Three small partner proteins facilitate the type VII-dependent secretion export of an antibacterial nuclease

Yaping Yang¶[1], Eleanor Boardman¶[1], Justin Deme[2], Felicity Alcock[1], Susan Lea[2] and Tracy Palmer*[1]

¶ These authors contributed equally

[1] Newcastle University Biosciences Institute, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK.

[2] Center for Structural Biology, Center for Cancer Research, National Cancer Institute, NIH, Frederick, MD 21702.

* To whom correspondence may be addressed. Email: tracy.palmer@newcastle.ac.uk

Keywords: Type VII secretion system; *Staphylococcus aureus*; nuclease toxin; effector protein; pre-secretion complex

Summary

The type VIIb protein secretion system (T7SSb) plays a role in interbacterial competition in Gram-positive Firmicute bacteria and secretes various toxic effector proteins. The mechanism of secretion and the roles of numerous conserved genes within T7SSb gene clusters remain unknown. EsaD is a nuclease toxin secreted by the *Staphylococcus aureus* T7SSb, which forms a complex with its cognate immunity protein, EsaG, and chaperone EsaE. Encoded upstream of EsaD are three small secreted proteins, EsxB, EsxC and EsxD. Here we show that EsxBCD bind to the transport domain of EsaD and function as EsaD export factors. We report the first structural information for a complete T7SSb substrate pre-secretion complex. Cryo-EM of the EsaDEG trimer and the EsaDEG-EsxBCD hexamer shows that incorporation of EsxBCD confers a conformation comprising a flexible globular cargo domain attached to a long narrow shaft that is likely to be crucial for efficient toxin export.

Introduction

In mixed bacterial communities interbacterial antagonism plays a central role in determining the composition of the microbiota.[1] Bacteria have evolved a diverse array of toxins, secretion systems, and defense mechanisms, for survival in what can be highly competitive environments.[2] The T7SS is a protein secretion system that is frequently found in Gram-positive bacteria and is classified into two distantly related subtypes.[3] The T7SSa is found across actinobacterial species and was first described in pathogenic mycobacteria where it is essential for virulence.[4-6] The T7SSb system is found in firmicutes, in both pathogenic (e.g. *Staphylococcus aureus*) and non-pathogenic (e.g. *Bacillus subtilis*) species.[7,8] The primary role of the T7SSb is inhibition of competitor bacteria via secretion of antibacterial toxins.[9-15] These toxin substrates usually have a transport-associated N-terminal LXG domain (IPR006829), and a highly polymorphic C-terminal toxin domain. T7b-associated toxin domains identified to date include nucleases, membrane depolarizing toxins, lipid II phosphatases, NADases and ADP ribosyltransferases.[9-11,13]

The *S. aureus* T7SSb is encoded at the *ess* locus.[8] The secretion system components are encoded at the 5' end of the gene cluster and are conserved across T7^{+} *S. aureus* strains. The central component of the machinery is EssC, a membrane-bound AAA+ ATPase of the SpoIIIE/FtsK family. Based on similarity to the EccC component of the T7SSa, an EssC hexamer likely forms the T7b transport channel.[16,17] The C-terminal region of EssC shows significant diversity across *S. aureus* strains and can classified into one of four types, EssC1 to EssC4.[18] The first gene in the *ess* locus is *esxA*, encoding a WXG100 protein (IPR036689). The WXG100 family is associated with T7-secretion, and EsxA is critical to the function of the core machinery in both the T7SSa and T7SSb. While EsxA is essential for substrate export, it is itself secreted by the T7 apparatus.[8,19] The 3' region of the *ess* locus is more variable, with sequences diverging at the 3' end of the *essC* gene (Fig. 1A). Variant-specific genes found downstream of *essC* include secreted substrates, immunity proteins, and a number of genes of as yet unknown function.

EsaD is an antibacterial nuclease found in *essC1* strains of *S. aureus* and is related to the *B. subtilis* YeeF toxin.[9] It has helical N-terminal region, a central PT-TG domain (IPR027797) and a C-terminal nuclease domain (IPR044927) (Fig. 1B). EsaG, the EsaD immunity protein is encoded downstream of *esaD*. Sandwiched between *essC1* and *esaD*

are four genes: *esxC*, *esxB*, *esaE* and *esxD* (Fig. 1A). EsaE is a chaperone which binds to EsaD and targets it to the T7 apparatus[9], and EsxB is small α-helical hairpin of the WXG100 family. EsxC and EsxD, while not annotated as WXG100 proteins, are similarly predicted to fold as small α-helical hairpins[20], and EsxD has a putative T7SS targeting signal (YxxxD/E) at its C-terminus (Fig. 1B).[21,22] Prior work has shown that EsxB, EsxC and EsxD are all secreted by the *S. aureus* T7SS and both EsxB and EsxC have been implicated as effectors involved in the pathogenesis of *S. aureus* murine abscesses.[8,20,23]

Here we have further investigated the functions of the small WXG100-like Esx proteins EsxB, EsxC and EsxD. We show that they form a complex with EsaDEG, and that they stabilize and dramatically boost EsaD secretion. To gain insight into the secretion mechanism we undertook structural analysis of EsaDEG and EsaDEG-EsxBCD complexes using cryo-electron microscopy. Although the flexible nature of the complexes precluded high-resolution structure determination, the low-resolution maps show that incorporation of EsxBCD gives the complex a distinctive long, narrow conformation, similar to the PE/PPE substrates of the mycobacterial T7SSa.[24] We propose that this is a common structural feature of T7-secreted substrates from both T7a and T7b systems.

Results

EsxB, EsxC and EsxD are export/stability factors for EsaD

Prior work has implicated the small, secreted substrates, EsxB and EsxC in *S. aureus* virulence.[8,23]. EsxD has also been identified as a T7SS substrate, and an interaction partner of EsxB.[20] Gene neighborhood analysis reveals that genes encoding homologs of these three small WXG100-like proteins always co-occur with *esaD*, and with conserved synteny (*esxC-esxB-esaE-esxD*). While the gene cluster is frequently found at the T7SS locus in firmicutes, in some *Bacillus* species the cluster is found elsewhere on the genome (Fig. 1C). Of the three small proteins, EsxC is the most divergent in sequence, although structural predictions indicate a conserved fold (Fig. S1).

As EsxB, EsxC and EsxD are also encoded by non-pathogenic bacteria such as *Listeria aquatica*, they are unlikely to have a conserved role in pathogenesis. Instead, based on the co-occurrence with EsaD and EsaE, we hypothesized that they might function in secretion of EsaD by the T7SS. We have previously shown that EsaD is secreted, albeit inefficiently, when EsaD, EsaE and EsaG are overproduced from a plasmid in *S. aureus essC1* strains.[9] To test our hypothesis, we constructed a plasmid carrying the whole genomic region from *esxC* to *esaG* under the control of the P(*xyl*/*tet*) promoter (pCBED-DG) (Fig. 2A). The *esaD* gene in this construct was modified to encode a C-terminal HA tag for immunodetection and carried the catalytic mutation H528A to reduce toxicity. We then made four additional constructs, each with a single gene deletion of *esxB*, *esxC*, *esxD* or *esaE.* These plasmids were transferred into the *S. aureus essC1* strain COL, in which the same genomic region (*esxC-esaG*) had been deleted (COLΔ*esxC-esxG*). Following induction of gene expression, the whole cell and culture supernatant fractions were isolated and analyzed for the presence of EsaD^{H528A}HA. Immunoblot of these fractions showed that in the presence of all four accessory proteins, EsaD^{H528A}HA was readily detected in both cell and supernatant fractions indicating secretion of EsaD^{H528A}HA (Fig. 2B). Deletion of either *esxB* or *esxC* from the plasmid vastly reduced the EsaD^{H528A}HA detected in the secreted fraction. Deletion of *esxD* or *esxE* produced an even more drastic effect, with EsaD^{H528A}HA not detected (*ΔesxD*) or barely detected (*ΔesaE*) in the culture supernatant. In the case of all four deletions, the EsaD^{H528A}HA signal in the cell fractions was also significantly reduced - in the case of Δ*esxD* to below the

109 detection limit in this particular experiment, though a longer exposure showed that
110 EsaD^{H528A}HA was still minimally produced and secreted in this sample (Fig. 2B). A control
111 immunoblot with antibodies against EsxA confirmed that the T7SS secretion efficiency
112 was unaffected by the presence or absence of EsxB, EsxC and EsxD (Fig. 2B, middle
113 panel). A similar approach showed that the addition of *esxB, esxC* and *esxD* to a plasmid
114 carrying *esaE*, *esaD* and *esaG* vastly improves the secretion efficiency of EsaD (Fig. S2).
115 These results clearly demonstrate that EsxB, EsxC, EsxD and EsaE all contribute to both
116 the stability and secretion of EsaD, but not EsxA, and so can be considered EsaD-specific
117 export factors.
118
119 ***Composition of the EsaD pre-secretion complex***
120 We have previously shown that prior to secretion, EsaD forms a ternary complex with its
121 immunity protein EsaG and targeting chaperone EsaE.[9] Given our observation that EsxB,
122 EsxC and EsxD facilitate EsaD secretion and/or stability, and the fact that EsxB, EsxC
123 and EsxD are themselves secreted by the T7SS[8,11,23,25], we explored whether the
124 WXG100-like proteins could interact with the EsaDEG complex. All six genes were
125 assembled on an *E. coli* expression plasmid where EsaD^{H528A} is encoded with a C-terminal
126 his$_6$ tag, and EsxD with a C-terminal twin-strep tag (Fig. 3A). Protein complexes were
127 purified sequentially via the his-tag on EsaD^{H528A} and the twin-strep tag on EsxD and were
128 analyzed by size exclusion chromatography (SEC) (Fig. 3B). A complex was eluted at a
129 size corresponding to ~400 kDa, significantly larger than the predicted molecular weight
130 of a hetero-hexamer containing one copy of each protein (~150 kDa). This complex was
131 visualized by Coomassie staining (Fig. 3C) and all six components were identified by
132 immunoblot (EsaD, EsxC, EsxD and EsxB) (Fig. 3D) or mass spectrometry (EsaE, EsaG).
133
134 Previously it was noted that other toxin families with homology to *B. subtilis* YeeF / *S.*
135 *aureus* EsaD were encoded in *Listeria monocytogenes* strains[13], and Fig. 1D shows that
136 they are also found in other firmicute species. These toxin sequences diverge prior to the
137 nuclease domain, instead harboring other C-terminal toxin domains such as nucleotide
138 deaminase or NAD$^+$ glycohydrolase domains. Despite this, each of these toxin families is
139 also encoded at a locus with *esxC-esxB-esaE-esxD* genes (Fig. 1D).[13] We therefore
140 reasoned that these accessory proteins may interact with the conserved N-terminal region
141 of EsaD, rather than the nuclease domain. To address this, we repeated the co-expression

