

# Direct interaction of a chaperone-bound type III secretion substrate with the export gate

Dominic Gilzer <sup>1</sup>, Madeleine Schreiner<sup>1</sup> & Hartmut H. Niemann <sup>1</sup>✉

Several gram-negative bacteria employ type III secretion systems (T3SS) to inject effector proteins into eukaryotic host cells directly from the bacterial cytoplasm. The export gate SctV (YscV in *Yersinia*) binds substrate:chaperone complexes such as YscX:YscY, which are essential for formation of a functional T3SS. Here, we present structures of the YscX:YscY complex alone and bound to nonameric YscV. YscX binds its chaperone YscY at two distinct sites, resembling the heterotrimeric complex of the T3SS needle subunit with its chaperone and co-chaperone. In the ternary complex the YscX N-terminus, which mediates YscX secretion, occupies a binding site within one YscV that is also used by flagellar chaperones, suggesting the interaction's importance for substrate recognition. The YscX C-terminus inserts between protomers of the YscV ring where the stalk protein binds to couple YscV to the T3SS ATPase. This primary YscV–YscX interaction is essential for the formation of a secretion-competent T3SS.

<sup>1</sup>Department of Chemistry, Bielefeld University, Universitaetstrasse 25, 33615 Bielefeld, Germany. ✉email: [hartmut.niemann@uni-bielefeld.de](mailto:hartmut.niemann@uni-bielefeld.de)

Type III secretion systems (T3SSs; also: injectisomes, virulence-associated  $\nu$ T3SS) are a common device of bacterial virulence that is present in gram-negative bacteria such as *Salmonella*, *Pseudomonas*, or *Yersinia*<sup>1</sup>. The macromolecular complex is composed of approximately 20 different protein species, many of which are present in multiple copies, and spans both bacterial membranes. Injectisomes share high structural similarity and likely a common ancestor with the secretion system found in bacterial flagella ( $f$ T3SS). With the needle-like character of purified injectisomes<sup>2</sup> and their capability to secrete effector proteins directly into a host cell, the T3SS can also be described as a ‘molecular syringe’ and may be a useful protein delivery tool<sup>3</sup>.

Secretion through the *Yersinia* T3SS is strictly regulated by the inner membrane-anchored export apparatus containing YscR-STUV and can be divided into three distinct phases: (i) ‘early’ substrates, which self-assemble into the extracellular needle; (ii) ‘intermediate’ substrates, hydrophobic translocator proteins that insert themselves into the host membrane, thereby forming a direct connection between bacterial and host cytosols; and (iii) ‘late’ substrates or effector proteins that are then directly funneled through the injectisome into the host cytoplasm<sup>4</sup>. Many substrates of the T3SS are transported to the export apparatus in complex with a type III secretion chaperone (T3SC), which keeps the substrate in a partially unfolded and secretion-competent conformation. The ATPase complex then separates the complex under ATP-hydrolysis<sup>5</sup> and allows for the secretion of the substrate in an unfolded or partially helical state<sup>6</sup>.

T3SCs can further be subdivided into three classes. Class I chaperones bind effector proteins before injection. Class II binds to hydrophobic translocator proteins. Class III is a heterogeneous group and contains chaperones of flagellin in  $f$ T3SS or the injectisomal needle subunit SctF (YscF). Class II chaperones contain three tetratricopeptide repeat (TPR) motifs<sup>7</sup> where one TPR is made up of two antiparallel  $\alpha$ -helices. The tandem arrangement of TPRs generates a convex outer side that forms extensive interactions with the translocator and a concave binding groove where a short stretch of the substrate molecule binds in an extended conformation<sup>8–11</sup>. Additionally, a gatekeeper protein is involved to ensure timely secretion of translocators before effector proteins<sup>12</sup>. For the class III chaperone of the needle component SctF, two structures (YscEFG from *Yersinia* and PscEFG from *Pseudomonas*) have been published<sup>13,14</sup>. Both show heterotrimeric complexes of SctF with the two chaperones YscG/PscG and YscE/PscE, where SctF binds into the concave groove of the TPR protein YscG/PscG via a short  $\alpha$ -helix. Interestingly, the second chaperone YscE/PscE, which mainly consists of two  $\alpha$ -helices, caps the N-terminus of YscG/PscG but barely interacts with the substrate.

Injectisomes can be divided into seven families<sup>15</sup>, such as the Inv-Mxi-Spa family present in *Salmonella enterica*, *Shigella* sp. or *Escherichia coli*, or the Ysc family associated with *Yersinia* sp., *Pseudomonas aeruginosa*, *Aeromonas* sp. or *Phototrabdus luminescens*. Ysc family T3SS produce two proteins, SctX and SctY (YscX and YscY in *Yersinia*), that are not found in other T3SS families. Functionally, little is known about these two proteins except that they are required for the export of early substrates and hence the assembly of a functional T3SS<sup>16–18</sup>. YscX and its *Aeromonas* homolog AscX are secreted with YscY acting as chaperone for YscX prior to secretion<sup>17,19</sup>. YscX is not an early substrate, since its secretion occurs after the YscP-mediated substrate specificity switch<sup>18</sup>. It has been shown that the N-terminus of YscX harbors its secretion signal, while its C-terminus is essential for the formation of a secretion-competent T3SS<sup>17</sup>. While no experimental structures are available for either protein, YscY can be assigned a role as T3SC based

on its sequence which predicts 3 TPRs<sup>20</sup>. The structure of YscX remains elusive, given that its only homologs are other uncharacterized SctX proteins. The binding mode appears to be conserved between different SctX and SctY, as heterologous complexes from different species have been observed via yeast two-hybrid assays<sup>21</sup>. After separation from YscX, YscY probably serves a regulatory function within the bacterial cell due to its interaction with SycD<sup>22,23</sup>, the chaperone escorting the hydrophobic translocator proteins YopB and YopD.

The interaction between YscX, YscY, and the major export apparatus protein YscV has been established to be a prerequisite for the timely secretion of early substrates and the tripartite complex SctVXY is conserved between different T3SS-carrying species<sup>18,21</sup>. SctVs belong to the FlhA superfamily and as such are structurally similar to their flagellar FlhA counterparts. The export gate may receive substrates previously bound by a cytosolic sorting platform<sup>24,25</sup>. Over the years, several structures of FlhA and SctVs have been published (FlhA:<sup>26–28</sup> SctV:<sup>29–34</sup>), recently amended by the electron microscopy (EM) structure of *Yersinia* YscV as a homo-nonamer<sup>35</sup>. In general, SctVs contain an N-terminal transmembrane domain followed by a linker and a large cytosolic domain. It has been shown that SctVs form oligomeric structures in vivo that are consistent with the nonameric ring observed in different crystal and EM structures<sup>36</sup>. The cytosolic domain can further be subdivided into four subdomains (SD1–SD4), where SD2 is an insertion in SD1. SD3 and SD4 are necessary for the formation of the nonameric ring<sup>30,37</sup>. NMR data of a ternary complex composed of the cytosolic domain of *Salmonella* FlhA, the substrate FlhC and its chaperone FlhS has demonstrated that for a flagellar T3SS, recognition of the substrate:chaperone complex is facilitated via the chaperone and a conserved region in SD2 of FlhA<sup>38</sup>. It is noted, however, that the recognition site is not conserved in injectisomal SctVs and deletion of the entire SD2 did not abolish secretion of translocator and tip proteins for *Shigella* MxiA<sup>30</sup>. Therefore, the binding of substrate:chaperone complexes might occur at a different position in virulence-associated T3SS than in  $f$ T3SS.

Here, we report the structure of the YscX:YscY complex alone, and as a tripartite complex with the cytosolic domain of YscV. Our results demonstrate that the binding mode of this injectisomal substrate:chaperone complex is distinct from the binding of flagellar chaperones to the export gate and they explain why the C-terminus of YscX is essential for its function as injectisomal component.

## Results

**Structure determination of the YscX:YscY complex.** To structurally characterize the two proteins YscX and YscY, we first attempted to express them independently as YscX fused to maltose-binding protein (MBP-YscX) and His<sub>6</sub>-YscY in *E. coli*. While YscY was expressed insolubly, MBP-YscX was observed in the soluble fraction but showed heavy precipitation after cleaving off the solubility-promoting MBP-tag. We, therefore, tried co-expression of the two constructs, which yielded soluble protein. Each protein could pull down the other during affinity chromatography and YscY and YscX co-eluted in a symmetrical peak when subjected to size exclusion chromatography. Since stored samples of the complex showed clear signs of proteolytic digestion of YscX, limited proteolysis coupled with N-terminal sequencing and mass spectrometric analysis was conducted to identify a more stable N-terminally truncated version of YscX (YscX<sub>32</sub>). Neither the C-terminal region of YscX nor YscY were sensitive towards proteolysis. A study by Day and Plano<sup>17</sup> further suggested that only the region YscX<sub>50–110</sub> is necessary for it to bind YscY, so further constructs truncating the N-terminus

(YscX<sub>50</sub>), the C-terminus (YscX<sub>32–110</sub>), or both (YscX<sub>50–110</sub>) were created.

Our crystallization constructs contained N-terminally truncated YscX<sub>50</sub> and full-length YscY with an N-terminal hexahistidine tag. The complex crystallized in space group *P2*<sub>1</sub>*2*<sub>1</sub>*2* with the asymmetric unit (AU) containing two heterodimers related via translational non-crystallographic symmetry. We solved the structure by molecular replacement (MR) using a model generated by AlphaFold 2 (Supplementary Fig. 1). Data collection and refinement statistics are reported in Supplementary Table 1. The two dimers in the AU are very similar with an r.m.s.d. value of 0.744 Å between the Ca residues of the complex. Differences lie mostly within YscX—where the r.m.s.d. between the two molecules is 0.904 Å—caused by a slight shift in the second helix ( $\alpha$ 2) of the protein.

