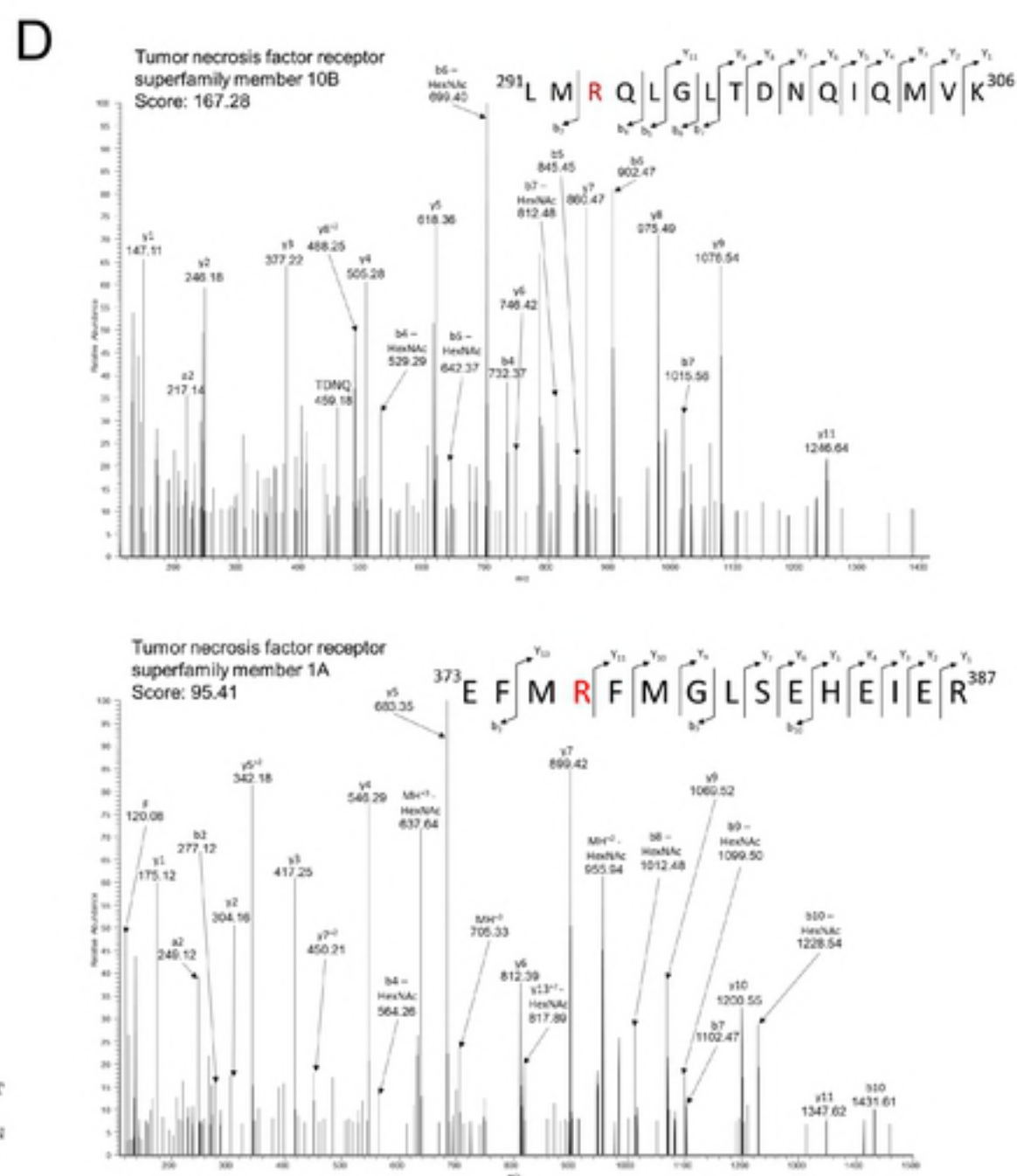
A**B**



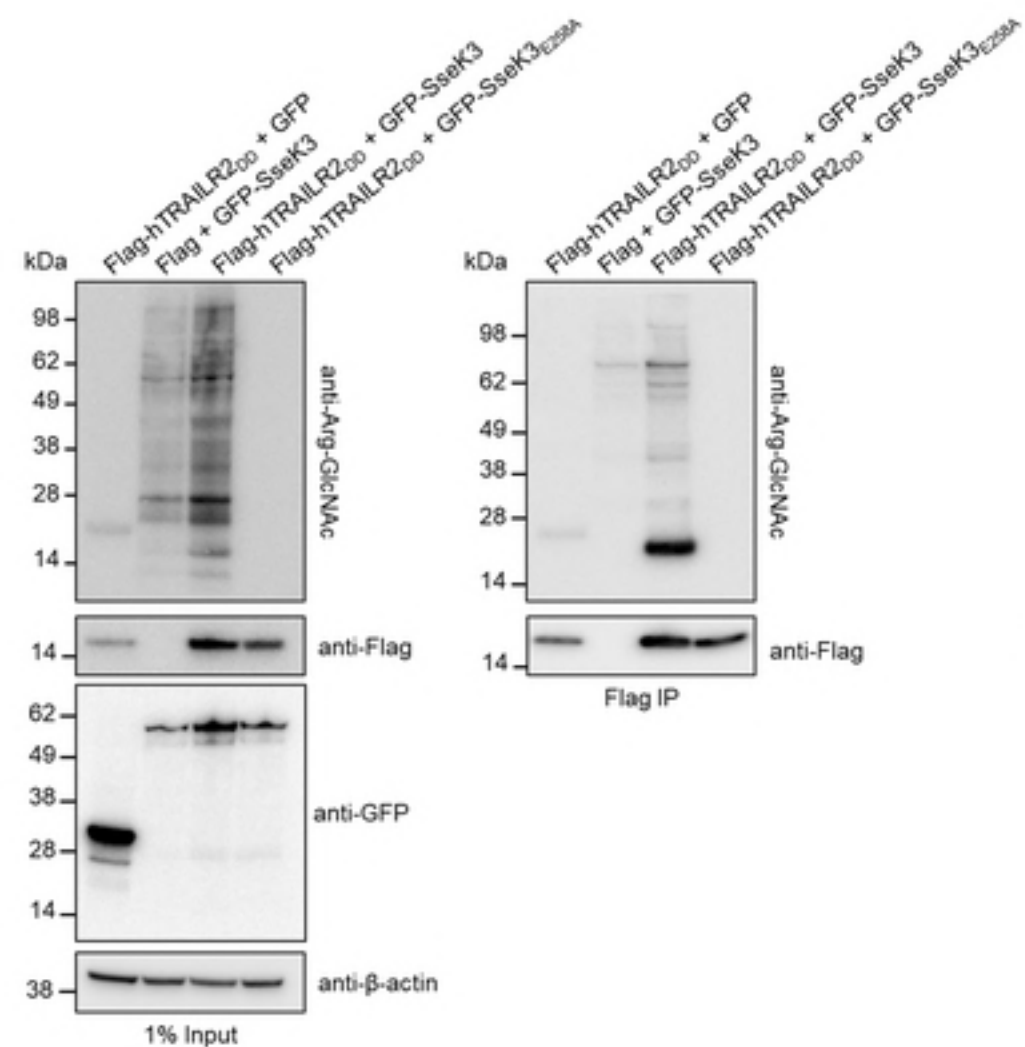
C

Tnfrsf10b DIVPFD~~SW~~RLMRQLGLTDNQIQMVKAET
D VP W MR +GL++ +I+ ++ +