142 of all six genes but with a truncated *esaD* lacking the C-terminal toxin sequence. As shown
143 in Fig. 4A, EsxB, EsxC, EsxD and EsaE still co-purified with $EsaD_{1-477}$-his by nickel affinity
144 chromatography, indicating that these four partners interact with the transport-associated
145 N-terminal domain of EsaD. When the N-terminal part of EsaD was removed, the EsaG
146 immunity protein remained bound to the $EsaD_{421-614}$-his nuclease as expected, but all other
147 partners were lost (Fig. 4B).
148
149 ***Structural analysis of the EsaD pre-secretion complex***
150 To gain further insight into the EsaD secretion mechanism we analyzed the structures of
151 EsaD-containing complexes. EsaD alone is unstable to purification, however in the
152 presence of EsaE and EsaG a trimeric complex could be isolated (Fig. 5A). Most of the
153 purified protein was eluted from the SEC column as a major peak at a volume
154 corresponding to ~400 kDa, much larger than the theoretical molecular weight of ~100
155 kDa for a 1:1:1 heterotrimer (Fig. 5A, peak b, black trace). An additional smaller peak was
156 eluted at an even larger molecular weight (Fig. 5A, peak a, black trace). Initial cryo-EM
157 analysis of the ~400 kDa peak fractions for yielded poor, undefined 2D class averages
158 due to sample heterogeneity (not shown). Therefore, the EsaDEG trimer and the
159 hexameric EsxBCD-EsaDEG complex were each subject to crosslinking with
160 glutaraldehyde to reduce flexibility and improve particle alignment (Fig. 5A, B).
161 Crosslinking of EsaDEG resulted in a later elution volume of EsaDEG, more consistent
162 with the predicted molecular weight, and SDS-PAGE revealed the presence of ~100 kDa
163 EsaDEG trimers (Fig. 5A, peak b) and ~200 kDa dimers-of-trimers (Fig 5A, peak a). Some
164 additional higher molecular weight species were also present in the earliest fractions.
165 Fractions corresponding to crosslinked EsaDEG trimers were analyzed by cryo-EM,
166 discussed below. Glutaraldehyde crosslinking of the EsaDEG-EsxBCD hexamer (Fig 5B)
167 produced two species which correspond in size to the heterohexamer (approximately 150
168 kDa) and dimers-of-hexamers (approximately 300 kDa), as judged by SDS-PAGE. Larger
169 species might also be present but were not resolved on this gel (Fig 5B, right hand panel
170 Crosslinked fractions corresponding to the ~150 kDa hexamer (peak c) were selected for
171 analysis by cryo-EM.
172
173 Analysis of crosslinked EsaDEG trimer by single particle cryo-EM yielded reference-free
174 2D class averages that lacked secondary structure detail but yielded a bilobed shape of

175 approximately 100 Å in diameter. Generation of 3D classes *ab initio* from these particles
176 gave volumes consistent in shape and particle number (21.4 – 29.2 % across four
177 classes), suggesting considerable conformational heterogeneity within the sample (Fig.
178 S3A). 3D reconstructions of the two most populated classes generated low resolution
179 volumes consistent with the 2D class averages (Fig. S3A). An Alphafold model of EsaDEG
180 (Fig. S3C) could be partially docked into one of these volumes (Fig. 6A). In this model
181 EsaE is sandwiched between the N- (residues 1-388) and C-terminal (residues 444-614)
182 domains of EsaD and EsaG interacts almost exclusively with an orthogonal face on the
183 C-terminal domain of EsaD. A recent study showed that a complex between the EsaD
184 nuclease domain and EsaG adopts an alternative conformation to that generated by
185 Alphafold, where the nuclease domain is split into two parts with EsaG sandwiched
186 between them.[26] However, the overall volume of this domain is similar in both
187 conformations.

188 Single particle cryo-EM was also performed on crosslinked EsaDEG-EsxBCD hexamer.
189 Reference-free 2D class averages included cane-like structures consisting of a helical
190 elongated shaft and lobed head not previously seen with the crosslinked EsaDEG trimer
191 sample (Fig. S3B). One of the four volumes generated *ab initio* from these particles was
192 consistent with the bilobed volume generated from the EsaDEG trimer sample (Fig. S3B,
193 grey volume). The other three *ab initio* classes yielded volumes consistent with cane-like
194 structures, but the size and morphology of the lobed head differed across the classes (Fig.
195 S3B). Refinement of the most populated class yielded a low-resolution volume consisting
196 of a 25 Å x 140 Å shaft and 50 Å by 90 Å lobed head (Fig. 6A). An Alphafold model of
197 $EsaD_{1-220}$EsxBCD matched the dimensions of the shaft and could be docked into this map
198 (Fig 6B); we therefore propose that the shaft of the cane is assembled from these proteins.
199 An alternate volume generated from the second most populated class showed additional
200 density within the cane head (Fig. 5E) that swings back down to contact the shaft. This
201 bilobed domain is grossly consistent in shape and size with the EsaDEG trimer volume.
202 We propose that these two prominent volumes represent separate states of the EsaDEG-
203 EsxBCD hexamer where the toxin-immunity domain is either tethered to the shaft (latter
204 volume) or are partially disordered or disassembled (former volume). These experiments
205 provide the first structural information for a complete pre-secretion complex for a T7b toxin
206 substrate.

207

208 ***EsaE is structurally homologous to the T7SSa chaperone EspG***

209 In the absence of high-resolution structural data, we examined Alphafold models of
210 individual components of the pre-secretion complex. EsxB, EsxC and EsxD all fold as α-
211 helical hairpins (Fig. S1). For EsaD on its own (Fig S1) or in complex with EsaE and EsaG
212 (Fig S3C) only the nuclease domain could be modelled with high confidence. However,
213 the N-terminal region (residues 1-220) could be confidently modelled in complex with
214 EsxB, EsxC and EsxD, generating a structure that fits well with the cryo-EM determined
215 volume of the shaft (discussed above) (Fig. S3C, S4 and 6B). In this model $EsaD_{1-220}$
216 extends along the length of the shaft. EsxC binds to the first half of the EsaD helical
217 domain at the bottom end of the shaft and EsxB and EsxD are positioned at the top,
218 adjacent to the globular head. EsxB and EsxD, which have been shown experimentally to
219 interact[20], form an antiparallel heterodimer within the complex, with the same architecture
220 as known WXG100 dimer structures.[27,28] The central WXG motif of EsxB is located at the
221 top end of the complex, and the C-terminus of EsxD, carrying the putative targeting signal
222 YxxxE, protrudes from the shaft at the same end (Fig. S4). This dimer interacts with the
223 second half of the EsaD helical domain, with the widest part comprising a bundle of five
224 α-helices: two each from EsxB and EsxD, and one from EsaD. While the structure is
225 largely α-helical, midway down the shaft are two short antiparallel β-strands contributed
226 by the N-termini of EsxB and EsaD. Each is four residues long, and they are sandwiched
227 between the central loop regions of EsxC and EsxD.

228

229 Intriguingly, a search of the PDB using the Dali server revealed the Alphafold model for
230 EsaE to be structurally homologous to EspG chaperones utilized by PPE substrates of the
231 mycobacterial T7SSa[29-31], despite lacking any sequence homology to these proteins. The
232 top three hits from the Dali search have Z-scores of 10.3 (EspG3, PDB 5SXL), 10.0
233 (EspG5, PDB 4W4L) and 9.9 (EspG1, PDB 5VBA), and Figure S1C shows a structural
234 alignment of the EsaE model with the EspG3 crystal structure, with an RMSD of 3.2 Å.
235 EspG proteins are found in mycobacterial T7a systems and, like EsaE, function as
236 targeting chaperones for T7 substrates. EspG substrates belong to the PE/PPE protein
237 families which, like LXG proteins, have a conserved N-terminal domain and a highly
238 polymorphic C-terminal cargo domain.[32] In the published crystal structure of a trimer of
239 PE25, PPE41 and EspG3 from *M. tuberculosis*, the PE/PPE dimer is a long narrow bundle