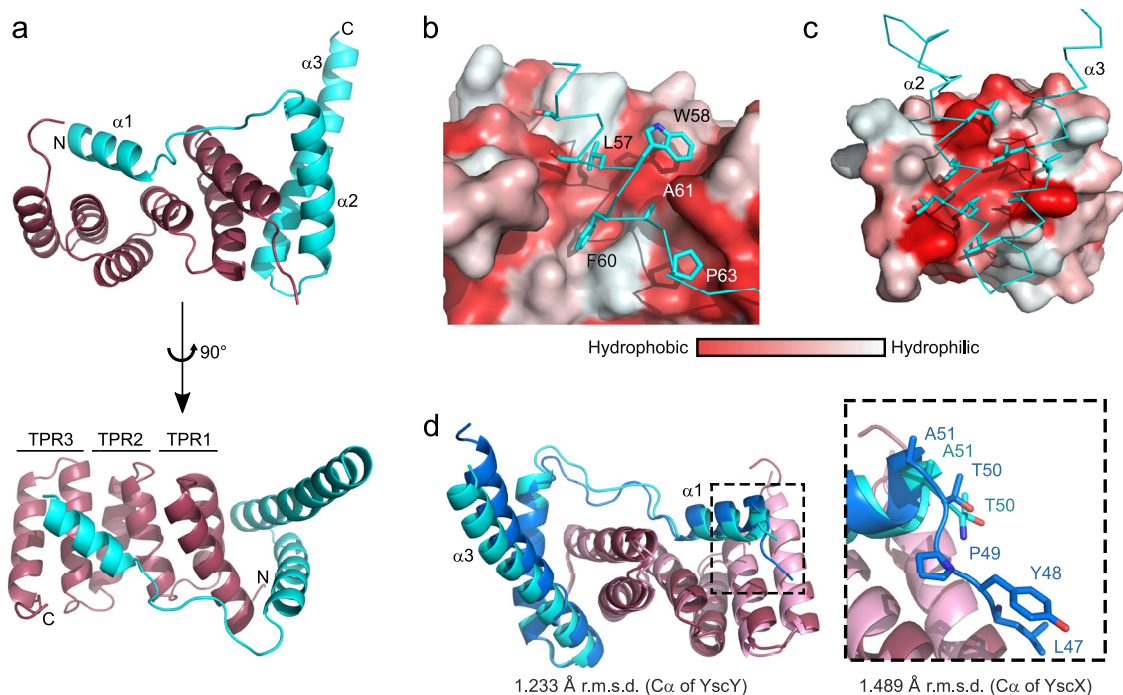
**YscX and YscY form an entwined heterodimer.** The AlphaFold 2 model correctly predicted the overall fold of the complex with YscX binding to YscY at two distinct sites connected via a short linker, resulting in YscX entwining its chaperone (Fig. 1a). Similar to other T3SCs the concave face of YscY forms an interaction site to which YscX binds via a short helix  $\alpha$ 1 (S53–A61). The area of interaction for this site is approximately 500 Å<sup>2</sup>. Moreover, this groove is largely lined by hydrophobic residues and allows for the binding of YscX's L57, W58, F60, A61, and P63 into distinct binding pockets. Hydrophilic residues such as K54, R55, and D59 point toward the solvent (Fig. 1b). S53 and S56 of YscX are 3.8 and 3.9 Å apart from Y101 and Y104 of YscY, respectively, making their contribution to the binding interface via hydrogen bonds unlikely. The importance of the amphipathic character of the binding helix in YscX is further underlined by the high degree of sequence conservation for the hydrophobic residues that interact with the chaperone when compared to the otherwise low

conservation of helix  $\alpha$ 1 (Fig. 2a). The linker between the binding helix and the C-terminal helices of YscX is only resolved in one of the two heterodimers in the AU, indicating the flexibility of the loop.

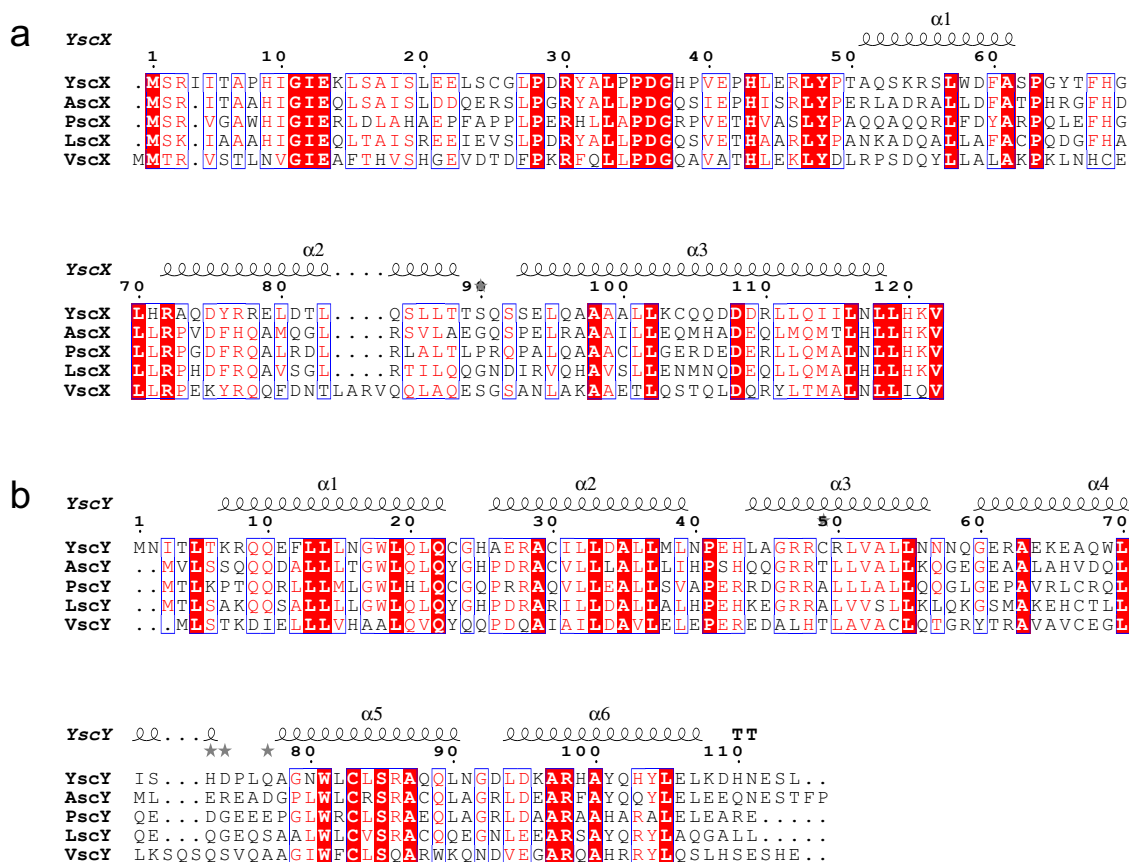
A second interface is formed via the interaction of YscY's N-terminal TPR and the two longer C-terminal  $\alpha$ -helices ( $\alpha$ 2 and  $\alpha$ 3) of YscX. The helices of YscX are rotated by about 90° relative to YscY with the binding site containing the hydrophobic residues I3, L5, F12, L13, I31, L32, A35, and L39 for YscY as well as L80, L83, L86, L87, L95, A98, L101 and L102 for YscX (Fig. 1c). The total area of interaction averages out at 1436.7 Å<sup>2</sup> for both YscXY heterodimers (PISA<sup>39</sup>).

The C-terminal helix of YscX extends about 26 Å further out from where it last interacts with its chaperone, as measured from L102 to H120. A hydrogen bond of Q106 to D81 is the last intramolecular interaction between  $\alpha$ 2 and  $\alpha$ 3 of YscX. Therefore, the remaining residues (about 20 Å of the C-terminal helix) are not forming interactions within the YscX:YscY heterodimer. In the crystal, this helix is stabilized by interacting with helix  $\alpha$ 3 of the neighboring molecule as well as the same helix from symmetry-related molecules, forming a 4-helix interface. B-factors for the protruding part of helix  $\alpha$ 3 are relatively high in chain B, interestingly however not in chain D despite its generally higher B-factors, especially in the poorly defined loop and helix  $\alpha$ 2 (Supplementary Fig. 2). Furthermore, K121 and V122 are not ordered in the crystal, suggesting that the C-terminus of YscX is somewhat flexible.

**The YscX N-terminus is disordered.** In addition to our structure of YscX<sub>50</sub>:YscY, we also crystallized a longer construct containing residues 32–122 of YscX (YscX<sub>32</sub>:YscY). Crystallization of this complex required the reductive methylation of primary amines<sup>40</sup>. The resolution was considerably poorer, with the protein



**Fig. 1 Overall structure of the YscX:YscY heterodimer.** **a** Cartoon representation of YscX (cyan) and YscY (red). YscX binds to YscY at two independent sites. **b, c** Surface hydrophobicity gradient of YscY shown from hydrophobic (red) to hydrophilic (white) using the normalized consensus hydrophobicity scale<sup>65</sup>. **b** YscX binds the groove of the TPR protein YscY via the hydrophobic side of a short amphipathic helix. **c** Two leucine-rich helices of YscX interact with the largely hydrophobic surface of the N-terminal TPR in YscY. **d** Superposition of YscX<sub>50</sub>:YscY (cyan and red) and YscX<sub>32</sub>:YscY (blue and pink) models. The alignment was calculated using the Ca atoms of YscY. Zoom-in shows the YscX N-terminus before  $\alpha$ 1 for both structures when aligned using the Ca atoms of YscX.