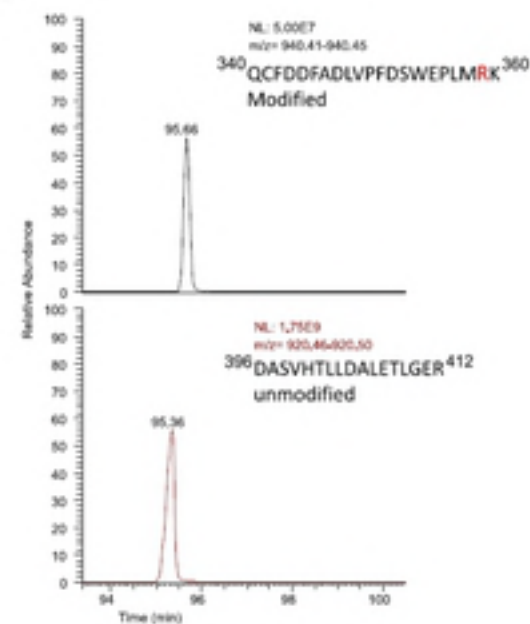
Tnfrsf1a DGVPPARWKEFMRFMGLSEHEIERLEMQN



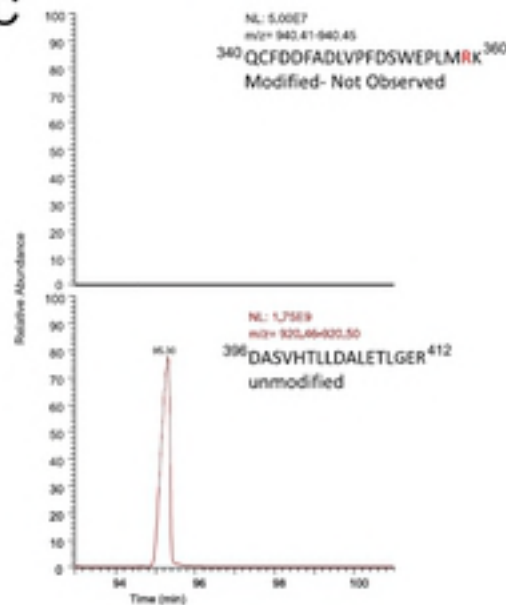
A