240 of four α-helices to which EspG3 binds at one end[30] and bears a striking resemblance to
241 our cryo-EM structure (Fig S4). Binding of EsaE to EsaD might therefore mimic the
242 interaction between EspG and the PE/PPE substrates of the T7SSa.
243
244 **Discussion**
245 Previously it has been shown that the EsaD nuclease toxin forms a ternary complex with
246 a chaperone, EsaE, and an immunity protein, EsaG.[9] Here we show that EsaD has three
247 additional protein partners, EsxB, EsxC and EsxD that interact with the EsaD N-terminal
248 region, facilitating its secretion. Cryo-EM analysis of the EsaDEG and EsaDEG-EsxBCD
249 complexes indicate that incorporation of the small partner proteins results in the formation
250 of a complex with a long helical shaft and a globular head. The shaft has a diameter of
251 approximately 25 - 30 Å, similar to the diameter of the EsxA homodimer that is known to
252 be secreted in a folded form.[28,33] A model for EsaD secretion is shown in Fig. 7. It should
253 be noted that current high-resolution structures of the hexameric T7SSa show a closed
254 central pore.[16,17] While the size of the open channel is not clear, modelling studies, based
255 at least in part on high resolution structures of the related hexameric DNA translocase
256 FtsK, suggest it may be in the region of 30 Å.[34] We propose that the helical shaft is
257 secreted in a folded conformation (Fig. 7), which would account for the co-dependent
258 secretion of these proteins observed previously.[20] Intriguingly, our Cryo-EM analysis
259 revealed that the EsaD cargo domain (comprising the EsaD nuclease domain, EsaG and
260 EsaE) exists in at least two different states, one of which is partially disordered or
261 disassembled (Fig. 6BC, Fig. S3B), supporting the view that this domain is not tightly
262 folded. As the dimensions of the cargo domain are significantly wider than the predicted
263 pore size, there would likely need to be some unfolding of this part of the complex to be
264 accommodated within the secretion channel. The EsaG immunity protein may be liberated
265 from the complex at this point, as it is not secreted with EsaD.[9]
266
267 WXG100-like protein partners have previously been identified for T7SSb substrates with
268 N-terminal LXG domains from *Streptococcus intermedius*.[10,35] Two small WXG100-like
269 proteins designated Lap1 (a DUF3130 family protein) and Lap2 (a DUF3958 protein) are
270 encoded upstream of each *S. intermedius* LXG toxin, and these function as export factors
271 forming a pre-secretion complex with the LXG region of their cognate toxin.[35] Lap1 carries
272 a C-terminal FxxxD targeting signal, similar to the YxxxD/E targeting signal in T7a

273 substrates.[22] While the N-terminal helical domain of EsaD is not recognized as belonging
274 to the IPR006829 LXG domain family, a PFAM LXG domain signature (PF04740) is
275 detected from EsaD residue 1-71, and an HHPred search of the PFAM database predicts
276 a PF04740 domain from residues 2-170 with 98% probability. We therefore conclude that
277 the first 170 residues of EsaD do encompass an LXG domain, but this is more divergent
278 than those categorized as IPR0044927, such as the LXG domains of the *S. intermedius*
279 TelABCD toxins. While there are no DUF3130- or DUF3958-encoding genes at the *esaD*
280 locus, we show here that EsxB, EsxC and EsxD function analogously to Lap1 and Lap2
281 in EsaD secretion. Furthermore, a C-terminal arm including a YxxxE motif is found in EsxD
282 and likely functions in EsaD targeting to the translocon. The reason for the differing
283 requirement for two or three WXG100-like proteins in a pre-secretion complex is not clear,
284 although this may correlate with structural differences between different types of LXG
285 domain.
286
287 The T7SSa system, found in actinobacteria, and T7SSb system present in firmicutes
288 share limited similarity at the sequence level. Both utilize an FtsK/SpoIIIE family ATPase
289 for substrate recognition and protein transport, albeit the T7SSb EssC carries two N-
290 terminal forkhead-associated (FHA) domains that are absent from the T7SSa homologue
291 EccC. Additionally, both systems encode at least one WXG100 family protein, although
292 the sequence similarity between these components is low. Beyond EssC/EccC and
293 EsxA/ESAT-6, the two systems utilize a number of additional core components which are
294 completely unrelated. We have shown here, from a combination of cryo-EM and Alphafold
295 modelling, that the structure of the T7SSb EsaD pre-secretion complex shares a common
296 architecture with the PE-PPE-EspG pre-secretion complexes of mycobacterial T7SSa
297 systems, despite a complete lack of sequence similarity. This includes a long α-helical
298 shaft with a diameter of ~25 Å. A long α-helical shaft has also been seen in the crystal
299 structure of the T7SSa substrate EspB.[36] This striking structural similarity suggests an
300 unexpected level of mechanistic overlap, which will contribute to our understanding of both
301 systems. Within each type of complex, targeting information is supplied by both a
302 WXG100-like partner and an EsaE/EspG partner. These targeting signals presumably
303 interact with the EssC ATPase domains which are conserved in both systems. It will be
304 interesting to see whether the structural features we describe here are common to other
305 substrates of the T7a and T7b systems.

Acknowledgements

This study was supported by the Wellcome Trust (through Investigator Awards 10183/Z/15/Z and 224151/Z/21/Z to TP), the Intramural Research Program of the NIH, NCI, Center for Cancer Research (awarded to SL), the China Scholarship Council (through award of a PhD studentship to YY) and the University of Newcastle. We thank the bioscience technology facility at the University of York for mass spectrometry and analysis.

Author contributions

Conceptualization (YY, EB, FA, TP); Methodology (YY, EB, JD, FA, SL, TP); Investigation (YY, EB, JD); Writing – original draft (YY, EB, JD, FA, SL, TP) and funding acquisition (TP and SL).

Declarations of interest

The authors declare no competing interests.

Figure legends

Figure 1. *esxB*, *esxC* and *esxD* co-occur with *esaD*.

A. Schematic of the T7 locus in the *essC-1* strain COL. The 5' region from *essH* to approximately two thirds of the way into *essC* is highly conserved across all strains. The 3' end of the locus encodes different variant-specific proteins, which for *essC-1* strains includes the *esaD-esaG* toxin-immunity pair.

B. Schematic representation of *essC-1* specific proteins. The locations of EsaD LXG, PT-TG, disordered ('Dis') and endonuclease domains are indicated. The approximate positions of the EsxB WXG motif and EsxD targeting signal (YxxxE) are indicated.

C and D. Representative loci from gene neighborhood analysis of *esaD*. Genes corresponding to *esxB*, *esxC* and *esxD* occur in a conserved four-gene cluster (*esxC-esxB-esaE-esxD*) directly upstream of either *esaD* (C), or an *esaD*-like gene with a different toxin domain (D). N.B. *esxC* genes were designated as two distinct groups by WebFlaGs (closed/open pink arrows). However, the conserved gene size, synteny and predicted fold (Fig. S1A) all indicate that both groups are *esxC*. Accessions for the EsaD-like protein from each cluster are as follows: (1) WP_000159025, (2) WP_224186791, (3)

339 WP_244964453, (4) WP_112384664, (5) WP_251523463, (6) WP_077309997, (7)
340 WP_151522567, (8) WP_194333108, (9) WP_088085129, (10) WP_185646797, (11)
341 WP_256797855, (12) WP_077916384.
342
343 **Figure 2. Contribution of EsxB, EsxC, EsaE and EsxD to EsaD secretion.**
344 A. Schematic representation of the coding region of plasmid pRab11-CBED-DG used in
345 (B). 'HA' denotes a C-terminal HA epitope tag on $EsaD^{H528A}$.
346 B. COLΔ*esxC-esaG* cells carrying empty vector (pRab11), the complete pRab11-CBED-
347 DG plasmid shown in (C) (CBED-DG) or derivatives of that plasmid with one of *esxB*,
348 *esxC*, *esxD* or *esaE* deleted (*ΔesxB*, *ΔesxC*, *ΔesxD*, *ΔesaE*, respectively) were cultured
349 in TSB growth medium. Following induction of plasmid-encoded gene expression, whole
350 cell (wc) and culture supernatant (sn) fractions were isolated, prepared for immunoblot,
351 and separated by SDS PAGE. EsaD was detected with antibodies against the C-terminal
352 HA tag (top panel). A longer exposure of the anti-HA blot for the *ΔesxD* sample is shown
353 below this panel. The T7-secreted control EsxA and cytoplasmic control TrxA were
354 detected with anti-EsxA and anti-TrxA antibodies (middle and bottom panels). N.B. the
355 supernatant fractions analyzed here correspond to 6 x more culture volume than the cell
356 fractions.
357
358 **Figure 3. Purification of an EsaD hetero-hexamer**
359 A. Schematic representation of the pQE70-based plasmid used for co-expression of *esxB,*
360 *esxC, esxD, esaD, esaE* and *esaG*. '6xHis' denotes a C-terminal His_6 affinity tag on
361 $EsaD^{H528A}$ and 'TS' denotes a C-terminal twin-strep affinity tag on EsxD.
362 B. *E. coli* M15 pREP4 cells carrying pQE70-GGEDCBED (shown in (A)) were grown in LB
363 medium. Following induction of protein expression with IPTG, complexes were
364 sequentially purified by nickel- and streptactin- affinity purification, and the resulting
365 sample was analyzed by size-exclusion chromatography (SEC).
366 C. Fractions corresponding to the major SEC peak at 10-12 ml in (A) were separated by
367 SDS PAGE and visualized by Coomassie-staining. EsaG and EsaE were identified by
368 mass spectrometry.
369 D. Fractions corresponding to the major SEC peak at 10-12 ml in (A) were analyzed by
370 antibodies against the his-tag (for EsaD-his), EsxC, the strep-tag (for EsxD-TS) and EsxB.
371

