Biogenesis and structure of a type VI secretion baseplate

Yassine Cherrak, Chiara Rapisarda, Riccardo Pellarin, Guillaume Bouvier, Benjamin Bardiaux, Fabrice Allain, Christian Malosse, Martial Jean-Pierre Rey, Julia Chamot-Rooke, E. Cascales, et al.

Biogenesis and structure of a type VI secretion baseplate

Yassine Cherrak[1#], Chiara Rapisarda[2,3#], Riccardo Pellarin[4], Guillaume Bouvier[4], Benjamin Bardiaux[4], Fabrice Allain[4], Christian Malosse[5,6], Martial Rey[5,6], Julia Chamot-Rooke[5,6], Eric Cascales[1], Rémi Fronzes[2,3]* and Eric Durand[1,7]*

[1]Laboratoire d'Ingénierie des Systèmes Macromoléculaires (LISM), Institut de Microbiologie de la Méditerranée (IMM), UMR7255, Aix-Marseille Université - CNRS, Marseille, France.
[2]Institut Européen de Chimie et Biologie, University of Bordeaux, 2 rue Robert Escarpit, 33607, Pessac, France.
[3]CNRS UMR 5234 Microbiologie Fondamentale et Pathogénicité, Paris, France.
[4]Institut Pasteur, Structural Bioinformatics Unit, Department of Structural Biology and Chemistry, CNRS UMR 3528, C3BI USR 3756, Paris, France.
[5]USR 2000, CNRS, Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France.
[6]Mass Spectrometry for Biology Unit, Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France
[7]Laboratoire d'Ingénierie des Systèmes Macromoléculaires (LISM), Institut de Microbiologie de la Méditerranée (IMM), UMR7255, INSERM - Marseille, France.

[#] these authors contributed equally to the work
*corresponding author: edurand@imm.cnrs.fr & edurand@inserm.fr and r.fronzes@iecb.u-bordeaux.fr

ABSTRACT

To support their growth in a competitive environment and cause pathogenesis, bacteria have evolved a broad repertoire of macromolecular machineries to deliver specific effectors and toxins. Among these multiprotein complexes, the type VI secretion system (T6SS) is a contractile nanomachine that targets both prokaryotic and eukaryotic cells. The T6SS comprises two functional sub-complexes: a bacteriophage-related tail structure anchored to the cell envelope by a membrane complex. As in other contractile injection systems, the tail is composed of an inner tube wrapped by a sheath and built on the baseplate. In the T6SS, the baseplate is not only the tail assembly platform, but also docks the tail to the membrane complex and hence serves as an evolutionary adaptor. Here we define the biogenesis pathway and report the cryo-EM structure the wedge protein complex of the T6SS from Enteroaggregative Escherichia coli (EAEC). Using an integrative approach, we unveil the molecular architecture of the whole T6SS baseplate and its interaction with the tail sheath, offering detailed insights into its biogenesis and function. We discuss architectural and mechanistic similarities but also revealed key differences with the T4 phage and Mu phage baseplates.

INTRODUCTION

The bacterial Type VI secretion system (T6SS) is one of the key players for microbial competition, and an important virulence factor during bacterial infections. This versatile nanomachine delivers a wide arsenal of effector proteins directly into prokaryotic and eukaryotic target cells [1–4]. T6SS anti-bacterial activities promote privileged access to the niche, to nutrients or to DNA. In most cases, T6SS causes damage within competitor bacterial cells and therefore participates in the reshaping of bacterial communities such as the

microbiota [5,6]. In addition, some T6SS confer anti-host capabilities, *e.g.* phagocytosis inhibition, by remodelling the host cell cytoskeleton [7–10].

The T6SS belongs to the broad family of contractile injection systems (CIS), that includes bacteriophages, high-molecular-weight tailocins such as R-pyocins, and specific *apparati* necessary for the establishment of symbiosis or for the induction of morphological changes [11–16]. All these structures comprise a common core: the tail. CIS tails are composed of an inner tube wrapped by a sheath built under an extended, metastable conformation on an assembly platform, the baseplate. The T6SS tail tube/sheath is a hundred-nanometer-long cytoplasmic structure. It is made of TssB/C subunits that polymerize to form the contractile sheath [17,18], which surrounds the attacking arrow composed of an inner tube of stacked Hcp hexameric rings [19,20] tipped by the trimeric VgrG puncturing spike [21]. Various signals, such as contact with the target cell, chemical signals released by competitor or kin cells, response to attacking cells or conjugative transfer, induce structural rearrangements of the sheath leading to its contraction and to the propulsion of the Hcp-VgrG arrow into the target cell [22–25]. Assembly of the tail tube/sheath is initiated on the baseplate. In addition to controlling sheath extension, the baseplate also serves to trigger sheath contraction. During T6SS biogenesis, the baseplate docks to a trans-envelope complex [17,26–28] composed of TssJ, TssL and TssM [29,30]. By connecting the tail to the membrane complex and initiating tail tube/sheath polymerization, the baseplate is a central piece of the T6SS machinery. In addition, by binding cargo effectors through VgrG, the T6SS baseplate also serves as an effector-sorting platform [2,3,31].

CIS baseplates comprise a minimal core of five proteins that share homology with the prototypical T4 phage gp6, gp7, gp25, gp53, and gp27 proteins [11]. Gp6, gp7, gp25 and gp53 assemble into a unit called wedge [32]. Biogenesis of the baseplate occurs by the polymerization of six wedges around the central gp27 hub [32,33]. The T6SS baseplate is composed of five essential subunits: TssE, TssF, TssG, TssK and VgrG [27]. TssE is a structural homologue of gp25 [34] and has been recently identified as the sheath initiator [35]; TssF shares a homology with the N-terminal region of gp6, whereas TssG has been proposed to fulfill the role of gp7 or gp53 [27,36]. VgrG is a chimeric protein in which the gp27 hub is fused to the OB-fold/β-helix needle of gp5 [7,21]. TssF and TssG interact tightly and stabilize each other [27]. TssK interacts with the TssFG complex [27,37]. Taylor *et al.* recently reported the purification of the TssKFG complex bound to TssE [36]. Hence, it is proposed that TssFG, TssKFG and TssKFGE are assembly intermediates of the T6SS baseplate and have structural and functional homologies to the bacteriophage wedges. In agreement with this hypothesis, contacts between the TssFG complex and VgrG have been identified [27], suggesting that as in the T4 phage biogenesis pathway, TssKFGE wedges could polymerize around the VgrG hub to form a hexagonal baseplate. The baseplate is docked to the membrane complex mainly by interactions between TssK and both the cytoplasmic domains of TssL and TssM [26,27,38,39]. The crystal structure of TssK recently revealed that it shares a similar structural architecture with Siphoviridae phage receptor binding proteins and uses the membrane complex as a receptor to anchor the tail [28].

Besides the critical role of the T6SS baseplate complex, we still lack crucial information on its biogenesis and architecture. Recently, the 8.5-Å-resolution structure of an assembled T6SS baseplate revealed its overall shape [35], but did not provide detailed information on the molecular organization of the subunits. Here we used a hybrid approach combining cryo-electron microscopy (cryo-EM), biochemical analysis, native mass spectrometry, evolutionary covariance, and molecular modelling to unveil the assembly pathway of the T6SS baseplate and report the detailed structure of the TssKFGE baseplate wedge complex from the model organism Enteroaggregative *Escherichia coli* (EAEC). The structure highlights unanticipated structural and functional conservation with orthologous bacteriophage proteins. Finally, we fit the atomic model of the T6SS wedge complex into the

recent reconstruction of the fully assembled baseplate [35] to provide an unprecedented structural and functional understanding of the T6SS baseplate.

RESULTS

Biogenesis of the T6SS baseplate.

Previous fluorescence microscopy studies have monitored T6SS baseplate dynamics using a chromosomally-encoded and functional fusion protein between TssK and the super-folder GFP (TssK$_{sf\text{GFP}}$) in EAEC [27]. TssK$_{sf\text{GFP}}$ assembles fluorescent foci that form independently of the tail sheath and that are recruited to the membrane complex [27]. To gain further insights into the genetic requirements for TssK$_{sf\text{GFP}}$ foci formation, we observed TssK$_{sf\text{GFP}}$ in T6SS non-polar gene deletion mutants corresponding to T6SS wedge potential components (see Strains construction in Methods). Fluorescence microscopy recordings demonstrate that TssK$_{sf\text{GFP}}$ foci formation depends on TssF and TssG but is independent of TssE (**Fig. 1a & Supplementary Fig. 1**). When monitored into additional baseplate mutants, we observed that these foci form independently of TssA but require VgrG (**Fig. 1a**). Therefore, these foci likely represent TssK$_{sf\text{GFP}}$FG complexes bound to VgrG.

To provide further information on the composition of the T6SS wedge, we investigated the formation of stable TssK$_{sf\text{GFP}}$-containing subcomplexes in EAEC cells by native PAGE. Native PAGE profiles immunodetected with anti-GFP antibodies revealed the presence of a high-molecular weight complex (HMWC) with a size of ~ 800 kDa (**Fig. 1b**). This complex does not contain TssE, TssA, VgrG and TssM and likely corresponds to TssK$_{sf\text{GFP}}$FG since (1) it disappears in the absence of *tssF* or *tssG,* (2) a HMWC of a comparable size is observable upon pull-down of TssK$_{sf\text{GFP}}$ co-produced with TssF and TssG in the heterologous host *E. coli* BL21(DE3), and (3) analysis of this HMWC on denaturing SDS-PAGE reveals the presence of TssK$_{sf\text{GFP}}$, TssF and TssG (**Fig. 1b**).

Taken together, the fluorescence microscopy and native-PAGE results, and the previous reports of TssKFG and TssKFGE complex purifications in *Serratia marcescens* and *E. coli* [36,37], suggest that the TssKFG complex represents a stable intermediate during T6SS baseplate biogenesis. We therefore propose that T6SS baseplate biogenesis starts with the formation of the TssKFG complex and then proceeds with the polymerization of TssKFG building units around the VgrG hub. The observation that TssE is not required for TssKFG-VgrG complex formation, further suggests that TssE binds to the TssKFG either prior to or after its polymerization. This assembly pathway is comparable to that of the minimal phage baseplate, in which gp25 attaches to the baseplate either after completion of the gp10-7-8-6-53 complex [40] or at a later stage, triggering the polymerization of the contractile sheath [41].

Interaction network within the T6SS baseplate

To gain further insight into the connectivity network between the T6SS baseplate components, we performed a systematic biochemical pull-down assay (**Supplementary Fig. 2a-e**). This approach confirmed or revealed a number of contacts including interactions between TssG and TssF, TssE, and TssK (**Fig. 1c).** We then tested whether intermediate sub-complexes, suggested by the assembly pathway defined above, could be purified. In agreement with the proposed assembly pathway, we succeeded to pull-down biogenesis intermediate complexes consisting of TssFG, TssKFG and TssKFGE (**Supplementary Fig. 2a-b**). Based on these data, we propose that the TssKFGE sub-complex represents the T6SS equivalent of the bacteriophage wedge unit (TssFGE), bound to the TssK membrane complex adaptor.

Purification, stoichiometry and cryo-EM structure of a T6SS wedge complex

To biochemically and structurally characterize T6SS baseplate building units, purification tags were positioned on the *EAEC* TssK, TssF, TssG and TssE proteins, at locations that do not interfere with their function. Upon overproduction in BL21(DE3), the TssKFG and TssKFGE complexes were isolated by affinity chromatography followed by gel filtration (**Fig. 1d** & **Supplementary Fig. 3a**). To gain further insight into the architecture of the T6SS baseplate wedge assembly, the isolated TssKFGE complex was visualized by cryo-electron microscopy (cryo-EM). The 4.6-Å resolution, three-dimensional reconstruction (**Fig. 2a**) shows that the TssKFGE complex displays an intricate architecture with no apparent symmetry. It can be, however, divided into distinct parts: (1) two wing-like structures wrapping a central backbone and (2) a root-like structure with two identical entities displaying apparent C3 symmetry linked to the first part by a thin stalk (**Fig. 2b**).
In order to properly interpret the density map, we first sought to determine the stoichiometry and stability of the TssKFGE complex using native mass spectrometry (see **Supplementary data for more details)** (**Supplementary Fig. 3b, 3c and 3d)**. We determined that the TssKFGE complex comprises 6 TssK, 2 TssF, 1 TssG and 1 TssE subunits ($TssK_6F_2G_1E_1$; theoretical mass: 498,905 Da).
Based on this stoichiometry, we identified densities that could accommodate the two TssK trimers, for which the crystal structure has been recently determined [28]. The two trimers of TssK readily fit in the density map corresponding to the root-like structure **(Fig 2c)**.. The remaining densities corresponding to the stalk and wing-like domains would contain TssF, TssG and TssE. Careful inspection and segmentation of the map lead us to determine that the wing-like domains are formed by two similar densities (**Fig. 2d-f**) that would correspond to TssF. The remaining density, bridging TssK and TssF, would be attributed to TssG **(Fig. 2d and Fig. 2g)**, while TssE would be located at the tip of the complex (**Fig. 2h**) (**Supplementary data**).

Structural analysis of the T6SS wedge complex

Biochemical and evolutionary covariance analyses establish TssG as a central structural component of the T6SS wedge complex – Data described above unveil TssG as the central component of the baseplate. Unfortunately, no structural information is available for TssG. We therefore used residue contact predictions based on evolutionary covariance [42] to determine the TssG domain organization. This analysis identified two putative independent domains in TssG, TssG-D1 (amino-acids 15-140) and TssG-D2 (amino-acids 180-300) (**Supplementary Fig. 4a**). Pull-down assays with these two domains demonstrated that TssG-D1 interacts with TssE, whereas TssG-D2 interacts with TssK and TssF (**Fig. 1c & Supplementary Fig. 4b-c**).
To further characterize the relative importance of these two domains *in vivo*, we performed a dominant-negative approach by "small domain interference" (SDI) [43,44] **(Supplementary data) (Supplementary Fig. 5a-e and Supplementary Fig 6a-b)** and confirmed that TssG-D2 has a central role for T6SS wedge assembly. We then used the EVcomplex program [45] to predict inter-molecular contacts between TssG and TssK using evolutionary covariance analysis (see Methods section) (**Supplementary Fig. 6a**). Two TssG-D2 residues, Pro-240 and Leu-255, corresponding to predicted TssK-TssG-D2 interfacial residue pairs with the highest scores, were substituted, to alanines (P240A and L255A and P240A-L255A) and assayed by SDI in interbacterial competition experiments (**Supplementary Fig. 6b**)

Structure of the TssK trimers – The resolution of the density map corresponding to the two TssK trimers was between 3.8 and 33 Å (**Fig. 2a**). This level of detail allowed us to obtain a full atomic model of TssK (**Fig. 3a**) (**Supplementary Fig. 7a-b**) (see Methods for details). As described previously, TssK can be divided into four parts from its N- to C-terminus: an N-

terminal α-helix, a β-sandwich domain (also named shoulder domain), a 4 α-helix bundle domain (also named neck domain) and a C-terminal α/β domain (also named head domain) (**Fig. 3a**). When compared with the published TssK crystal structure, with the exception of the relative position of the C-terminal domain, the overall structure of the TssK protomer is conserved in the T6SS wedge. The cryo-EM and crystal structures could be superimposed with an RMSD of 1.14 Å (**Supplementary Fig. 7a-b**).
In the wedge complex, the two TssK trimers are in contact with each other and interact with TssG. The TssK inter-protomer contacts define a large interacting surface of 2,700-2,800 $Å^2$ stabilized by hydrogen bonds and salt bridges [46]. The newly built loops 1-18 and 130-143 participate to this interface by forming contacts with the neighbouring protomer. Remarkably the three loops located between residues 105 to 145 define a triangle that encompasses the loop 1-18 and α1 helix bundle at the centre of the trimer (**Supplementary Fig. 7c**), forming a flat triangular surface at the top of each TssK trimer, which contacts the rest of the wedge complex. This triangular surface is delineated by a polar scaffold made of the strands 105-143 and loops around a hydrophobic patch made of part of the loop 1-18 (residues 12-14).

Structure of TssG and TssF –The resolution of the densities corresponding to TssG and TssF varied between 4.3 and 8 Å. Since there was no homologous structure of TssG and TssF available, we built the structure of these proteins de novo helped by a priori knowledge on their topology, secondary structure, and intra-molecular contacts predicted from evolutionary covariance. We devised an iterative pipeline to integrate all this data (prior data and pipeline are described in Methods sections "Evolutionary Covariance Analysis", "TssKFGE model fitting and de novo tracing", as well as **Supplementary Fig. 12**). Eventually, we were able to obtain an atomistic model of this part of the complex, in which most of the sequence of the proteins could be assigned to the cryo-EM density and secondary structure elements could be identified (**Fig. 3c-d**).
TssG is made of two globular domains, head and body/feet, separated by a neck domain (**Fig. 3c**). The N-terminal neck domain, corresponding to TssG-D1, is made of two short helices and loops, whereas the C-terminal body domain, corresponding to TssG-D2, folds as an α/β domain comprising a four-strand β-sheet and three helices (**Fig. 3c**). On each side of the β-sheet, two loops extend to form the two-foot domains (foot1 and foot2). The last strand of the sheet extends into a C-terminal extension of 17 residues.
TssF is a globular protein with an N-terminal elongated extension called antenna (**Fig. 3d**). The antenna is made of 2 helices while the C-terminal globular domain can be divided into 5 sub-domains: domain 1 (TssF-D1) is a β-sandwich flanked by loops containing short helices; domain 2 (TssF-D2) is a β-sandwich; domain 3 (TssF-D3) is an α-helical domain and comprises 3 short helices; domain 4 (TssF-D4) is an α/β domain composed of one helix and a 4-strand β-sheet. The last domain, named branching domain (TssF-BD), is a β-sandwich that is formed by strands that link the antenna with TssF-D1 (three strands of the fold), TssF-D1 with TssF-D2 (one strand of the fold) and TssF-D2 with TssF-D3 (two strands of the fold). As previously noted from the two wing densities, the structures of the two TssF proteins (TssFa and TssFb) are superimposable with the exception of the two antennas, which are in two distinct orientations (**Supplementary Fig. 7d**).

Structure of the TssFG complex – TssG and the two copies of TssF (TssFa and TssFb) are assembled to form the pyramidal cap of the T6SS wedge complex (**Fig. 3e**). TssG-D1 forms an heterotrimeric helical bundle with the two TssF antennas (**Fig. 3e-f**). Interestingly, the fold of the TssG body is similar to the TssF-D1 domain (**Supplementary Fig. 7e**), and these three domains define a triangular assembly at the base of the T6SS wedge complex (**Fig. 3f**). Together with the heterotrimeric helical bundle, this structure forms the heterotrimeric

scaffold of the TssFG complex (**Fig. 3g)**. The D2, D3, D4 and BD domains from TssFa and TssFb form the wing-like domains on both sides of the trimeric scaffold (**Fig. 3G)**. TssG is the central backbone of the wedge complex: it interacts with both TssFa and TssFb all along its structure, whereas TssFa and TssFb have very few points of contact between each other.

The TssK-TssFG interface – The interaction of the two TssK trimers with the TssFG complex is mainly mediated through the two TssG-D2 foot domains (**Fig. 3h-i)**. The TssG foot1 domain (residues 227-242) interacts with residues 10-15 of the three TssK protomers of the TssK1 trimer (**Fig. 3i, upper panel; Supplementary Fig. 7f**). The situation is more complex for the TssK2 trimer: residues 10-15 of the three TssK2 protomers and 138-143 of one monomer (green) make contacts with the TssG foot2 domain (residues 303-322), and the TssG C-terminal extension (residues 345-347) makes contact with residues 116-117 of one TssK monomer (green in **Fig. 3i, lower panel; Supplementary Fig. 7g**), whereas the TssG foot1 domain (residues 221-224) also contributes to the stabilization of the edifice by interacting with the loop 116-120 in TssK2.

Molecular model of the T6SS baseplate

Recently, an 8-Å-resolution cryo-EM structure of the *V. cholerae* T6SS baseplate associated to a non-contractile sheath was reported [35]. Although no density could be attributed to specific baseplate components, densities corresponding to the Hcp tube and VgrG spike are clearly visible [35]. We used the deposited map (EMD-3879) to build a molecular model of the entire T6SS baseplate from EAEC (See Methods for more details) (**Fig. 4a)** (**Supplementary Fig. 8a-b**).

The fully assembled T6SS baseplate is 337 Å in diameter and 180 Å in height (**Fig. 4a**). These dimensions are compatible with the densities attributed to the baseplate complex in the cryotomogram of the *Myxococcus xanthus* T6SS [47]. The rings formed by TssFG (wedge ring) and TssK (connector ring) are 100 Å and 110 Å in height respectively (**Fig. 4a**). Within the wedge ring, the individual wedge complexes are organized side by side. Their main axis, along the helical bundle, makes a 30° angle with the symmetry axis of the ring (**Fig. 4b)**. The lateral contacts are mediated by interactions between TssFa and TssFb from two adjacent wedge complexes (**Fig. 4b**). Overall, the two protomers are perpendicular to each other, TssFa wrapping the adjacent TssFb (**Fig. 4c**). In detail, the TssFa D1 and BD domains interact with the antenna of the adjacent TssFb whereas TssFa D3 and D4 domains interact with the adjacent TssFb D3 and BD domains (**Fig. 4c**). To a lesser extent, contacts also exist between two adjacent TssFb, two adjacent TssFa, and TssG and TssFb (**Fig. 4b**). In the TssK ring, one TssK protomer, belonging to the TssK1 trimer (grey density), interacts with two TssK protomers belonging to the TssK2 trimer in the adjacent wedge complex (green and brown densities) (**Fig. 4b**). Finally, there are contacts between TssFa and two TssK protomers from the TssK2 trimer from the adjacent wedge complex (**Fig. 4b**).

The D1, D2 and BD domains of the TssFa proteins delineate the inner surface of the wedge ring, defining a chamber named TssF chamber (**Fig. 4a**). Together with the wedge ring, the connector ring defines another chamber, named the TssK chamber. Both chambers are separated by a central constriction of ~ 40 Å in diameter due to TssFa D2 and BD domains (**Fig. 4a**). Fitting of the VgrG crystal structure in the assembled baseplate reveals that the TssFa chamber accommodates perfectly the gp27-like hub domain of VgrG while the VgrG gp5-like spike crosses the TssFa constriction and extends into the TssK chamber (**Supplementary Fig. 8b**). In agreement with previous bacterial two-hybrid data [27], this reconstruction suggests that the interactions between VgrG and the baseplate are exclusively mediated by TssFa-VgrG interactions (**Fig. 4c; Supplementary Fig. 8b)**.

Surprisingly, there is no density in the cryo-EM map of the *V. cholerae* baseplate that could accommodate TssE. All the densities seen in the baseplate region of this map could be attributed to TssK, TssG or TssF. The rest of the map corresponds to the VipAB sheath, the Hcp tube, and the VgrG spike. The fact that TssE is invisible in this map could be explained by some degree of flexibility of the protein within the structure after sheath assembly. However, based on the location of TssE in the cryo-EM of the EAEC TssKFGE complex, TssE elegantly fits in a space between the wedge helical bundle and TssB in the model of the assembled baseplate (**Fig. 4d)**.

DISCUSSION

The T6SS baseplate - sheath connection

The TssFGE wedge ring interacts directly with the sheath structure. A homology model of the EAEC extended sheath was fitted into the molecular model of the baseplate attached to the sheath. Our model suggests that the main contacts are established between the TssFb D2 domain and the TssBC N-terminal antenna. In addition, the TssG head domain and TssE likely stabilize the interaction between the sheath and the baseplate (**Fig. 4c-d**) by interacting with the C-terminal domain of TssB as recently proposed [48].

The T6SS baseplate - membrane complex connection

A number of contacts have been identified between the baseplate components and the cytoplasmic domains of the TssL and TssM inner membrane proteins [26–28,38,39]. The main contacts involve binding of TssK to both TssL and TssM [26,28,39] but additional contacts, notably between TssL and TssE, and between TssM and TssG have been reported [27,39]. While these contacts could not be explained by the current structure, it is known that structural rearrangements occur in the bacteriophage T4 upon sheath contraction [49–51]. If such conformational changes occur in the T6SS baseplate, different contacts may stabilize the interaction of the baseplate with the membrane complex after sheath contraction.

However, the position of the TssK connector ring confirms that TssK is the major determinant for mediating baseplate docking to the membrane complex. We have recently shown that the $TssK_S$ domain shares homology with siphophage receptor-binding proteins, whereas it has evolved a specific C-terminal head domain, $TssK_H$, to use the membrane complex as a receptor [28]. In agreement with these results, the orientation of the TssK trimers in the baseplate places the $TssK_S$ domains in contact with the TssFG cap complex (**Fig. 2c**), whereas the $TssK_H$ domains extend in the opposite direction compared to the sheath, at the predicted location of the membrane complex (**Fig. 4a**).

Comparison between T6SS and bacteriophage baseplates.