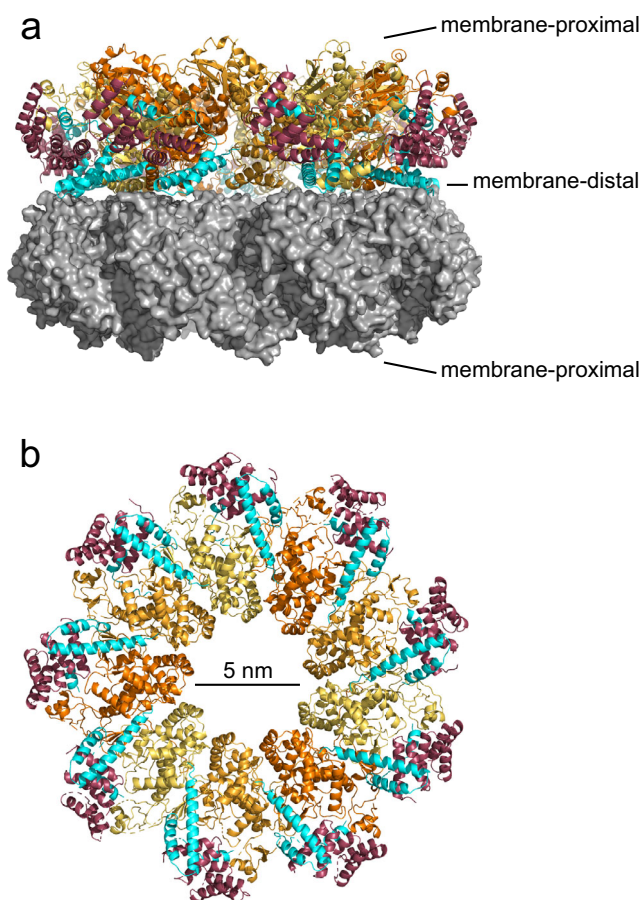
**Fig. 2** Sequence alignment of SctX and SctY. **a** Alignment of SctX. **b** Alignment of SctY. Alignments were calculated using ClustalW<sup>66</sup> and visualized with ESPrnt3 (esprnt3.ibcp.fr,<sup>67</sup>). SctX and SctY sequences from *Yersinia enterocolitica* (YscX: AAD16820.1, YscY: AAK69217.1), *Aeromonas hydrophila* (AscX: AAS91814.1, AscY: AAS91815.1), *Pseudomonas aeruginosa* (PscX: AAY17109.1, PscY: AAY17110.1), *Photobacterium luminescens* (LscX: CAE16135.1, LscY: CAE16134.1), and *Vibrio parahaemolyticus* (VscX: UJX28051.1, VscY: EDM61220.1) were used to generate the alignment. Definition of colors and symbols: red box, white character: strict identity; red character: similarity in a group; blue frame: similarity across groups; gray star: residues with alternate conformations.

adopting a loose packing in space group  $I4_122$  with a solvent content of 64.3% (Supplementary Table 1). The overall fold is identical to that of the shorter construct and the structures align well with r.m.s.d. values of 2.033 Å between YscX<sub>32</sub>:YscY and chain A and B of YscX<sub>50</sub>:YscY or 1.388 Å with chain C and D of the latter. This is compared to an r.m.s.d. of 0.744 Å between the two complexes in the AU of the YscX<sub>50</sub>:YscY crystal. In general, the r.m.s.d. values were a bit higher for YscX (1.489 Å or 1.383 Å for chain B or chain D, respectively) than for YscY (1.233 Å or 1.078 Å) and the models differ the most in the position of the loop between helices  $\alpha 1$  and  $\alpha 2$  of YscX and in a slightly increased bending at the C-terminus of  $\alpha 3$  in the model for YscX<sub>32</sub>:YscY (Fig. 1d). Even though this YscX construct is 18 residues longer than YscX<sub>50</sub>, only L47–P49 of these are resolved, suggesting that the N-terminus is largely disordered. Moreover, the C-terminus in this model is fully resolved and is helical albeit with high B-factors for the last residues (Supplementary Fig. 2).

**Structure determination of YscX:YscY bound to nonameric YscV.** We then attempted to co-purify the cytosolic domain of YscV (YscV<sub>C</sub>) with complexes of YscX and YscY from *E. coli*. Yields were considerably higher for the ternary complex (30 mg per liter of culture) than for the heterodimer (1–5 mg per liter). During gel filtration, the complex eluted as a species of high molecular weight. The assembly crystallized readily and conditions were identical regardless of which YscX construct (YscX<sub>full-length</sub>, YscX<sub>32</sub>, or YscX<sub>50</sub>) was used. Diffraction measurement of the

ternary complex containing YscV<sub>C</sub>, YscX<sub>32</sub>, and His<sub>6</sub>-YscY gave data to a resolution of 4.1 Å (Supplementary Table 1). We tested approximately 100 crystals of this complex with most not diffracting further than 8 Å if at all. The phase problem was solved via MR with an EM model of nonameric YscV (PDB: 7ALW). Two nonamers are stacked upside-down with their membrane-distal sides interacting in an assembly probably lacking biological significance also described by Kuhlen et al.<sup>35</sup> (Fig. 3a). We then attempted to place the model of YscX<sub>50</sub>:YscY via MR as well, which failed most likely due to its relatively low scattering mass. Instead, 18 binary complexes were placed manually into helical density present outside of the nonameric YscV rings (Supplementary Fig. 3). The crystal described here is packed loosely with a solvent content of 67.2% that originates both from the central 5 nm-wide pore of the cyclic YscV<sub>C</sub> nonamer and the fact that the stacked nonamers hardly interact with one another, leaving a large volume of bulk solvent between the 18-mers (Supplementary Fig. 4). Instead, almost all crystal contacts occur between YscY molecules or between an YscY molecule and SD2 of YscV<sub>C</sub> which also points out from the ring.

**Overall structure of nonameric YscV<sub>C</sub>.** In general, the electron density was reasonably well defined for subdomains 3 and 4 (SD3 and SD4), which are mostly buried within the ring. The average area of the YscV-YscV interaction is approximately 1600 Å<sup>2</sup> and is stabilized by four salt bridges between E431-R560, D511-R592, E517-R563, and R519-E531. The electron density of SD1 was



**Fig. 3 Overall structure of the ternary YscV<sub>C</sub>:YscX<sub>32</sub>:YscY complex.**

**a** Side View of the two stacked nonameric rings of YscV<sub>C</sub> (orange, three shades for clarity). Half of the chains are shown as gray surface. YscX<sub>32</sub> (cyan) and YscY (red) bind on the periphery of the nonamer. **b** View on the membrane-distal side of one nonamer. The C-terminal helix of YscX protrudes into the ring between two adjacent YscV<sub>C</sub> protomers.

interpretable, even though the central 4-stranded  $\beta$ -sheet was only visible as a continuous sheet of electron density. In contrast, SD2 showed the lowest map quality with broader variations between the protomers with regard to connectivity of the density. The cleft formed between SD2 and SD4 is in an open conformation, with a distance of roughly 10 Å between the two domains at the cleft's entrance. Parts of the linker between transmembrane and cytoplasmic domain are also visible (residues 359–371) and fold onto the neighboring protomer, where it runs along a defined groove, as previously observed<sup>35</sup>. We confirmed the sequence register of YscV<sub>C</sub> via anomalous difference density from a selenomethionine derivative (Supplementary Fig. 5).

The individual chains of the complex superimpose well (Supplementary Fig. 6a), especially the rigid SD3 of YscV<sub>C</sub>. More flexibility can be observed for SD2 and at the outer regions of SD1 and SD4. YscX chains are very similar except for the linker between helices  $\alpha$ 1 and  $\alpha$ 2, which was poorly defined (Supplementary Fig. 6e). The binding of YscX:YscY to YscV<sub>C</sub> seems to allow some leeway, since an alignment of all substrate or all chaperone chains results in fanning out of the YscV<sub>C</sub> chains (Supplementary Fig. 6d–g).

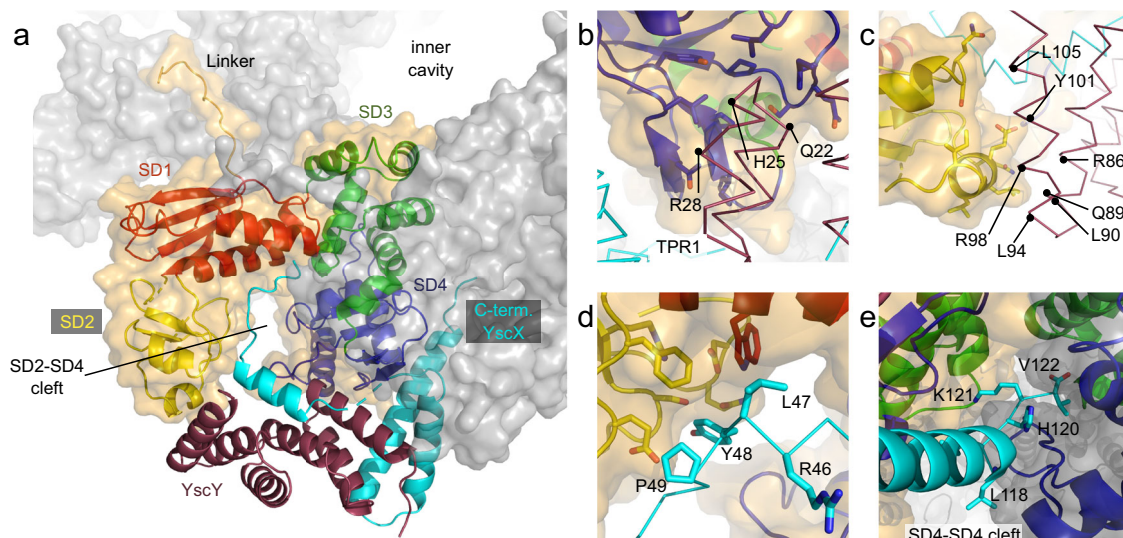
**YscY and the YscX N-terminus bind the SD2-SD4 cleft.** Each YscX:YscY complex binds to the SD2-SD4 cleft of an individual YscV<sub>C</sub> protomer. The PISA server<sup>39</sup> reports interactions between

(A) YscY's first TPR and SD4, (B) the third TPR and SD2, as well as (C) a putative interaction between the extended N-terminus (L47 and Y48) of YscX with SD1 from inside the cleft between SD2 and SD4 (Fig. 4b–d). The latter interaction was only observed for chains FB, GB, KB as well as LB and may be biologically relevant since these residues are conserved between SctX. Due to the poor resolution, the exact nature of these interactions could not be determined as density was usually absent for side chains in YscX and YscY. For the aforementioned site (A) hydrophobic interactions could originate from I604, P606, L652, I655, P659, and L699 of YscV, and hydrogen bonds could arise between Y608, Q656, S657, and Q697 of YscV with Q22, H25, and R28 of YscY. Several leucines and isoleucines can be found in site (B), which may be complemented by polar interactions between Q452, E454, S486, and Q487 of SD2 as well as R86, Q89, R98, and Y101 of YscY. In case of site (C), the conserved Y48 of YscX might be interacting with aromatic residues of YscV in vicinity, such as F491 and W499 of SD1. Furthermore, Y48 is pointing towards a patch of different possible hydrogen bonding partners, such as R447, S449, D469, Q470, and E489 of YscV.