372 **Figure 4. EsxB, EsxC, EsxD and EsaE co-purify with the EsaD N-terminal domain.**
373 A and B. *E. coli* M15 pREP4 cells carrying pQE70-GGED(1-477)-his-CBED (A) or pQE70-
374 GGED(478-614)-his-CBED (B) were grown in LB medium, and protein expression was
375 induced with IPTG. Complexes were purified by nickel affinity chromatography and
376 analyzed by SEC (left hand panels). Fractions corresponding to the indicated elution
377 volumes were analyzed by SDS-PAGE and Coomassie staining (right hand side, upper
378 panels), or by immunoblot with antibodies against the his-tag (for EsaD-his), EsxC or EsxB
379 ((A), right hand side, lower panel). EsxD in (A) and EsaG in (B) were identified by mass
380 spectrometry.
381
382 **Figure 5. Purification and crosslinking of EsaD complexes for cryoEM**
383 A. Left hand panel: SEC analysis of EsaDEG complexes before (black trace) and after
384 (red trace) glutaraldehyde crosslinking. Right hand panel: SDS PAGE and Coomassie
385 staining of SEC fractions of the crosslinked complexes shows EsaDEG trimers (~130 kDa)
386 in peak b and $(EsaDEG)_2$ hexamers (~260 kDa) in peak a.
387 B. Left hand panel: SEC analysis of EsaDEG-EsxBCD complexes before (black trace) and
388 after (red trace) glutaraldehyde crosslinking. Right hand panel: SDS PAGE and
389 Coomassie staining of SEC fractions of the crosslinked complexes shows EsaDEG-
390 EsxBCD hexamers (~150 kDa) in peak c, and a mixture of hexamers and (EsaDEG-
391 $EsxBCD)_2$ dodecamers (~300 kDa) in peaks a and b.
392
393 **Figure 6. Structural analysis of EsaD complexes by cryo-EM**
394 A. Left, 2D class averages of crosslinked EsaDEG. Middle, low resolution cryo-EM
395 reconstruction of EsaDEG (grey). Right, Alphafold model of EsaDEG partially fit into the
396 cryo-EM volume. Cryo-EM map of EsaDEG shown in transparent grey, with overlayed
397 ribbon models of EsaD (purple), EsaE (pink) and EsaG (green).
398 B. Left, 2D class averages of crosslinked EsaDEG-EsxBCD. Middle, low resolution cryo-
399 EM reconstruction of EsaDEG-EsxBCD (blue). Right, alphafold model of $EsaD_{1-220}$-
400 EsxBCD fit into the cryo-EM volume. Cryo-EM map of EsaDEG-EsxBCD shown in
401 transparent blue, with overlayed ribbon models of $EsaD_{1-220}$ (purple), EsxB (red), EsxC
402 (lime green) and EsxD (beige).

403 C. Left, 2D class averages of an alternate volume of crosslinked EsaDEG-EsxBCD. Right,
404 low resolution cryo-EM reconstruction of alternate EsaDEG-EsxBCD volume shown in
405 light pink.
406
407 **Figure 7 – Model for EsaD secretion by the T7SS.**
408 A. The EsaD pre-secretion complex engages the ATPase domains of EssC mediated
409 through interaction of both the chaperone EsaE, and the EsxD C-terminus with EssC. B.
410 Interaction of the pre-secretion complex with the EssC ATPase domains promotes a
411 conformational change to a 'contracted' conformation where the ATPase domains from
412 adjacent protomers are brought into close proximity to form a cage.[16] C. The membrane
413 pore opens to allow passage of the secretion complex. We hypothesize that the cage/pore
414 is too narrow to allow passage of the folded EsaD nuclease domain with bound EsaG and
415 that the nuclease domain unfolds around the active site, as described previously[26],
416 releasing the immunity protein. EsaE many also be released at this stage. D. Secreted
417 EsaD with its EsxB, EsxC and EsxD partner proteins are released at the surface of the
418 cell, potentially facilitated by the T7SSb component EsaA[37] (not shown). Note that ATP
419 binding/hydrolysis by EssC and interaction with EsxA are also expected to be required for
420 at some stages the secretion process, but these have been omitted for clarity.

421

422 ## Materials and Methods

423 ***Bacterial strains, plasmids and growth conditions***

424 *E. coli* was cultured in LB medium (Melford) with ampicillin (110 ug/ml) where required; *S.*
425 *aureus* was cultured in tryptic soy broth (TSB) (Oxoid) with chloramphenicol (10 ug/ml)
426 where required.

427 All strains, plasmids and primers used in this study are listed in Tables S1, S2 and S3
428 respectively. Plasmids for expression in *S. aureus* for secretion assays were based on
429 pRab11[38] and were constructed as follows: *esaG* was amplified from COL genomic DNA
430 and cloned into pRab11 KpnI and EcoRI restriction sites to make pRab11-esaG. *esaD-*
431 *HA* was then amplified and inserted into upstream of *esaG* using HiFi DNA assembly
432 (NEB). The catalytic residue H528 of *esaD* was mutated to encode an alanine using a Q5
433 site-directed mutagenesis kit (NEB). *esxCBED* amplified from COL genomic DNA was

434 then inserted upstream of *esaD* by HiFi assembly, creating pRab11-CBED-DG. Individual
435 genes were then deleted from this construct using a Q5 site-directed mutagenesis kit.

436 Plasmids used for protein purification were based on pQE70 (Qiagen). For purification of
437 EsaDEG, genes for SAOUHSC_00267 (*esaE*), SAOUHSC_00268 (*esaD*) and
438 SAOUHSC_00269 (*esaG*) were ordered as gblocks (IDT) codon optimized for *E. coli*
439 expression with a strong RBS attached to each gene. PCR fragments were assembled by
440 HiFi DNA assembly to include two copies of *esaG* followed by one copy of *esaE* and one
441 copy of *esaD* with a C-terminal hexahistidine tag and the H528A substitution, generating
442 pQE70-GGED. The *esaD* gene was amplified from the EsaD gblock using the
443 combinations of codoptDnostop_fwd and codoptH528A_rev or codoptH528A_fwd and
444 codoptDnostop_rev to create fragments that were then joined together by short overlap
445 extension PCR to introduce the H528A mutation using the outside codoptDnostop_fwd
446 and codoptDnostop_rev primers. To make pQE70-GGED-his-CBED-tstrep a genomic
447 fragment encompassing SAOUHSC_00264 (*esxC*), SAOUHSC_00265 (*esxB*),
448 SAOUHSC_00266 (*esaE*) and SAOUHSC_00267 (*esxD*), including the *esxC* RBS, was
449 amplified from *S. aureus* COL genomic DNA. A twin-strep tag was included at the EsxD
450 C-terminus to allow for an additional affinity chromatography step and immunodetection.
451 This fragment was inserted downstream of *esaD* in pQE70-GGED by HiFi assembly. To
452 avoid EsaD toxicity during cloning, all cloning steps were performed in DH5α carrying
453 plasmid pREP4 (Qiagen) for expression of the Lac repressor. The completed plasmids
454 were transformed into *E. coli* M15 pREP4 for protein expression.

455 Chromosomal deletion of *esxC-esaG* was accomplished using plasmid pIMAY-Z and
456 following the published protocol.[39] pIMAY-Z_CBEDDG, carrying regions flanking *esxC-*
457 *esaG*, was constructed by HiFi assembly with primers listed in Table S3. The COLΔ*esxC-*
458 *esxG* genotype was verified by sequencing of the relevant region.
459

460 ***Cell fractionation and Western blotting***

461 To monitor EsaD secretion, pRab11-based plasmids encoding EsaD with the catalytic
462 inactivating H528A substitution and a C-terminal HA tag for immunodetection, alongside
463 different combinations of EsxBCD, EsaE and EsaG were transformed into *S. aureus* strain
464 COLΔ*esxC-esxG* by electroporation. Strains were subcultured at 1/25 from an overnight-
465 grown preculture into fresh TSB medium. At OD600nm = 2, cells were collected and the

466 supernatant samples precipitated with trichloroacetic acid in the presence of
467 deoxycholate, as described previously.[24] Samples were normalized to OD600 = 2 in PBS,
468 and lysed by the addition of 50 µg ml–1 lysostaphin (Ambi) with incubation at 37 °C for 30
469 min. All samples were mixed with an equal volume of NuPAGE LDS sample buffer and
470 boiled for 10 min. Antibodies against EsxA, TrxA, EsxB, EsxC used for immunoblotting
471 have been previously described.[25,40,41] Antibodies against the HA tag, strep tag and myc
472 tag were obtained from Sigma, IBA and Abcam, respectively.
473
474 ***Protein purification***
475 Protein complexes were produced from the relevant pQE70 plasmid (see Table S2) in
476 strain M15 pREP4 (Qiagen) as follows: Cells were cultured in LB broth supplemented with
477 ampicillin (125 ug/ml) and kanamycin (50 ug/ml) at 37°C with aeration until an OD_{600nm} of
478 0.6-1.0 was reached, before induction with 0.5 mM IPTG and growth overnight at 18°C.
479 Cells were harvested by centrifugation and lysed by sonication in binding buffer (50 mM
480 HEPES pH 7.5, 150 mM NaCl and 50 mM imidazole) with cOmplete protease inhibitor
481 tablets (Roche). Cell lysates were clarified by centrifugation and applied to a 5ml Histrap
482 FF column (Cytiva) equilibriated in the same buffer. The column was washed with binding
483 buffer and bound proteins were eluted using a 50-500 mM imidazole gradient. Eluted
484 protein was concentrated using a 10 kDa MWCO spin concentrator (Vivaspin) before
485 injection onto a Superdex S200 10/300 gl column equilibrated in 50 mM HEPES pH 7.5,
486 150 mM NaCl. The hexameric complex containing full length EsaD carried an additional
487 twin-strep tag on EsxD, and this complex was further purified by strep-tactin affinity
488 purification, followed by a second round of SEC.