(A detailed version of this section is provided in supplementary information)

While our data confirms a strong analogy between the T6SS baseplate and the "simple contractile baseplate" from the Mu phage, it is also clear that T6SS and T4 phage wedge components are structurally related to each other. Indeed, the structures of TssF and TssG reavealed that they are gp6 and gp7 counterparts in the T4 phage baseplate (**Fig. 5a-b**) (**Supplementary Fig. 9a-b).** Interestingly, a detailed analysis of both baseplate architectures revealed that, while inter-wedge contacts are different in the T4 and the T6SS baseplates **(Fig. 4b & Fig. 5c),** interactions with the central spike and the contractile sheath are quite well conserved (**Fig. 5c-d, Supplementary Fig. 9c-e**). Finally, this comparison revealed how the

apical part of the each baseplate is structurally specialized to interact with different targets (**Fig. 5c-e)**.

Assembly mechanism and stability of of the T4 and T6SS baseplates
(A detailed version of this section is provided in supplementary information)

While the T4 bacteriophage wedge complex appears to be transient, the T6SS wedge complex is stable, as shown by the isolation of TssKFG or TssKFGE complexes in EAEC, *S. marcescens* and uropathogenic *E. coli* (this work; [36,37]). By contrast, the fully assembled T4 phage baseplate is much more stable than the T6SS baseplate, since we did not succeed to purify the T6SS hexagonal baseplate. This higher stability of a preformed T6SS baseplate intermediate may reflect an adaptation to the secretion process. Delayed polymerization of the wedge around the VgrG hub and fast recycling of the wedge complexes might be necessary during each secretion cycle

Concluding remarks

In this work, we provide an unprecedented functional and structural study of the T6SS baseplate building block, the wedge complex. Due to the conservation of T6SS wedge complexes among pathogenic bacteria, the atomic model of the TssKFGE complex will facilitate the design of anti-T6SS compounds targeting hot spots of the baseplate assembly, paving the way towards new therapeutic avenues to replace or help classical antibiotherapies.

References:

1. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475,** 343–347 (2011).
2. Durand, E., Cambillau, C., Cascales, E. & Journet, L. VgrG, Tae, Tle, and beyond: the versatile arsenal of Type VI secretion effectors. *Trends Microbiol.* **22,** 498–507 (2014).
3. Alcoforado Diniz, J., Liu, Y.-C. & Coulthurst, S. J. Molecular weaponry: diverse effectors delivered by the Type VI secretion system. *Cell. Microbiol.* **17,** 1742–1751 (2015).
4. Russell, A. B., Peterson, S. B. & Mougous, J. D. Type VI secretion system effectors: Poisons with a purpose. *Nature Reviews Microbiology* **12,** 137–148 (2014).
5. Chassaing, B. & Cascales, E. Antibacterial Weapons: Targeted Destruction in the Microbiota. *Trends Microbiol.* **26,** 329–338 (2018).
6. Sana, T. G., Lugo, K. A. & Monack, D. M. T6SS: The bacterial 'fight club' in the host gut. *PLoS Pathogens* **13,** (2017).
7. Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D. & Mekalanos, J. J. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 15508–15513 (2007).
8. Durand, E. *et al.* Crystal Structure of the VgrG1 Actin Cross-linking Domain of the Vibrio cholerae Type VI Secretion System. *J. Biol. Chem.* **287,** 38190–38199 (2012).
9. Sana, T. G. *et al.* Internalization of Pseudomonas aeruginosa Strain PAO1 into Epithelial Cells Is Promoted by Interaction of a T6SS Effector with the Microtubule Network. *MBio* **6,** e00712 (2015).

10. Ma, A. T., McAuley, S., Pukatzki, S. & Mekalanos, J. J. Translocation of a Vibrio cholerae Type VI Secretion Effector Requires Bacterial Endocytosis by Host Cells. *Cell Host Microbe* **5,** 234–243 (2009).
11. Leiman, P. G. & Shneider, M. M. Contractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.* **726,** 93–114 (2012).
12. Ge, P. *et al.* Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat. Struct. Mol. Biol.* **22,** 377–382 (2015).
13. Böck, D. *et al.* In situ architecture, function, and evolution of a contractile injection system. *Science (80-.).* **357,** 713–717 (2017).
14. Heymann, J. B. *et al.* Three-dimensional Structure of the Toxin-delivery Particle Antifeeding Prophage of Serratia entomophila. *J. Biol. Chem.* **288,** 25276–25284 (2013).
15. Shikuma, N. J. *et al.* Marine tubeworm metamorphosis induced by arrays of bacterial phage tail-like structures. *Science (80-.).* **343,** 529–533 (2014).
16. Cascales, E. Microbiology: And Amoebophilus Invented the Machine Gun! *Curr. Biol.* **27,** R1170–R1173 (2017).
17. Brackmann, M., Wang, J. & Basler, M. Type VI secretion system sheath inter‐subunit interactions modulate its contraction. *EMBO Rep.* **19,** 225–233 (2018).
18. Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483,** 182–186 (2012).
19. Ballister, E. R., Lai, A. H., Zuckermann, R. N., Cheng, Y. & Mougous, J. D. In vitro self-assembly of tailorable nanotubes from a simple protein building block. *Proc. Natl. Acad. Sci.* **105,** 3733–3738 (2008).
20. Brunet, Y. R., Henin, J., Celia, H. & Cascales, E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. *EMBO Rep.* **15,** 315–321 (2014).
21. Leiman, P. G. *et al.* Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci.* **106,** 4154–4159 (2009).
22. Basler, M., Ho, B. T. & Mekalanos, J. J. Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152,** 884–94 (2013).
23. Ho, B. T., Basler, M. & Mekalanos, J. J. Type 6 Secretion System-Mediated Immunity to Type 4 Secretion System-Mediated Gene Transfer. *Science (80-.).* **342,** 250–253 (2013).
24. LeRoux, M. *et al.* Kin cell lysis is a danger signal that activates antibacterial pathways of Pseudomonas aeruginosa. *Elife* **4,** 1–65 (2015).
25. Vettiger, A. & Basler, M. Type VI Secretion System Substrates Are Transferred and Reused among Sister Cells. *Cell* **167,** 99–110.e12 (2016).
26. Zoued, A. *et al.* TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. *J. Biol. Chem.* **288,** 27031–27041 (2013).
27. Brunet, Y. R., Zoued, A., Boyer, F., Douzi, B. & Cascales, E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization. *PLOS Genet.* **11,** e1005545 (2015).
28. Nguyen, V. S. *et al.* Type VI secretion TssK baseplate protein exhibits structural similarity with phage receptor-binding proteins and evolved to bind the membrane complex. *Nat. Microbiol.* **2,** 17103 (2017).
29. Aschtgen, M.-S., Gavioli, M., Dessen, A., Lloubès, R. & Cascales, E. The SciZ protein anchors the enteroaggregative Escherichia coli Type VI secretion system to the cell

wall. *Mol. Microbiol.* **75,** 886–899 (2010).
30. Durand, E. *et al.* Biogenesis and structure of a type VI secretion membrane core complex. *Nature* **523,** 555–560 (2015).
31. Flaugnatti, N. *et al.* A phospholipase A 1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol. Microbiol.* **99,** 1099–1118 (2016).
32. Yap, M. L., Mio, K., Leiman, P. G., Kanamaru, S. & Arisaka, F. The Baseplate Wedges of Bacteriophage T4 Spontaneously Assemble into Hubless Baseplate-Like Structure In Vitro. *J. Mol. Biol.* **395,** 349–360 (2010).
33. Leiman, P. G. *et al.* Morphogenesis of the T4 tail and tail fibers. *Virology Journal* **7,** (2010).
34. Bingle, L. E., Bailey, C. M. & Pallen, M. J. Type VI secretion: a beginner's guide. *Curr. Opin. Microbiol.* **11,** 3–8 (2008).
35. Nazarov, S. *et al.* Cryo‐EM reconstruction of Type VI secretion system baseplate and sheath distal end. *EMBO J.* **37,** e97103 (2018).
36. Taylor, N. M. I. *et al.* Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533,** 346–352 (2016).
37. English, G., Byron, O., Cianfanelli, F. R., Prescott, A. R. & Coulthurst, S. J. Biochemical analysis of TssK, a core component of the bacterial Type VI secretion system, reveals distinct oligomeric states of TssK and identifies a TssK–TssFG subcomplex. *Biochem. J.* **461,** 291–304 (2014).
38. Logger, L., Aschtgen, M. S., Guérin, M., Cascales, E. & Durand, E. Molecular Dissection of the Interface between the Type VI Secretion TssM Cytoplasmic Domain and the TssG Baseplate Component. *J. Mol. Biol.* **428,** 4424–4437 (2016).
39. Zoued, A. *et al.* Structure–Function Analysis of the TssL Cytoplasmic Domain Reveals a New Interaction between the Type VI Secretion Baseplate and Membrane Complexes. *J. Mol. Biol.* **428,** 4413–4423 (2016).
40. Yap, M. L. *et al.* Sequential assembly of the wedge of the baseplate of phage T4 in the presence and absence of gp11 as monitored by analytical ultracentrifugation. *Macromol. Biosci.* **10,** 808–813 (2010).
41. Arisaka, F., Yap, M. L., Kanamaru, S. & Rossmann, M. G. Molecular assembly and structure of the bacteriophage T4 tail. *Biophys. Rev.* **8,** 385–396 (2016).
42. Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLOS Comput. Biol.* **13,** e1005324 (2017).
43. Onoue, Y. *et al.* Construction of functional fragments of the cytoplasmic loop with the C-terminal region of PomA, a stator component of the Vibrio Na+driven flagellar motor. *J. Biochem.* **155,** 207–216 (2014).
44. Rowe, H. M. *et al.* Modification of the CpsA protein reveals a role in alteration of the Streptococcus agalactiae cell envelope. *Infect. Immun.* **83,** 1497–1506 (2015).
45. Hopf, T. A. *et al.* Sequence co-evolution gives 3D contacts and structures of protein complexes. *Elife* **3,** (2014).
46. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
47. Chang, Y., Rettberg, L. A., Ortega, D. R. & Jensen, G. J. *In vivo* structures of an intact type VI secretion system revealed by electron cryotomography. *EMBO Rep.* **18,** 1090–1099 (2017).
48. Douzi, B. *et al.* Structure–Function Analysis of the C-Terminal Domain of the Type VI Secretion TssB Tail Sheath Subunit. *J. Mol. Biol.* **430,** 297–309 (2018).
49. Kostyuchenko, V. A. *et al.* Three-dimensional structure of bacteriophage T4 baseplate.

Nat. Struct. Biol. **10,** 688–693 (2003).
50. Kostyuchenko, V. A. *et al.* The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat. Struct. Mol. Biol.* **12,** 810–813 (2005).
51. Leiman, P. G., Chipman, P. R., Kostyuchenko, V. A., Mesyanzhinov, V. V. & Rossmann, M. G. Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118,** 419–429 (2004).

Figures legends.

Figure 1. Composition of the T6SS wedges complex. a. Fluorescence microscopy recordings showing $TssK_{sfGFP}$ localisation in the absence of the TssF, TssG, TssE, VgrG and TssA proteins. The positions of foci corresponding to fully-assembled baseplates are indicated by arrowheads. Microscopy analyses were performed independently three times, each in technical triplicate, and a representative experiment is shown. Scale bars, 1 μm. b. Native 4-16% gel analysed by Coomassie staining (upper left panel) or immunobloting using anti-GFP antibodies (right panel) or anti-GFP, anti-Streptag and anti-FLAG antibodies (lower left panels). The $TssK_{sfGFP}$-6 × His and $TssK_{sfGFP}$-6 × His-FG produced and purified from BL21(DE3) cells show the positions of two high-molecular weight complexes (HMWC, indicated by * and **). Formation of the higher HMWC is monitored in different *tss* mutant backgrounds, revealing that the * and ** complexes correspond to a $TssK_{sfGFP}$ trimer and $TssK_{sfGFP}$-TssF-TssG complex, respectively. Native gel experiment was performed independently three times and a representative experiment is shown. c. Summary of protein-protein interactions within the TssKFGE complex, as defined by pair-wise pull-down experiments (see Supplementary Fig. 2 and 4). Arrows indicate interactions between the two proteins or domains. d. Purification and biochemical characterization of the TssKFGE wedge complex. Analytical size-exclusion chromatography analysis of the purified TssKFGE complex (continuous line) on a Superose 6 column, calibrated with 43-, 75-, 158-, 440- and 660-kDa molecular mass markers (dotted lines). The molecular mass of each marker (in kilodaltons) is indicated on the top of the corresponding peak. An arrow indicates the position of the peak fraction corresponding to the TssKFGE complex. Inset: Purified TssKFGE complex subjected to sodium dodecyl sulfate 12.5%-acrylamide PAGE and Coomassie staining. The different proteins are indicated on the right, whereas molecular weight markers are indicated on the left. TssKFGE complex purification and analytical size-exclusion chromatography analysis experiments were performed at least three times and a representative result is shown.

Figure 2. Cryo-EM density map of the TssKFGE wedge complex. a. Surface representation of the composite cryo-EM density map of the TssKFGE complex. The maps corresponding to the root and stalk/wings regions were refined separately (see material and methods section). Dimensions and labelling of the various densities of the complex are shown. **b**. Schematic representation of the different densities of the TssKFGE map. **c**. Two copies of the TssK trimer crystal structure (PDB:5M30; [28]) can be fitted in the density corresponding to the "roots". **d.** Segmentation of the stalk and wing regions of the density map (shown in the same orientation than B). Two regions forming the wings (in cyan and blue) interact with a central backbone (in yellow). **e-f.** Densities corresponding to the two wing sub-regions, positioned in the same orientation (**e**) and superimposed (**f**). Each of these densities corresponds to one TssF subunit. **g.** The density corresponding to the central backbone is displayed alone. It corresponds to one TssG subunit. **h.** At lower density threshold, a density appears at the tip of the TssFG density. A TssE homology model could be

fitted into this new density with a correlation of 0.870. The scale bars correspond to 25 Å for each panel.

Figure 3. Structure of the TssKFGE complex. a. Ribbon diagram of one TssK protomer. The structure can be divided into 4 parts, from N- to C-terminus: N-terminal helix (residues 1-42, blue), shoulder domain (residues 43-185, cyan), neck domain (residues 186-312, green) and head domain (residues 313-444, red). **b.** Ribbon diagram and surface representation (transparent) of one TssK trimer viewed from the side (top panel) and from the bottom (Bottom panel). Each TssK protomer is colored in beige, dark green and brown. **c.** Ribbon diagram of the TssG protomer. The structure can be divided into 6 parts, from N- to C-terminus: N-terminal head domain (residues 8-144, blue), neck domain (residues 145-192, green), body domain (residues 146-215, 252-300 and 331-342, red) and C-terminal extension (residues 342-356, purple, foot1 (residues 216-252) and foot2 (residues 300-330), in yellow. **c.** Ribbon diagram of the TssF protomer. The structure can be divided into 6 parts, from N- to C-terminus: N-terminal antenna (residues 4-82, blue), domain 1 (residues 140-295, green), domain 2 (residues 304-416, red), domain 3 (residues 453-502, magenta) and domain 4 (residues 503-587, purple). A branching domain connects domains 1, 2 and 3 (residues 83-139, 296- 303 and 417-452, yellow). **e**. Ribbon diagram of the TssFG complex. It contains one copy of TssG (yellow) and two copies of TssF, named TssFa (cyan) and TssFb (blue). **f.** Ribbon diagram of the TssG-TssFa-TssFb trimer scaffold. TssG, TssFa and TssFb are represented in yellow, cyan and blue respectively. The triangular organization is highlighted by the red dotted-line triangle. **g.** The TssG-TssFa-TssFb trimer scaffold is decorated by TssF wing and TssG N-terminal head domains. The trimer scaffold is represented as ribbons (magenta). The decorations are represented as ribbons and transparent surfaces. The colour code is the same as in panel e. **h**. Ribbon diagram of the whole TssKFGE wedge complex. The complex made of TssG (yellow), TssFa (cyan) and TssFb (blue) interacts with two TssK trimers (TssK1 and TssK2), same color code as in B. **i.** Interaction between TssG foot domains and TssK trimers. The same color code is used as in B. and E. TssK1 and TssK2 are depicted in surface rendition and ribbon diagram respectively. TssG antenna and body are represented as ribbons while the feet are represented as surfaces.

Figure 4. Structural model of the EAEC T6SS baseplate.
a. Overall view of the interaction between the T6SS baseplate and the TssB/C sheath. The color code used to identify the various subunits is shown on the left of the panel. Left panel: surface representation of the model of the EAEC T6SS baseplate in interaction with a model EAEC sheath. While TssB/C protomers are colored in grey, the TssB/C protomers interacting with the baseplate are colored in orange. The different parts of the tail (connector ring, wedge ring, extended sheath) are indicated, as well as the putative location of the membrane complex. Middle panel: central slice of the complex seen in the left panel. The different parts of the tail, as well as the two chambers delimitated by the TssFa and TssK subunits (TssF and TssK chamber, respectively) are indicated. Right panel: surface representation of the baseplate viewed from the membrane complex (top view) and from the sheath (bottom view). **b.** Interaction surface between two wedge complexes as they are assembled in the baseplate. The color code used to identify the various subunits is shown on the left of the panel. The predicted interacting residues between two wedge complexes are colored in red. Left panel: the two bound wedge complexes as they are assembled in the baseplate. The red dotted lines represent the baseplate symmetry axis (vertical line) and the wedge complex main axis (tilted line). Right panel: the same two wedge complexes are split open to reveal the interaction surfaces between them. The main interfaces are identified. **c.** Ribbon diagram and transparent surface representation of TssFa-TssFb belonging to two adjacent wedge complexes (main

interface between wedge complexes within the T6SS baseplate). TssFa and TssFb are represented in cyan and blue respectively. Within each subunit, the regions interacting with VgrG and TssBC are colored in magenta and orange respectively. **d.** Ribbon diagram and surface representation of the wedge-TssB/C interaction. The color code used to identify the various subunits is shown on the left of the panel. Top panel: surface representation views (rotated 180° along the axis) of TssBC bound to the TssFa-TssFb-TssG-TssE complex. The ribbon diagram of the same complex is shown on bottom.

Figure 5. Comparison between the T4 bacteriophage and T6SS baseplates.
In all panels, T6SS and T4 bacteriophage representations are shown on left and right, respectively. For panels **c-e**, the color code used to identify the various subunits or domains is shown on the left of each panel. **a.** T6SS and T4 bacteriophage trimer scaffolds share the same structural organization. In the T4 baseplate, the backbone of the wedge domain is made of a heterotrimeric helical bundle and a trifurcation unit (in purple), which are made of the gp6/gp7 antennas and of the gp6/gp7 α/β domains, respectively. As seen in Fig. 3F, this organization is conserved in the T6SS baseplate (purple). In gp6, two consecutive β-sandwich domains following the N-terminal antenna are called the wing domains (green). These domains resemble the TssF D2, and BD (green). **b.** Same assembly as in **a.**, seen from the bottom. Both T6SS and T4 bacteriophage wedge complexes contain a trifurcation unit made of the α/β domains in TssFa-TssFb-TssG and gp6a-gp6b-gp7 respectively. A red triangle delineates this trifurcation unit. **c.** Top views of the T6SS and T4 bacteriophage baseplates. For the T4 bacteriophage baseplate, the inner and intermediate baseplate are separated by a red dotted line. The inter-wedge gp6a/gp6b main interface is indicated. **d.** Bottom views of the T6SS and T4 bacteriophage baseplates. The domains interacting with the sheath are indicated and colored in orange. **e.** Side views of the T6SS and T4 bacteriophage baseplates.

Figure 6. Schematic representation of the T6SS assembly pathway.
Upper panel: schematic representation of the T6SS assembly pathway, starting with the initial positioning of the membrane complex (MC, blue) (stage 1), the assembly of the wedge complexes (grey), their polymerization around the VgrG spike (pink), the recruitment of effectors (light orange skull), and the recruitment of the baseplate (BP) to the MC (stage 2); and the polymerization of the tail tube/sheath complex (TTC, salmon). A surface representation of the modeled T6SS baseplate with the extended sheath bound to the membrane complex is shown on right. OM, outer membrane; PG, peptidoglycan; IM, inner membrane. Lower panel: from left to right, the TssKFGE wedge structure (same color code as in Fig. 3); the structure of the assembled T6SS BP (same color code as in Fig. 4) and a schematic represenation of the protein-protein contacts and topology of the fully-assembled baseplate.

SUPPLEMENTARY INFORMATION

Supplementary Information includes 19 figures and 3 tables.

ACKNOWLEDGEMENTS

We thank members of the E.C., R.F. research groups, and James Sturgis for helpful discussions and support, Prof. M. Weigt and Dr. C. Feinauer for useful discussion on evolutionary covariance, Tru Huynh and Gérald Canet for assistance with the computer

clusters at the Institut Pasteur and IECB respectively, Artemis Kosta (Plateforme de microscopie, IMM, Marseille, France) for providing access to the IMM EM facility, and for checking the quality of samples by negative stain EM, Armel Bezault for support at the cryo-EM facility at IECB. Thierry Doan et Leon Espinosa for their helpful support with the fluorescence microscopy device and analysis. We thank Prof. M. Nilges (Institut Pasteur) for access to the computer cluster of the ERC project "BayCellS". This work was supported by the CNRS, the Aix-Marseille Université, the Institut Pasteur, and the INSERM and by grants from the Agence Nationale de la Recherche (ANR-14-CE14-0006 to E.C.; ANR-17-CE11-0039 to E.C., ANR-11-EQPX-008 to JCR). Work of Y.C. was supported by an "Ecole Doctorale" PhD fellowship from the "Fondation pour la Recherche Médicale" (FRM-ECO20160736014). RF and CR were supported by IDEX Bordeaux through a "chaire d'excellence" to RF. We acknowledge the European Synchrotron Radiation Facility for provision of beamtime on CM01 and thank Gregory Effantin and Eaazhisai Kandiah for assistance. F.A. thanks the French Institute of Bioinformatics (IFB; ANR-11-INBS-0013) for financial support.

AUTHORS CONTRIBUTION

E.D., R.F. and E.C. designed research, assembled results, and wrote the paper with input from all authors. E.D. and Y.C. performed biochemical experiments and initial negative stain EM observations. C.R. performed sample preparation for cryo-EM, *de novo* reconstruction, and structure refinement. C.R. and R.F. analyzed the cryo-EM data and the final structure. C.M. and M.R. performed MS experiments. Y.C. performed all *in vivo* experiments. R.P. and G.B. performed method development, and structural modeling; B.B. and F.A. performed structural modeling. J.C.R. performed design and analysis of MS experiments.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Methods

A key resource table (supplementary table 1) provides details about the reagents, strains, and software used in this study.

Strains, media and chemicals

The strains, plasmids and oligonucleotides used in this study are listed in Supplementary Table 1. The *E. coli* K-12 DH5α strain was used for cloning procedures; *E. coli* K-12 BL21(DE3) strain was used for protein expression and purification; *E. coli* K-12 W3110 bearing the pUA66-*rrnB* vector (Kan^R and GFP^+, [1]) was used as recipient for antibacterial competition assays. Strains were routinely grown in lysogeny broth (LB) rich medium or in Sci-1-inducing medium (SIM; M9 minimal medium, glycerol 0.2%, vitamin B1 1 $\mu g.mL^{-1}$, casaminoacids 100 $mg.mL^{-1}$, LB 10%, supplemented or not with bactoagar 1.5%) [2] with shaking at 37°C. Plasmids were maintained by the addition of streptomycin (100 $\mu g.mL^{-1}$), kanamycin (50 $\mu g.mL^{-1}$), chloramphenicol (30 $\mu g.mL^{-1}$) or ampicillin (100 $\mu g.mL^{-1}$). Expression of genes from pCDF, pRSF and pETDuet vectors was induced with 1 mM of isopropyl-β-D-thio-galactopyrannoside (IPTG, Eurobio) for 16h at 16 ˚C.

Strains construction

Gene deletion into the enteroaggregative *E. coli* 17-2 $tssK_{sfgfp}$ strain [3] was achieved by using a modified one-step inactivation procedure [4] as previously described [5] using plasmid pKOBEG [6].

Briefly, a kanamycin cassette was amplified from plasmid pKD4 using oligonucleotide pairs carrying 5' 50-nucleotide extensions homologous to regions adjacent to the gene to be deleted. After electroporation of 600 ng of column-purified PCR product, kanamycin-resistant clones were selected and verified by colony-PCR. The kanamycin cassette, inserted at the gene locus on the bacterial chromosome, was then excised using plasmid pCP20, leaving an FRT scar [4]. Gene deletions were confirmed by colony-PCR and sequencing. All the mutations have been previously trans-complemented for the T6SS-dependent bacterial competition or Hcp secretion phenotype by plasmid expressing the WT version of the T6SS genes TssFG [7] and TssK[8].