**The C-terminus of YscX binds between YscV protomers.** The C-terminal  $\alpha$ 3 helix—which is the extended helix seen in the YscX:YscY structures—is inserted between two YscV<sub>C</sub> protomers (Fig. 3b) at the cleft formed between two adjacent SD4 (hereafter referred to as SD4–SD4 cleft). This is the major binding interface between YscV and YscX and involves SD3 and SD4 of YscV. Except for a salt bridge formed between D108 of YscX and R701 of YscV<sub>C</sub> directly at the entry point for the helix, the interaction is largely hydrophobic (Fig. 5a) with most residues involved being highly conserved (Fig. 2a). The C-terminal helix of YscX does not extend all the way into the inner hole of the YscV nonamer but instead ends contacting the neighboring YscV molecule. The negative charge of the carboxy-terminus at V122 is presumably neutralized by R551 and R669 of a second YscV molecule (Figs. 4e and 5). Mutating R551 to alanine (R551A) or glutamate (R551E) reduced or abolished the ability of YscV<sub>C</sub> to be pulled down by MBP-YscX<sub>50</sub>:His<sub>6</sub>-YscY in MBP pull-down assays (Fig. 6a), indicating the importance of charge neutralization for recognition of YscX by the export gate.

**YscV requires the C-terminus of YscX for its recognition.** From the ternary complex structure, it became clear that the major mode of binding between the YscX:YscY heterodimer and the nonameric YscV is via an insertion of YscX's C-terminus between two YscV monomers within the ring (Figs. 3 and 4a) with further interactions involving the chaperone and SD2 or SD4 as well as residues L47 and Y48 of YscX. To test whether the inserting helix of YscX is necessary for binding to YscV, an MBP pull-down assay was employed, where MBP-YscX:YscY complexes were immobilized from lysate either with or without YscV<sub>C</sub> present. The SDS-PAGE (Fig. 6a) clearly demonstrated that the N-terminal truncation YscX<sub>50</sub> has no impact on binding, whereas deletion of the C-terminal 12 amino acids (YscX<sub>32–110</sub>:YscY or YscX<sub>50–110</sub>:YscY) abolished the interaction. This was underlined by analytical gel filtrations done with mixtures of either YscV<sub>C</sub> or a truncated version lacking SD3 and SD4 (YscV<sub>SD12</sub>, residues 356–513)—which therefore does not contain the binding site for the C-terminal helix of YscX—with YscX<sub>32</sub>:YscY. Since YscX<sub>32</sub>:YscY contains the putative binding region <sup>46</sup>RLYP<sup>49</sup>, it should be able to bind to SD1 via YscX and to SD2 via YscY.

On its own, YscX<sub>32</sub>:YscY tends to produce a three-peak pattern during gel filtration, most likely due to the formation of oligomers in solution (Fig. 6b). At a concentration of 10  $\mu$ M, YscV<sub>C</sub> forms a single peak at around 60 kDa apparent molecular mass. As



**Fig. 4 Interactions of YscV, YscX, and YscY.** One YscV<sub>C</sub> protomer is highlighted (orange surface) within the nonameric ring (gray surface). The cartoon representation of YscV<sub>C</sub> is divided into the four subdomains: SD1 (red, residues 372–438 and 493–513), SD2 (yellow, residues 439–492), SD3 (green, residues 514–600), and SD4 (blue, residues 601–704). For clarity, YscX (cyan) and YscY (dark red) are shown only for this heterotrimer and are depicted as ribbon in (b–d). **a** The overall structure of the heterotrimer within the assembly shows that all interactions, except for the insertion of YscX's C-terminus, are confined within one heterotrimer, i.e., YscY does not interact with a neighboring YscV subunit. **b** YscY interacts with SD2 via the loop of its first TPR (site A) and **c** with SD2 via the face of its C-terminal TPR (site B). **d** In chains FB, GB, KB, and LB, YscX exhibits density for an extended N-terminus which binds to SD1 and SD2 via a conserved <sup>46</sup>RLYP<sup>49</sup> motif (site C). Note that the exact positions of the side chains could not be determined at this resolution and are only included for completeness. **e** The C-terminus of YscX is inserted into the SD4–SD4 cleft. Both involved YscV<sub>C</sub> subunits are shown as cartoon colored by subdomain. YscX is shown as ribbon with sidechains starting from L118.

discussed later, the shift from the expected weight of 40.5 kDa was attributed to dynamic assembly and disassembly of the nonameric ring during the run (Fig. 7). With a mixture of YscV<sub>C</sub> and the binary complex, a clear shift towards higher apparent molecular masses can be seen as a peak around 10 mL retention volume (Fig. 6b top, Supplementary Table 2, Supplementary Fig. 8), which we attributed to a nonameric ring of YscV<sub>C</sub> with an unknown quantity of YscX:YscY bound to it. For the ternary complex, two more peaks were observed at approximately 14 and 15 mL, the latter of which could be attributed to free YscV<sub>C</sub> in solution or a 1:1:1 ternary complex for which a molecular mass of 65 kDa would be expected. The large and broad signal at around 14 mL might also indicate other oligomeric forms and a dynamic equilibrium between different states. The shortened YscV<sub>SD12</sub> construct elutes after approximately 16 mL, which corresponds to around 36 kDa, potentially representing a dimer (Fig. 6b, bottom). Here, no shift in mass was observed when the protein was mixed with YscX<sub>32</sub>:YscY in an equimolar fashion and thus we concluded that SD1 and SD2 alone do not bind YscX:YscY with high affinity. A construct of YscV<sub>C</sub> only containing SD3 and SD4 (residues 514–704) could not be analyzed as it was insoluble upon expression in *E. coli*.

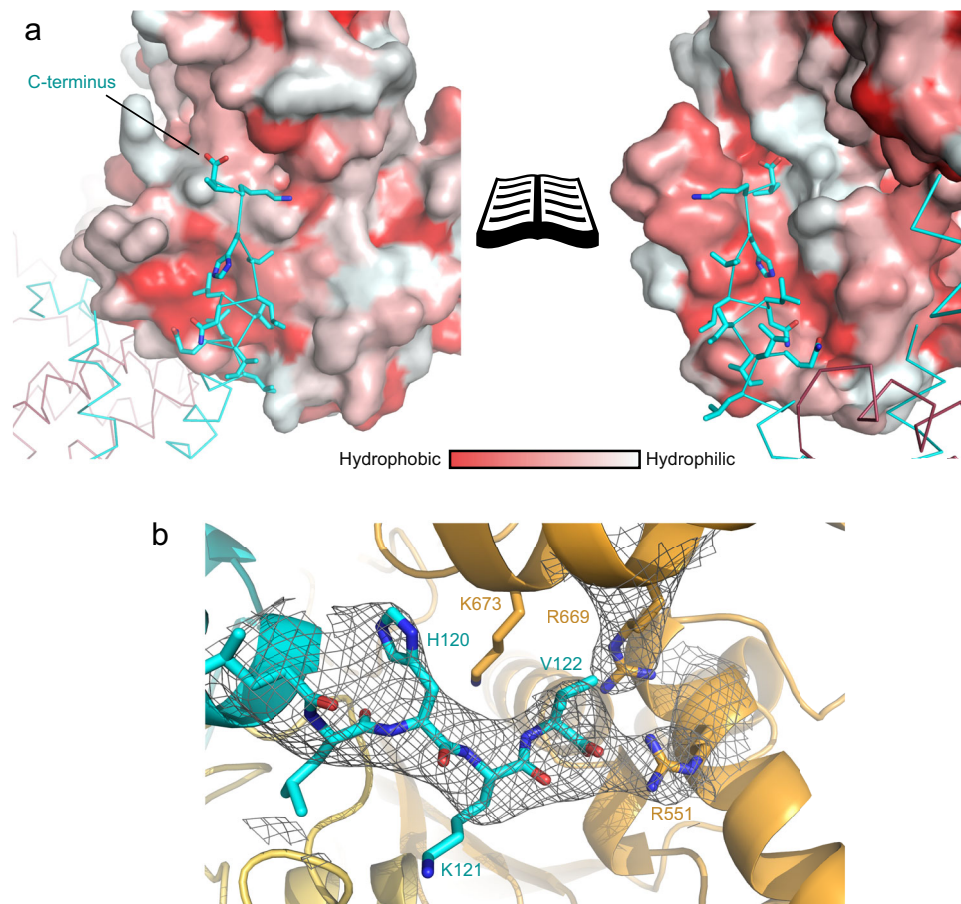
**The termini of YscX oppositely influence nonamerization of YscV<sub>C</sub>.** YscV forms higher oligomers in *Yersinia* regardless of whether YscX and YscY are present<sup>36,37</sup> and micelle-embedded YscV assembles into nonamers in vitro<sup>35</sup>. For the isolated cytosolic domain of other SctV proteins, both monomeric, as well as nonameric species, have been observed in solution<sup>31,32,34</sup>. With the major recognition site of YscX:YscY and YscV<sub>C</sub> being located at the far C-terminus of YscX and between two YscV protomers, we aimed at investigating a possible influence of YscX on the nonamerization of the export gate's cytosolic domain. Size exclusion chromatography of YscV<sub>C</sub> at different concentrations revealed a concentration-dependent dynamic equilibrium

between different states with the nonameric form being stabilized at higher concentrations (Fig. 7, top). At 2.5 μM initial concentration of YscV<sub>C</sub>, the protein appears to be largely monomeric as indicated by a symmetrical peak at around the expected molecular mass of 40.5 kDa. At 10 or 25 μM, we observed a shift in the apparent mass and significant fronting, indicating the dynamic association and dissociation of higher oligomeric species on the column. At 100 μM the main peak is shifted to a high apparent molecular mass corresponding to a nonamer, but there is still notable tailing, i.e., incomplete conversion to the nonameric ring. We then compared the oligomerization behavior in the presence of two different YscX:YscY constructs, which were co-purified with YscV<sub>C</sub>. At 10 μM, YscV<sub>C</sub>X<sub>50</sub>Y showed almost-complete conversion to higher masses. Interestingly, however, the full-length YscX-containing ternary complex (YscV<sub>C</sub>X<sub>fl</sub>Y) favored the monomeric species at the same molar concentration of protein (Fig. 7, middle and bottom).