489 For glutaraldehyde crosslinking the SEC fractions corresponding to major peaks were
490 diluted to 0.5 mg/ml in 50 mM HEPES pH 7.5, 150 mM NaCl and crosslinked using 0.1%
491 glutaraldehyde for 30 minutes at room temperature, before quenching with at least a
492 tenfold excess of tris. The crosslinked protein was concentrated again using a 10 kDa
493 MWCO spin concentrator before injection on Superdex S200 10/300 gl. Eluted protein
494 was again concentrated by spin concentrator (10 kDa MWCO) and snap frozen with liquid
495 nitrogen for subsequent cryo-EM analysis.
496
497 ***Cryo-EM sample preparation and data acquisition***

498 Crosslinked EsaDEG and EsaDEG-EsxBCD samples were diluted to 0.05 mg/ml and 0.2
499 mg/ml in SEC buffer, respectively. Three microliters of sample were applied onto glow-
500 discharged (60 s, 30 mA) 300-mesh R1.2/1.3 Quantifoil Au grid. Grids were blotted for 2
501 s in 100 % humidity at 8-12 ℃ and plunge frozen in liquid ethane using a Vitrobot Mark IV
502 (Thermo Fisher Scientific). Data were collected in counted mode in EER format on a
503 CFEG-equipped Titan Krios G4 (Thermo Fisher Scientific) operating at 300 kV with a
504 Selectris imaging filter (Thermo Fisher Scientific) with slit width of 10 e^-V and Falcon 4
505 direct detection camera (Thermo Fisher Scientific) at ×165,000 magnification, with a
506 physical pixel size of 0.723 Å (EsaDEG) or 0.693 Å (EsaDEG-EsxBCD). Movies were
507 collected with a total dose of 57.0 e^-/Å^2.
508
509 ***Cryo-EM data processing***
510 Patched motion correction, CTF parameter estimation, particle picking, extraction, and
511 initial 2D classification were performed in SIMPLE 3.0.[42] All downstream processing was
512 carried out in cryoSPARC v3.3.1[43], including Gold-standard Fourier shell correlations
513 (FSCs) using the 0.143 criterion. For crosslinked EsaDEG trimer (Fig. S3A), 3,493,119
514 particles were extracted from 5,824 motion-corrected micrographs and subjected to
515 reference-free 2D classification (k=500). An additional round of reference-free 2D
516 classification was performed on 678,695 pruned particles in cryoSPARC (k=200). 2D-
517 cleaned particles (278,959) were then used to generate four volumes *ab initio*. Particles
518 belonging to the two most populated classes were individually selected for non-uniform
519 refinement against their 30 Å lowpass-filtered volume counterparts, yielding volumes of
520 moderate resolution from which only gross conformational information could be assessed.
521 For crosslinked EsaDEG-EsxBCD hexamer (Fig. S3B, 2,615,890 particles were isolated
522 after an initial round of reference-free 2D classification (k=300). These particles were
523 further cleaned with another two rounds of reference-free 2D classification within
524 cryoSPARC (k=200 each round). Particles (343,131) belonging to classes with strong
525 structural features were selected and used to generate four volumes *ab initio*, of which
526 one class was similar to the volume generated from the crosslinked EsaDEG trimer. The
527 other three classes demonstrated cane-like structures that consisted of a shaft and head.
528 The head varied in size across the three classes, and the two most well-defined classes
529 were further selected for non-uniform refinement against their own 15 Å lowpass-filtered
530 volumes as references. Refinements produced moderate resolution volumes of cane-like

structures consisting of a shaft and either a large or bilobed head. Further extensive classification in 2D and 3D space for both the EsaDEG trimer and EsaDEG-EsxBCD hexamer datasets did not significantly improve map quality, indicating the samples suffered from conformational and/or compositional heterogeneity.

Modelling and map fitting

Alphafold was used to generate models of individual proteins.[44] Alphafold models of EsaDEG and $EsaD_{1-220}$EsxBCD (Fig. S3C) were generated using the default settings of Alphafold-Multimer[45] using UniProt entries Q2G179 (EsaD), Q2G181 (EsaE), Q2G179 (EsaG), Q2G182 (EsxB), Q2G183 (EsxC), Q2G180 (EsxD). The top ranked Alphafold models were fit into the cryo-EM maps using the fitmap function of USCF ChimeraX (43). The superose program within ccp4 was used to align EsaE with EspG3. Map and model figures were prepared within ChimeraX.[46]

Flanking gene analysis (FlaGs)

Gene neighborhood analysis was carried out using the WebFlaGs server.[47] Input sequences were generated by running a BLASTP search on bacilli, excluding *Staphylococcus*, with *S. aureus* EsaD (Q2G179) as a query. Sequences lacking the LXG domain (residues 2-120) were manually removed, and the resulting sequence list was input into WebFlaGs with ten flanking genes selected and other parameters set to default.

References

1. Peterson, S.B., Bertolli, S.K., and Mougous, J.D. (2020). The central role of interbacterial antagonism in bacterial life. Curr. Biol. *30*, R1203-R1214.

2. Ruhe, Z.C., Low, D.A., and Hayes, C.S. (2020). Polymorphic toxins and their immunity proteins: diversity, evolution, and mechanisms of delivery. Annu. Rev. Microbiol. *74*, 497–520.

3. Abdallah, A.M., Gey van Pittius, N.C., DiGiuseppe Champion, P.A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C.M.J.E., Appelmelk, B.J., and Bitter, W. (2007). Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5**, 883-891.

4. Hsu, T., Hingley-Wilson, S.M. Chen, B., Chen, M., Dai, A.Z., Morin, P.M., Marks, C.B., Padiyar, J., Goulding, C. Gingery, M. et al. (2003). The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function

required for invasion of lung interstitial tissue. Proc. Natl. Acad. Sci. USA *100*, 12420-12425.

5. Pym, A.S., Brodin, P., Majlessi, L., Brosch, R., Demangel, C., Williams, A., Griffiths, K.E., Marchal, G., Leclerc, C., and Cole, S.T. (2003). Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. Nat. Med. *9*, 533-539.

6. Stanley, S.A., Raghavan, S., Hwang, W.W., and Cox, J.S. (2003). Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. Proc. Natl. Acad. Sci. USA *100*, 13001-13006.

7. Pallen, M.J. (2002). The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? Trends Microbiol. *10*, 209-212.

8. Burts, M.L., Williams, W.A., DeBord, K., and Missiakas, D.M. (2005). EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. Proc. Natl. Acad. Sci. USA *102*, 1169-1174 .

9. Cao, Z., Casabona, M.G., Kneuper, H., Chalmers, J.D., and Palmer, T. (2016). The type VII secretion system of *Staphylococcus aureus* secretes a nuclease toxin that targets competitor bacteria. Nat. Microbiol. *2*, 16183.

10. Whitney, J.C., Peterson, S.B., Kim, J., Pazos, M., Verster, A.J., Radey, M.C., Kulasekara, H.D., Ching, M.Q., Bullen, N.P., Bryant, D. et al. (2017). A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. Elife *6*, e26938.

11. Ulhuq, F.R., Gomes, M.C., Duggan, G.M., Guo, M., Mendonca, C., Buchanan, G., Chalmers, J.D., Cao, Z., Kneuper, H., Murdoch, S., et al. (2020). A membrane-depolarizing toxin substrate of the *Staphylococcus aureus* type VII secretion system mediates intraspecies competition. Proc. Natl. Acad. Sci. USA *117*, 20836-20847.

12. Chatterjee, A., Willett, J.L.E., Dunny, G.M., and Duerkop, B.A. (2021). Phage infection and sub-lethal antibiotic exposure mediate *Enterococcus faecalis* type VII secretion system dependent inhibition of bystander bacteria. PLoS Genet. *17*, e1009204.

13. Bowran, K., and Palmer, T. (2021). Extreme genetic diversity in the type VII secretion system of *Listeria monocytogenes* suggests a role in bacterial antagonism. Microbiology *167*, mic.0.001034. doi: 10.1099/mic.0.001034.

14. Kobayashi, K. (2021). Diverse LXG toxin and antitoxin systems specifically mediate intraspecies competition in *Bacillus subtilis* biofilms. PLoS Genet. *17*, e1009682.

15. Tassinari, M., Doan, T., Bellinzoni, M., Chabalier, M., Ben-Assaya, M., Martinez, M., Gaday, Q., Alzari, P.M., Cascales, E., Fronzes, R., and Gubellini F. (2022).

The antibacterial type VII secretion system of *Bacillus subtilis*: structure and interactions of the pseudokinase YukC/EssB. mBio *13*, e0013422.

16. Bunduc, C.M., Fahrenkamp, D., Wald, J., Ummels, R., Bitter, W., Houben, E.N.G., and Marlovits, T.C. (2021). Structure and dynamics of a mycobacterial type VII secretion system. Nature *593*, 445-448.

17. Beckham, K.S.H., Ritter, C., Chojnowski, G., Ziemianowicz, D.S., Mullapudi, E., Rettel, M., Savitski , M.M., Mortensen, S.A., Kosinski, J., and Wilmanns, M. (2021). Structure of the mycobacterial ESX-5 type VII secretion system pore complex. Sci. Adv. *7*, eabg9923.

18. Warne, B., Harkins, C.P., Harris, S.R., Vatsiou, A., Stanley-Wall, N., Parkhill, J., Peacock, S.J., Palmer, T., and Holden, M.T.G. (2016). The Ess/Type VII secretion system of *Staphylococcus aureus* shows unexpected genetic diversity. BMC Genomics *17*, 222.

19. Huppert, L.A., Ramsdell, T.L., Chase, M.R., Sarracino, D.A., Fortune, S.M., and Burton, B.M. (2014). The ESX system in *Bacillus subtilis* mediates protein secretion. PLoS One *9*, e96267.

20. Anderson, M., Aly, K.A., Chen Y.H., and Missiakas, D. (2013). Secretion of atypical protein substrates by the ESAT-6 secretion system of *Staphylococcus aureus*. Mol. Microbiol. *90*, 734-743.

21. Digiuseppe Champion, P.A., Stanley, S.A., Champion, M.M., Brown, E.J., and Cox, J.S. (2006). C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. Science *313*, 1632-1636.

22. Daleke, M.H., Ummels, R., Bawono, P., Heringa, J., Vandenbroucke-Grauls, C.M.J.E., Luirink, J., and Bitter, W. (2012). General secretion signal for the mycobacterial type VII secretion pathway. Proc. Natl. Acad. Sci. USA *109*, 11342-11347.

23. Burts, M.L., DeDent, A.C., and Missiakas, D.M. (2008). EsaC substrate for the ESAT-6 secretion pathway and its role in persistent infections of *Staphylococcus aureus*. Mol. Microbiol. *69*, 736-746.

24. Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA *103*, 8060-8065.

25. Kneuper, H., Cao, Z., Twomey, K.B., Zoltner, M., Jäger, F., Cargill, J.S., Chalmers, J., van der Kooi-Pol, M.M., van Dijl, J.M., Ryan, R.P. et al. (2014). Heterogeneity in *ess* transcriptional organization and variable contribution of the Ess/Type VII protein secretion system to virulence across closely related *Staphylocccus aureus* strains. Mol. Microbiol. *93*, 928-943.