Plasmid construction

PCRs were performed using the Phusion DNA polymerase (Thermo Scientific). Restriction enzymes were purchased from New England Biolabs and used according to the manufacturer's instructions. Custom oligonucleotides were synthesized by Sigma Aldrich and are listed in Supplementary Table 1. Enteroaggregative *E. coli* 17-2 chromosomal DNA was used as a template for all PCRs. Construction of pCDF-TssK^{H}-^{S}F-G^{F} has been previously described [9]. Plasmids pCDF-TssK^{H}-^{S}F-G^{Flag}-E^{HA} and pETDuet-VgrG^{Ha} were engineered by restriction cloning. Briefly, the sequences encoding the full-length *tssE* and *vgrG* were PCR-amplified using primers 5-pRSF-E^{HA} (ATAAAGCTTAAGGAGATATACATATGCCGCGTCCTTCCCTTTATGAAATTCTCTATGGC) and 3-pRSF-E^{HA} (ATAGCGGCCGCTCA*AGCGTAATCTGGAACATCGTATGGGTA*CGTCTGCACGTAGCGCTGCT GTTTCAGATGGC), and 5-pETDuet-VgrG^{HA} (ATAGGATCCAAGGAGATATACATATGAATCTCACTGACTCCCTGCAAAATGTTTTATCCG G) and 3-pETDuet-VgrG^{HA} (TATAAGCTTTCA*AGCGTAATCTGGAACATCGTATGGGTA*TTCTGTTTCTCCATGAATTTTTAC CTTCCCAAACTC), respectively. Primers introduced a C-terminal HA epitope tag extension (italicized in the primer sequences), and HindIII/NotI and BamHI/HindIII restriction sites (underlined in the primer sequences) respectively. All other plasmids were constructed by restriction-free cloning [10]. Briefly, the gene of interest was amplified with oligonucleotides carrying 5' extensions annealing to the target vector. The product of the first PCR was then used as oligonucleotide for a second PCR using the target vector as template. All constructs have been verified by DNA sequencing (Eurofins Genomics).

Interbacterial competition assay

The antibacterial growth competition assay was performed as previously described [11]. Wild-type *E. coli* K-12 strain W3110 bearing the pUA66-*rrnB* plasmid (conferring kanamycin resistance and constitutive GFP fluorescence (*gfp* gene under the control of the ribosomal *rrnB* promoter, [12] was used as recipient. Attacker and recipient cells were grown for 16 h in LB medium, diluted in SIM to allow maximal expression of the *sci-1* gene cluster [2]. Once the culture reached A_{600nm}~ 0.8, cells were harvested and normalized to A_{600nm} = 0.5 in SIM. Attacker and recipient cells were mixed to a 4:1 ratio and 15-µl drops of the mixture were spotted in triplicate onto a pre-warmed dry SIM agar plate supplemented or not with arabinose 0.5 mg.mL^{-1}. After incubation for 4 h at 37°C, the bacterial spots were resuspended in LB and bacterial suspensions were normalized to A_{600nm} = 0.5. For the enumeration of viable prey cells, bacterial suspensions were serially diluted and spotted onto kanamycin LB plates. The assays were performed from at least three independent cultures, with technical triplicates and a representative technical triplicate is shown.

TssKFGE complex production and purification

The tags were rationally positioned at specific locations in the TssKFGE complex: Introduction of the tags (i) still permits protein-protein interaction as assayed by bacterial two hybrid (BTH) and (ii) allows trans-complementation of the T6SS-dependent "interbacterial competition" phenotype in a mutant deleted for a specific gene with a plasmid encoding a tagged version of this same gene. The pCDF–TssK^{H}-^{S}F-G^{F}-^{Ha}E plasmid was transformed into the *E. coli* BL21(DE3) expression strain. Cells were grown at 37°C in lysogeny broth (LB) to A_{600nm} ~ 0.6 and the expression of the *tssKFGE* genes

was induced with 1.0 mM IPTG for 16 h at 16 °C. Cell pellets were resuspended in ice-cold 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA supplemented with DNase I (100 mg.mL^{-1}), lysozyme (100 mg.mL^{-1}), $MgCl_2$ (10mM) and EDTA-free protease inhibitor (Roche) to an A_{600nm} of 125. Cells were broken using an Emulsiflex-C5 (Avestin) and clarified by ultracentrifugation at 20,000×*g* for 30 min. The supernatant was loaded onto a 5-mL HisTrap HP (GE Healthcare) column equilibrated in affinity buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl) supplemented with 20 mM imidazole. The column was then washed using the affinity buffer supplemented with 50 mM imidazole and the TssKFGE complex was eluted in the same buffer supplemented with 250 mM imidazole. Peak fractions were pooled and loaded onto a Superose 6 10/300 column (GE Healthcare) equilibrated in 50 mM HEPES pH 7.5, 150 mM NaCl. The complex eluted as a single monodisperse peak and the sample was used for EM sample preparation.

Protein production and purification for interaction studies

Plasmids expressing the genes combination of interest were co-transformed into *E. coli* BL21(DE3) and cells were treated as described before. For His-tag affinity, the supernatant was loaded, washed and eluted as above. For strep-tag affinity, the supernatant was loaded onto a 5-mL StrepTrap HP column (GE Healthcare), washed with affinity buffer and eluted in affinity buffer supplemented with 2.5 mM desthiobiotin (IBA Technologies). The lysate, flow through, wash and elution fractions were collected, resuspended in Laemmli loading buffer supplemented with 300 mM 2-Mercaptoethanol, heated for 10 min at 96°C prior to analyses by SDS-PAGE and immunoblotting.

SDS–PAGE, protein transfer, immunostaining and antibodies

SDS–PAGE was performed on Bio-Rad Mini-PROTEAN® systems using standard protocols. For immunostaining, proteins were transferred onto 0.2-µm nitrocellulose membranes (Amersham Protran). Immunoblots were probed with primary antibodies and goat secondary antibodies coupled to alkaline phosphatase, and developed in alkaline buffer in presence of 5-bromo-4-chloro-3-indolylphosphate and nitro-blue tetrazolium. The anti-HA (HA-7 clone, Sigma Aldrich), anti-Flag (M2 clone, Sigma Aldrich), anti-StrepII (Sigma Aldrich), anti VSV-G (Sigma Aldrich) and anti-5His (Sigma Aldrich) monoclonal antibodies, and mouse secondary antibodies (Millipore) were purchased as indicated.

Native polyacrylamide gel electrophoresis

After overnight cultures in LB, the enteroaggregative *E. coli* EAEC strain 17-2-*tssK*$_{sfgfp}$ and its mutant variants were diluted 1/100 in 500 mL of SIM and grown at 37°C to an A_{600nm} ~ 1,2. Cells were harvested, resuspended in ice-cold 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA to an A_{600nm} = 120 and broken using an Emulsiflex-C5 (Avestin). After clarification by ultracentrifugation at 20,000×*g* for 30 min, lysates were loaded on a native 4-16% gel (Mini-PROTEAN® TGX, Bio-Rad). After migration, proteins and protein complexes were transferred onto a nitrocellulose membrane and immunoblotted as described above.

Mass spectrometry

Purified TssKFG and TssKFGE protein complexes were first buffer exchanged with 500 mM ammonium acetate by size exclusion chromatography on Superdex 200 increase (3.2 / 300) using a ÄKTAmicro System (GE healthcare) at isocratic flow of 50 µL.min^{-1}. Samples were then nano-electrosprayed using a TriVersa NanoMate (Advion Biosciences, Ithaca, USA) coupled to a Synapt G2-Si mass spectrometer (Waters Corporation, Manchester, UK). The instrument was calibrated from 1,000 m/z to 12,000 m/z with CsI (50 mg/mL^{-1}) with an accuracy of 6 ppm. Native mass measurements were recorded between 2,000 and 20,000 m/z with sensitivity mode activated. The following settings were chosen: sampling cone 150 V, source offset 45 V, source temperature 40°C, trap gas flow 5 mL.min^{-1}, helium cell gas flow 180 mL.min^{-1}. Sub-complexes were obtained with the same parameters and additional collisional activation up to 60 NCE (Normalized Collisional Energy). Data were accumulated several minutes, averaged and smoothed with the Mass Lynx smoothing Algorithm (30 cycles, 30 amu large channel). Theoretical masses were calculated with the algorithm embedded within MassLynx using the protein sequence of the constructs. Measured masses were

obtained averaging the mass calculated for the most intense charge states of the complex with a minimum of 3 charge states.

Fluorescence microscopy, image treatment and analyses

Fluorescence microscopy experiments were performed as described [8,13]. Briefly, cells were grown overnight in LB medium and diluted to A_{600nm} ~ 0.04 in SIM. Exponentially growing cells (A_{600nm} ~ 0.8–1) were harvested, washed in phosphate-buffered saline buffer (PBS), resuspended in PBS to A_{600nm} ~ 50, spotted on a 1.5% agarose pad and covered with a cover slip. For domain interference, the *E. coli* 17-2 *tssK-sfgfp* or *tssB-mCherry* strains expressing TssK or TssG domains were cultured as described above, except that 0.05% (w/v) arabinose was added in the culture once reached A_{600nm} ~ 0,6 for 30 min. Fluorescence and phase contrast micrographs were captured using AxioImager M2 microscope (Zeiss) equipped with an OrcaR2 digital camera (Hamamatsu). Fluorescence images were acquired with a minimal exposure time to reduce bleaching and phototoxicity effects, typically 500 ms for TssK-sfGFP and 200 ms for TssB-mCherry. Noise and background were reduced using the 'Subtract Background' (20 pixels Rolling Ball) and Band plugins of imageJ (Image J, National Institutes of Health). The sfGFP foci were automatically detected using the microbeJ plugin (http://www.microbej.com/index.html). Box plots representing the number of detected foci for each strain were made using microbeJ. The number of sheath per cells was measured manually. Microscopy analyses were performed at least three times, each in technical triplicate, and a representative experiment is shown.

Cryo-EM grids preparation and data acquisition

For cryo-EM analyses, the buffer of the purified TssKFGE complex was exchanged for 50 mM HEPES pH 7.5, 150 mM NaCl to a final protein concentration of 0.2 mg.mL^{-1}. 3.5 µL of the protein solution was deposited on Lacey grids and vitrified using a Vitrobot (Thermo Fisher, Waltham, MA, USA) (parameters: blotting 4 s, temperature 4°C, humidity 100 %). Micrographs (Supplementary Fig. 11a) were recorded at a specimen temperature of 85 K in a Titan Krios electron microscope (Thermo Fisher, Waltham, MA, USA) at 300 kV and a nominal magnification of 130,000 on a K2 summit direct electron detector mounted on a Bioquantum LS/967 energy filter (Gatan, Pleasanton, California) in counting mode with a pixel size of 1.1 Å, at an electron flux of about 12.35 e-/px/s. Dose-fractionated movie frames (30 in total) were acquired for 4 s with 0.13 s exposure time per frame. The total electron dose was ~45 e-/Å^2 (1.5 e-/Å^2/frame). The defocus range chosen for the automatic collect was 0.7 to 2 µm, which resulted in an actual range between 0.4 to 3 µm.

Cryo-EM image processing

750 micrographs were processed (Supplementary Fig. 10a). Subframes were divided into 5×5 patches and corrected with MotionCor2, with dose weighting (1.5 e-/Å^2/frame) to dampen the high-resolution signal in later frames [14]. CTF parameters were estimated by gCTF [15]. Particles on micrographs were picked manually in box size of 450 pixels and classified into 2D class averages. Selected classes were used as references for autopicking in RELION 2.1 [16]. The total number of initial extracted particles (167,825) was reduced to 52,069 by subsequent rounds of 2D classifications (Supplementary Fig. 10b) and an initial model (Supplementary Fig. 10c) of what appears to be 2 TssKFGE full complexes (12×TssK, 4×TssF, 2×TssG, 2×TssE) was generated on cryoSPARC 0.6 [17]. The selected particles from the 2D classifications were converted back to RELION 2.1 using the script csparc2star.py [18] and then subjected to an additional round of 2D and 3D classification (Supplementary Fig. 10c), with the initial model low-pass filtered to 60Å. A final cleaner dataset of 32,504 particles was selected for further processing. The 40-Å low-pass filtered 3D class was then used as an initial model for 3D refinement with a solvent mask corresponding to most well defined half of the larger TssKFGE complex (6×TssK, 2×TssF, 1×TssG, 1×TssE). No symmetry was applied during any of the 3D classifications and refinements. To obtain the density map of the single repeating unit of the assembly of the two complexes (hereafter called TssKFGE), the complex with the least defined wing was subtracted, and a soft mask was applied to the remaining region for 3D auto-refinement. The final resolution was 4.6 Å, calculated with masked post-process by refining two half-maps independently,

according to the "gold standard" FSC 0.143 criterion. The B factor applied of -136 gave a local resolution range between 3.9 and 18 Å (Supplementary Fig. 11a-c).
To obtain the best density map for the root-like domains that correspond to TssK, the regions corresponding to the wings were subtracted and a soft mask was applied to the TssK trimers during autorefinement, producing a map with a 4.3-Å overall resolution. The calculated B factor was -111 and it gave a local map resolution of 3.8-33 Å (Supplementary Fig. 11d-f, j-k).
Masking was not sufficient to improve the density map of the flexible regions. A re-centering of the particles in the wings region was thus performed using the REP algorithm [19] and the box size was reduced to 200 pixels. A masked 3D auto-refinement to exclude the root-like domains was performed to obtain a resolution of 4.7 Å (4.3-8 Å local with B factor of -202; Supplementary Fig.11g-i).
All of the densities obtained were subjected to Autosharpen [20] in the Phenix software package. All the models were built on autosharpened densities. Subsequent molecular graphics and analyses were performed using UCSF Chimera [21].

Evolutionary Covariance Analysis.

Residue–residue contacts can be predicted based on sequence information alone through the evolutionary covariance analysis [22] Essentially, the prediction of residue–residue contacts is linked to strong evolutionary constraints, such as the presence of functionally important structures, and is measured by the covariance of contacting residues. Evolutionary constraints can be detected at a sequence level by aligning thousands of homologous protein sequences. Statistical probability models can separate direct from indirect residue–residue couplings, increasing the signal-to-noise ratio in the predicted contact map. Therefore, contacts with the strongest signal, indicated by the highest global statistical scores, are most likely to represent the true residue interactions in a protein. In this work we use two software packages that uses distinct statistical probability models: EVcomplex [23] and RaptorX [24] for inter-molecular and intra-molecular contact prediction, respectively. EVcomplex computes co-evolution between proteins, pairing up protein sequences and assuming proximity of the two interacting partners on the genome, with the goal of reducing incorrect pairings. The paired sequences are concatenated and statistical co-evolution analysis is performed using EVcoupling [25]. The RaptorX algorithm predicts intra-molecular contacts by integrating evolutionary coupling, pairwise potentials, and sequence conservation information through an ultra-deep neural network.

TssKFGE model fitting and de novo tracing

Two copies of the trimeric TssK unit (PDB: 5M30; [9] were docked into the EM density map of the TssKFGE complex and fit as rigid bodies in Chimera [21]. Missing regions were manually built using Coot [26]. The final model was refined by multiple rounds of manual refinement in Coot [26], Rosetta refine [27] and the real-space refine function of Phenix [28].
Owing to the lack of structural information for TssF and TssG and the limited resolution of the density map in that region, we devised an iterative semi-automatic protocol that employed several sources of information and pieces of software (Supplementary Fig. 12). The sources of information included the EM density map, the stoichiometry and the symmetry of the TssFG subcomplex, the position of bulky residues in the sequence, a homology model of TssF based on the T4 bacteriophage gp6 (generated using i-TASSER, consensus secondary structure obtained using PSIPRED [29], Rosetta [27], Phyre2 [30], DeepCNF [25], and i-TASSER [31], and intra-molecular contact predictions for TssF and TssG using the RaptorX contact prediction tool [31] (Supplementary Fig. 12b**)**. Contact maps are binary two-dimensional matrices that represent the proximity between all residue pairs for a given protein [27]. The RaptorX algorithm predicts contacts by integrating evolutionary coupling, pairwise potentials, and sequence conservation information through an ultra-deep neural network. Predicted contact maps were obtained by filtering RaptorX contacts with a score higher than 0.4. Model contact maps, which are compared with the predicted ones, were obtained from a structure calculating the pairwise distances between Cα atoms and considering as a contact any distance below 12 Å. We tested the quality the predicted contact maps by calculating the accuracy on a known structure. Using the enteroaggregative *E. coli* TssK structure as model [9]; PDB: 5M30) (Supplementary Fig. 12c), the ratio between predicted contacts consistent with the structure and the total contacts within the structure gave an accuracy of 0.97 (Supplementary Table S3).

The protocol described below iterates between structural refinement and sequence-structure registration based on contact prediction. It is organized into four steps (Supplementary Fig. 12a):

Step 1. Initial segmentation and manual tracing. The TssFG map was segmented using Segger v1.9.4 in UCSF Chimera [21] and the density of one half of the wing superimposed with the other half to identify repeating patterns due to the presence of two TssF monomers in the complex. The segmented map identified densities corresponding to TssG and the two TssF subunits. An initial Cα tracing of the TssFG complex was based on the three density segments. The sequence registering was guided by sequence position of bulky amino-acids and consensus prediction of secondary structures. The model building was aided by the use of the Coot-trimmings script [32] and sharpening of the map was modulated by varying the resolution limit.

Step 2. Sequence Registration using predicted contact maps. The model was validated and registered using residue contact prediction [33]. The contact map of the obtained model was computed and aligned to the predicted contact-map using the MapAlign software [34]. Using the resulting contact-map alignment we identified regions of the model that agree and disagree with the predicted contact-map. The boundaries of these regions were used to divide the structure into several fragments. The connectivity and the sense (N- to C-terminus orientation) of the fragments were globally optimized using dynamic programming and the MapAlign scoring function.

Step 3. De novo modeling of inconsistent regions. The model generated in the previous step was again validated comparing its contact-map with the predicted contact-map. Regions in disagreement (TssG-D1, residues 1-144; TssG-D2 body, residues 180-300; and TssF-D1, residues 495-585) were modeled de novo using ARIA with predicted contacts as distance restraints [35]. Inter-residue contacts for the TssG-D1, TssG-D2 and TssF-D1 domains were predicted from co-evolution analysis using RaptorX [31] and the top scoring contacts were converted to distance restraints. Secondary structure predictions from DeepCNF [25,36] were converted to canonical dihedral angle restraints for residues predicted to be in α-helical and β-strand conformation. Additionally, hydrogen bond restraints were generated between residues *i* and *i*+4 in regions predicted as α-helical (same face of the helix). Atomic models were calculated by successive rounds of restrained molecular dynamics simulated annealing with CNS [37] using the iterative ARIA approach [35,38]. At each iteration, 100 conformations were produced using inter-residue distance, dihedral angle and hydrogen bond restraints and clustered from the coordinates of Cα atoms. The 15 best conformations of the lowest-energy cluster were then used to refine the list of predicted contacts on the basis of their structural consistency. A conformational database potential term was also used in the energy function during simulated annealing [39].

Step 4. Refinement. All the models of different domains independently generated in step 2 and 3 were merged together in a single chain. The coordinates of the obtained single-chain model were modified manually using Coot [26] and refined with repeated rounds of Coot, Rosetta refine [40] and the Phenix real-space refine function [28]. Steps 2-4 were repeated until a model maximally consistent with the EM map (Supplementary Fig 12e) and the predicted contact-maps was obtained.

For TssK and TssFGE the EMRinger [41] scores were 1.74 and 0.23 and the Molprobity [42] scores were 1.9 and 2.36 respectively. Final accuracies of the TssF and TssG contact maps generated from the structures were 0.85 and 0.87, respectively (Supplementary Table 2). The models built in TssK and TssFGE were then fit as rigid bodies in the TssKFGE structure with EMRinger and Molprobity scores of 1.62 and 1.94 respectively (Supplementary Table 2).

FSC curves were calculated between the model and the map using Phenix after real space refinement. To assess the presence of overfitting, as described before [43], the FSC computed between the model and the autosharpened map (FSC-sum) was compared to the FSC calculated between the "shaken" model (applying a random pertubation of 0.5Å to the atomic coordinates, with the module pdbtools of Phenix, Adams et al., 2010) refined against the first half map (FSC-work) and the resulting model refined with the second half map (FSC-free, Supplementary Figure 11c,f,i). The overlapping between the FSC-free and FSC-work curves demonstrates the absence of overfitting. Interaction surfaces were analysed using the PISA software [45]. For illustration purposes, TssFGE secondary structures were predicted using Cablam [46].

Data and software availability

The cryo-EM structure of the full complex TssKFG , TssK and TssFGE have been deposited in the EMDB under ID codes EMD-0008, 0010, 0009. The TssKFG, TssK and TssFGE model have been deposited in the PDB under ID codes PDB 6GIY, 6GJ3, 6GJ1. Raw cryo-EM data are available on request.

References:

1. Zaslaver, A. *et al.* A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat. Methods* **3,** 623–628 (2006).
2. Brunet, Y. R., Bernard, C. S., Gavioli, M., Lloubès, R. & Cascales, E. An Epigenetic Switch Involving Overlapping Fur and DNA Methylation Optimizes Expression of a Type VI Secretion Gene Cluster. *PLoS Genet.* **7,** e1002205 (2011).
3. Logger, L., Aschtgen, M. S., Guérin, M., Cascales, E. & Durand, E. Molecular Dissection of the Interface between the Type VI Secretion TssM Cytoplasmic Domain and the TssG Baseplate Component. *J. Mol. Biol.* **428,** 4424–4437 (2016).
4. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97,** 6640–6645 (2000).
5. Aschtgen, M.-S., Bernard, C. S., De Bentzmann, S., Lloubes, R. & Cascales, E. SciN Is an Outer Membrane Lipoprotein Required for Type VI Secretion in Enteroaggregative Escherichia coli. *J. Bacteriol.* **190,** 7523–7531 (2008).
6. Chaveroche, M. K., Ghigo, J. M. & d'Enfert, C. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. *Nucleic Acids Res.* **28,** E97 (2000).
7. Brunet, Y. R., Zoued, A., Boyer, F., Douzi, B. & Cascales, E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization. *PLOS Genet.* **11,** e1005545 (2015).
8. Zoued, A. *et al.* TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. *J. Biol. Chem.* **288,** 27031–27041 (2013).
9. Nguyen, V. S. *et al.* Type VI secretion TssK baseplate protein exhibits structural similarity with phage receptor-binding proteins and evolved to bind the membrane complex. *Nat. Microbiol.* **2,** 17103 (2017).
10. Van Den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67,** 67–74 (2006).
11. Flaugnatti, N. *et al.* A phospholipase A 1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol. Microbiol.* **99,** 1099–1118 (2016).
12. Gueguen, E. & Cascales, E. Promoter Swapping Unveils the Role of the Citrobacter rodentium CTS1 Type VI Secretion System in Interbacterial Competition. *Appl. Environ. Microbiol.* **79,** 32–38 (2013).
13. Brunet, Y. R., Espinosa, L., Harchouni, S., Mignot, T. & Cascales, E. Imaging Type VI Secretion-Mediated Bacterial Killing. *Cell Rep.* **3,** 36–41 (2013).
14. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14,** 331–332 (2017).
15. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193,** 1–12 (2016).
16. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).
17. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14,** 290–296

(2017).
18. Asarnow, D. pyem. (2016).
19. Sanchez, R. M. Recentering and subboxing of particles (REP). (2017).
20. Terwilliger, T. C., Sobolev, O., Afonine, P. V & Adams, P. D. Automated map sharpening by maximization of detail and connectivity. *bioRxiv* (2018).
21. Pettersen, E. F. *et al.* UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* **25,** 1605–1612 (2004).
22. Simkovic, F., Ovchinnikov, S., Baker, D. & Rigden, D. J. Applications of contact predictions to structural biology. *IUCrJ* **4,** 291–300 (2017).
23. Hopf, T. A. *et al.* Sequence co-evolution gives 3D contacts and structures of protein complexes. *Elife* **3,** (2014).
24. Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLOS Comput. Biol.* **13,** e1005324 (2017).
25. Marks, D. S. *et al.* Protein 3D structure computed from evolutionary sequence variation. *PLoS One* **6,** (2011).
26. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).
27. Leaver-Fay, A. *et al.* Rosetta3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* **487,** 545–574 (2011).
28. Afonine, P. V *et al.* Real-space refinement in Phenix for cryo-EM and crystallography. *bioRxiv* (2018).
29. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **16,** 404–405 (2000).
30. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10,** 845–858 (2015).
31. Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLOS Comput. Biol.* **13,** e1005324 (2017).
32. Clarke, O. B. Coot Trimmings. (2017).
33. Bouvier, G., Bardiaux, B. & Nilges, M. Automatic Building of Protein Atomic Models from Cryo-EM Maps. *Biophys. J.* **114,** 190a–191a (2018).
34. Ovchinnikov, S. *et al.* Protein structure determination using metagenome sequence data. *Science (80-.).* **355,** 294–298 (2017).
35. Rieping, W., Bardiaux, B., Bernard, A., Malliavin, T. E. & Nilges, M. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23,** 381–382 (2007).
36. Wang, S., Peng, J., Ma, J. & Xu, J. Protein Secondary Structure Prediction Using Deep Convolutional Neural Fields. *Sci. Rep.* **6,** (2016).
37. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2,** 2728–2733 (2007).
38. Mareuil, F., Malliavin, T. E., Nilges, M. & Bardiaux, B. Improved reliability, accuracy and quality in automated NMR structure calculation with ARIA. *J. Biomol. NMR* **62,** 425–438 (2015).
39. Kuszewski, J., Gronenborn, A. M. & Clore, G. M. Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. *Protein Sci.* **5,** 1067–1080 (2008).
40. Wang, R. Y. R. *et al.* Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* **5,** (2016).