Using the same concentration series as for YscV<sub>C</sub>, the nonameric ring of YscV<sub>C</sub>X<sub>50</sub>Y is stable even at an initial concentration of 2.5 μM, whereas full-length YscX seems to inhibit ring formation even at high concentrations (Fig. 7, bottom). We concluded that the N-terminus of YscX is not necessary for binding YscV, but it decreases the propensity of YscV<sub>C</sub> to nonamerize and therefore must interact in some other way.

## Discussion

YscX and YscY are largely uncharacterized proteins from the T3SS of *Yersinia* with a yet unknown function despite being essential for a functional injectisome. With YscX belonging to a unique and elusive protein family without any homologous structures, we were unable to solve the phase problem via MR. After laborious attempts at experimental phasing, we attempted to predict the structure of the two proteins computationally using I-TASSER<sup>41</sup>, tFold (to be published), RoseTTAfold<sup>42</sup>, and



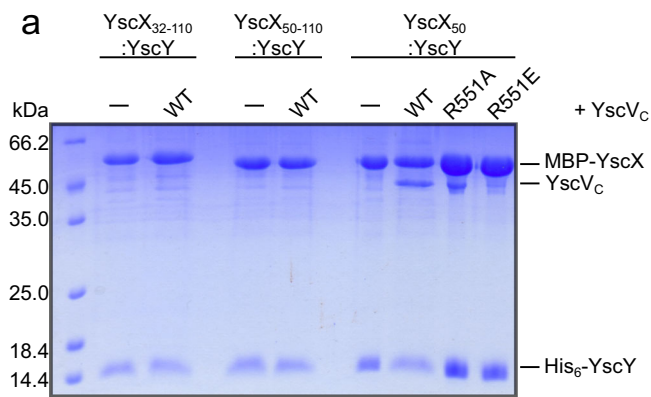
**Fig. 5 The C-terminus of YscX binds between two YscV protomers.** **a** Surface hydrophobicity of YscV (left) and the neighboring protomer (right) from hydrophobic (red) to hydrophilic (white) with a ribbon and stick representation of YscX (cyan). Helix  $\alpha 3$  inserts between two YscV molecules and binds via hydrophobic interactions. The C-terminus is pointing towards a positively charged patch of the neighboring subunit, which likely stabilizes the carboxylate. **b**  $2mF_o - DF_c$  density contoured at  $1\sigma$  at the C-Terminus of YscX suggests positioning of the carboxylate towards R551 of YscV (orange). The density of the arginines is exemplary and varies between the chains.

AlphaFold 2<sup>43</sup>. YscY was always predicted as a TPR protein and the general fold was very similar between programs (Supplementary Fig. 1b). YscX on the other hand produced models that were only similar for the C-terminal  $\alpha 2$  and  $\alpha 3$  helices (Supplementary Fig. 1a). Finally, we attempted to predict the complex structure using the same programs but supplying both sequences fused together by a 33-residue linker. While the other predictions were disregarded, we were intrigued by how compatible the AlphaFold 2 model was to our biochemical and crystallographic observations. Indeed, we could solve the phase problem for our YscX:YscY crystals with a truncated version of the AlphaFold 2 complex model. In retrospect, models of YscY were also sufficient for MR when generated by AlphaFold 2, tFold or RoseTTAfold, but not I-TASSER.

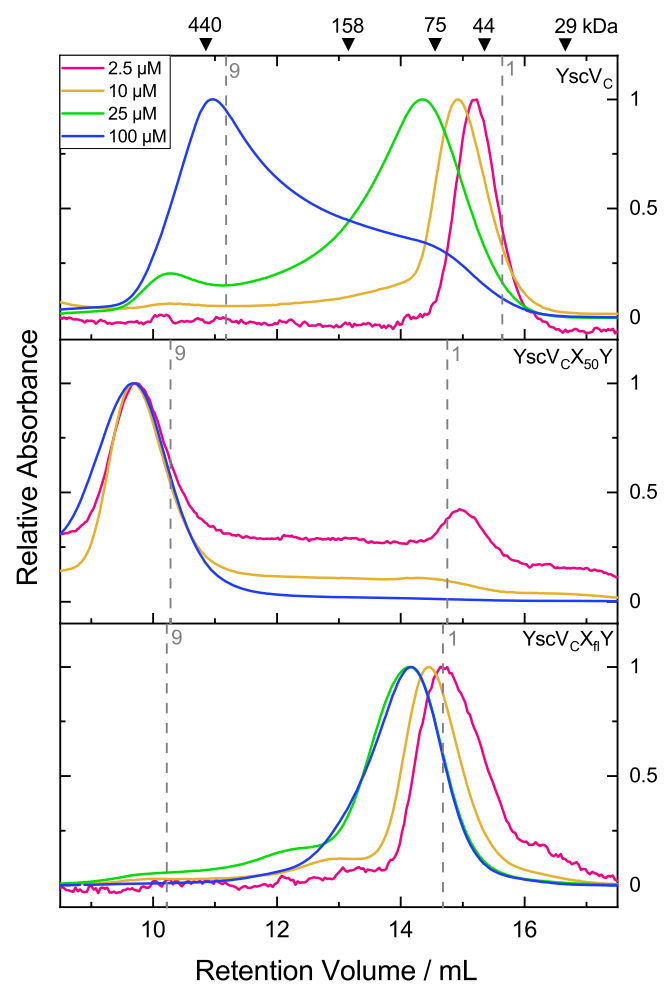
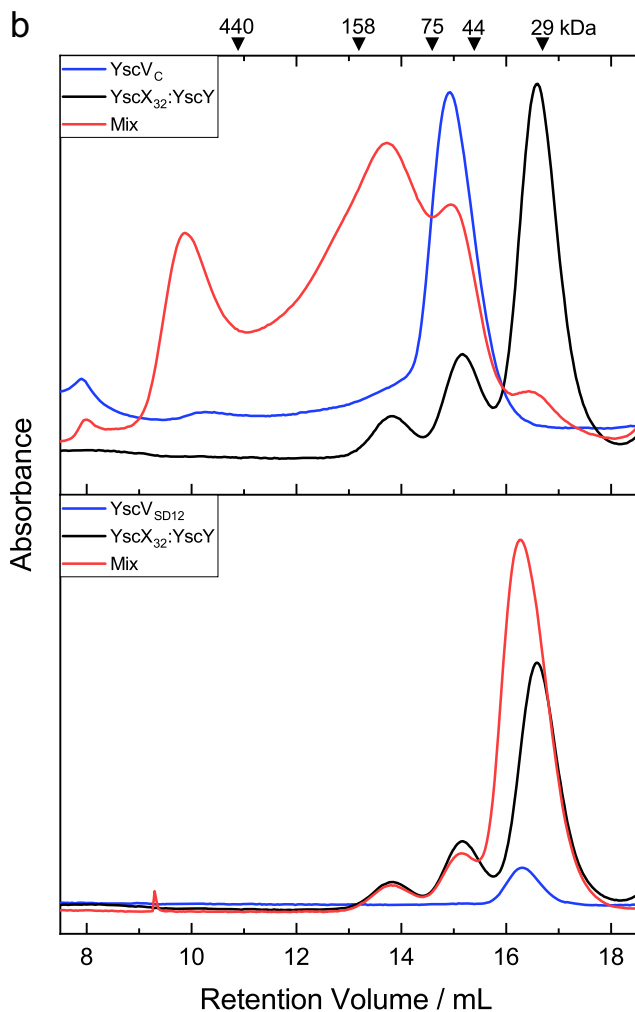
Here, we provide the structure of a SctX-family protein. YscX and YscY depend on each other for soluble expression in *E. coli*, which is understandable given how entwined the two proteins are in their complex. Lacking YscX, YscY would present a large hydrophobic face at its N-terminal TPR that could explain why it is insoluble in isolation. After secretion of its substrate, this face might be responsible for recognizing SycD, an interaction previously observed<sup>22,23</sup>. We found that the N-terminal 31 residues of YscX are prone to proteolysis, indicating that they may not exhibit a tertiary structure. It is possible that the N-terminus carries a proteinaceous signal for YscX secretion, especially considering how conserved the sequence is among the SctX family (Fig. 2a). The smallest YscX fragment to direct secretion (amino

acids 1–60) contains helix  $\alpha 1$ <sup>17</sup>, leaving the possibility that YscY bound to helix  $\alpha 1$  of YscX is part of the secretion signal. This does not seem to be the case for the structurally similar YscEFG complex, for which the transport of YscF to the injectisome depends on its co-chaperone YscE, which structurally fulfills the role of  $\alpha 2$  and  $\alpha 3$  in YscX<sup>44</sup>. Alternatively, the structurally unresolved N-terminus of YscX might bind other T3S proteins and only fold in their presence, as was observed for YopN in presence of TyeA in the same secretion system<sup>45</sup>. It is unclear whether YscX serves a hitherto unknown function outside the bacterial cell after its secretion or if its absence from the export apparatus signals a switch in substrate recognition, as was postulated by Day, et al.<sup>17</sup>.