26. Wang, Y., Zhou, Y., Shi, C., Liu, J., Lv, G., Huang, H., Li, S., Duan, L., Zheng, X., Liu, Y., et al. (2022). A toxin-deformation dependent inhibition mechanism in the T7SS toxin-antitoxin system of Gram-positive bacteria. Nat. Commun. *13*, 6434.

27. Renshaw, P.S., Lightbody, K.L., Veverka, V., Muskett, F.W., Kelly, G., Frenkiel, T.A., Gordon, S.V., Hewinson, R.G., Burke, B., Norman, J., et al. (2005). Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. EMBO J. *24*, 2491-2498.

28. Sundaramoorthy, R., Fyfe, P.K., and Hunter, W.N. (2008). Structure of *Staphylococcus aureus* EsxA suggests a contribution to virulence by action as a transport chaperone and/or adaptor protein. J. Mol. Biol. *383*, 603-614.

29. Daleke, M.H., van der Woude, A.D., Parret, A.H.A., Ummels, R., de Groot, A.M., Watson, D., Piersma, S.R., Jiménez, C.R., Luirink, J., Bitter, W., and Houben, E.N.G. (2012). Specific chaperones for the type VII protein secretion pathway. J. Biol. Chem. *287*, 31939-31947.

30. Korotkova, N., Freire, D., Phan, T.H., Ummels, R., Creekmore, C.C., Evans, T.J., Wilmanns, M., Bitter, W., Parret, A.H.A., Houben, E.N.G., and Korotkov, K.V. (2014). Structure of the *Mycobacterium tuberculosis* type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer. Mol Microbiol *94*, 367-382.

31. Ekiert, D.C., and Cox, J.S. (2014). Structure of a PE-PPE-EspG complex from *Mycobacterium tuberculosis* reveals molecular specificity of ESX protein secretion. Proc. Natl. Acad. Sci. USA *111*, 14758-14763.

32. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E. et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature *393*, 537-544.

33. Sysoeva, T.A., Zepeda-Rivera, M.A., Huppert, L.A., and Burton, B.M. (2014). Dimer recognition and secretion by the ESX secretion system in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *111*, 7653-7658.

34. Zoltner, M., Ng, W.M.A.V., Money, J.J., Fyfe, P.K., Kneuper, H., Palmer,T., and Hunter, W.N. (2016). EssC: domain structures inform on the elusive translocation channel in the Type VII secretion system. Biochem. J. *473*, 941-952.

35. Klein, T.A., Grebenc, D.W., Shah, P.Y., McArthur, O.D., Dickson, B.H., Surette, M.G., Kim, Y., and Whitney, J.C. (2022). Dual Targeting Factors Are Required for LXG Toxin Export by the Bacterial Type VIIb Secretion System. mBio *13*, e0213722.

36. Korotkova, N., Piton, J., Wagner, J.M., Boy-Röttger, S., Japaridze, A., Evans, T.J., Cole, S.T., Pojer, F., and Korotkov, K.V. (2015). Structure of EspB, a secreted substrate of the ESX-1 secretion system of *Mycobacterium tuberculosis*. J. Struct. Biol. *191*, 236-244.

37. Klein, T.A., Grebenc, D.W., Gandhi, S.Y., Shah, V.S., Kim, Y., and Whitney, J.C. (2021). Structure of the Extracellular Region of the Bacterial Type VIIb Secretion System Subunit EsaA. Structure *29*, 177-185.

38. Helle, L., Kull, M., Mayer, S., Marincola, G., Zelder, M.-E., Goerke, C., Wolz, C., and Bertram, R. (2011). Vectors for improved Tet repressor-dependent gradual gene induction or silencing in *Staphylococcus aureus*. Microbiology *157*, 3314-3323.

39. Monk, I.R., and Stinear, T.P. (2021). From cloning to mutant in 5 days: rapid allelic exchange in *Staphylococcus aureus*. Access Microbiol**.** *3*, 000193.

40. Casabona, M.G., Kneuper, H., Alferes de Lima, D., Harkins, C.P. Zoltner, M., Hjerde, E., Holden, M.T.G., and Palmer, T. (2017). Haem-iron plays a key role in the regulation of the Ess/type VII secretion system of *Staphylococcus aureus* RN6390. Microbiology *163*, 1839-1850.

41. Miller, M., Donat, S., Rakette, S., Stehle, T., Kouwen, T.R.H.M., Diks, S.H., Dreisbach, A., Reilman, E., Gronau, K., Becher, D., et al. (2010). Staphylococcal PknB as the first prokaryotic representative of the proline-directed kinases. PLoS ONE *5*, e9057.

42. Caesar, J., Reboul, C.F., Machello, C., Kiesewetter, S., Tang, M.L., Deme, J.C. Johnson, S., Elmlund, D., Lea, S.M, and Elmlund, H. (2020). SIMPLE 3.0. Stream single-particle cryo-EM analysis in real time. J. Struct. Biol. X *4*, 10004.

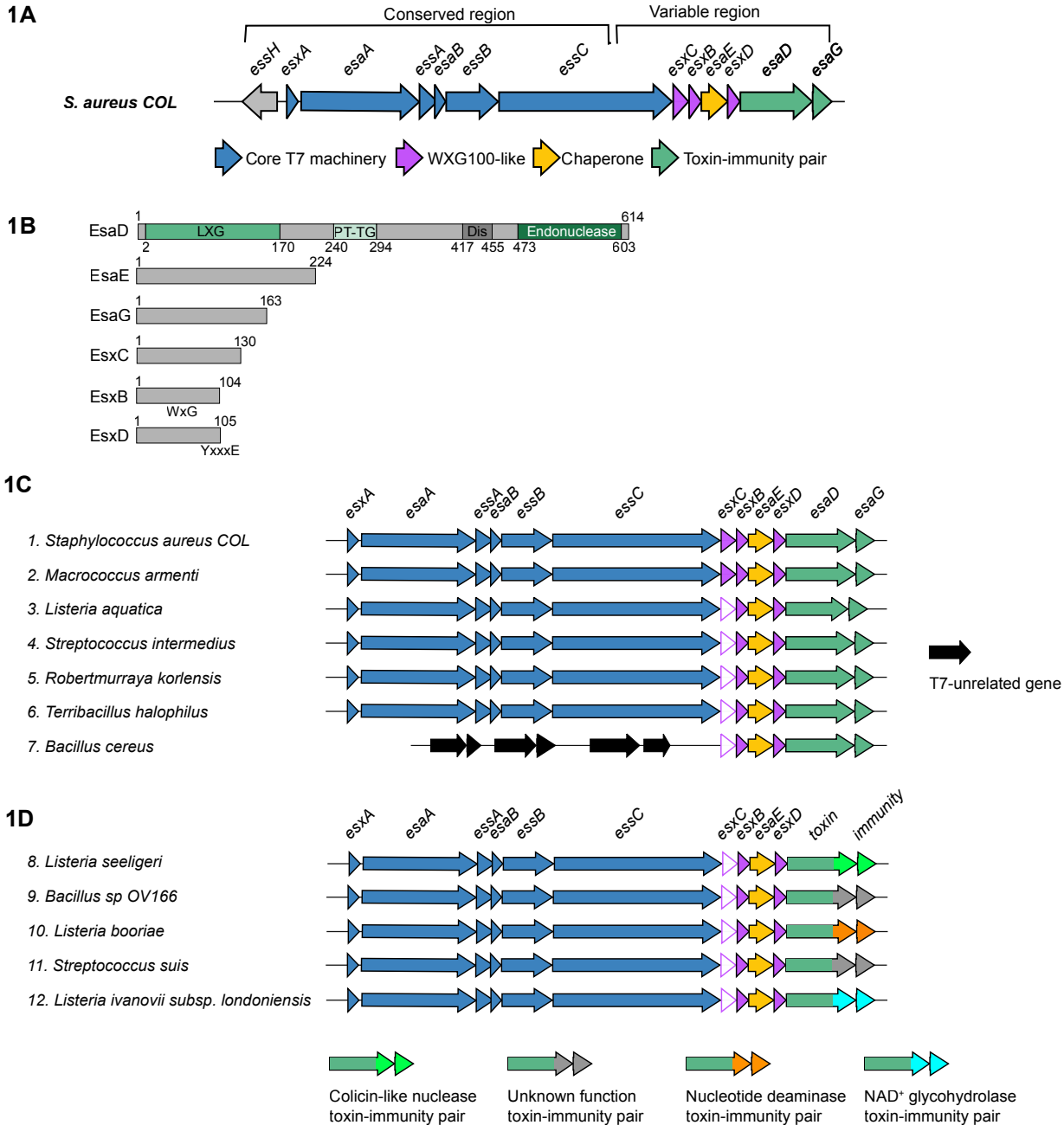
43. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290-296.

44. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko,A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589.

45. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. *BioRxiv* https://doi.org/10.1101/2021.10.04.463034.

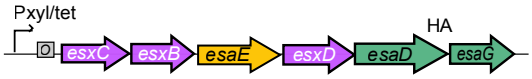
46. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70-82.

47. Saha, C.K., Sanches Pires, R., Brolin, H., Delannoy, M., and Atkinson, G.C. (2021). FlaGs and webFlaGs: discovering novel biology through the analysis of gene neighbourhood conservation. Bioinformatics *37*, 1312-1314.