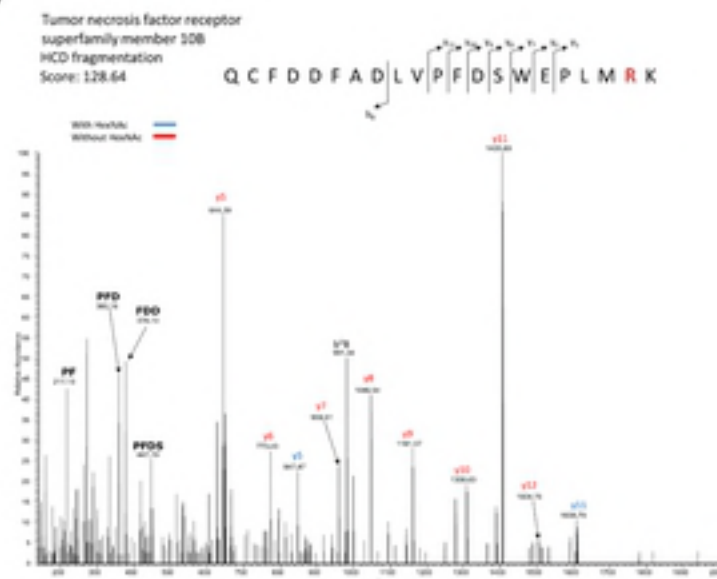
B



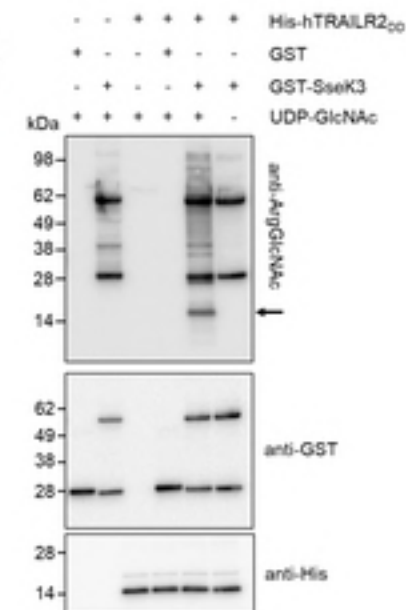
C



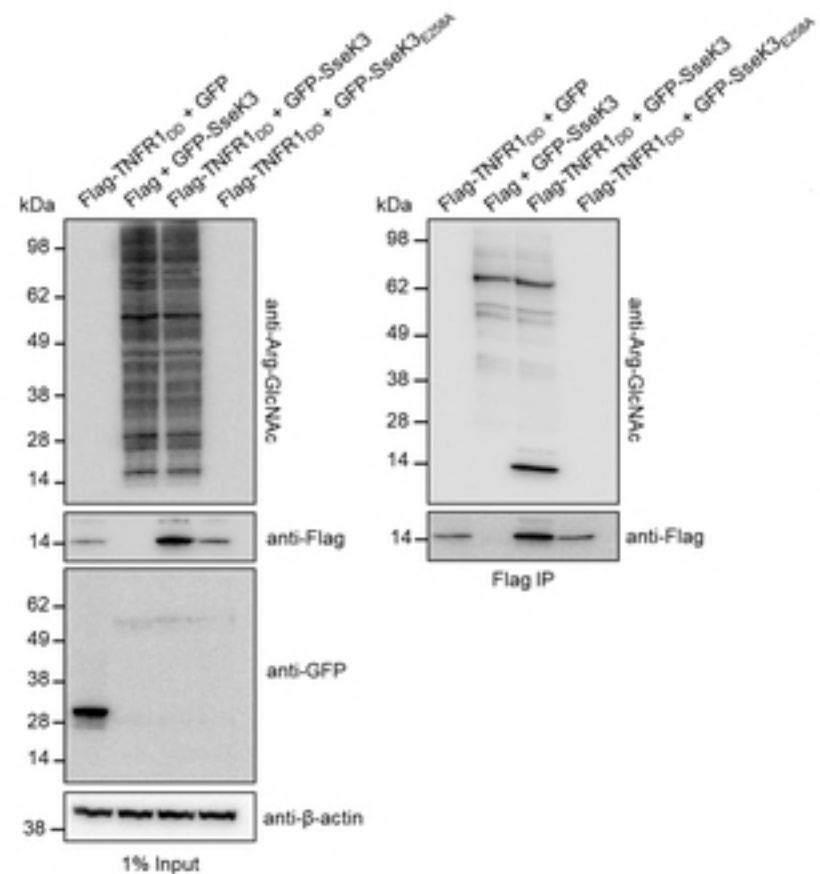
D



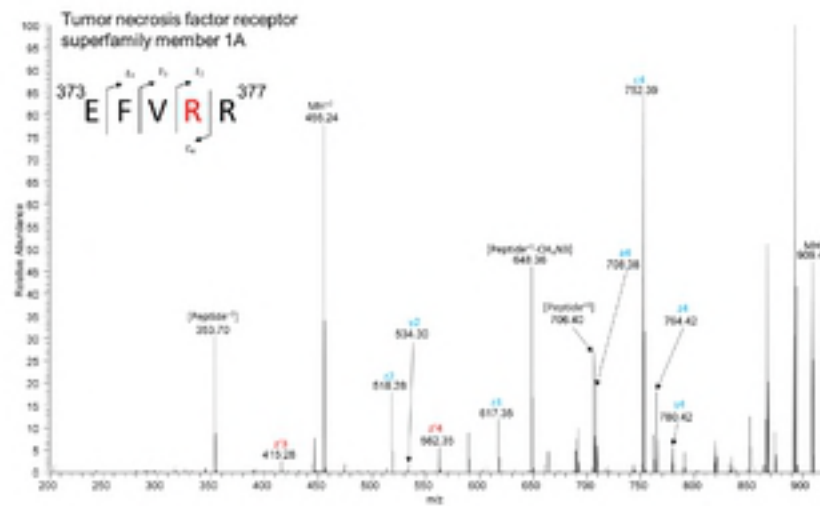