When compared to other T3S substrate:chaperone complexes, the structural similarities between YscXY and the complex of the needle protein with its two chaperones become obvious (Fig. 8a). Two structures of such a complex, one from *Pseudomonas* (PscEFG, PDB: 2UW)<sup>14</sup> and the other from *Yersinia* (YscEFG, PDB: 2P58<sup>13</sup>) have been published, both revealing an amphipathic SctF helix that binds into a hydrophobic groove formed by the concave side of the TPR protein YscG/PscG. The co-chaperone YscE/PscE was found to bind the hydrophobic face of the N-terminal TPR of YscG/PscG with its two helices that contribute several leucine residues to the binding interface. Thus, it appears as if YscX fulfills the roles of both “co-chaperone” and “substrate” in this assembly reminiscent of the YscEFG complex. In contrast, class II substrate:chaperone complexes—which are



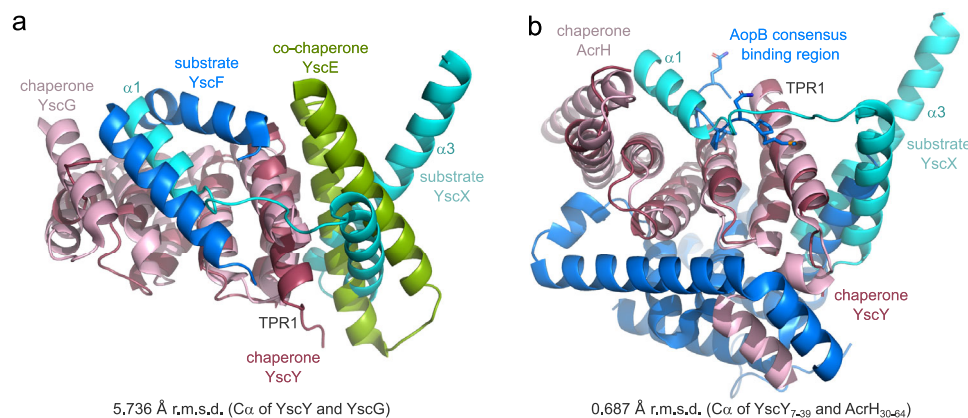
**Fig. 6 Recognition of YscXY by YscV<sub>C</sub> depends on the C-terminus of YscX as well as SD3 and SD4 of YscV<sub>C</sub>.** **a** Coomassie-stained SDS-PAGE of MBP pull-downs. N- and C-terminally truncated MBP-YscX constructs were co-expressed with His<sub>6</sub>-YscY and YscV<sub>C</sub>. Lysates were loaded onto amylose resin and the washed matrix loaded onto the gel. The pull-downs generated consistent results over three replications. Source data are provided as a Source Data file. **b** Analytical size exclusion chromatography with 10 μM YscV<sub>C</sub>, YscV<sub>SD12</sub> (both blue), YscX<sub>32</sub>:YscY (black), or mixtures (red) of the proteins. Absorbance at 280 nm shows a shift in retention volume for a mixture of YscV<sub>C</sub> with YscX<sub>32</sub>:YscY, but not YscV<sub>SD12</sub> with YscX<sub>32</sub>:YscY. Chromatographies involving YscV<sub>SD12</sub> or YscV<sub>C</sub> were performed twice or four times, respectively, with consistent results. A SDS-PAGE and Western blot analyzing contents of gel filtration peaks are shown in Supplementary Fig. 8.



**Fig. 7 Concentration-dependent nonamerization of YscV<sub>C</sub>.** For comparability, absorbances were normalized to the highest peak. From top to bottom: YscV<sub>C</sub>, a ternary complex of YscV<sub>C</sub> co-expressed with YscX<sub>50</sub> and His<sub>6</sub>-YscY, and the same ternary complex but with full-length YscX (YscX<sub>fl</sub>) were loaded at 2.5 μM (pink) to 100 μM (blue) onto a Superdex 200 column. For reference, the expected retention volumes for the monomeric species (or 1:1:1 heterotrimer) and nonameric species (assuming 9:9:9 stoichiometry and a globular shape) are shown as dashed gray lines. The weak auto-nonamerization of YscV<sub>C</sub> is enhanced by the presence of YscX<sub>50</sub> but severely impeded by full-length YscX. Experiments were performed three times with consistent results.

hydrophobic translocator proteins and their chaperones—form soluble 1:1 heterodimers like YscXY, but exhibit vastly different binding behaviors. A short stretch close to the N-terminus of the substrate binds to the groove formed by the chaperone's concave side in extended conformation (Fig. 8b, refs. 8–11,46). The structure of *Aeromonas* AcrH with translocator protein AopB shows further interactions between the two proteins. Analogous to the YscEFG and YscXY complex, AopB binds the N-terminal TPR of AcrH via an  $\alpha$ -helix. In addition, AopB extensively wraps around the convex side of the chaperone, a feature neither present in YscXY nor YscEFG<sup>11</sup>. Thus, despite its heterodimeric nature, the





**Fig. 8 Comparison of YscXY with other T3S substrate:chaperone complexes.** **a** Superposition of YscX<sub>50</sub>:YscY with YscEFG (PDB: 2P58). The alignment was done using the chaperones YscY (red) and YscG (pink) with an r.m.s.d. of 5.736 Å. The substrates YscX (cyan) and YscF (blue) as well as the co-chaperone YscE (green) are also shown. **b** Superposition of YscXY with the complex of class II chaperone AcrH (pink) and the major translocator protein AopB (blue), both from *Aeromonas* (PDB: 3WXX). The alignment was calculated using the first TPR of YscY (residues 7–39) and the first TPR of AcrH (residues 30–64) with an r.m.s.d. of 0.687 Å to obtain the right register. For clarity, sticks are shown for side chains of AopB where it binds the TPR groove of AcrH.

YscXY complex more closely resembles the heterotrimeric YscEFG complex, suggesting structural similarity to class III chaperones.

Since structures of FlhA were solved in 2010, several flagellar and injectisomal export gates were deposited in the PDB (FlhA:<sup>26–28</sup> SctV:<sup>29–33,35</sup>). Some discussion is dedicated to elucidating differences in how the subdomains are arranged relative to each other, especially considering the cleft between SD2 and SD4 that is responsible for the recognition of the flagellar FliD:FliT and FliC:FliS substrate:chaperone complexes<sup>38</sup>. The structure of YscV<sub>C</sub> was previously reported as adopting the “open” conformation with space between SD2 and SD4<sup>35</sup> when present without additional binding partners. Our structure of the ternary complex YscV<sub>C</sub>:YscX<sub>32</sub>:YscY exhibits the same open conformation showing that no major domain rearrangements are necessary to accommodate YscXY (Fig. 9a). Minor movements are observed at the binding site between YscY and SD4 as well as helix  $\alpha 3$  of YscX and SD4 but these do not seem to affect the export gate globally (Supplementary Fig. 7). Our model also establishes that a highly conserved <sup>46</sup>RLYP<sup>49</sup> motif of YscX binds within the aforementioned SD2–SD4 cleft. While we could model this region only in some of the chains, considerable difference electron density is present in the remaining chains indicating that this binding is a consistent feature. Interestingly, the binding site coincides with the recognition site of flagellar FliT and FliS chaperones (Fig. 9b, c). The binding of YscXY to YscV as seen in our structure might block access of other chaperone-bound secretion substrates to the SD2–SD4 cleft and thereby impede their secretion.

The major binding site for YscX lies between the fourth subdomain of two adjacent subunits of the YscV nonamer, where the C-terminal  $\alpha 3$  helix of YscX inserts, pointing towards the inner cavity of the ring but not reaching it. While we initially thought that this hints at YscX stabilizing an otherwise dissociation-prone YscV<sub>C</sub> nonamer in vivo, our gel filtration results indicate that full-length YscX actually disfavors nonamerization of the export gate (Fig. 7 bottom). The importance of the YscX C-terminus for recognition by YscV explains why complementation of *yscX* deletion strains with C-terminally FLAG-tagged YscX does not produce an injectisome competent to secrete Yops<sup>17</sup>. The same construct, however, was secreted by a functional T3SS<sup>17</sup>, suggesting that the interaction of YscXY with the SD2–SD4 cleft is sufficient for the secretion of YscX.