41. Barad, B. A. *et al.* EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12,** 943–946 (2015).
42. Chen, V. B. *et al.* MolProbity : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).
43. Brown, A. *et al.* Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 136–153 (2015).
44. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).
45. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
46. Williams, C. Using C-alpha geometry to describe protein secondary structure and motif. *Dr. Diss.* (2016).

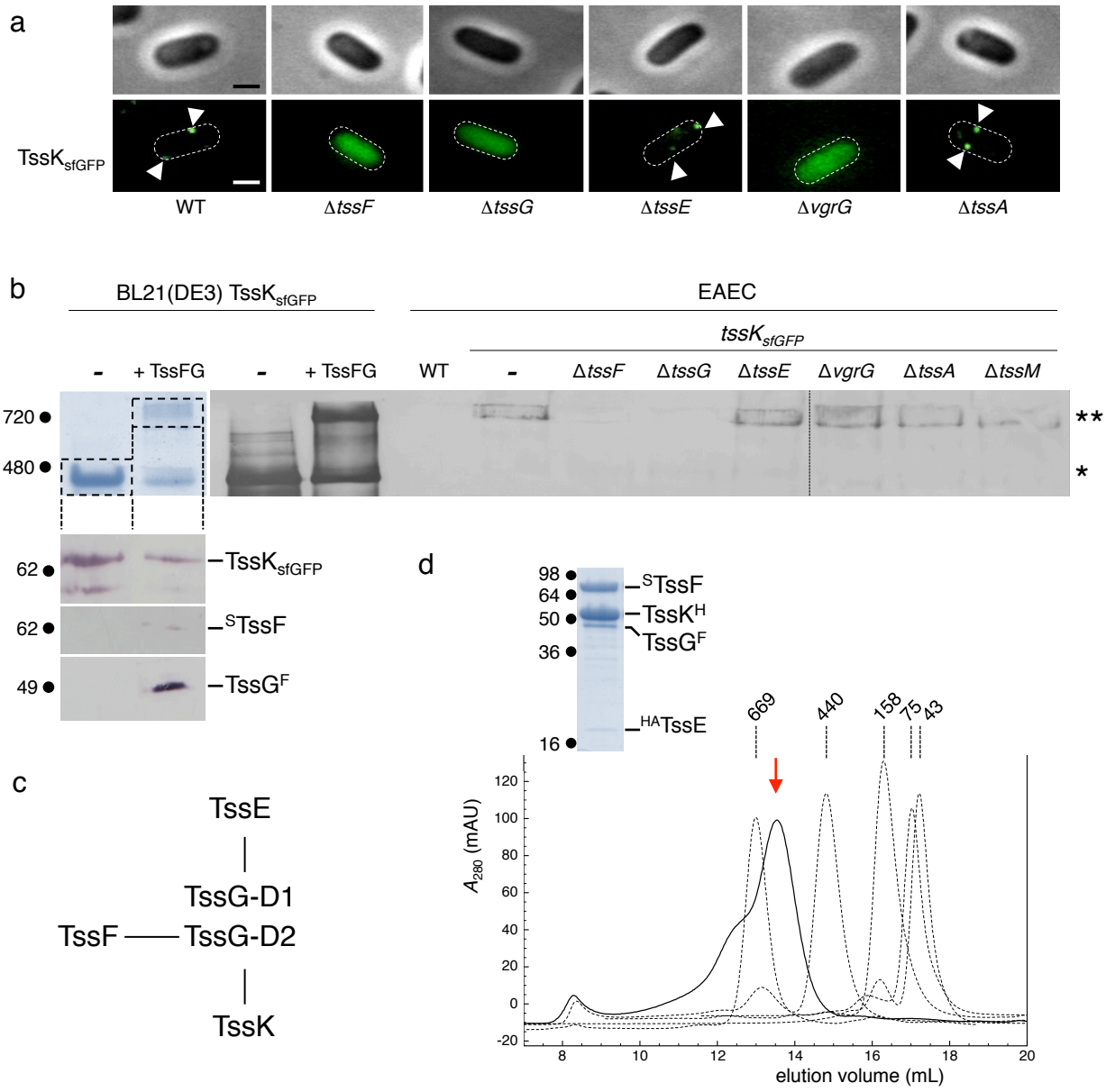

a
TssK$_{sfGFP}$
WT
ΔtssF
ΔtssG
ΔtssE
ΔvgrG
ΔtssA
b
BL21(DE3) TssK$_{sfGFP}$
EAEC
tssK$_{sfGFP}$
- + TssFG - + TssFG
WT - ΔtssF ΔtssG ΔtssE ΔvgrG ΔtssA ΔtssM
720
480
62
62
49
TssK$_{sfGFP}$
^{S}TssF
TssG^{F}
**
*
c
TssE
TssG-D1
TssF
TssG-D2
TssK
d
98
64
50
36
16
^{S}TssF
TssK^{H}
TssG^{F}
^{HA}TssE
669
440
158
75
43
A_{280} (mAU)
120
100
80
60
40
20
0
-20
8
10
12
14
16
18
20
elution volume (mL)

Figure 1

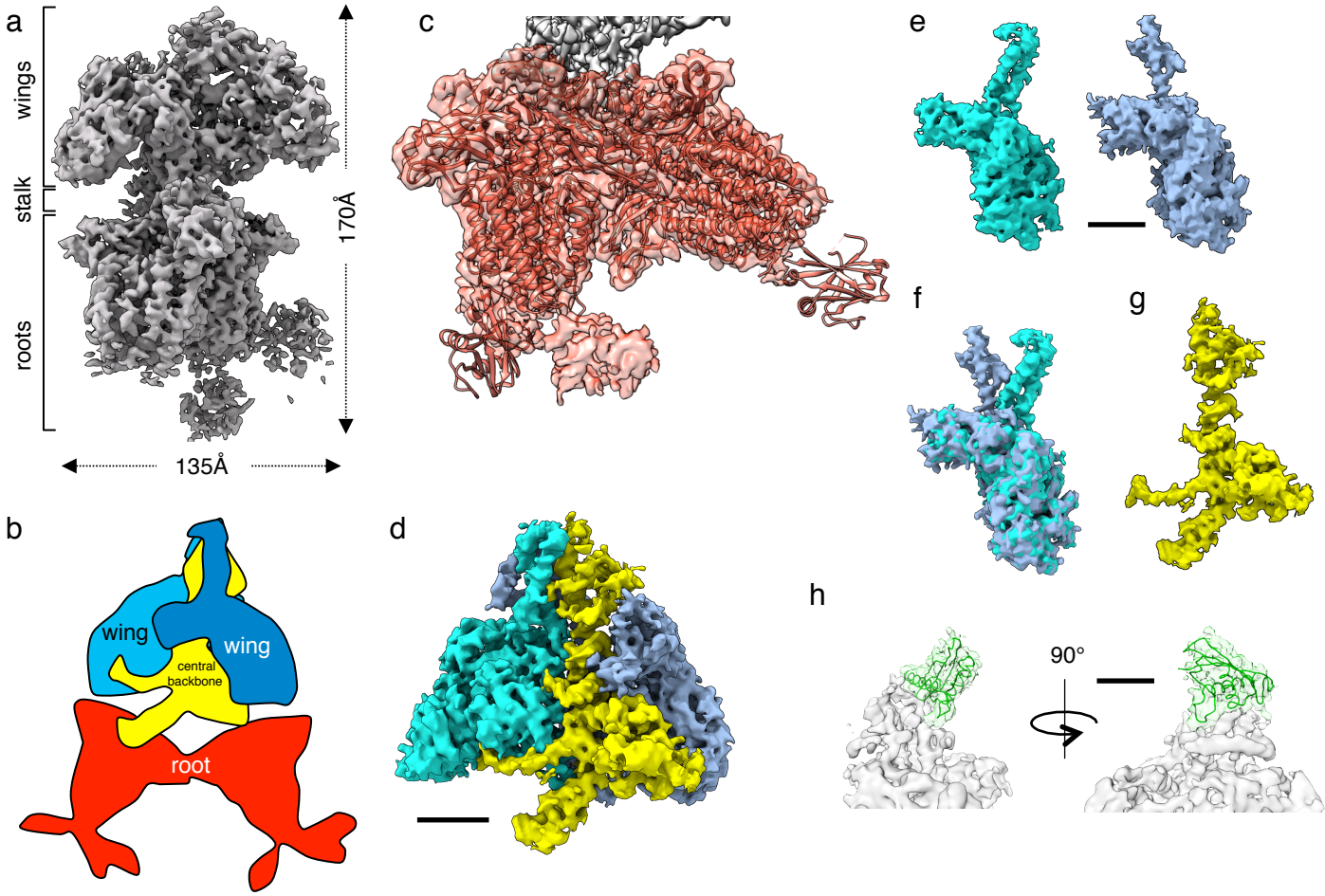

a
wings
stalk
roots
170Å
135Å
b
wing
wing
central backbone
root
c
d
e
f
g
h
90°

Figure 2

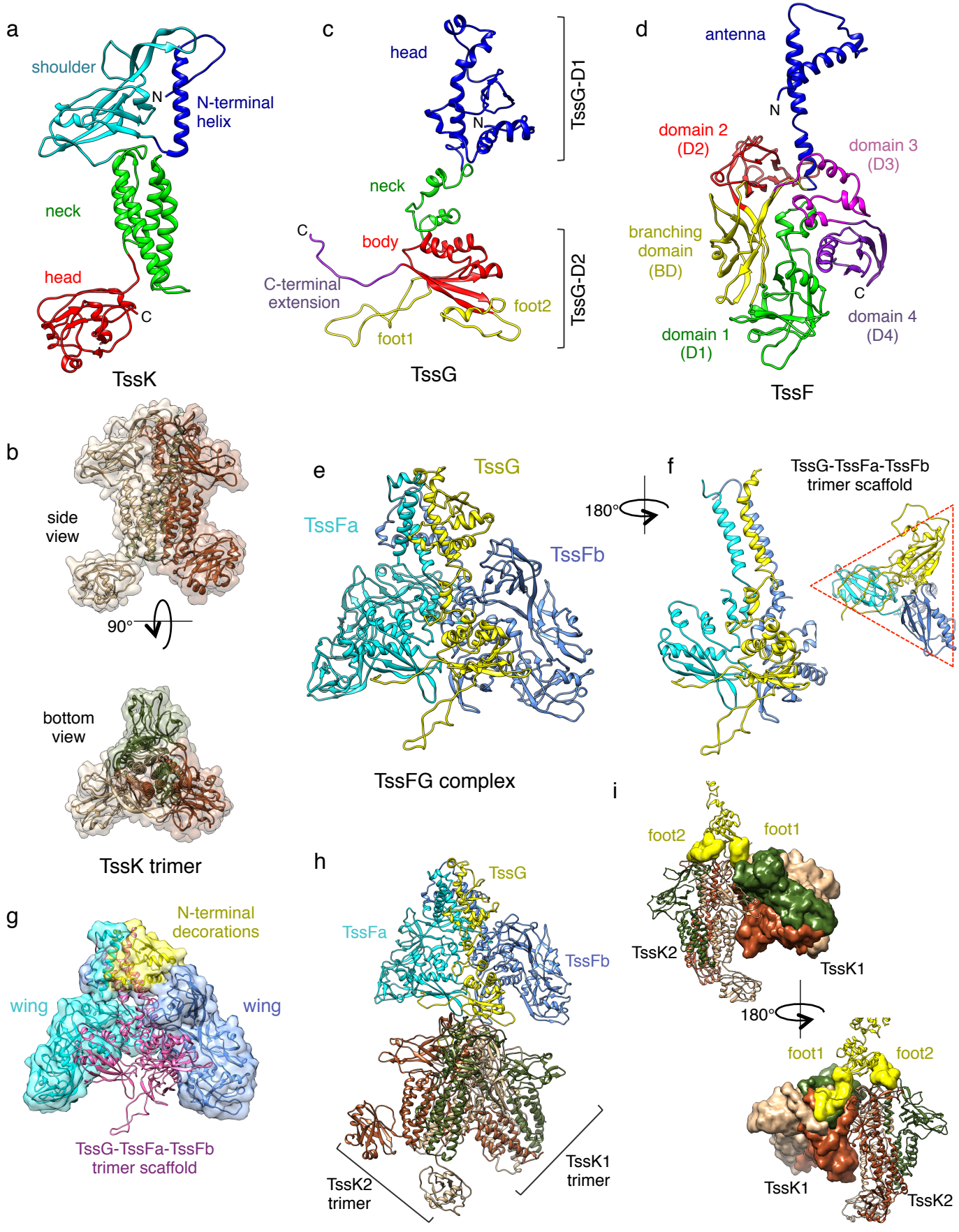

a
shoulder
N
N-terminal
helix
neck
head
C
TssK
b
side
view
90°
bottom
view
TssK trimer
c
head
N
TssG-D1
neck
C
body
C-terminal
extension
foot2
foot1
TssG-D2
TssG
d
antenna
N
domain 2
(D2)
domain 3
(D3)
branching
domain
(BD)
C
domain 4
(D4)
domain 1
(D1)
TssF
e
TssG
TssFa
TssFb
TssFG complex
f
180°
TssG-TssFa-TssFb
trimer scaffold
g
N-terminal
decorations
wing
wing
TssG-TssFa-TssFb
trimer scaffold
h
TssG
TssFa
TssFb
TssK2
trimer
TssK1
trimer
i
foot2
foot1
TssK2
TssK1
180°
foot1
foot2
TssK1
TssK2

Figure 3

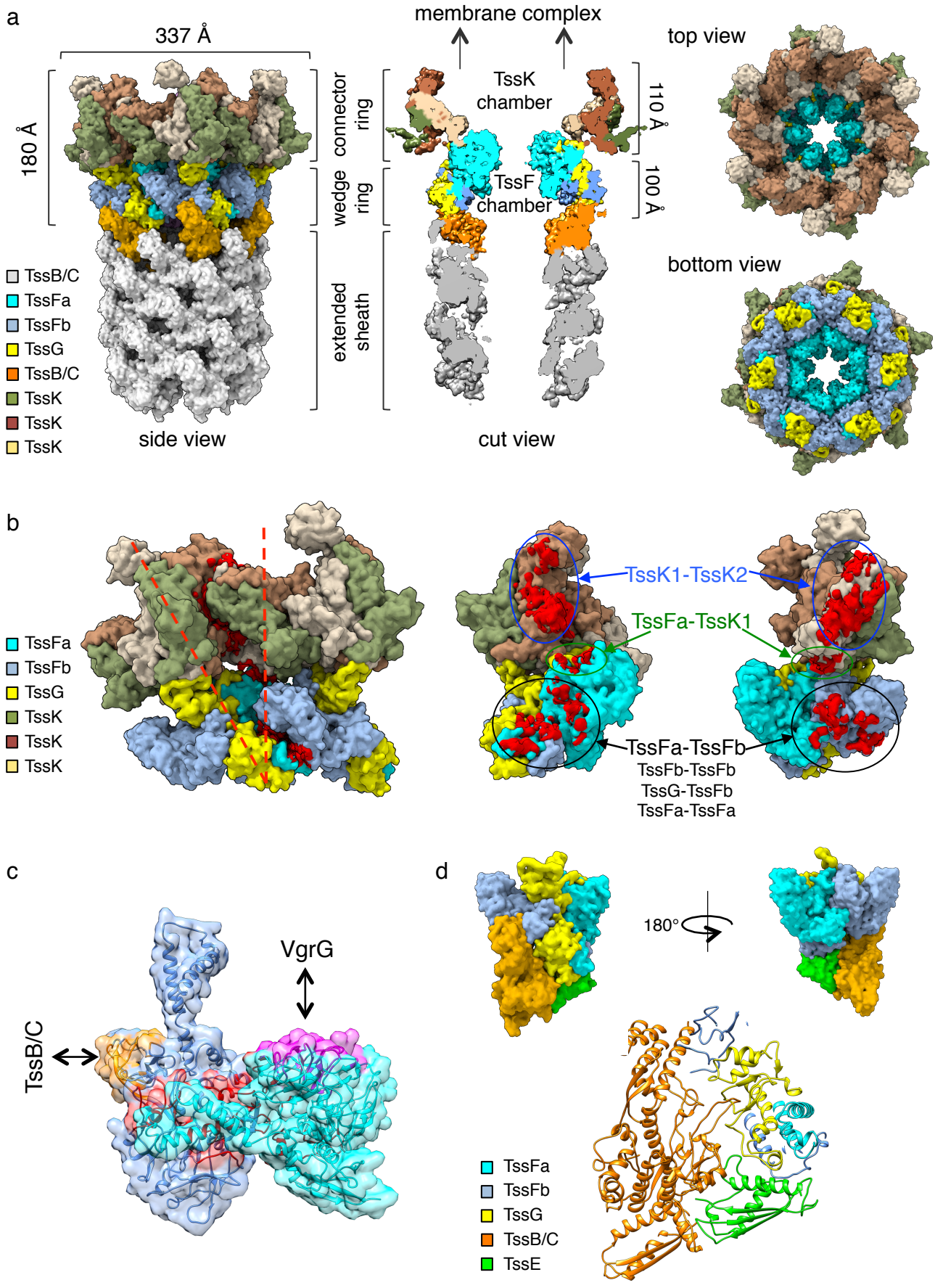

a
337 Å
180 Å
connector ring
wedge ring
extended sheath
membrane complex
TssK chamber
TssF chamber
110 Å
100 Å
top view
bottom view
TssB/C
TssFa
TssFb
TssG
TssB/C
TssK
TssK
TssK
side view
cut view
b
TssFa
TssFb
TssG
TssK
TssK
TssK
TssK1-TssK2
TssFa-TssK1
TssFa-TssFb
TssFb-TssFb
TssG-TssFb
TssFa-TssFa
c
VgrG
TssB/C
d
180°
TssFa
TssFb
TssG
TssB/C
TssE

Figure 4

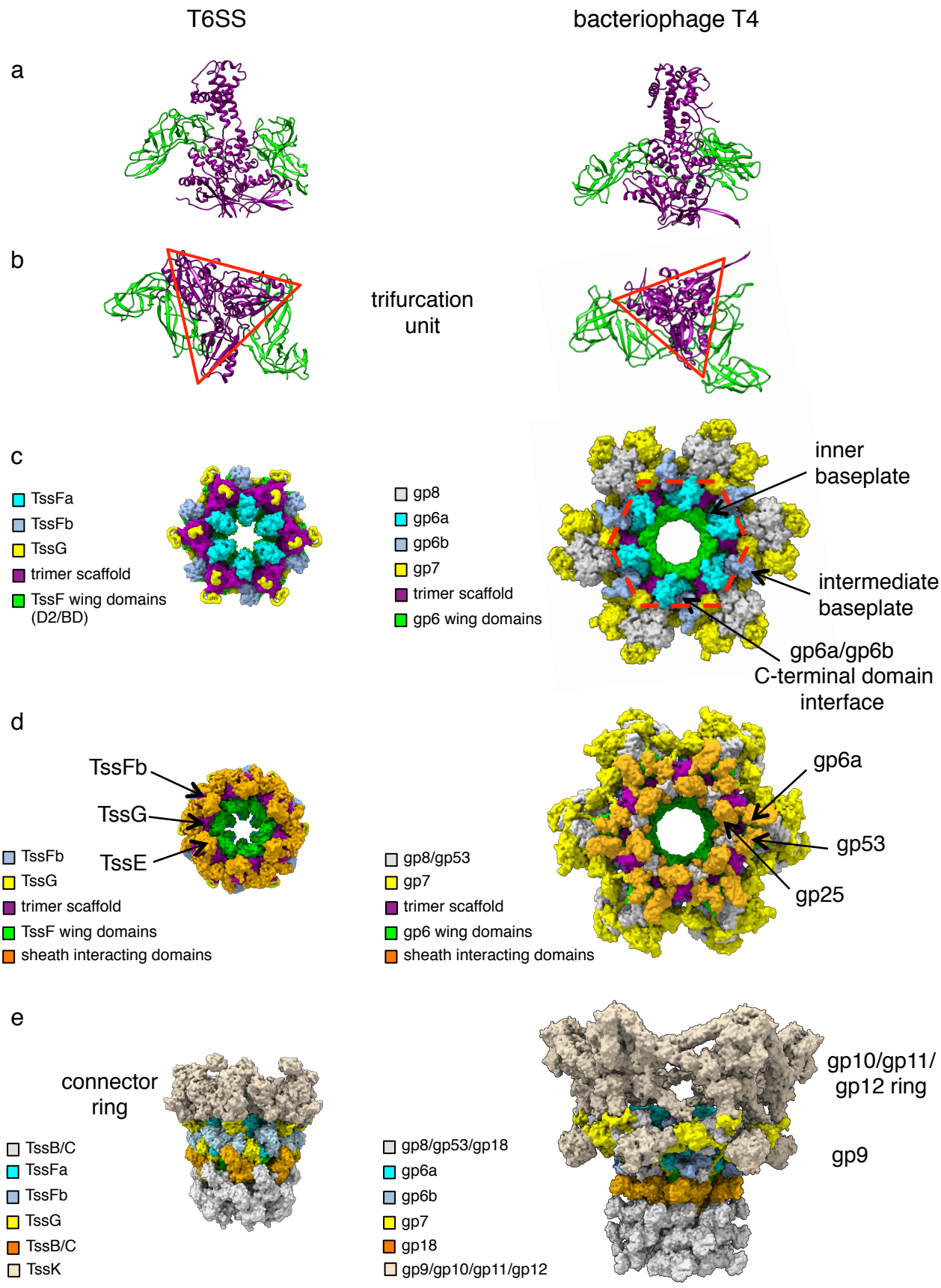

T6SS
bacteriophage T4
a
b
trifurcation unit
c
TssFa
TssFb
TssG
trimer scaffold
TssF wing domains (D2/BD)
gp8
gp6a
gp6b
gp7
trimer scaffold
gp6 wing domains
inner baseplate
intermediate baseplate
gp6a/gp6b C-terminal domain interface
d
TssFb
TssG
TssE
TssFb
TssG
trimer scaffold
TssF wing domains
sheath interacting domains
gp8/gp53
gp7
trimer scaffold
gp6 wing domains
sheath interacting domains
gp6a
gp53
gp25
e
connector ring
TssB/C
TssFa
TssFb
TssG
TssB/C
TssK
gp8/gp53/gp18
gp6a
gp6b
gp7
gp18
gp9/gp10/gp11/gp12
gp10/gp11/ gp12 ring
gp9

Figure 5

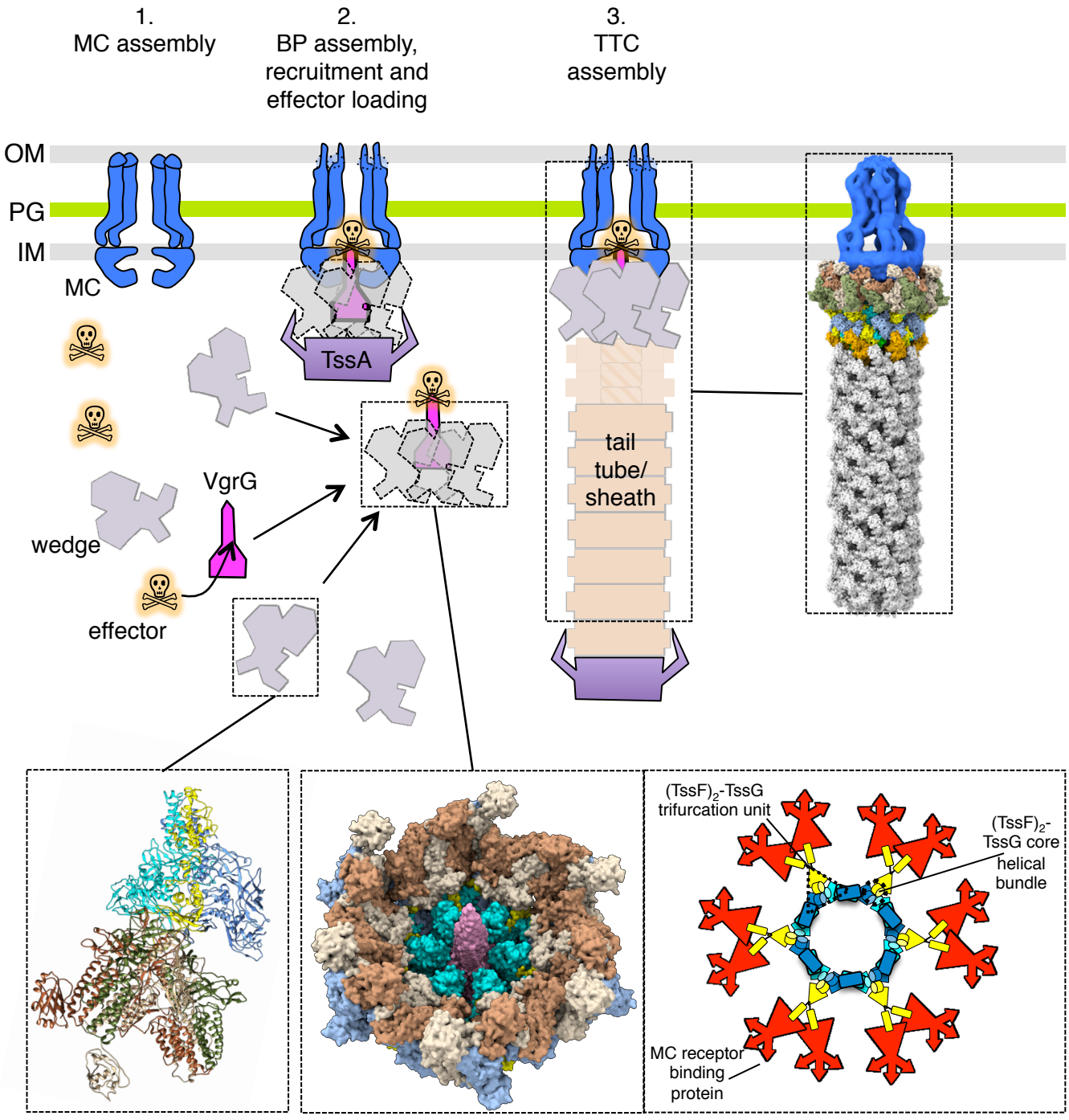

1.
MC assembly
2.
BP assembly,
recruitment and
effector loading
3.
TTC
assembly
OM
PG
IM
MC
TssA
VgrG
wedge
effector
tail
tube/
sheath
(TssF)2-TssG
trifurcation unit
(TssF)2-
TssG core
helical
bundle
MC receptor
binding
protein

Figure 6

Supplementary Results

Native mass spectrometry.