E



A



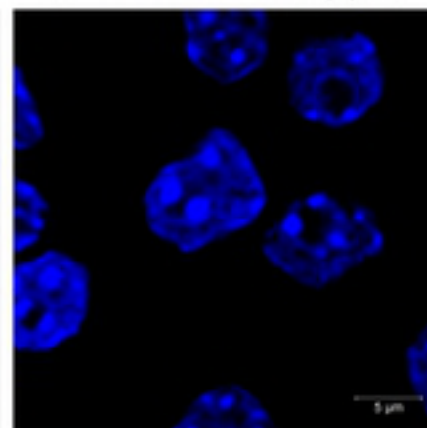
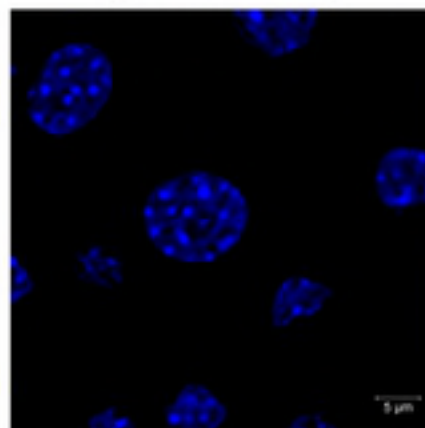
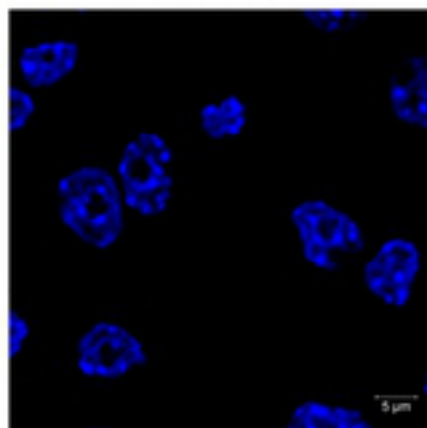
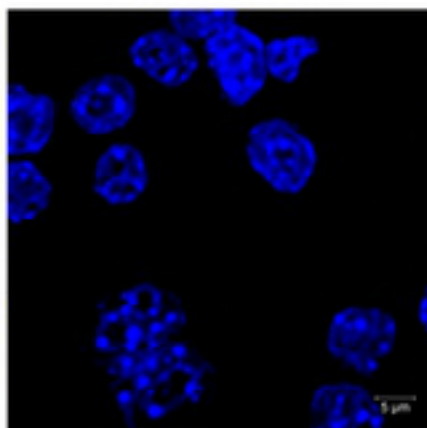