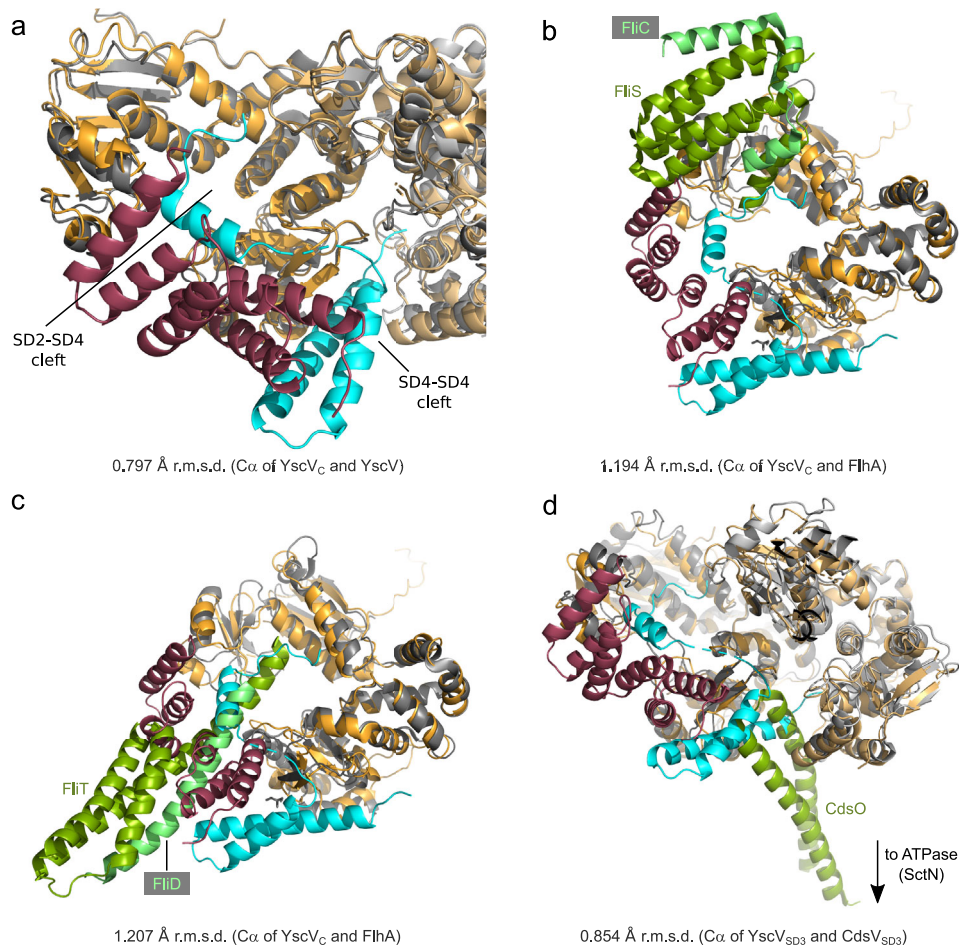
The binding site of the ATPase stalk protein SctO, as published for the complex of *Chlamydia* CdsV:CdsO<sup>32</sup>, is likewise located in the SD4–SD4 cleft and would not allow for simultaneous binding of SctO and SctXY (Fig. 9d). It should be noted that the *Chlamydia* T3SS does not belong to the Ysc family and therefore lacks SctX and SctY. It is still poorly understood how different factors work together in the regulation of SctV. These include the binding of interacting proteins such as SctO, SepL:SepD in enteropathogenic *E. coli*<sup>47</sup>, or substrate:chaperone complexes as well as local pH<sup>33</sup>, and the linker between transmembrane and cytosolic domain<sup>34,48,49</sup>. As such, further studies should investigate the potential interplay between YscV, YscO, and YscXY. We argue that, based on the lack of rearrangement after YscXY binding and the nonamer destabilization in vitro, YscX (and its chaperone YscY) may contribute to those dynamics by destabilizing the nonameric ring of YscV, most likely via YscX’s yet uncharacterized N-terminus. This may contribute to facilitating the substrate switch from needle components to hydrophobic translocators.

## Methods

**Cloning.** The vectors pETM-40\_yscX and pACYCDuet-1\_yscY were generated for the expression of YscX:YscY binary complexes via PCR and restriction cloning using *NcoI* and *NotI*. Furthermore, pACYCDuet-1 containing both *yscY* in its first and *yscV356* in its second multiple cloning site was prepared using the same method but employing *NdeI* and *XhoI*. For the expression of YscV<sub>C</sub> alone, pETM-11\_yscV356 was used. From this vector, a deletion mutant lacking SD3 and SD4 was generated (YscV<sub>SD12</sub>). All genes were cloned from the *Yersinia enterocolitica* pYVe227 virulence plasmid (G. Cornelis, Basel) via PCR. Deletions and mutations were introduced via ‘Round the horn’ mutagenesis. For a full overview of all constructs used, refer to Supplementary Table 3.

**Expression and purification of YscX:YscY complexes.** Complexes of YscX with its chaperone YscY were expressed in *E. coli* BL21 (DE3) transformed with both pETM-40\_yscX and pACYCDuet-1\_yscY vectors. Cell cultures were grown at 37 °C to OD<sub>600</sub> = 0.5 in LB medium and cooled to 20 °C before induction with 0.25 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside at OD<sub>600</sub> = 0.8–0.9. After incubation overnight, cells were spun down at 4000  $\times$  g and pellets stored at –20 °C. Cell pellets were resuspended in 50 mM Tris pH 8, 150 mM NaCl and 10 mM  $\beta$ -mercaptoethanol supplemented with one tablet of cOmplete protease inhibitor cocktail (Roche) and 0.6 mg DNase I per liter of expression culture and then passed through a pressure cell homogenizer (Stansted FPG12800) at 120 MPa.

Lysates were cleared by centrifugation (60 min, 30,000  $\times$  g, 4 °C). Amylose affinity chromatography (10 mL Amylose Resin HighFlow, New England Biolabs) was employed to capture MBP-YscX:His<sub>6</sub>-YscY. After 60 min of light pivoting at 4 °C, the resin was washed by raising the NaCl concentration to 300 mM. The protein-loaded amylose was re-equilibrated with 50 mM Tris pH 8, 150 mM NaCl,



**Fig. 9 Superposition of export gate complexes with YscVXY.** YscV, YscX, and YscY are shown in orange, cyan, and red, respectively. **a** Superposition with the cryo-EM structure of YscV<sub>C</sub> (gray, PDB: 7ALW), which is shown as cartoon representation. YscV<sub>C</sub> protomers are aligned with an r.m.s.d. of 0.797 Å. **b, c** Superposition with FlhA (gray) bound to the flagellar chaperone:substrate complexes (green:light green). **b** FlhA:FliS:FliC (PDB: 6CH3). **c** FlhA:FliT:FliD (PDB: 6CH2). Alignments were obtained using SD3 of YscV<sub>C</sub> (residues 514–600) with an r.m.s.d. of 1.194 Å (PDB: 6CH2) and 1.207 Å (PDB: 6CH3). **d** YscVXY superimposed with *Chlamydia pneumoniae* CdsV in complex with the stalk protein CdsO (PDB: 6WA9). The alignment was calculated using SD3 of YscV<sub>C</sub> with an r.m.s.d. of 0.854 Å. Two subunits of YscV and CdsV (gray) are shown, CdsO is colored in green.

and 1 mM dithiothreitol. To remove the MBP-tag of YscX, 1 mg of TEV-protease was added per 50 mg of target protein and the resin was incubated at 20 °C over night on a seesaw shaker. The flow-through and successive washing fractions were collected, pooled and imidazole was added to a final concentration of 15 mM before being loaded onto 5 mL of NiNTA resin (Protino Ni-NTA Agarose, Macherey-Nagel).

After 60 min of incubation, unbound protein was washed from the column by increasing the imidazole and NaCl concentrations to 100 and 300 mM, respectively. YscX:YscY eluted from the resin when adding elution buffer (20 mM Tris pH 8, 150 mM NaCl, 250 mM imidazole, 1 mM DTT). Finally, the protein was purified via size exclusion chromatography on a HiLoad 16/60 Superdex 75pg (Cytiva) where the buffer was exchanged to 20 mM Tris pH 8, 150 mM NaCl. After concentrating the complex to approximately 10 mg/mL, 5 mM tris(2-carboxyethyl) phosphine (TCEP) was added and the solution was flash-frozen at  $-80^{\circ}\text{C}$ .

**Reductive methylation of YscX:YscY.** NiNTA-purified protein was first dialyzed against 2 × 2 L 50 mM Hepes pH 8, 300 mM NaCl, 2 mM DTT and then diluted to approximately 1 mg/mL. Borane dimethylamine complex (97%, Alfa Aesar) as well as formaldehyde were added to 1.2 mg/mL and 1.2% (v/v), respectively, and the mixture incubated on ice. After three and six hours, the same amounts of the two reagents were added and the protein was further incubated over night. The reaction was stopped with the addition of 250 mM Tris pH 8, the protein was concentrated and purified via gel filtration as described above.

**Expression and purification of YscV<sub>C</sub>XY.** For the ternary complex of YscX<sub>32</sub>:YscY with YscV<sub>C</sub>, pETM-40\_yscX32 and pACYCDuet-1\_yscY\_yscV356 were co-transformed into *E. coli* BL21 (DE3). Expression and purification steps were identical until the protein was eluted from the amylose resin after TEV-digestion. Instead of successive NiNTA affinity chromatography, the complex was loaded

onto 5 mL amylose resin to remove the remaining MBP-tagged protein. The flow-through and washing fractions were combined, concentrated, and purified via size exclusion chromatography on a HiLoad 16/60 Superdex 200 pg (Cytiva). Before freezing the concentrated protein complex, 5 mM TCEP was added.

**Expression and purification of His<sub>6</sub>-YscV constructs.** For the stand-alone purification of the cytosolic domain of YscV (YscV<sub>C</sub>, residues 356–704) or YscV<sub>SD12</sub> (residues 356–513), the proteins were expressed as described above from pETM-11 with an N-terminal and TEV-cleavable hexahistidine tag. The cleared lysate was loaded onto 10 mL NiNTA resin and incubated for 60 min with 10 mM imidazole. The flow-through was collected and the resin was washed first with 50 mM Tris pH 8, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol and then with 30 mM imidazole in the same buffer. An increase to 250 mM imidazole eluted the target protein, which was then dialyzed against 20 mM Tris pH 8, 150 mM NaCl, 1 mM DTT after adding 1:50 w/w TEV protease. Re-chromatographic NiNTA cleared the target of uncleaved His<sub>6</sub>-YscV before the protein was concentrated and polished via size exclusion chromatography on a HiLoad 16/60 Superdex 200pg (YscV<sub>C</sub>) or HiLoad 16/60 Superdex 75pg (YscV<sub>SD12</sub>) (Cytiva) with 20 mM Tris pH 8, 150 mM NaCl as running buffer. After concentration to 10–15 mg/mL, 5 mM TCEP was added and the protein was flash-frozen.

**MBP-pull-down assay.** To test the binding of different YscX truncations to YscV<sub>C</sub>, both proteins and YscY were co-expressed as described for the ternary complex but at a scale of 100 mL culture. The cleared lysate was applied to 1 mL of amylose resin and unbound protein was washed off. Detection of amylose-bound protein was achieved by resuspending the resin in washing buffer and loading it on an SDS-PAGE.