Native mass spectrometry (Native MS) was used to measure the mass of the intact protein complex from which the stoichiometry can be deduced [1]. The molecular mass of each individual subunit was firstly checked in denaturing conditions. A molecular mass of 499,490 ± 200 Da was obtained for the complex indicating that the TssKFGE complex comprises six TssK, two TssF, one TssG and one TssE subunits ($TssK_6F_2G_1E_1$; theoretical mass: 498,905 Da) (**Supplementary Fig. 3b**). The difference from the measured and theoretical masses arises from the presence of numerous salt adducts, which is a common phenomenon in native MS. A similar stoichiometry was obtained for the purified TssKFG complex ($TssK_6F_2G_1$; measured mass: 481,916 Da ± 200 Da, theoretical mass: 481,520 Da; **Supplementary Fig. 3c**). For TssF, TssG and TssE, this stoichiometry is in agreement with previous studies, indicating the conserved nature of the T6SS wedge among various species. Interestingly, native MS analyses confirmed the peripheral and weaker association of TssE in the TssKFGE complex, as TssE dissociates first when the TssKFGE complex is submitted to collisional activation (**Supplementary Fig. 3d).** Only 3 TssK monomers per wedge complex were found in *Serratia marcescens* and uropathogenic *E. coli* [2–4]. These discrepancies could be explained by the different approaches used to purify or assess the stoichiometry of the complex, or by differences in T6SS baseplate composition in these various bacterial species.

Segmentation of the CryoEM map

Our native MS and biochemical analyses of the purified TssKFGE complex unambiguously showed that TssE is present with a stoichiometry of one TssE per complex. At lower density thresholds, the cryo-EM map reveals a density that could be attributed to TssE at the tip of the TssFG. In agreement with this hypothesis, a model of the EAEC TssE generated based on the gp25 crystal structure (PDB: 5IW9; [2]) fits into the identified density (**Fig. 2h**).

Dominant-negative approach by "small domain interference" (SDI)

First, to validate this approach, we performed SDI with TssK sub-domains. It was previously shown that the TssK N-terminal shoulder domain is anchored into the baseplate whereas the TssK C-terminal head domain binds the membrane complex [5]. We reasoned that overproduction of TssK shoulder ($TssK_S$) or TssK shoulder and neck ($TssK_{SN}$) sub-domains should interfere with T6SS wedges assembly, and by lacking the head domain will prevent proper association with the membrane complex. *In vivo*, overproduction of $TssK_{SN}$ and, to a lower extent, $TssK_S$ in wild-type cells impacts the assembly of T6SS wedges, blocks T6SS sheath dynamic and prevents T6SS activity (**Supplementary Fig. 5a-c**). The difference in the negative effect observed between $TssK_{SN}$ and $TssK_S$ could be explained by the stabilization of the TssK trimer by the helix packing of the neck domains [5] that results in a better incorporation of $TssK_{SN}$ compared to $TssK_S$ into the TssKFGE complex *in vitro* (**Supplementary Fig. 5d**). We then hypothesized that overproduction of the TssG domain which mediates the recruitment of TssK, should have a dramatic negative effect on T6SS function. Our competition experiments showed that only the overproduction of TssG-D2 has a negative effect on the T6SS-dependent killing of *E. coli* target cells (**Supplementary Fig. 5e**). This result confirms that TssG-D2 has a central role for T6SS wedge assembly. To provide further evidence for the position of TssG-D2 at the interface between TssF and the two TssK trimers, we used the EVcomplex program (Hopf et al., 2014) to predict inter-molecular interactions between TssG and TssK using evolutionary covariance analysis (see Methods section). First, the algorithm confirmed our previous findings, showing that the TssG-D2 domain is the major interface with the TssK shoulder domain. Second, the algorithm predicted a number of potential $TssK_S$-TssG-D2 interfacial residue pairs with high scores (**Supplementary Fig. 6a**). The corresponding TssG-D2 residues, Pro-240 and Leu-255 were substituted, and these TssG-D2 variants (P240A and L255A and P240A-L255A) were assayed by SDI in interbacterial competition experiments. Although these TssG-D2 variants were produced at levels comparable to the wild-type TssG-D2 (see insets in **Supplementary Fig. 6b**), the TssG-D2 P240A and L255A mutants were drastically affected in their ability to inhibit T6SS-function. Combining the P240A and L255A mutations decreases even further the T6SS-inhibitory effect of TssG-D2 overproduction (**Supplementary Fig. 6b**).

The TssK-TssFG interface

To rationalize the results of the Alanine mutation analysis discussed above, we mapped the positions of TssG Pro-240 and TssG Leu-255 on the TssKFG structure. Only Pro-240 is proximal to the TssKG interface, being directly in contact to TssK Leu-14. Thus, its mutation is compatible with a TssKG interface destabilization. On the other hand, TssG Leu-255 is far from the interface, and its mutation probably destabilized the fold of the C-terminal region. When compared with the structure, the Alanine mutational study above has revealed some key residues important for the stability of the TssK-G interaction stability, although the molecular mechanism remains to be addressed.

Supplementary Discussion

The T6SS baseplate - membrane complex connection

As mentioned in recent publications, there is a symmetry mismatch between the 5-fold symmetric TssJLM membrane complex and the 6-fold symmetric baseplate [4–6]. In this configuration, the 36 copies of TssK (6 per individual wedge) could contact the 10 copies of TssLM. However, our cryo-EM structure shows that the $TssK_H$ domains within each TssK trimer are in different orientations. We thus suggest that the flexibility of the $TssK_H$ domains is key to adapt to the symmetry mismatch and that only a subset of these $TssK_H$ domains engages with the membrane complex. This symmetry mismatch is likely to be important for T6SS function. It may maintain the system under a metastable conformation, allowing a higher turnover rate of ejection of the Hcp tube. It may also accommodate the large conformational changes expected between the baseplate and the membrane complex during sheath contraction. Finally, it may play a critical role in the controlled dissociation of the baseplate after firing, as suggested for the dissociation of V-ATPases [7].

Comparison between T6SS and bacteriophage baseplates.

Analogy between the T6SS baseplate and the "simple contractile baseplate" from the Mu phage – The architecture and stability of the TssKFGE wedge complex is reminiscent to the Mu phage baseplate wedge [8]. Although sequence identity between the T6SS and Mu phage protein is very low, their respective overall structural organization suggests that the T6SS and Mu phage wedge complexes are highly similar. The Mu phage wedge is formed by the proteins Mup46, Mup47, and Mup48. Mup46 is a gp25 homolog, like TssE. Mup47 is homologous to the gpJ wedge protein of *E. coli* phage P2, like TssF [9]. Mup48 is homologous to the gpI wedge protein of *E. coli* phage P2, like TssG [9]. Mup47 and Mup48 form a stable complex, a behavior comparable with the T6SS TssF-TssG complex [9]. However, the relative stoichiometry of both wedge complexes differs significantly: 2:2:2 for Mup46/47/48 and 6/2/1/1 for TssK/F/G/E. The whole Mu phage baseplate is formed by the attachment of dimers of the Mup46/47/48 wedge (BW1, BW2 and BW3, according to the standard nomenclature previously proposed [8]) on the Mup43/44 hub complex (BH1 and BH2). However, the predicted T6SS baseplate does not seem to contain such BH1/BH2 hub complex. The T6SS proteins share many structural characteristics with "simple" contractile baseplates, even though TssK has no equivalent in Mu phage proteins. We have previously demonstrated that TssK is homologous to lactococcal siphophage receptor binding proteins (RBP) proteins [5] and we have shown in this work that TssF and TssG are partly a T4 phage gp6 and gp7 homologs. Therefore, the T6SS wedge complex and baseplate seem to be an evolutionary "patchwork", rooted in T4 bacteriophage, siphophage and Mu phage.

TssF and TssG are gp6 and gp7 counterparts in the T4 phage baseplate – Previous structural studies on the T4 bacteriophage have reported the moderate to high-resolution structures of the whole baseplate [2,10–12]. Despite no sequence conservation, it was proposed that TssF and TssG could be the structural counterparts of gp6 and gp7 respectively [2,9]. Gp6 and gp7 are subunits of the T4 phage inner baseplate in which they interact tightly to form the T4 phage wedge complex. The TssG and gp7 structures are not readily superimposable. However, TssG and the central part of gp7 (residues 640-900) display the same domain organization with an N-terminal antenna followed by a helical neck domain and an α/β body (**Supplementary Fig. 9a**). The rest of the gp7 structure is specific to the T4 baseplate. By contrast, based on secondary structure predictions [9], TssF shares much higher structural homology with the N-terminal region of gp6 (**Supplementary Fig. 9b**). The domain from residue 1 to 410 of gp6 is composed of an N-terminal antenna followed by two consecutive β-sandwich domains (called the wing domains) and an α/β domain (called the trifurcation domain). These domains resemble the TssF antenna, D2, BD and D4

domains, respectively (**Supplementary Fig. 9b**). The C-terminal region of gp6 structure is not present in TssF and is not found in any protein of the T6SS baseplate. Strikingly, not only TssF and TssG adopt a similar structure to portions of gp6 and gp7, but the structural organization of the T4 $(gp6)_2$–gp7 heterotrimeric unit is conserved in the T6SS wedge complex (**Fig. 5a-b**). In the T4 baseplate, the backbone of the wedge domain is made of a heterotrimeric helical bundle and a trifurcation unit, which are made of the gp6/gp7 antennas and of the gp6/gp7 α/β domains, respectively. This trimeric scaffold, which tightly connects gp7 and the two gp6, is very similar to that of the TssFG complex (**Fig. 5a-b**). Although the organization is identical, it is interesting to note that the T4 phage trifurcation unit comprises three equivalent domains made of two helices and one three-strands β sheet while the TssFG trimeric scaffold contains two helices and one four-strands β sheet (**Fig. 5a-b**).

Inter-wedge contacts are different in the T4 and T6SS baseplates – The multimerization of T6SS wedges into the hexagonal baseplate diverge from the bacteriophage counterpart. In the T6SS baseplate, we showed that the inter-wedge contacts are mostly mediated by TssFa-TssFb interactions that spread in different domains of both protomers **(Fig. 4b)**. In the T4 baseplate, the inter-wedge contacts are also mediated by gp6a/gp6b interactions. However, these interactions are restricted to the dimerization of a T4-specific gp6 C-terminal domain (residues 412-659) (**Fig. 5c)**. Within each T4 wedge complex, these C-terminal domains are directly connected to the trifurcation unit. Therefore, a continuous ring of alternating trifurcation units and dimers of gp6 C-terminal domains stabilize the T4 inner baseplate [2,12,13]. The T4 phage baseplate also comprises an additional subunit, absent from the T6SS baseplate, gp8, that interacts with the N- and C- terminal domains of gp7 and thus reinforces the overall structure.

Interaction of the T4 and T6SS baseplates with the spike – Both T4 and T6SS baseplates contain a constriction at the center of the wedge ring that is delineated by gp6b and TssFa respectively. In the T6SS baseplate, TssFa wing domains exclusively cover the cavity that receives the VgrG gp27-like hub domain (**Fig. 5c-d, Supplementary Fig. 9c**). In the T4 baseplate, the gp6a wing domain mostly contributes to the central constriction but the gp27 hub position is stabilized in the cavity by a T4-specific component, gp53 (**Fig. 5c, Supplementary Fig. 9c**).

Interaction of the T4 and T6SS baseplates with the contractile sheath – The structure of the sheath component of the T4 phage gp18 is known (PDB: 3FOH; [14]). Based on the cryo-EM structure of the native T4 phage tail in the extended conformation (EMD-1126; [15]) determined at 12 Å, it is possible to fit gp18 in the sheath density and to reconstitute the extended sheath structure. The contacts between the sheath and the baseplate subunits can be inferred after docking to the T4 phage baseplate structure [2]. These contacts are in agreement with previously published biochemical and genetics data [12,16,17]. Overall the structures of the T4 and T6SS sheath building blocks (gp18 and TssBC complex respectively) are different (**Supplementary Fig. 9d-e)**. However, they can be split in similar domains: an N-terminal antenna domain, a central globular domain and an α/β domain that protrudes from the central domain (**Supplementary Fig. 9d-e)**. Strikingly, TssE and gp25 interact similarly with the sheath α/β domain and share the same fold as these TssBC and gp18 domains (**Supplementary Fig. 9d-e)**. The baseplate-sheath connections in phage T4 and T6SS have conserved features but are stabilized by specific contacts. The N-terminal domains of gp18 and of TssBC interact with one of the wing domains of gp6b or TssFb respectively (**Supplementary Fig. 9d-e)**. In the T4 phage, the central domain of gp18 interacts with a T4-specific protein, gp53 (**Supplementary Fig. 9e)**. In the T6SS, this interaction is replaced by the interaction with the N-terminal head domain of TssG, which is T6SS-specific (**Supplementary Fig. 9d)**.

Specialization of the apical part of the T4 and T6SS baseplates – The T4 phage and the T6SS baseplates are specialized molecular machines that interact with different targets. The T4 phage baseplate interacts with the bacterial surface prior to infection, while the T6SS baseplate interacts with the T6SS membrane complex. To achieve this specialization, the T4 and T6SS baseplates contain dedicated adaptor proteins recruited by gp7 and its TssG counterpart (**Fig. 5c)**. In T4, the N- and C- terminal domains of gp7 interact with gp9 and gp10, which in turn assemble the short tail fiber proteins gp12 (**Fig. 5c-f)** [17–20]. In T6SS, the TssG foot domains recruit two TssK trimers, which then interact with the T6SS membrane complex to anchor the T6SS contractile device to the bacterial cell envelope (**Fig. 5c-f)**.

Assembly mechanism and stability of the T4 and T6SS baseplates – In the T4 bacteriophage the wedge complex appears to be transient or less stable preventing its purification. By contrast, the T6SS wedge complex is much more stable and prone to purification, as shown by the isolation of TssKFG or TssKFGE complexes in EAEC, *S. marcescens* and uropathogenic *E. coli* (this work; [2,3]) This higher stability of a preformed T6SS baseplate intermediate may reflect an adaptation to the secretion process. One hypothesis is that a subset of toxin effectors needs to be loaded onto VgrG [21–23], hence delaying the polymerization of the wedge around the VgrG hub. By contrast, the fully assembled T4 phage baseplate is much more stable than the T6SS baseplate, since we did not succeed to purify the T6SS hexagonal baseplate. At the molecular level, TssF-TssF inter-domain interactions appear to be less tightly connected in the fully assembled T6SS baseplate, involving a limited contact surface area, as compared to bacteriophage T4 gp6-gp6 contacts. Consequently, the connection between individual T6SS wedges seems to be weaker, in agreement with the lower stability of the T6SS baseplate compared to that of bacteriophage T4. In addition, this could be explained by the presence of the T4-specific gp53 subunit that has been shown to promote polymerization of the wedges to form a hubless baseplate [19]. This difference may have functional implications. In phage T4, the hubless baseplate needs to remain stably associated during the transport of the phage genome. By contrast, after firing, the T6SS baseplate needs to dissociate, releasing individual wedges that will presumably assemble a new baseplate at the membrane complex. This reversible association would allow a recycling of the individual T6SS wedges.

Supplementary Table 1. Reagents and resources

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-VSVG, clone P5D4	Sigma-Aldrich	Cat# A5977; RRID:AB_439710
Mouse monoclonal anti-FLAG, clone M2	Sigma-Aldrich	Cat# F3165; RRID:AB_259529
Mouse monoclonal anti-HA, clone HA-7	Sigma-Aldrich	Cat# H3663; RRID:AB_262051
Mouse monoclonal anti-StrepII, clone GT661	Iba	Cat# 2-1507-001 RRID:AB_513133
Mouse monoclonal anti-polyHistidine, clone HIS-1	Sigma-Aldrich	Cat# H1029; RRID:AB_260015
Alkaline phosphatase-conjugated goat anti-mouse secondary antibody	Millipore	Cat# AP503A; RRID:AB_805353
Bacterial Strains		
DH5α	New England Biolabs	Cat# C2987I
W3110	Laboratory collection	N/A
BL21 (DE3	New England Biolabs	Cat# C2527I
Enteroaggregative *E. coli* strain 17-2	Laboratory collection	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP	[25]	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*tssF*	This paper	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*tssG*	This paper	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*tssE*	This paper	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*vgrG*	This paper	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*tssA*	This paper	N/A
Enteroaggregative *E. coli* strain 17-2- *tssK*-sfGFP-Δ*tssM*	[25]	N/A
Enteroaggregative *E. coli* strain 17-2 TssB-mcherry	[26]	N/A
Enteroaggregative *E. coli* strain 17-2 TssB-mcherry-Δ*tssM*	[9]	N/A
Chemicals, Peptides, and Recombinant Proteins		
HisTrap high performance (5mL)	GE Healthcare	Cat# GE17-5248-01
StrepTrap high performance (5mL)	GE Healthcare	Cat# GE28-9075-47
Superose 6 increase 10/300 GL	GE Healthcare	Cat# GE29-0915-96
Superdex 200 increase 3.2/300	GE Healthcare	Cat# GE28-9909-46
cOmplete™ ULTRA Tablets, EDTA-free, glass vials Protease Inhibitor Cocktail	Roche	Cat# 05892953001
DNase I	Roche	Cat# 10104159001
rLysozyme solution	Merck	Cat# 71110
Ammonium Acetate	Sigma-Aldrich	Cat# A1542
Imidazole	Sigma-Aldrich	Cat# 56750
Hepes	Sigma-Aldrich	Cat# H3375
NativeMark unstained protein standard	Invitrogen	Cat# LC0725
NativePAGE Sample Buffer (4X)	Invitrogen	Cat# BN2003
Mini-PROTEAN TGX Precast gel (4-15%)	Bio-Rad	Cat# 456-1084
4X Laemmli Sample Buffer	Bio-Rad	Cat# 161-0747
2-Mercaptoethanol	Sigma-Aldrich	Cat# M3148
Acrylamide/Bis-Acrylamide 37.5:1, 40%	Biosolve	Cat# 001422335BS
Nitro blue tetrazolium (NBT)	Apollo Scientific	Cat# BIMB1019
5-Bromo-4-chloro-3-indolyl phosphate disodium (BCIP)	Apollo Scientific	Cat# BIMB1018
Recombinant DNA		

pKD4	[27]	Addgene Plasmid #45605
pKOBEG	[28]	N/A
pBAD33	[29]	N/A
pBAD33-TssK$_S^{VSVG}$	[5]	N/A
pBAD33-TssK$_{SN}^{VSVG}$	[5]	N/A
pBAD33-TssK$_c^{VSVG}$	[5]	N/A
pBAD33-TssG-D1^{Iag}	This paper	N/A
pBAD33-TssG-D2^{Flag}	This paper	N/A
pCDF-Duet1	Novagen	Novagen Plasmid #71340-3
pRSF-Duet1	Novagen	Novagen Plasmid #71341-3
pET-Duet1	Novagen	Novagen Plasmid #71146-3
pACYC-Duet1	Novagen	Novagen Plasmid #71147-3
pCDF-TssK^{H}-^{S}F-G^{Flag}-^{HA}E	This paper	N/A
pCDF-TssK^{H}-^{S}F-G^{Flag}	[5]	N/A
pCDF-^{S}F-G^{Flag}-^{HA}E	This paper	N/A
pCDF-K^{His}-^{S}F-^{HA}E	This paper	N/A
pCDF-^{S}F-G^{Flag}	This paper	N/A
pCDF-K^{H}	This paper	N/A
pRSF-K^{H}	This paper	N/A
pCDF-K^{Ha}	This paper	N/A
pCDF-^{S}F	This paper	N/A
pCDF-G^{Flag}	This paper	N/A
pCDF-G$^{(Δ1-32)Flag}$	This paper	N/A
pRSF-G$^{(1-144)Flag}$	This paper	N/A
pRSF-G$^{(168-end)Flag}$	This paper	N/A
pCDF-^{HA}E	This paper	N/A
pRSF-^{HA}E	This paper	N/A
pRSF-^{H}E	This paper	N/A
pRSF-^{H}E$^{(Δ1-24)}$	This paper	N/A
pETDuet-VgrG^{HA}	This paper	N/A
pETDuet-^{S}VgrG	This paper	N/A
pACYC-^{Flag}TssA	This paper	N/A

Software and Algorithms		
ARIA	[30]	http://aria.pasteur.fr
Coot	[31]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Coot trimmings	[32]	https://github.com/olibclarke/coot-trimmings
Cryosparc 0.6	[33]	https://cryosparc.com/
Csparc2star.py	[34]	https://github.com/asarnow/pyem/blob/master/csparc2star.py
DeepCNF	[35]	http://raptorx.uchicago.edu/StructurePropertyPred/predict/
EMRinger	[36]	http://emringer.com/
EVcomplex program	[37]	https://evcomplex.hms.harvard.edu
gCTF	[38]	https://www.mrc-lmb.cam.ac.uk/kzhang/Gctf/
Gremlin	[39]	http://gremlin.bakerlab.org
i-TASSER	[24]	https://zhanglab.ccmb.med.umich.edu/I-TASSER/
ImageJ	[40]	https://imagej.net/ImageJ
MapAlign	[41]	https://github.com/sokrypton/map_align
MicrobeJ	[42]	http://www.microbej.com/index.html
MolProbity	[43]	http://molprobity.biochem.duke.edu/

MotionCor2	[44]	http://msg.ucsf.edu/em/software/motioncor2.html
Phenix	[45]	https://www.phenix-online.org/
Phenix real-space refine	[46]	https://www.phenix-online.org/
Phyre2	[47]	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index
PISA	[48]	http://www.ebi.ac.uk/pdbe/pisa/
RaptorX	[24]	http://raptorx.uchicago.edu/ContactMap/
RELION 2.1	[49]	http://www2.mrc-lmb.cam.ac.uk/relion/index.php/Download_%26_install
REP	[50]	https://github.com/rkms86/REP
Rosetta	[51]	https://www.rosettacommons.org/software
UCSF Chimera	[52]	https://www.cgl.ucsf.edu/chimera

Protein accession numbers		
TssE	WP_061358700.1	type VI secretion system baseplate subunit TssE [Escherichia coli]
TssF	WP_000342463.1	type VI secretion system baseplate subunit TssF [Escherichia coli]
TssG	WP_000553781.1	type VI secretion system baseplate subunit TssG [Escherichia coli]
TssK	WP_000708638.1	type VI secretion system baseplate subunit TssK [Escherichia coli]

Other		
Titan Krios	Thermo scientific	https://www.fei.com/krios-g3i/
K2 Summit camera	Gatan, Pleasanton, USA	http://www.gatan.com
Vitrobot	FEI - Thermo Fisher	https://www.fei.com/products/vitrobot/
ÄKTAmicro	GE Healthcare Life Sciences	https://www.gelifesciences.com/en/at/shop/aktamicro-p-02952
Synapt G2Si HDMS	Waters	www.waters.com
Triversa Nanomate	Advion	https://advion.com/products/triversa-nanomate/

Deposited Data		
EAEC TssKFG complex	This paper, deposited at EMdatabank	EMD-0008
EAEC TssK root-like domains	This paper, deposited at EMdatabank	EMD-0010
EAEC TssFG wings and stalk domain	This paper, deposited at EMdatabank	EMD-0009
EAEC TssKFG model	This paper, deposited at PDB	PDB 6GIY
EAEC TssK model	This paper, deposited at PDB	PDB 6GJ3
EAEC TssFG model	This paper, deposited at PDB	PDB 6GJ1

Supplementary table 2. Cryo-EM data collection, refinement and validation statistics