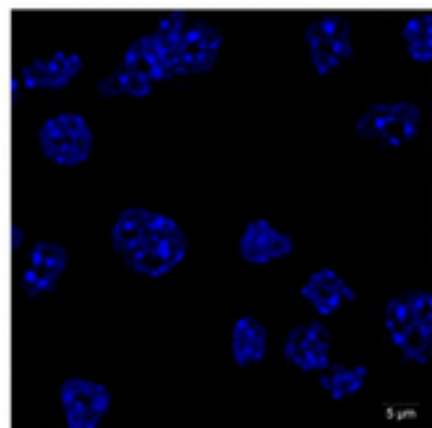
B



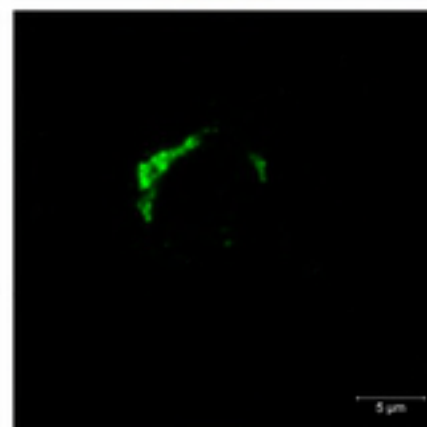
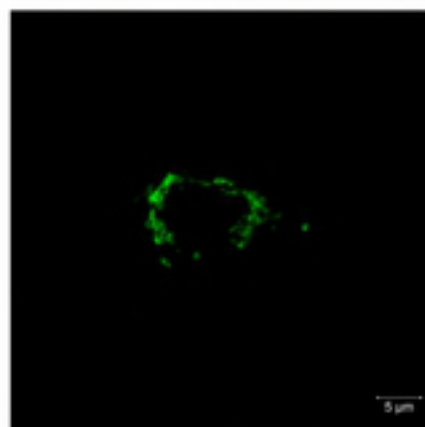
Hoechst