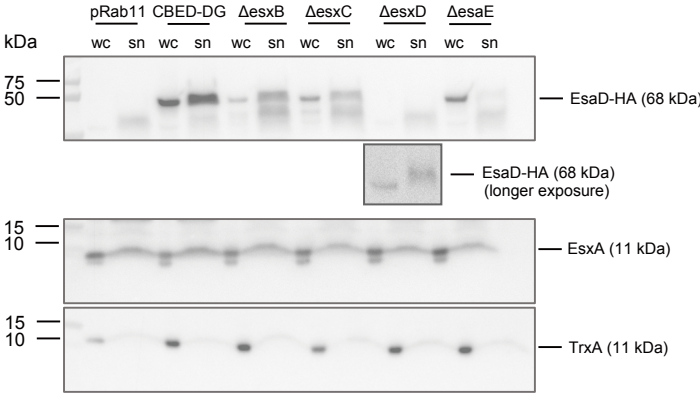

1A
Conserved region
Variable region
essH
esxA
esaA
essA
esaB
essB
essC
esxC
esxB
esaE
esxD
esaD
esaG
S. aureus COL
Core T7 machinery
WXG100-like
Chaperone
Toxin-immunity pair
1B
EsaD
1
614
LXG
PT-TG
Dis
Endonuclease
2
170
240
294
417
455
473
603
EsaE
1
224
EsaG
1
163
EsxC
1
130
EsxB
1
104
WxG
EsxD
1
105
YxxxE
1C
1. Staphylococcus aureus COL
2. Macrococcus armenti
3. Listeria aquatica
4. Streptococcus intermedius
5. Robertmurraya korlensis
6. Terribacillus halophilus
7. Bacillus cereus
T7-unrelated gene
1D
toxin
immunity
8. Listeria seeligeri
9. Bacillus sp OV166
10. Listeria booriae
11. Streptococcus suis
12. Listeria ivanovii subsp. londoniensis
Colicin-like nuclease toxin-immunity pair
Unknown function toxin-immunity pair
Nucleotide deaminase toxin-immunity pair
NAD+ glycohydrolase toxin-immunity pair

2A

pRab11-CBED-DG:

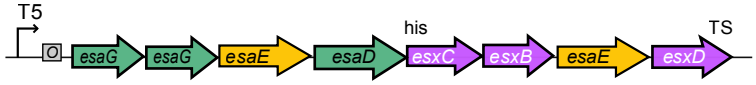


2B

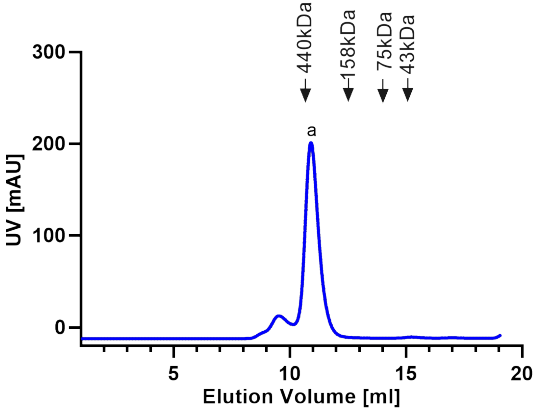


3A

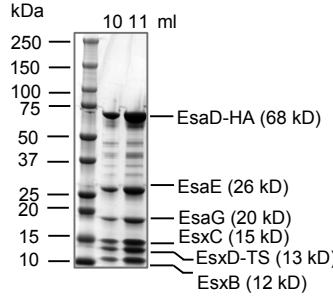
pQE70-GGED-his-CBED-tstrep:



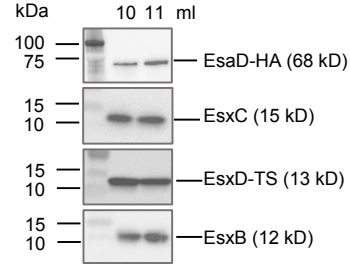
3B



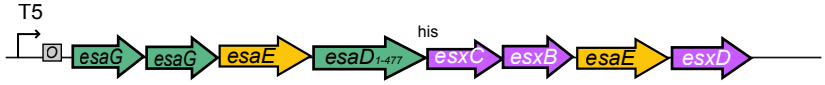
3C

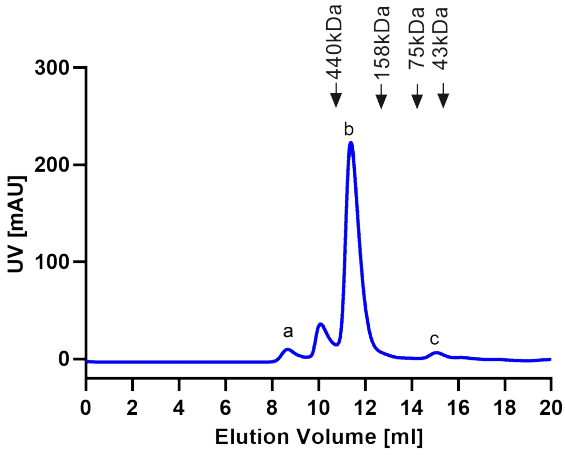


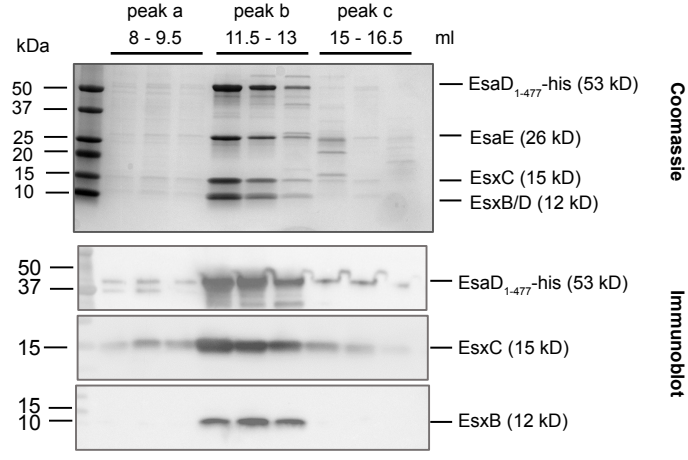
3D



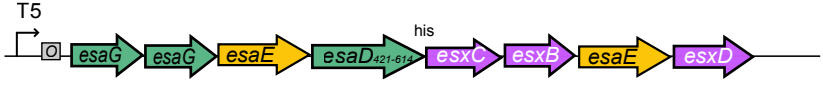
4A pQE70-GGED(1-477)-his-CBED-tstrep:

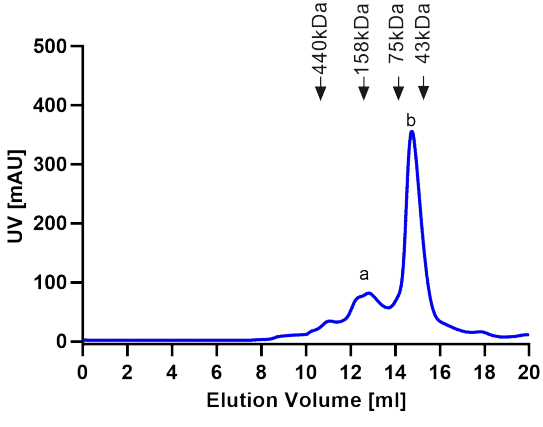


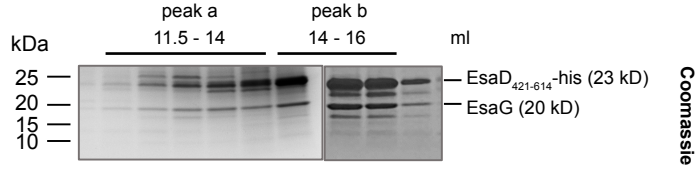




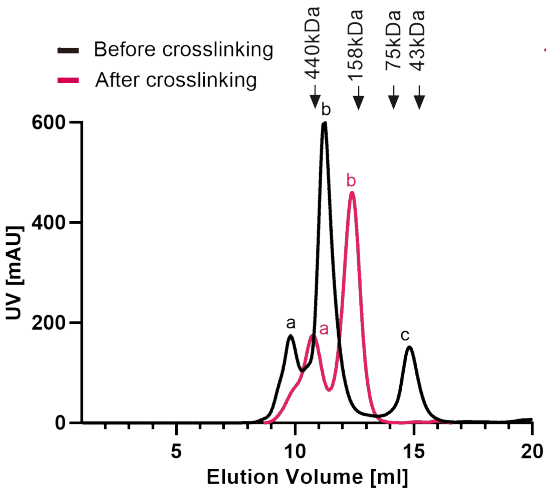
4B pQE70-GGED(478-614)-his-CBED-tstrep:

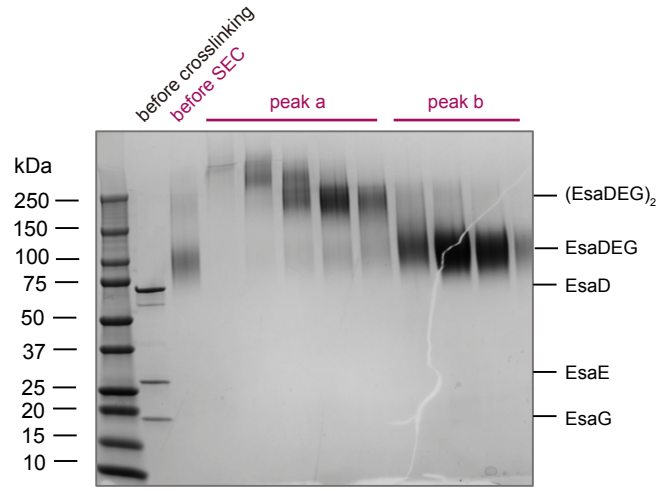




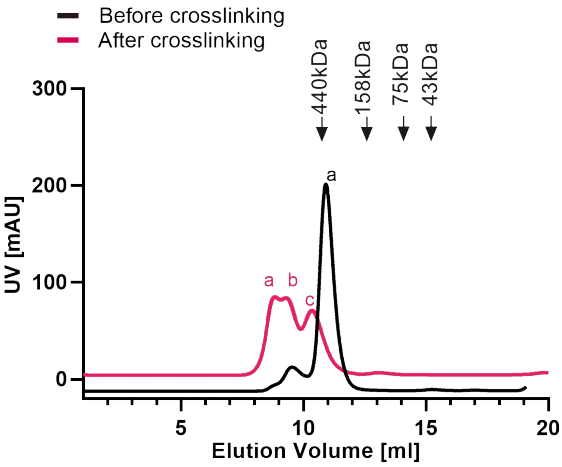


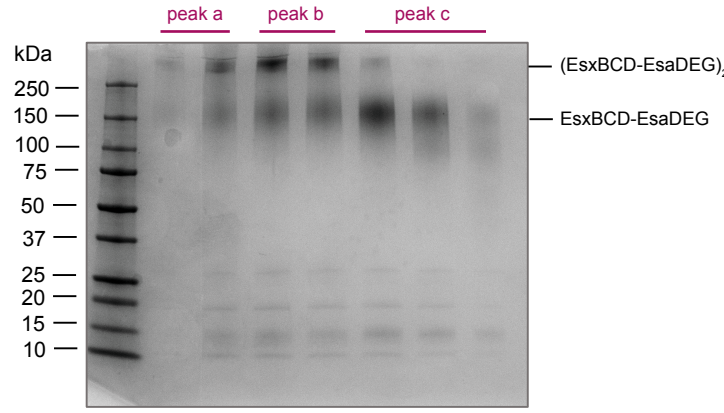
5A

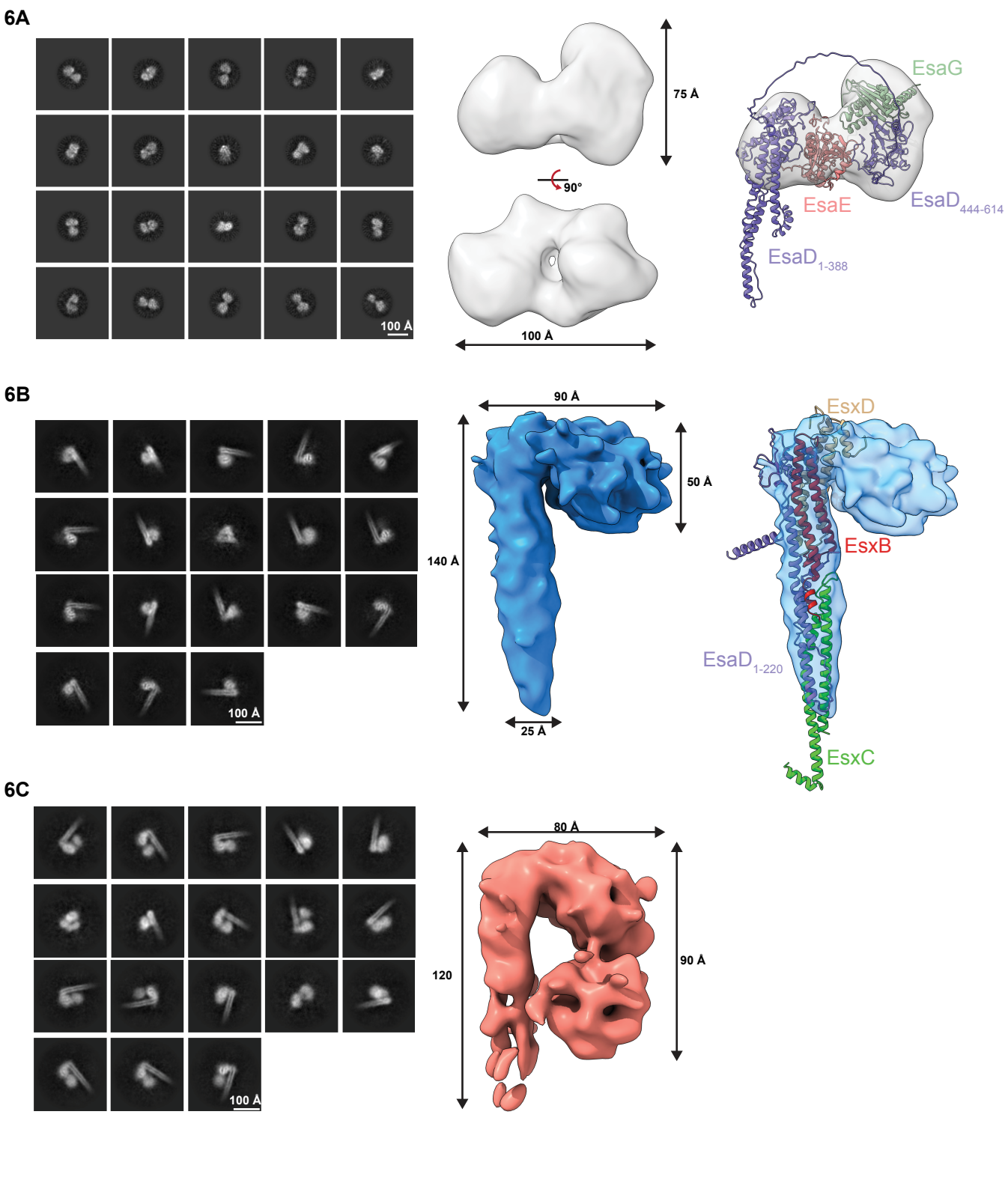

Before crosslinking
After crosslinking
440kDa
158kDa
75kDa
43kDa
UV [mAU]
Elution Volume [ml]
a
b
c

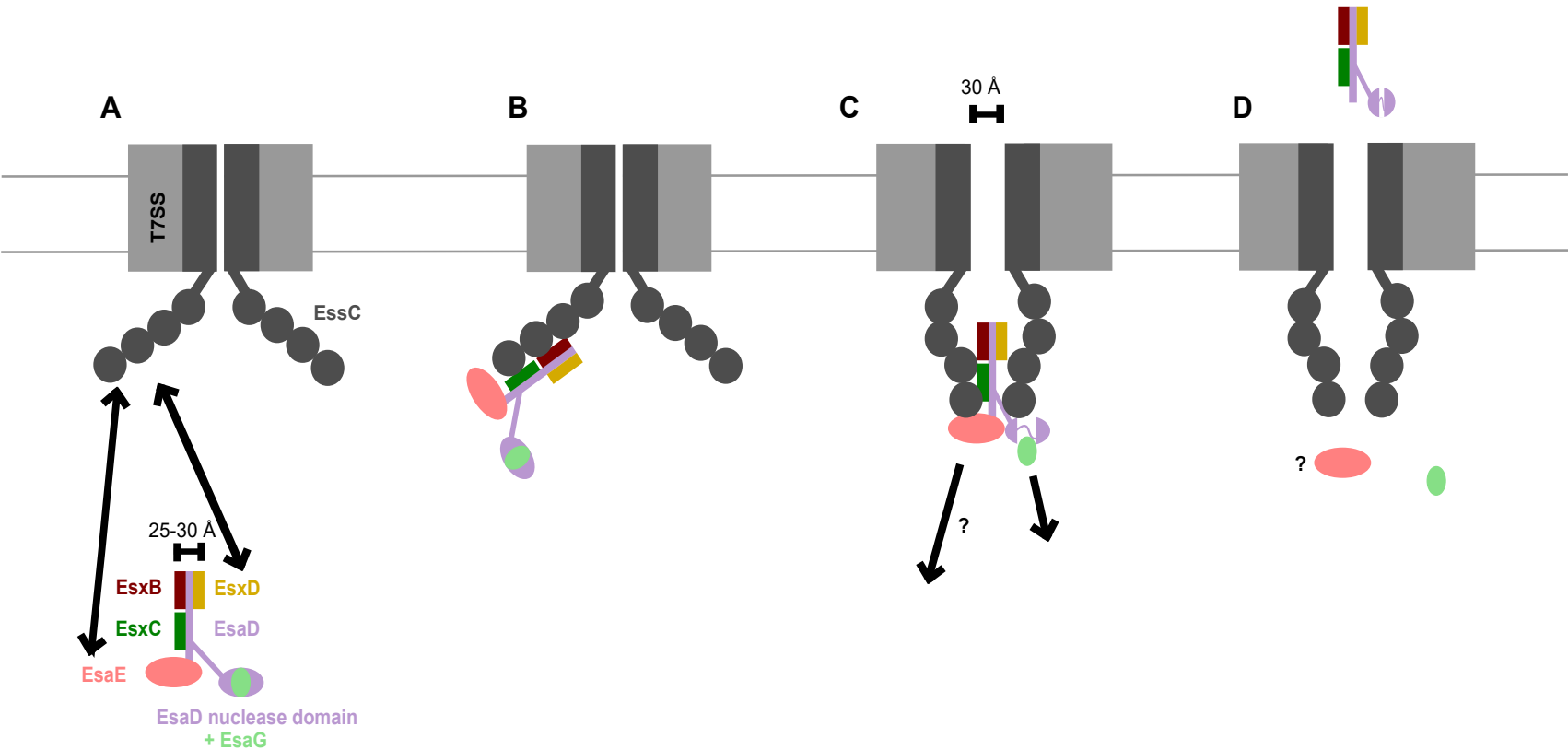

before crosslinking
before SEC
peak a
peak b
kDa
250
150
100
75
50
37
25
20
15
10
(EsaDEG)2
EsaDEG
EsaD
EsaE
EsaG

5B


Before crosslinking
After crosslinking
440kDa
158kDa
75kDa
43kDa
UV [mAU]
Elution Volume [ml]
a
b
c


peak a
peak b
peak c
kDa
250
150
100
75
50
37
25
20
15
10
(EsxBCD-EsaDEG)2
EsxBCD-EsaDEG


6A
100 Å
75 Å
90°
100 Å
EsaG
EsaE
EsaD444-614
EsaD1-388
6B
100 Å
90 Å
50 Å
140 Å
25 Å
EsxD
EsxB
EsaD1-220
EsxC
6C
100 Å
80 Å
120
90 Å


A
B
C
D
T7SS
EssC
30 Å
25-30 Å
EsxB
EsxD
EsxC
EsaD
EsaE
EsaD nuclease domain
+ EsaG
?
?