	TssKFG (EMDB-0008) (PDB 6GIY)	TssK (EMDB-0010) (PDB 6GJ3)	TssFG (EMDB-0009) (PDB 6GJ1)
Data collection and processing			
Magnification	130,000X	130,000X	130,000X
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	45	45	45
Defocus range (μm)	0.4 to 3 μm	0.4 to 3 μm	0.4 to 3 μm
Pixel size (Å)	1.1 Å	1.1 Å	1.1 Å
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	167,825	167,825	167,825
Final particle images (no.)	32,504	32,504	32,504
Map resolution (Å)	4.6Å	4.3Å	4.7Å
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	3.9 and 18 Å	3.8-33 Å	4.3-8Å
Refinement			
Initial model used (PDB code)	6GJ1 and 6GIY	5M30	-
Model resolution (Å)	4.6Å	4.3Å	4.7Å
FSC threshold	0.143	0.143	0.143
Model resolution range (Å)			
Map sharpening *B* factor (Å^2)	81.97	87.66	188.97
Model composition			
Non-hydrogen atoms	28969	17405	13013
Protein residues	3649	2194	1638
Ligands	N/A	N/A	N/A
B factors (Å^2)			
Protein	35.09-358.02	39.53-247.92	79.33-427.4
Ligand	N/A	N/A	N/A
R.m.s. deviations			
Bond lengths (Å)	0.28	0.39	0.52
Bond angles (°)	0.62	0.63	0.76
Validation			
MolProbity score	1.94	1.91	2.36
Clashscore	5.95	5.11	13
Poor rotamers (%)	0.32	0.37	0.31
Ramachandran plot			
Favored (%)	87.06	86.23	80.28
Allowed (%)	12.66	13.45	19.52
Disallowed (%)	0.28	0.32	0.2

Supplementary Table 3. Agreement of model contacts with predicted contact maps.
#Consistent: number of predicted contacts consistent with the final structure.
#Inconsistent: number of predicted contacts inconsistent with the final structure.
Accuracy: #consistent/(#inconsistent+#consistent)

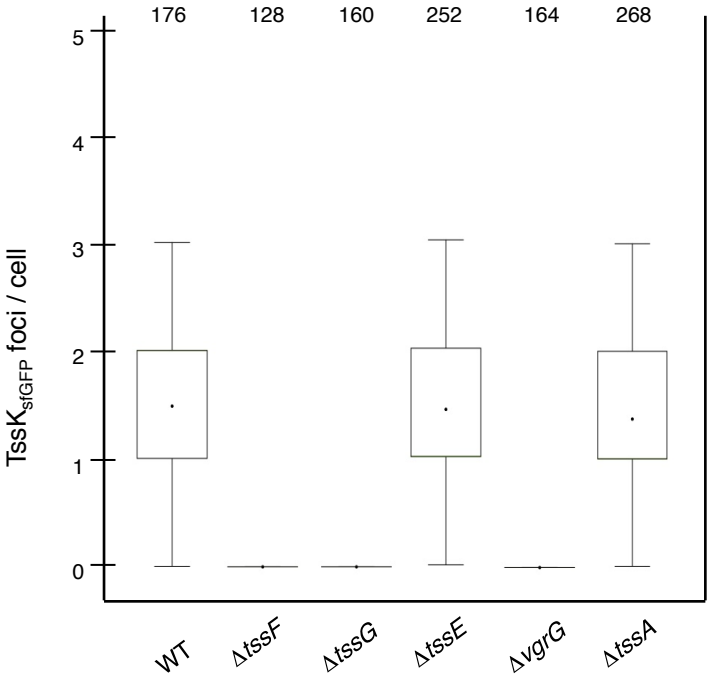
	#Consistent	#Inconsistent	Accuracy
TssK	1696	54	0.97
TssG	1180	180	0.87
TssF	2554	432	0.85

Supplementary References:

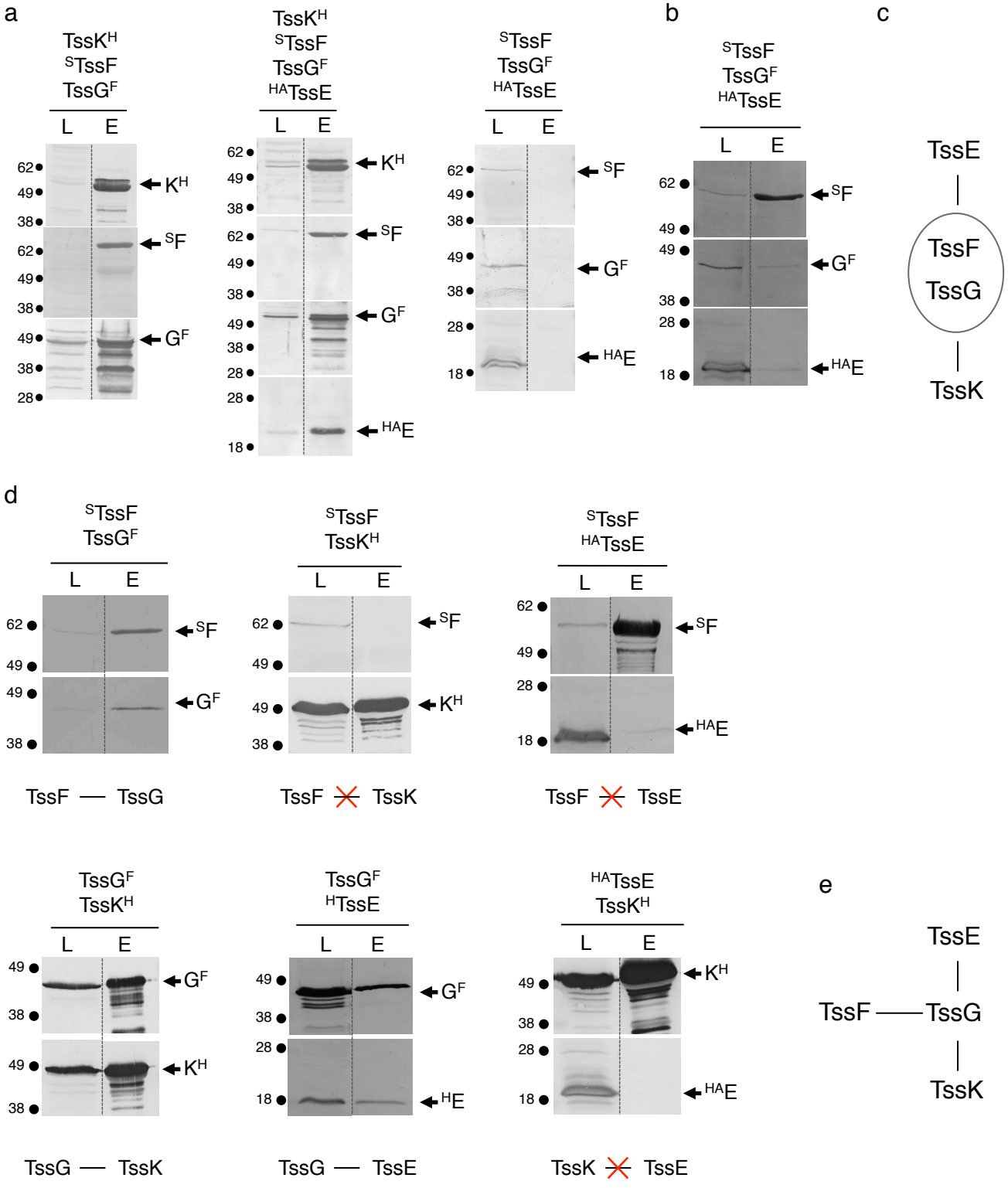
1. Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **28,** 5–13 (2017).
2. Taylor, N. M. I. *et al.* Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533,** 346–352 (2016).
3. English, G., Byron, O., Cianfanelli, F. R., Prescott, A. R. & Coulthurst, S. J. Biochemical analysis of TssK, a core component of the bacterial Type VI secretion system, reveals distinct oligomeric states of TssK and identifies a TssK–TssFG subcomplex. *Biochem. J.* **461,** 291–304 (2014).
4. Nazarov, S. *et al.* Cryo-EM reconstruction of Type VI secretion system baseplate and sheath distal end. *EMBO J.* **37,** e97103 (2018).
5. Nguyen, V. S. *et al.* Type VI secretion TssK baseplate protein exhibits structural similarity with phage receptor-binding proteins and evolved to bind the membrane complex. *Nat. Microbiol.* **2,** 17103 (2017).
6. Durand, E. *et al.* Biogenesis and structure of a type VI secretion membrane core complex. *Nature* **523,** 555–560 (2015).
7. Song, C. F. *et al.* Flexibility within the rotor and stators of the vacuolar H +-ATPase. *PLoS One* **8,** (2013).
8. Büttner, C. R., Wu, Y., Maxwell, K. L. & Davidson, A. R. Baseplate assembly of phage Mu: Defining the conserved core components of contractile-tailed phages and related bacterial systems. *Proc. Natl. Acad. Sci.* **113,** 10174–10179 (2016).
9. Brunet, Y. R., Zoued, A., Boyer, F., Douzi, B. & Cascales, E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization. *PLOS Genet.* **11,** e1005545 (2015).
10. Kostyuchenko, V. A. *et al.* Three-dimensional structure of bacteriophage T4 baseplate. *Nat. Struct. Biol.* **10,** 688–693 (2003).
11. Yap, M. L. *et al.* Structure of the 3.3MDa, in vitro assembled, hubless bacteriophage T4 baseplate. *J. Struct. Biol.* **187,** 95–102 (2014).
12. Arisaka, F., Yap, M. L., Kanamaru, S. & Rossmann, M. G. Molecular assembly and structure of the bacteriophage T4 tail. *Biophys. Rev.* **8,** 385–396 (2016).
13. Aksyuk, A. A., Leiman, P. G., Shneider, M. M., Mesyanzhinov, V. V. & Rossmann, M. G. The Structure of Gene Product 6 of Bacteriophage T4, the Hinge-Pin of the Baseplate. *Structure* **17,** 800–808 (2009).
14. Aksyuk, A. A. *et al.* The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. *EMBO J.* **28,** 821–829 (2009).
15. Kostyuchenko, V. A. *et al.* The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat. Struct. Mol. Biol.* **12,** 810–813 (2005).
16. Yap, M. L. & Rossmann, M. G. Structure and function of bacteriophage T4. *Future Microbiol.* **9,** 1319–1327 (2014).
17. Leiman, P. G. & Shneider, M. M. Contractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.* **726,** 93–114 (2012).
18. Yap, M. L. *et al.* Sequential assembly of the wedge of the baseplate of phage T4 in the presence and absence of gp11 as monitored by analytical ultracentrifugation. *Macromol. Biosci.* **10,** 808–813 (2010).
19. Yap, M. L., Mio, K., Leiman, P. G., Kanamaru, S. & Arisaka, F. The Baseplate Wedges of Bacteriophage T4 Spontaneously Assemble into Hubless Baseplate-Like Structure In Vitro. *J. Mol. Biol.* **395,** 349–360 (2010).
20. Leiman, P. G. *et al.* Morphogenesis of the T4 tail and tail fibers. *Virology Journal* **7,** (2010).
21. Flaugnatti, N. *et al.* A phospholipase A 1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol. Microbiol.* **99,** 1099–1118 (2016).
22. Alcoforado Diniz, J., Liu, Y.-C. & Coulthurst, S. J. Molecular weaponry: diverse effectors delivered by the Type VI secretion system. *Cell. Microbiol.* **17,** 1742–1751 (2015).
23. Bondage, D. D., Lin, J.-S., Ma, L.-S., Kuo, C.-H. & Lai, E.-M. VgrG C terminus confers the type

VI effector transport specificity and is required for binding with PAAR and adaptor–effector complex. *Proc. Natl. Acad. Sci.* **113,** E3931–E3940 (2016).
24. Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLOS Comput. Biol.* **13,** e1005324 (2017).
25. Brunet, Y. R., Henin, J., Celia, H. & Cascales, E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. *EMBO Rep.* **15,** 315–321 (2014).
26. Zoued, A. *et al.* TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. *J. Biol. Chem.* **288,** 27031–27041 (2013).
27. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97,** 6640–6645 (2000).
28. Chaveroche, M. K., Ghigo, J. M. & d'Enfert, C. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. *Nucleic Acids Res.* **28,** E97 (2000).
29. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177,** 4121–4130 (1995).
30. Rieping, W., Bardiaux, B., Bernard, A., Malliavin, T. E. & Nilges, M. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23,** 381–382 (2007).
31. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).
32. Clarke, O. B. Coot Trimmings. (2017).
33. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14,** 290–296 (2017).
34. Asarnow, D. pyem. (2016).
35. Wang, S., Peng, J., Ma, J. & Xu, J. Protein Secondary Structure Prediction Using Deep Convolutional Neural Fields. *Sci. Rep.* **6,** (2016).
36. Barad, B. A. *et al.* EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12,** 943–946 (2015).
37. Hopf, T. A. *et al.* Sequence co-evolution gives 3D contacts and structures of protein complexes. *Elife* **3,** (2014).
38. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193,** 1–12 (2016).
39. Ovchinnikov, S., Kamisetty, H. & Baker, D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *Elife* **2014,** (2014).
40. Rasband, W. ImageJ. *U. S. Natl. Institutes Heal. Bethesda, Maryland, USA* //imagej.nih.gov/ij/ (2012).
41. Ovchinnikov, S. *et al.* Protein structure determination using metagenome sequence data. *Science (80-.).* **355,** 294–298 (2017).
42. Ducret, A., Quardokus, E. M. & Brun, Y. V. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat. Microbiol.* **1,** (2016).
43. Chen, V. B. *et al.* MolProbity : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).
44. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14,** 331–332 (2017).
45. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).
46. Afonine, P. V *et al.* Real-space refinement in Phenix for cryo-EM and crystallography. *bioRxiv* (2018).
47. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10,** 845–858 (2015).
48. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
49. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).
50. Sanchez, R. M. Recentering and subboxing of particles (REP). (2017).
51. Leaver-Fay, A. *et al.* Rosetta3: An object-oriented software suite for the simulation and design of

macromolecules. *Methods Enzymol.* **487,** 545–574 (2011).
52. Pettersen, E. F. *et al.* UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* **25,** 1605–1612 (2004).

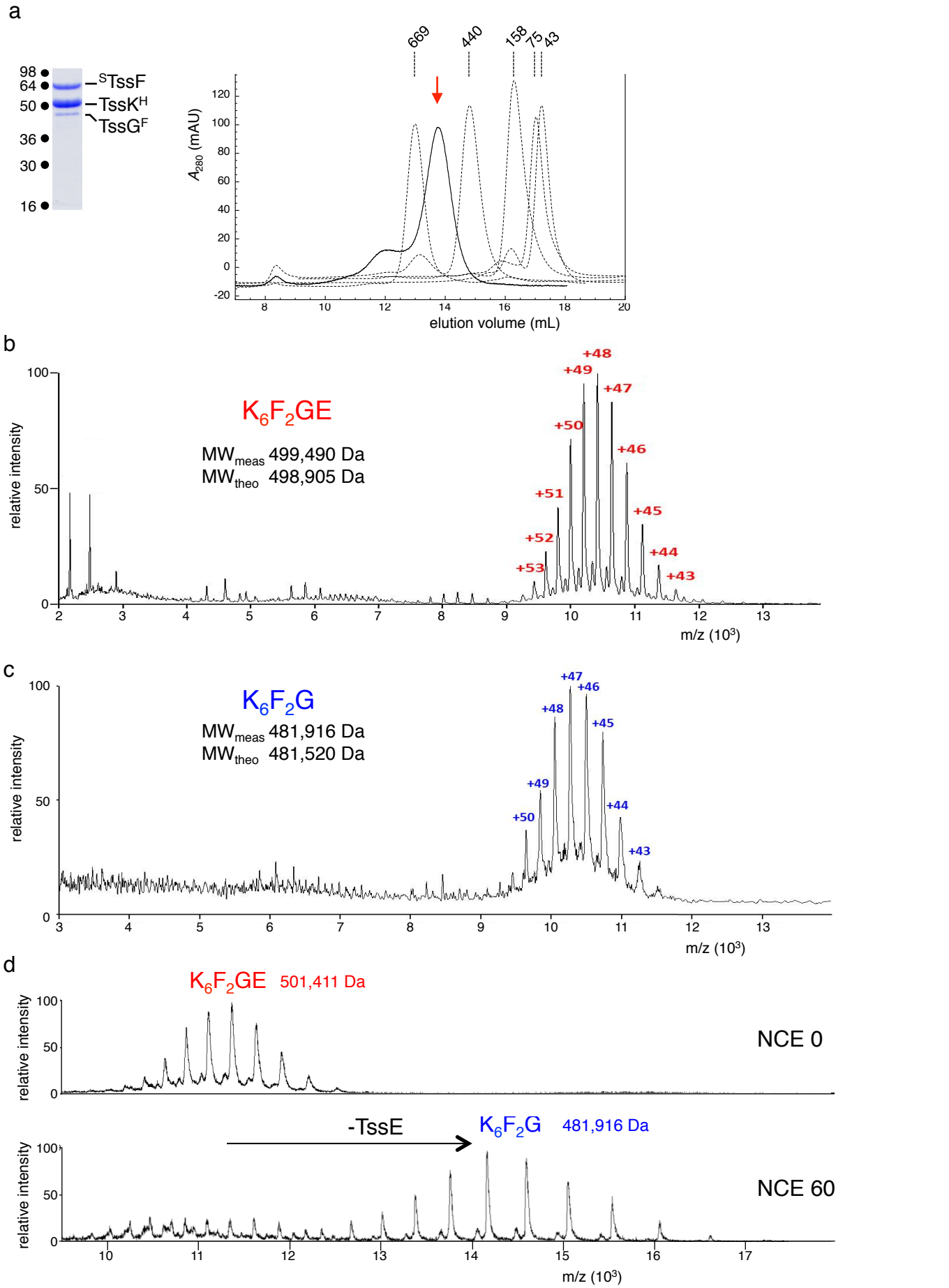


Supplementary Figure 1. Statistical analysis of TssK$_{sfGFP}$ foci formation. Related to Figure 1.
Statistical analysis of TssK$_{sfGFP}$ foci formation in various T6SS mutant backgrounds. Shown are the box-and-whisker plots of the measured number of TssK$_{sfGFP}$ foci per cell for each indicated strain with the lower and upper boundaries of the boxes corresponding to the 25thand 75th percentiles respectively (black dot, the median values for each strain; whiskers, the 10th and 90th percentiles); The number of cells analyzed for each strain is indicated on top.


a
TssK^H
^STssF
TssG^F
L
E
62
49
38
28
K^H
^SF
G^F
TssK^H
^STssF
TssG^F
^HATssE
^HAE
^STssF
TssG^F
^HATssE
18
b
^STssF
TssG^F
^HATssE
c
TssE
TssF
TssG
TssK
d
^STssF
TssG^F
TssF — TssG
^STssF
TssK^H
TssF TssK
^STssF
^HATssE
TssF TssE
TssG^F
TssK^H
TssG — TssK
TssG^F
^HTssE
^HE
TssG — TssE
^HATssE
TssK^H
TssK TssE
e
TssE
TssF — TssG
TssK

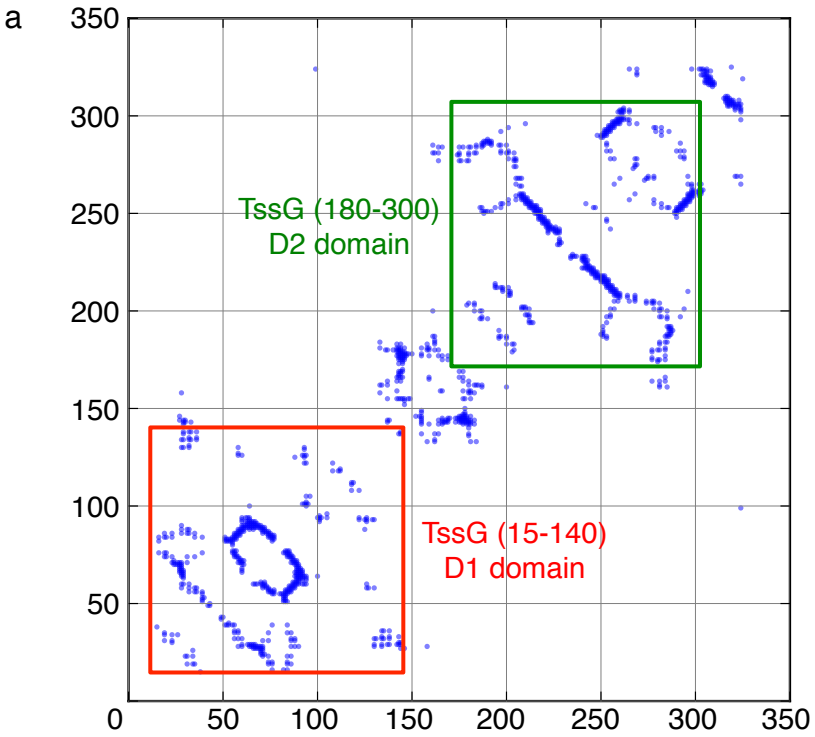
Supplementary Figure 2. Protein-protein interaction network between the TssKFGE complex. Related to Figure 1.

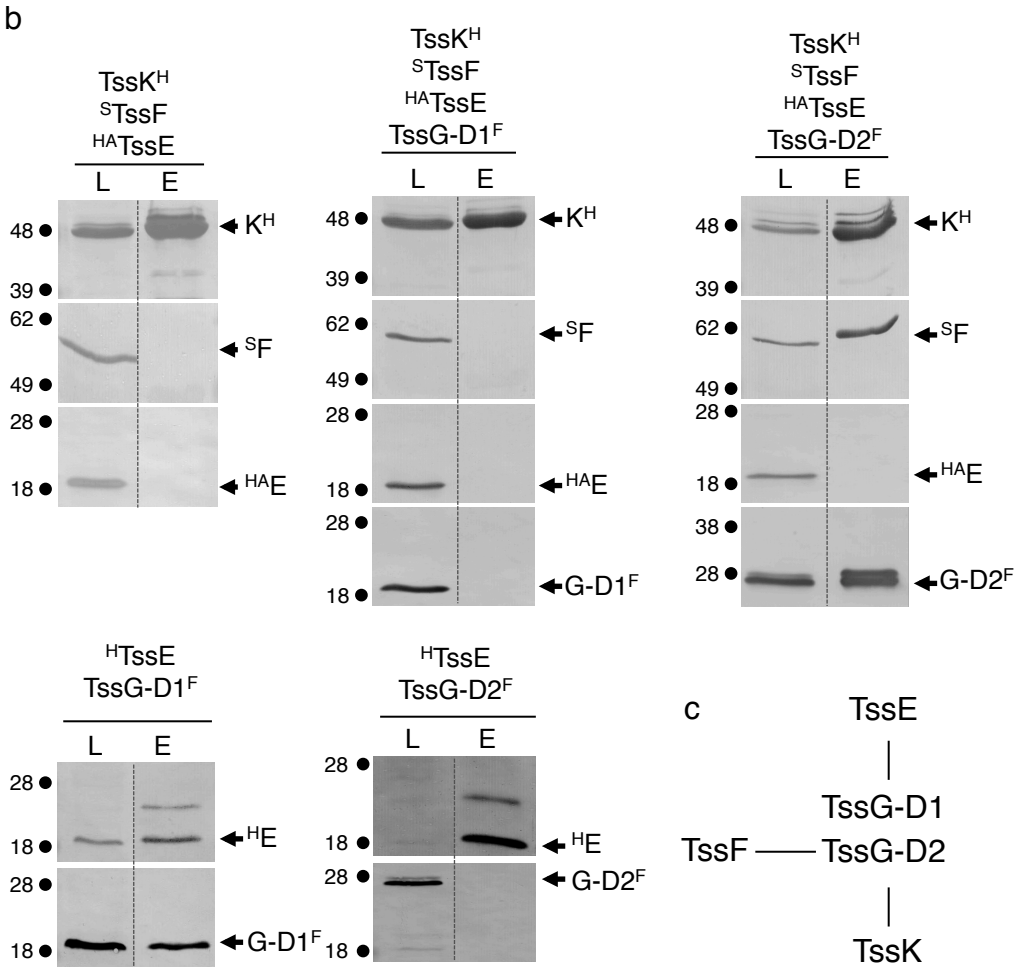
a, **b**, and **d**. Soluble extracts of *E. coli* BL21(DE3) cells producing the indicated protein were submitted to an affinity purification step on HisTrap or StrepTrap. The lysate (total soluble material, L), and eluate (E) were subjected to denaturing 12.5%-acrylamide polyacrylamide gel electrophoresis (PAGE) and immunodetected with the appropriate antibody. Immunodetected proteins are indicated in the right. Molecular weight markers (in kDa) are indicated in the left. Tags: H, 6×His; S, Strep-tag; F, FLAG; HA, hemagglutinin. In panel **d**, the result of the pairwise interaction is schematized (red cross indicates that no interaction is observed). Schematic summaries of panels **a**-**b** and **d** are shown in panels **c** and **e** respectively. Protein-protein interaction experiments have been performed independently twice and a representative result is shown.


a
98
64
50
36
30
16
^{S}TssF
TssK^{H}
TssG^{F}
669
440
158
75
43
A_{280} (mAU)
elution volume (mL)
b
K_6F_2GE
MW_{meas} 499,490 Da
MW_{theo} 498,905 Da
+48
+49
+47
+50
+46
+51
+45
+52
+44
+53
+43
relative intensity
m/z (10^3)
c
K_6F_2G
MW_{meas} 481,916 Da
MW_{theo} 481,520 Da
+47
+46
+48
+45
+49
+44
+50
+43
relative intensity
m/z (10^3)
d
K_6F_2GE 501,411 Da
NCE 0
-TssE
K_6F_2G 481,916 Da
NCE 60
relative intensity
m/z (10^3)

Supplementary Figure 3. Purification, biochemical characterization and stoichiometry analyses of the TssKFG and TssKFGE complexes. Related to Figure 1.