Uninfected

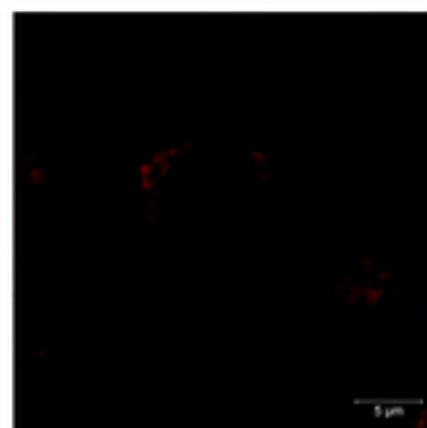
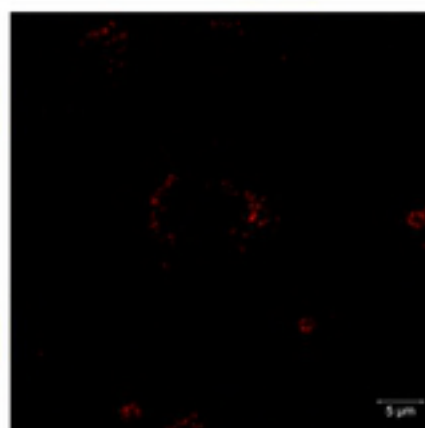
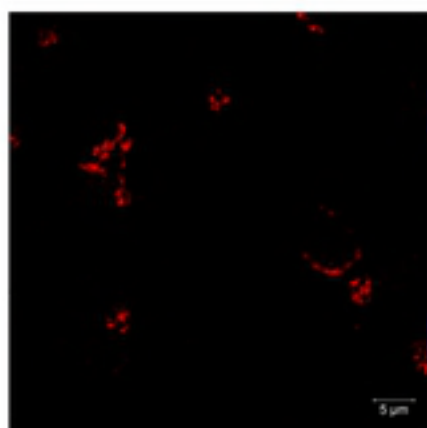
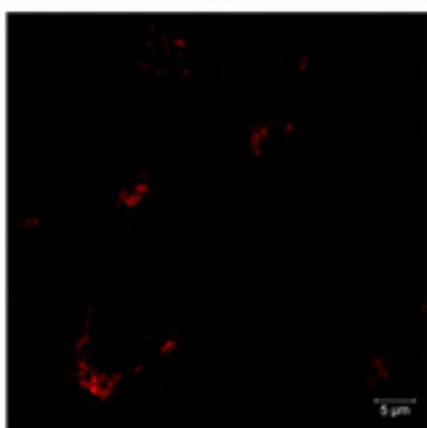
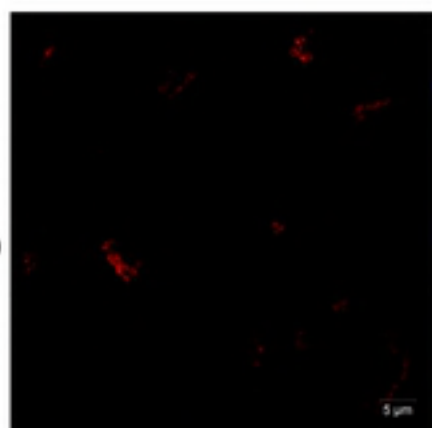
SL1344

 Δ sseK123 Δ sseK123
(pSseK3) Δ sseK123
(pSseK3_{E258A})

HA



Golgin-97



Merge