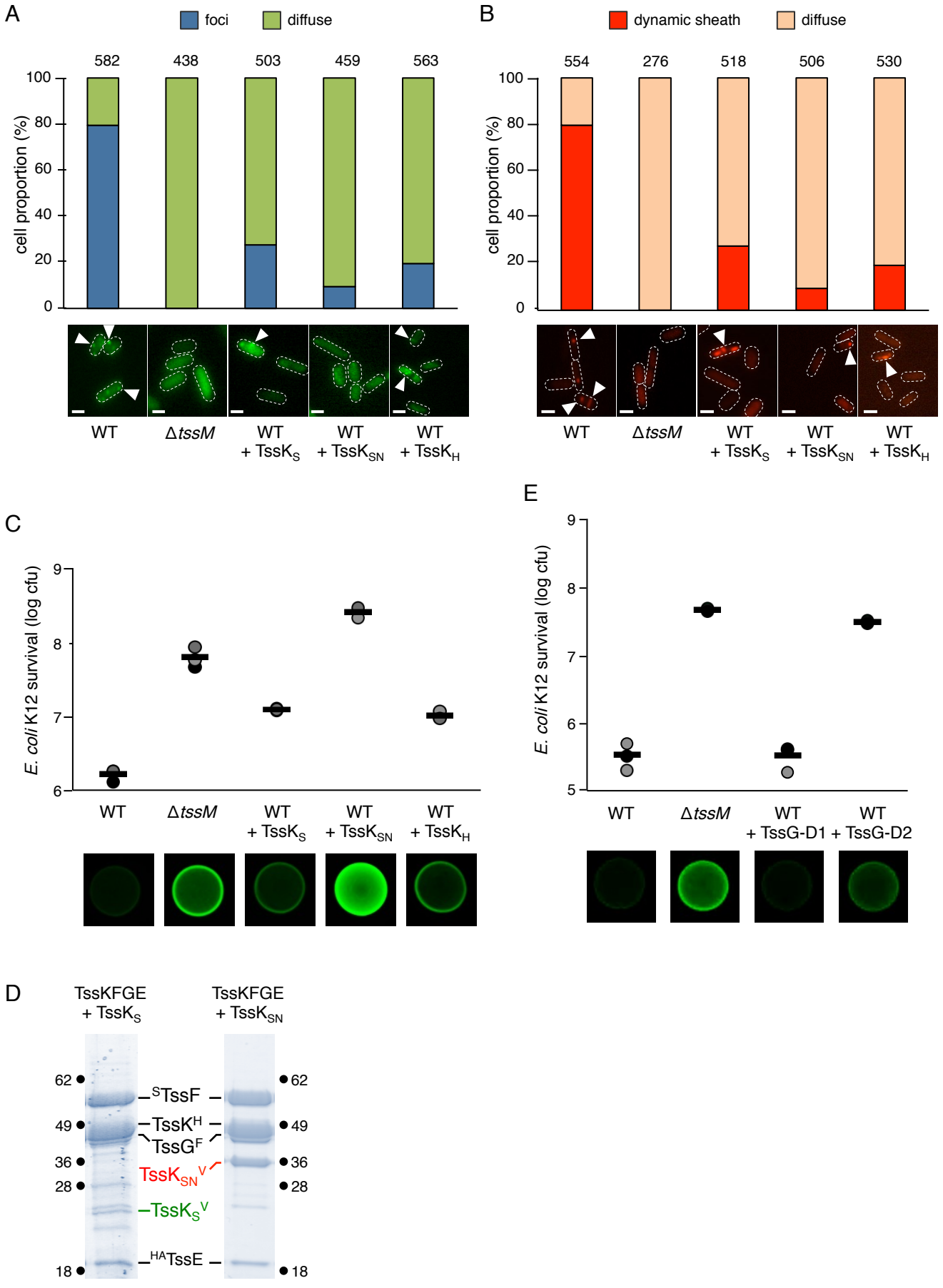
a. Left panel: Purified TssKFG complex subjected to sodium dodecyl sulfate 12.5%-acrylamide PAGE and Coomassie staining. The different proteins are indicated on the right, whereas molecular weight markers are indicated on the left. Right panel. Biochemical characterization of the TssKFG complex. Analytical size-exclusion chromatography analysis of the purified TssKFG complex (continuous line) on a Superose 6 column, calibrated with 43-, 75-, 158-, 440- and 660-kDa molecular mass markers (dotted lines). The molecular mass of each marker (in kDa) is indicated on the top of the corresponding peak. An arrow indicates the position of the peak fraction corresponding to the TssKFG complex. TssKFG complex purification and analytical size-exclusion chromatography analysis experiments were performed at least three times and a representative result is shown. **b-c.** Mass spectra of the TssKFGE (**b**) and TssKFG (**c**) complexes obtained in native conditions. The stoichiometry (indicated in red and blue for the TssKFGE and TssKFG complexes) were defined based on the comparison between theoretical (MW_{theo}) and measured (MW_{meas}) molecular weight markers. **d.** In-source dissociation (cone voltage: 150 V) of the TssKFGE complex obtained by collisional activation in native conditions. Native mass spectrometry and collisional activation experiments were performed 3 times and a representative result is shown.


a
350
300
250
200
150
100
50
0
50
100
150
200
250
300
350
TssG (180-300)
D2 domain
TssG (15-140)
D1 domain


b
TssK^H
^STssF
^HATssE
L
E
48
39
62
49
28
18
K^H
^SF
^HAE
TssK^H
^STssF
^HATssE
TssG-D1^F
L
E
48
39
62
49
28
18
28
18
K^H
^SF
^HAE
G-D1^F
TssK^H
^STssF
^HATssE
TssG-D2^F
L
E
48
39
62
49
28
18
38
28
K^H
^SF
^HAE
G-D2^F
^HTssE
TssG-D1^F
L
E
28
18
28
18
^HE
G-D1^F
^HTssE
TssG-D2^F
L
E
28
18
28
18
^HE
G-D2^F
c
TssE
TssG-D1
TssF
TssG-D2
TssK

Supplementary Figure 4. TssG domain organization and interactions. Related to Figure 1.

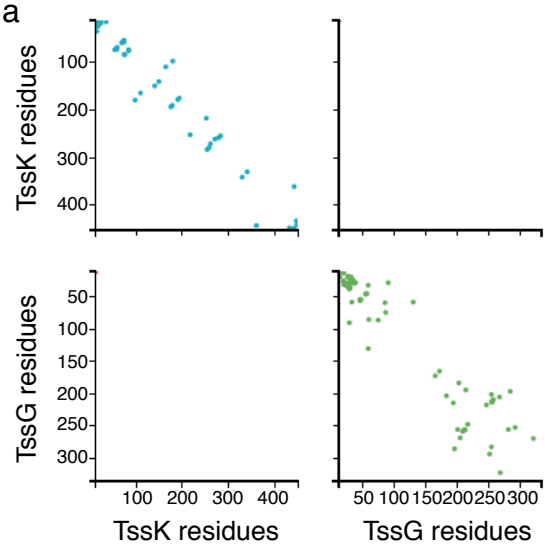
a. TssG residue contacts predicted by RaptorX [24] plotted into a contact map. X and Y axis are the residue indexes of TssG. Each blue dot is a predicted contact. The map is organized into two blocks of dense contacts located at the N- (TssG-D1, red, residues 15-140) and C-terminus (TssG-D2, green, residues 180-300). **b.** Soluble extracts of *E. coli* BL21(DE3) cells producing the indicated protein or domains were submitted to an affinity purification step on HisTrap. The lysate (total soluble material, L), and eluate (E) were subjected to denaturing 12.5%-acrylamide polyacrylamide gel electrophoresis (PAGE) and immunodetected with the appropriate antibody. Immunodetected proteins are indicated in the right. Molecular weight markers (in kDa) are indicated in the left. Tags: H, 6×His; S, Strep-tag; F, FLAG; HA, hemagglutinin. Protein-protein interaction experiments have been performed independently twice and a representative result is shown. **c.** Schematic summary of the interactions of TssG-D1 and -D2 with the TssKFE proteins.


A
foci
diffuse
582
438
503
459
563
cell proportion (%)
WT
ΔtssM
WT + TssK_S
WT + TssK_SN
WT + TssK_H
B
dynamic sheath
diffuse
554
276
518
506
530
C
E. coli K12 survival (log cfu)
WT + TssG-D1
WT + TssG-D2
D
TssKFGE + TssK_S
TssKFGE + TssK_SN
62
49
36
28
18
STssF
TssK^H
TssG^F
TssK_SN^V
TssK_S^V
HATssE

Supplementary Figure 5. Small Domain Interference analyses. Related to Figure 2.

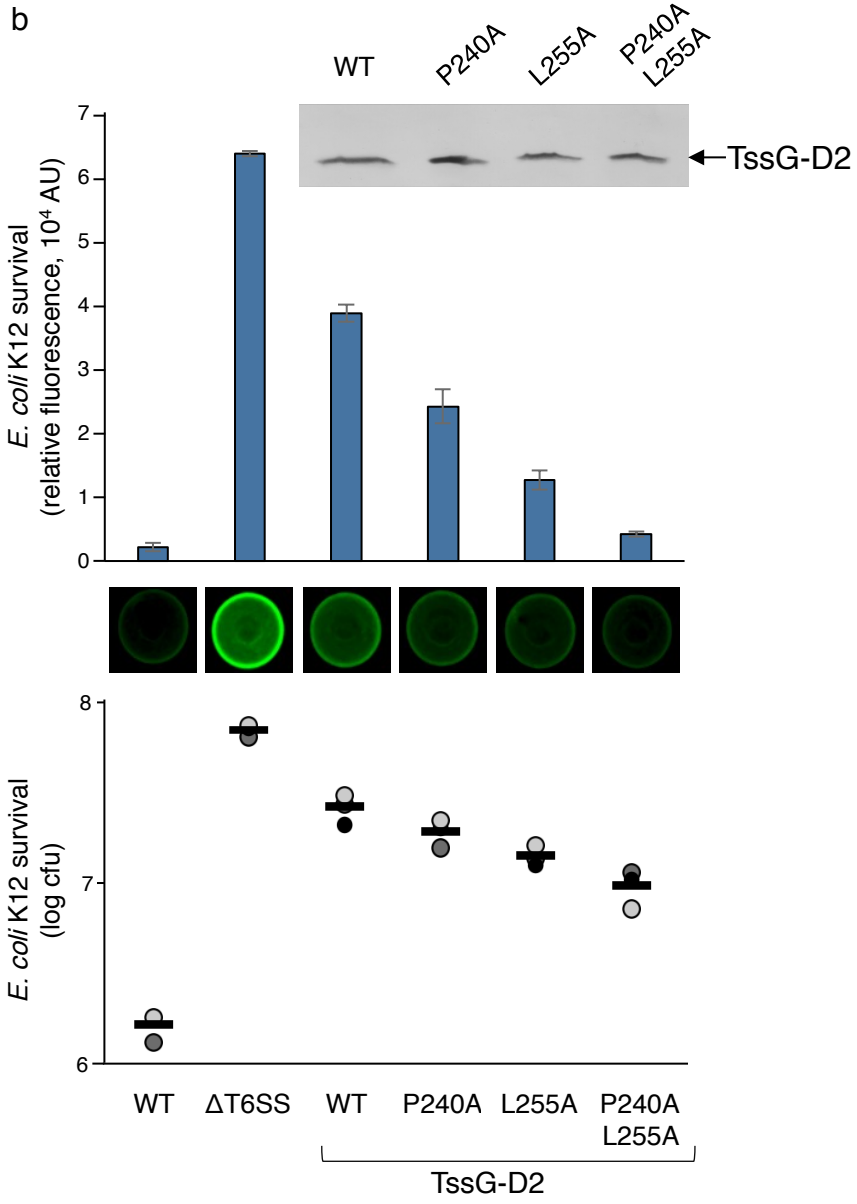
a-c. Overproduction of TssK domains inhibits T6SS activity. **a.** Fluorescence microscopy recordings and statistical analyses of $TssK_{sfGFP}$ foci formation upon overproduction of TssK domains ($TssK_S$, shoulder; $TssK_{SN}$, shoulder and neck; $TssK_H$, head). A representative image is shown on the bottom (white arrowheads indicate $TssK_{sfGFP}$ foci). The graph representing the percentage of cells with diffuse (green), and foci (blue) patterns is indicated on top (number of cells analyzed indicated on top of each bar). **b.** Fluorescence microscopy recordings and statistical analyses of $TssB_{mCherry}$ dynamics upon overproduction of TssK domains. A representative image is shown at the bottom (white arrowheads indicate dynamic $TssB_{mCherry}$ sheaths). The graph representing the percentage of cells with diffuse (light orange) and dynamic (orange) patterns is indicated on top (number of cells analyzed indicated on top of each bar). Microscopy analyses were performed independently three times, each in technical triplicate, and a representative experiment is shown. Scale bars, 1 μm. **c.** Antibacterial assay. Gfp^+ kan^R *E. coli* K-12 recipient cells were mixed with the indicated attacker cells, spotted onto Sci-1-inducing medium (SIM) agar plates supplemented with 0.05% L-arabinose, and incubated for 4 h at 37°C. The image of a representative bacterial spot is shown on the bottom and the number of recovered *E. coli* recipient cells is indicated in the upper graph (in log^{10} of colony forming units (cfu)). The assays were performed from at least three independent cultures, with technical triplicates and a representative technical triplicate is shown. The circles indicate values from the technical triplicate, and the average is indicated by the bar. **d.** Purification of the TssKFGE complex upon overproduction of the $TssK_S$ and $TssK_{SN}$ domains. TssKFGE complexes were purified from cells overproducing $TssK_S$ (left panel) or $TssK_{SN}$ (right panel). The different proteins are indicated, as well as molecular weight markers (in kDa). Purification and sodium dodecyl sulphate 12,5%-acrylamid**e** PAGE analysis were performed twice and a representative result is shown. **e.** The TssG-D2 domain interferes with T6SS antibacterial activity. Gfp^+ kan^R *E. coli* K-12 recipient cells were mixed with the indicated attacker cells overproducing the indicated TssG domain, spotted onto Sci-1-inducing medium (SIM) agar plates supplemented with 0.05% L-arabinose, and incubated for 4 h at 37°C. The image of a representative bacterial spot is shown on bottom and the number of recovered *E. coli* recipient cells is indicated in the upper graph (in log^{10} of colony forming units (cfu)). The circles indicate values from three independent assays, and the average is indicated by the bar.

a



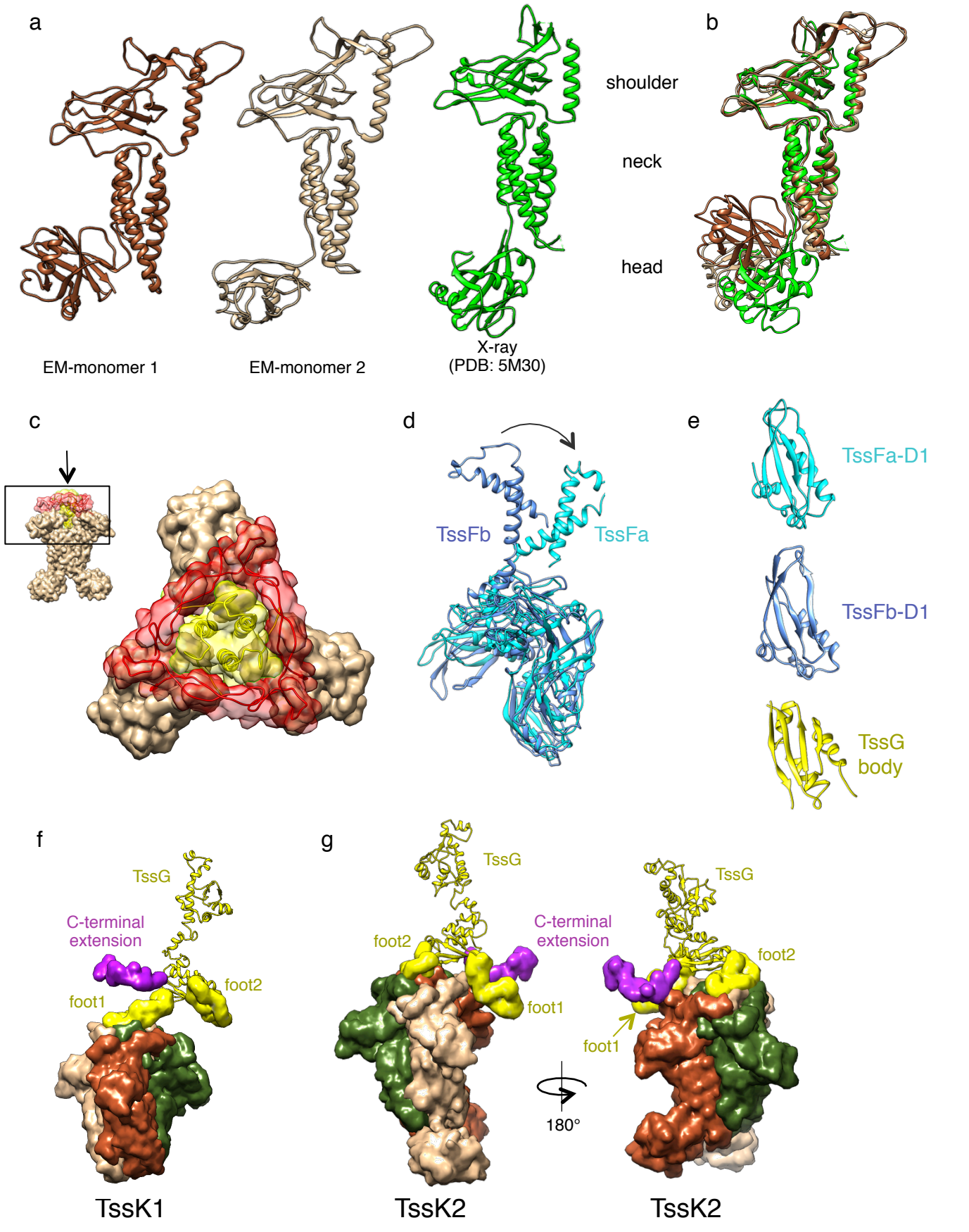
residues pairs		EV complex score
$TssK_S$	TssG-D2	
Leu-14	Pro-240	0.740
Arg-67	Pro-240	0.696
Phe-49	Leu-255	0.544

b



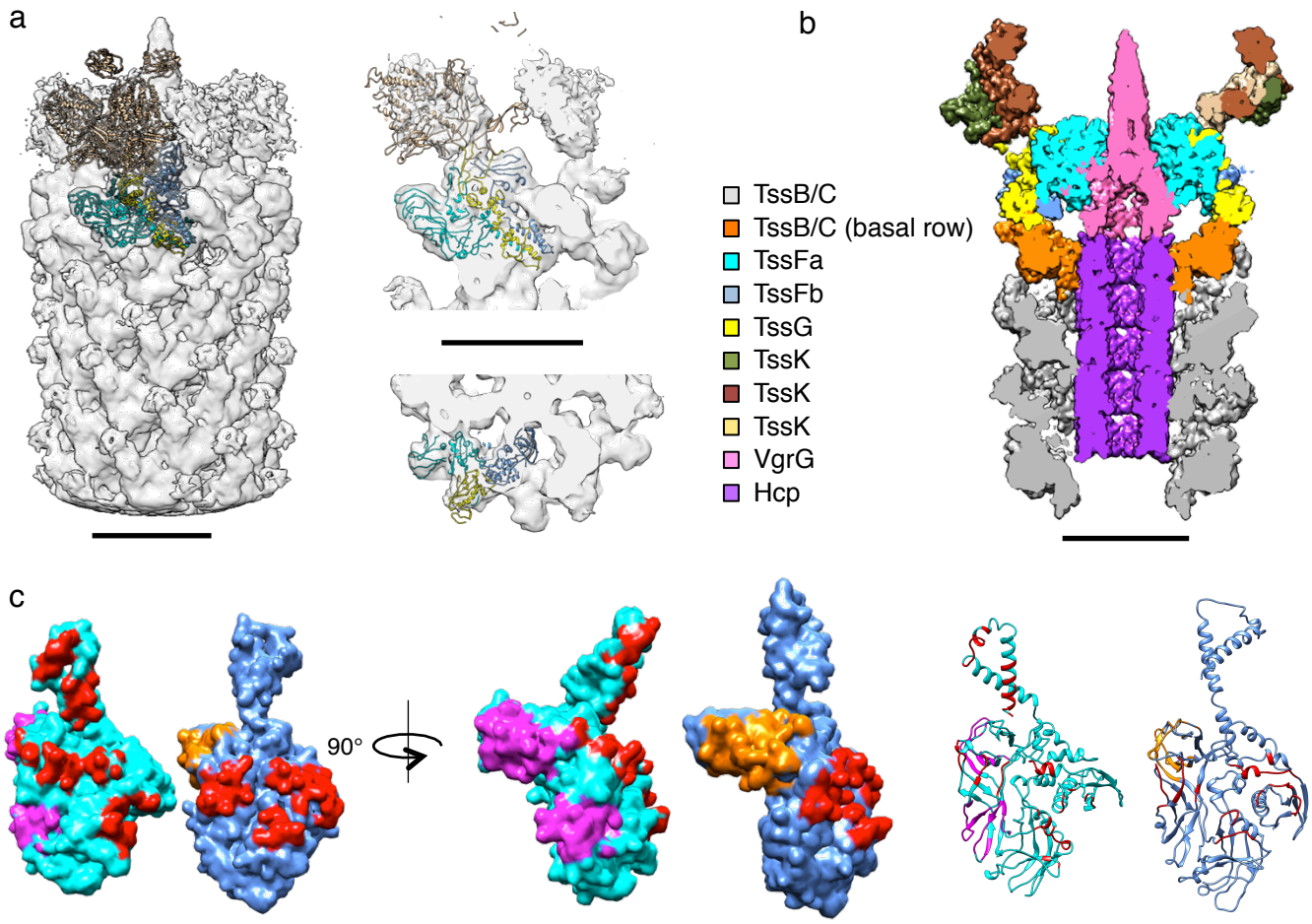
Supplementary Figure 6. TssG-TssK interaction. Related to Figure 2.

a. TssG-TssK residue contacts predicted by EVcomplex plotted into contact maps. X and Y axis are the residue indexes of TssG and TssK (as indicated). Each blue and green dot is a predicted contact in TssK and TssG respectively. The residue pairs with the higher EVcomplex scores are indicated on right. **b.** Gfp^{+} kan^{R} *E. coli* K-12 recipient cells were mixed with the indicated attacker cells overproducing the indicated TssG-D2 variants, spotted onto Sci-1-inducing medium (SIM) agar plates supplemented with 0.05% L-arabinose, and incubated for 4 h at 37°C. The image of a representative bacterial spot is shown on the middle, the relative fluorescence of the spot (in arbitrary units, AU) is shown on the top graph whereas the number of recovered *E. coli* recipient cells is indicated in the lower graph (in log^{10} of colony forming units (cfu)). The error bars represent the standard deviation across three independent assays. The circles indicate values from three independent assays, and the average is indicated by the bar. The levels of production of wild-type (WT) and mutant TssG-D2 variants is shown in the inset (SDS-PAGE and immunoblot analysis).

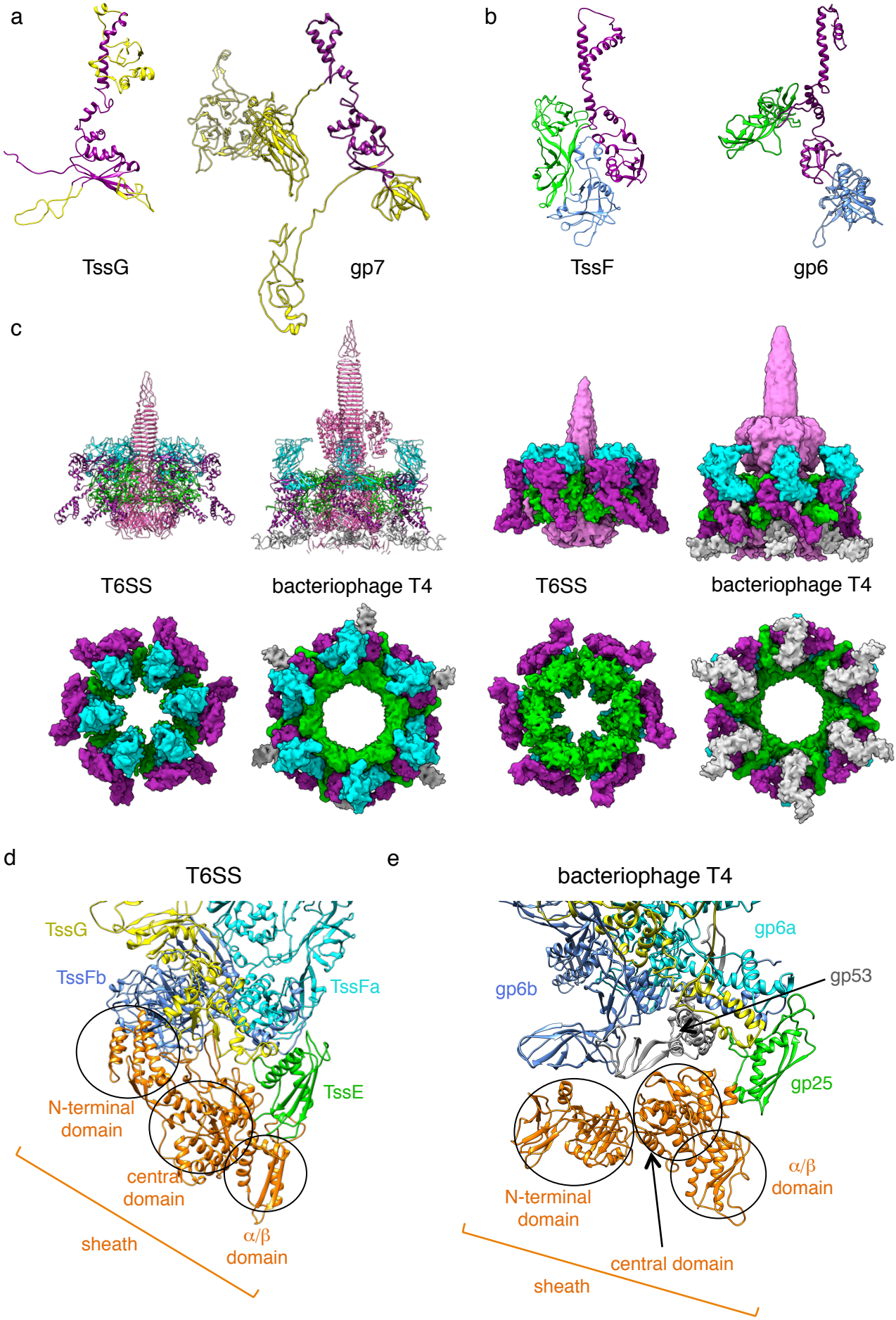

a
EM-monomer 1
EM-monomer 2
X-ray
(PDB: 5M30)
shoulder
neck
head
b
c
d
TssFb
TssFa
e
TssFa-D1
TssFb-D1
TssG
body
f
TssG
C-terminal
extension
foot2
foot1
TssK1
g
TssG
C-terminal
extension
foot2
foot1
180°
TssK2
TssK2

Supplementary Figure 7. Structures and domain interactions of the TssKFG proteins, Related to Figure 3

a. Comparison between cryo-EM and crystal structures of TssK. Each TssK protomer is represented in the same orientation. For the cryo-EM structure, two protomers with different orientations for the C-terminal domain are presented (in brown and beige). For the crystal structure, the protomer for which the terminal domain could be built is presented (in green). **b.** Superimposition of cryo-EM and crystal structures of TssK. The same protomers as in A. are superimposed. **c.** Triangular hub at the top of the TssK shoulder domains. The three loops located between residues 105 to 145 define a triangle (in red) that encompasses the loop 1-18 and α1 helix bundle at the centre of the trimer (in yellow). **d.** Superimposition of the TssFa and TssFb protomers. TssFa and TssFb are represented as ribbons in cyan and blue respectively. Both protomers differ only by the position of the N-terminal antenna (indicated by the black arrow). **e.** Comparison between the TssFa and TssFb D1 domains and TssG body structures, shown in the same orientation. **f-g.** Interactions between the TssK1 (**f**) or TssK2 (**g**) trimer and TssG. The TssK protomers (green, beige and brown) and TssG feet and C-terminal extension are represented as surfaces. TssG antenna and body are represented as ribbons. TssG is in yellow with the exception of its C-terminal extension that is coloured in purple.

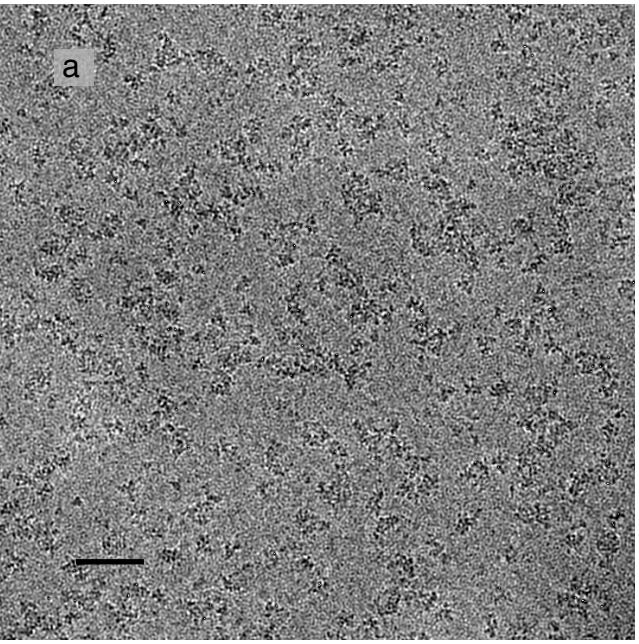


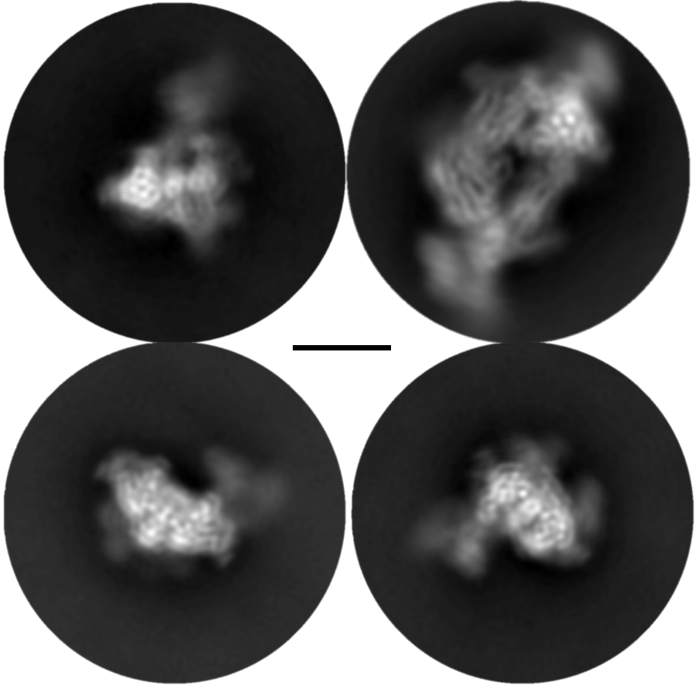
Supplementary Figure 8. T6SS extended sheath bound to the baseplate. Related to Figure 4
a. Fitting of the structure of the EAEC TssKFGE complex into the cryo-EM structure of the *V. cholerae* baseplate associated to a non-contractile sheath (EMD-3879). Left panel: side view highlighting one TssKFGE complex (ribbon diagram) fitted into the baseplate/sheath density (surface representation). Right panels: side (top) and bottom (bottom) slice of the TssKFGE complex fitted into the baseplate/ sheath density. The bars correspond to 100 Å **b.** Same representation of the EAEC baseplate and sheath models as in Figure 4a, with the location of the Hcp tube and VgrG spike from *V. cholerae*. The bar corresponds to 100 Å **c.** Ribbon diagram and surface representation of TssFa-TssFb belonging to two adjacent wedge complexes (main interface between wedge complexes within the T6SS baseplate). TssFa and TssFb are represented in cyan and blue respectively. The TssFa/Fb seen in Figure 4c is split open and the two protomers are placed in the same orientation. Within each subunit, the regions interacting with VgrG and TssBC are coloured in magenta and orange respectively. Left and middle panels display the same protomers rotated by 90° around their long axis. In the right panel, the protomers in ribbon diagram are shown in the same orientation as in the left panel.

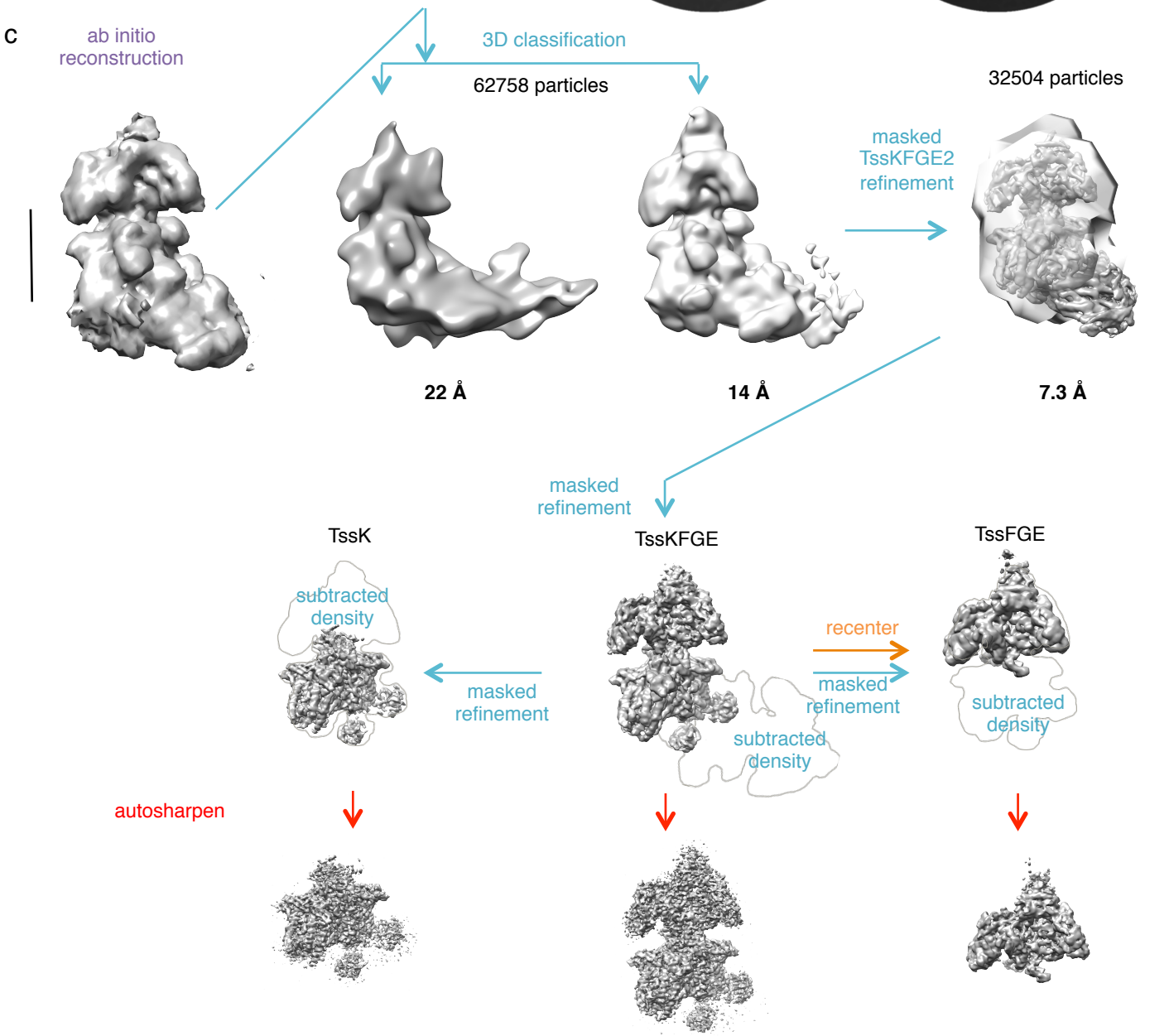

a
TssG
gp7
b
TssF
gp6
c
T6SS
bacteriophage T4
T6SS
bacteriophage T4
d
T6SS
TssG
TssFb
TssFa
TssE
N-terminal domain
central domain
α/β domain
sheath
e
bacteriophage T4
gp6a
gp6b
gp53
gp25
α/β domain
N-terminal domain
central domain
sheath

Supplementary Figure 9. Comparison of the bacteriophage T4 and T6SS baseplates. Related to Figure 5.

a. Comparison between TssG and gp7 structures (ribbon diagram). The structural features conserved in both structures are represented in purple. The structural elements specific to each protein are represented in yellow. **b.** Comparison between TssF and gp6 structures (ribbon diagram). The structural features conserved in both structures are represented in purple (trimer scaffold) and green (wing domain). The structural elements specific to each protein are represented in blue. **c.** Interaction between TssFa and gp6a with the spike. For TssFa and gp6a, the same colour code as in panel B is used. Top row: for the T6SS, TssFa and VgrG are represented. For the T4 bacteriophage, gp6a, gp53 (grey) and gp27-gp5 are represented. For each, a ribbon diagram and surface representation of the same view (side view) is displayed on the left and right part respectively. Bottom row: surface representation of the same assemblies but without the spike (Top and bottom views on left and right respectively). **d.** Ribbon diagram of the T6SS wedge-TssBC interaction. The colour code used to identify the various subunits is the same as in Figure 4D (TssFa, cyan; TssFb, blue; TssG, yellow; TssE, green; TssBC, orange). The different protomers and domains are identified. **e.** Ribbon diagram of the bacteriophage T4 wedge-gp18 interaction. The different protomers and domains are identified.

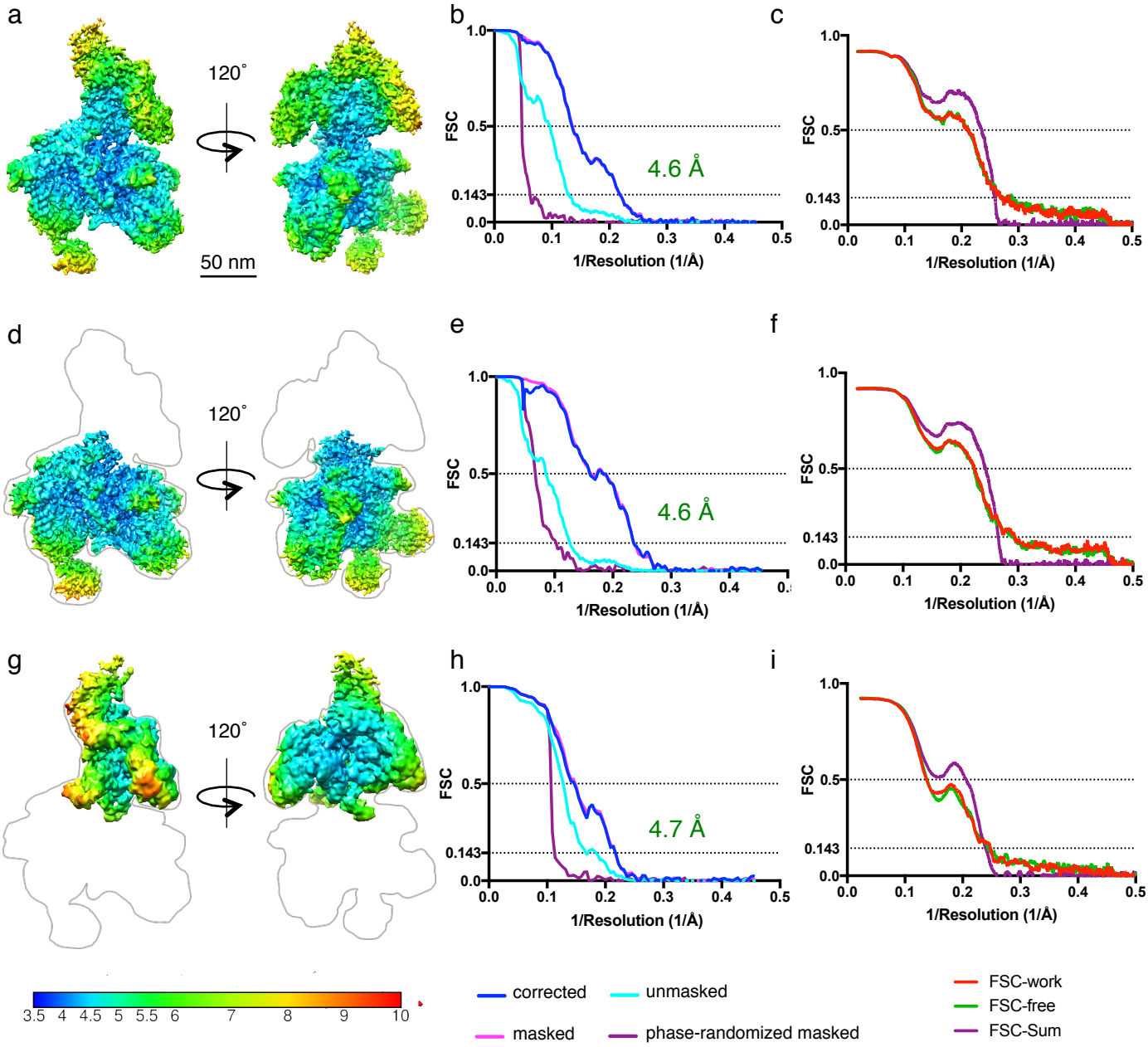

a


b

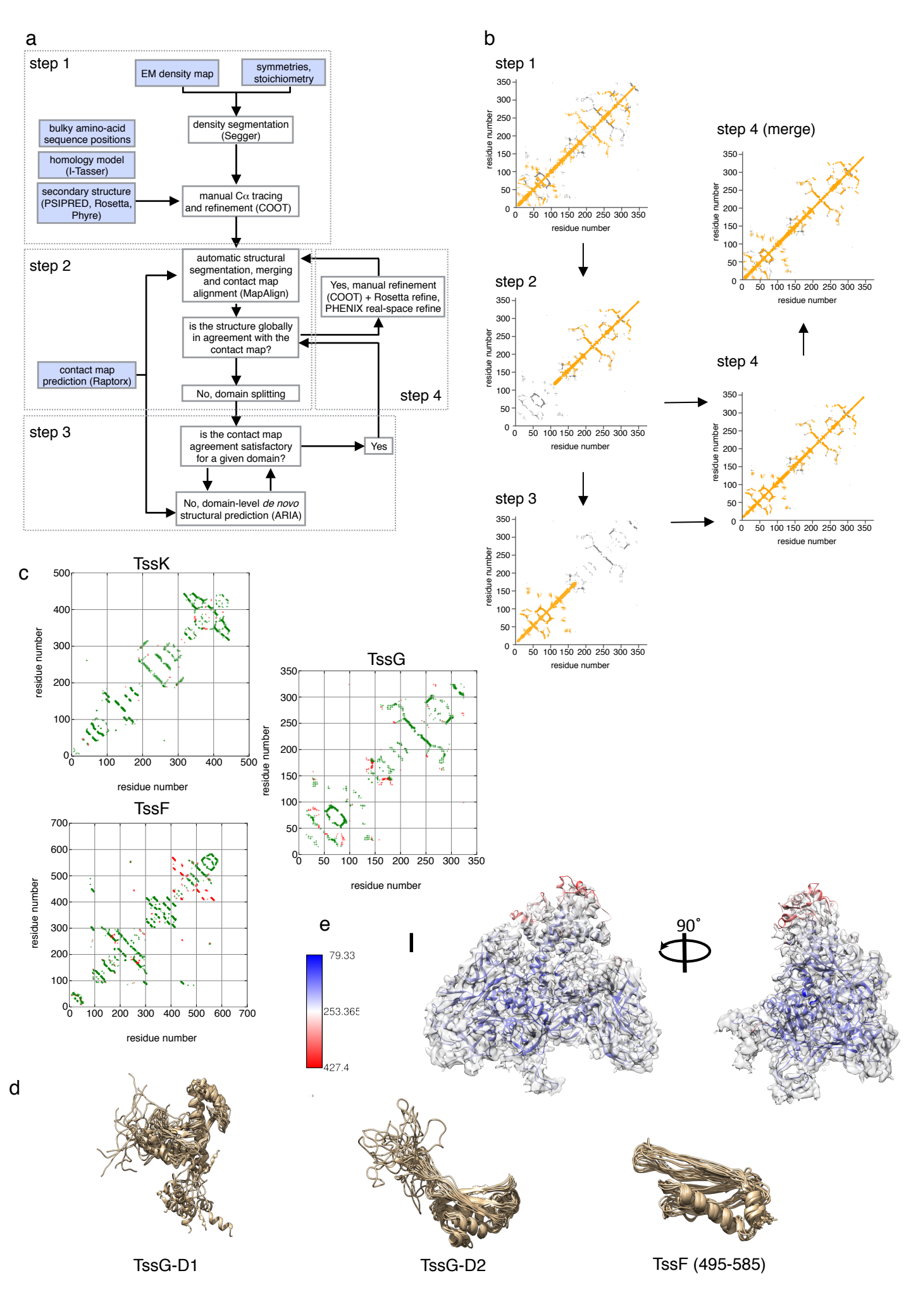

c
ab initio reconstruction
3D classification
62758 particles
32504 particles
masked TssKFGE2 refinement
22 Å
14 Å
7.3 Å
masked refinement
TssK
TssKFGE
TssFGE
subtracted density
masked refinement
recenter
masked refinement
subtracted density
subtracted density
autosharpen

Supplementary Figure 10. Structure determination. Related to Figure 2.

a. Representative cryo-micrograph of the purified TssKFGE complex. The scale bar corresponds to 50 nm **b.** Typical 2D classes of the 4 different views of the complex in ice (scale bar, 100 nm). **c.** Flowchart of the cryo-EM processing procedure. The colour of the arrows represents the different software packages used (light blue, Relion; red, Phenix; violet, Cryosparc; orange, REP).


a
120˚
50 nm
b
FSC
1.0
0.5
0.143
0.0
4.6 Å
0.0 0.1 0.2 0.3 0.4 0.5
1/Resolution (1/Å)
c
FSC
1.0
0.5
0.143
0.0
0.0 0.1 0.2 0.3 0.4 0.5
1/Resolution (1/Å)
d
120˚
e
FSC
1.0
0.5
0.143
0.0
4.6 Å
0.0 0.1 0.2 0.3 0.4
1/Resolution (1/Å)
f
FSC
1.0
0.5
0.143
0.0
0.0 0.1 0.2 0.3 0.4 0.5
1/Resolution (1/Å)
g
120˚
h
FSC
1.0
0.5
0.143
0.0
4.7 Å
0.0 0.1 0.2 0.3 0.4 0.5
1/Resolution (1/Å)
i
FSC
1.0
0.5
0.143
0.0
0.0 0.1 0.2 0.3 0.4 0.5
1/Resolution (1/Å)
3.5 4 4.5 5 5.5 6 7 8 9 10
corrected
unmasked
masked
phase-randomized masked
FSC-work
FSC-free
FSC-Sum
j
k

Supplementary Figure 11. Analysis of model quality and refinement. Related to Figure 2.
a, d, and **g**. Local resolution of the TssKFGE (**a**), TssFGE (**d**) and TssK (**g**) autoshapened densities. b, e, and h. Fourier shell correlation (FSC) of the final reconstruction of TssKFGE (**b**), TssFGE (**e**) and TssK (**h**). The resolution limit was calculated at the cut-off 0.143. **c, f**, and **i.** Model validation by comparison of the Fourier shell correlation (FSC) between model and half map 1 (work), model and half map 2 (free), and model versus full sharpened map (sum) plotted in red, green and purple respectively for of the TssKFGE (**c**), TssFGE (**f**) and TssK (**i**). **j.** Representative local density of an a-helix (residues 19-33 of TssK chain H). **k**. Representative local density of a b-strand (residues 147-152 of TssK chain H).


a
step 1
EM density map
symmetries, stoichiometry
density segmentation (Segger)
bulky amino-acid sequence positions
homology model (I-Tasser)
secondary structure (PSIPRED, Rosetta, Phyre)
manual Cα tracing and refinement (COOT)
step 2
automatic structural segmentation, merging and contact map alignment (MapAlign)
Yes, manual refinement (COOT) + Rosetta refine, PHENIX real-space refine
is the structure globally in agreement with the contact map?
contact map prediction (Raptorx)
No, domain splitting
step 4
step 3
is the contact map agreement satisfactory for a given domain?
Yes
No, domain-level de novo structural prediction (ARIA)
b
step 1
step 2
step 3
step 4
step 4 (merge)
residue number
c
TssK
TssG
TssF
residue number
d
TssG-D1
TssG-D2
TssF (495-585)
e
79.33
253.365
427.4
f
90°

Supplementary Figure 12. Structural modeling of TssF and TssG. Related to Figure 2.
a. Flowchart of the structural modelling of TssG and TssF. **b**. Example of TssG contact map fit progressing with the different steps of the pipeline. The grey and orange dots are the predicted and model contacts, respectively, as defined in the supplementary methods text. **c**. Contact maps of the final TssK, TssG, and TssF structures compared to the corresponding predicted contact maps of the modelling pipeline. Green and red dots are predicted contacts in agreement and in disagreement with the structure, respectively. **d**. Ensemble of models for TssG-D1, TssG-D2 and TssF-D1 obtained from the predicted contacts and the ARIA modelling software (Step 3 of the pipeline). **e**. Fitting of the TssF and TssG structures in the postprocessed cryo-EM density (TssFGE), coloured according to B-factor values. The scale bar corresponds to 10 Å.