**Limited proteolysis.** YscX<sub>full-length</sub>:YscY was subjected to limited proteolysis by trypsin, chymotrypsin, proteinase K, and thermolysin to identify stable constructs for expression. Purified protein (1 mg/mL) was mixed with 1:100 (w/w) protease in 50 mM Tris pH 8, 150 mM NaCl, and 1 mM DTT or 0.5 mM CaCl<sub>2</sub> (for thermolysin) and incubated on ice for 3 h. Samples were analyzed by SDS-PAGE and sequenced at their N-terminus via Edman degradation (Proteome Factory, Berlin, DE) after transfer to PVDF membrane (Immobilon, Merck) in a semi-dry set-up. Total mass was determined in-house using MALDI-TOF after dialysis against 10 mM acetate buffer pH 4.6.

**Analytical size exclusion.** Size exclusion chromatography on a Superdex 200/300GL was utilized to analyze protein oligomerization, dissociation, and association. The proteins were thawed on ice, combined for protein mixtures, or diluted for concentration gradients, and incubated for at least 30 min on ice before centrifugation and application to the column. All runs were carried out at 0.5 mL/min flowrate with 20 mM Tris pH 8, 150 mM NaCl as running buffer.

**Crystallization and data collection.** Crystals of YscX<sub>50</sub>:YscY grew in 0.1 M BisTris pH 6.0, 0.2 M MgCl<sub>2</sub> and 21% PEG 3,350 with 5 mg/mL protein and a drop ratio of 0.33  $\mu$ L reservoir to 0.66  $\mu$ L protein solution at 20 °C in a sitting drop setting over 60  $\mu$ L of reservoir. Microseeding of these into the same condition but with 13% PEG 3350 and 3–5 mg/mL protein was necessary to obtain single crystals. Drops were equilibrated for 6 h or over night before adding the seeds. Crystal growth was poorly reproducible and depended on purification batch. For cryoprotection, the crystals were transferred to a drop containing 0.1 M BisTris pH 6.0, 0.2 M CaCl<sub>2</sub>, 19% PEG 3350 supplemented with 5% glycerol. Measurements were carried out at the BL14.2. beamline at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin für Materialien und Energie<sup>50</sup>. Data were collected using the local installation of MxCube<sup>51</sup>.

For crystallization of the methylated complex YscX<sub>32</sub>:YscY, the protein buffer was exchanged to 10 mM Tris pH 8, 50 mM NaCl, 5% (w/v) sucrose, 2.5 mM TCEP. Crystals were obtained at 20–25 °C with 0.1 M sodium succinate pH 6–6.5, 4.2–4.4 M NaCl as reservoir solution using 5 mg/mL protein but could be optimized by the addition of 4% acetonitrile (Hampton Research Additive Screen). The drop ratio was 1  $\mu$ L protein plus 0.5 or 1  $\mu$ L reservoir solution. For harvesting, trays were transferred from 20 to 4 °C, and crystals were cryoprotected with 10% (w/v) sucrose. Data collection took place at beamline X06SA at SLS (Villigen, Switzerland) using 1.00 Å wavelength and 0.25° of oscillation between frames.

The ternary complex of YscV<sub>C</sub>, YscX<sub>32</sub>, and His<sub>6</sub>-YscY was crystallized in a sitting drop setting with 0.1 M Hepes pH 6.8–7.5, 1% PEG 2000 MME, and 0.7–1.0 M sodium succinate as reservoir solution. The drop contained 0.33  $\mu$ L reservoir solution and 0.66  $\mu$ L of 5 mg/mL protein. Large crystals between 200 and 1000  $\mu$ m in length grew after 2–3 days of incubation at 20 °C. As these crystals were very fragile and tended to crumble when touched, the plates were transferred to 4 °C the evening before harvesting. Cryoprotection was achieved by gradually transferring the crystals to solutions containing 0.1 M Hepes, 1% PEG 2000 MME, and 1.1 M sodium succinate as well as 5, 10, 15, and finally 22.5% glycerol. Data were collected through the local installation of MxCube<sup>51</sup> at beamline P13 at DESY (Hamburg, Germany) using a helical scan with 0.066 s exposure time per image and 100% transmission<sup>52</sup>.

**Data reduction, model building, and refinement.** Diffraction data were indexed, integrated, and scaled using XDS<sup>53</sup> or XDSAPP<sup>54</sup> for YscX<sub>50</sub>:YscY. The MR model was generated using a local installation of AlphaFold<sup>24,3</sup>, supplying the sequences of YscX and YscY with a 33 residue linker between them. The phase problem was initially solved for the YscX<sub>50</sub>:YscY data with a truncated model of the complex using Phaser<sup>55</sup> with the tNCS option disabled. After initial model building using BUCCANEER<sup>56</sup> in the CCP4i GUI<sup>57</sup> successive rounds of manual model building in Coot<sup>58</sup> and refinement in phenix.refine<sup>59,60</sup> generated the final model. TLS parameters were determined by the “find TLS” option in phenix.refine.

For YscX<sub>32</sub>:YscY a model of the shorter YscX<sub>50</sub>:YscY construct was employed as a search model during MR. Secondary structure restraints and reference model restraints from the aforementioned structure were used during refinement in phenix.refine. The ternary complex structure was solved via MR with a cryo-EM model of YscV<sub>C</sub><sup>35</sup> (PDB: 7ALW (YscV<sub>C</sub> nonamer)). After an initial round of refinement in phenix.refine, the YscX and YscY chains were placed manually into the density and refined with reference model restraints. To account for the low resolution data, group B-factors, secondary structure restraints, Ramachandran restraints, and torsion-angle NCS restraints were enabled in refinement. The resolution cutoff for the YscV<sub>C</sub>:X<sub>32</sub>:Y dataset was determined using paired refinements<sup>61</sup> as implemented in PAIREF<sup>62,63</sup>.

Molecular graphics and alignments were generated in PyMOL. The r.m.s.d. value of alignments was calculated in PyMOL using only Ca atoms without outlier rejection. The entire residue range was considered unless stated otherwise.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The atomic coordinates and X-ray intensities generated in this study have been deposited in the Worldwide Protein Data Bank (wwPDB) under accession codes 7QIH (YscX<sub>50</sub>:YscY), 7QII (YscX<sub>32</sub>:YscY), and 7QIJ (YscV<sub>C</sub>:YscX<sub>32</sub>:YscY). Diffraction images are available from the SB Grid Data Bank<sup>64</sup> (data.sbgriid.org) with Data ID 905 (YscX<sub>50</sub>:YscY), Data ID 906 (YscX<sub>32</sub>:YscY), and Data ID 907 (YscV<sub>C</sub>:YscX<sub>32</sub>:YscY). The structure of the YscV<sub>C</sub> nonamer used as search model to solve the phase problem of the YscVXY structure is available from the wwPDB under accession code 7ALW. The structures used to generate Figs. 8 and 9 are available from the wwPDB. Figure 8a: 2P58 (YscEFG); Fig. 8b: 3WXX (AcrH:AopB); Fig. 9a: 7ALW (YscV<sub>C</sub> nonamer); Fig. 9b: 6CH3 (FlhA:Flh:FliC); Fig. 9c: 6CH2 (FlhA:FliT:FliD); Fig. 9d: 6WA9 (CdsV:CdsO). Source data are provided with this paper.

Received: 17 December 2021; Accepted: 29 April 2022;

Published online: 02 June 2022

## References

- Dewoody, R. S., Merritt, P. M. & Marketon, M. M. Regulation of the Yersinia type III secretion system: traffic control. *Front. Cell. Infect. Microbiol.* **3**, 4 (2013).
- Kubori, T. et al. Supramolecular structure of the Salmonella typhimurium type III protein secretion system. *Science* **280**, 602–605 (1998).
- Bai, F., Li, Z., Umezawa, A., Terada, N. & Jin, S. Bacterial type III secretion system as a protein delivery tool for a broad range of biomedical applications. *Biotechnol. Adv.* **36**, 482–493 (2018).
- Portaliou, A. G., Tsois, K. C., Loos, M. S., Zorzini, V. & Economou, A. Type III secretion: Building and operating a remarkable nanomachine. *Trends Biochem. Sci.* **41**, 175–189 (2016).
- Akeda, Y. & Galán, J. E. Chaperone release, and unfolding of substrates in type III secretion. *Nature* **437**, 911–915 (2005).
- Miletic, S. et al. Substrate-engaged type III secretion system structures reveal gating mechanism for unfolded protein translocation. *Nat. Commun.* **12**, 1–14 (2021).
- Büttner, C. R., Sorg, I., Cornelis, G. R., Heinz, D. W. & Niemann, H. H. Structure of the Yersinia enterocolitica type III secretion translocator chaperone SycD. *J. Mol. Biol.* **375**, 997–1012 (2008).
- Schreiner, M. & Niemann, H. H. Crystal structure of the Yersinia enterocolitica type III secretion chaperone SycD in complex with a peptide of the minor translocator YopD. *BMC Struct. Biol.* **12**, 13 (2012).
- Discola, K. F. et al. Membrane and chaperone recognition by the major translocator protein PopB of the type III secretion system of Pseudomonas aeruginosa. *J. Biol. Chem.* **289**, 3591–3601 (2014).
- Lunelli, M., Lokareddy, R. K., Zychlinsky, A. & Kolbe, M. IpaB-IpgC interaction defines binding motif for type III secretion translocator. *Proc. Natl. Acad. Sci. USA* **106**, 9661–9666 (2009).
- Nguyen, V. S. et al. Structure of AcrH-AopB chaperone-translocator complex reveals a role for membrane hairpins in type III secretion system translocon assembly. *Structure* **23**, 2022–2031 (2015).
- Archuleta, T. L. & Spiller, B. W. A gatekeeper chaperone complex directs translocator secretion during type three secretion. *PLoS Pathog.* **10**, e1004498 (2014).
- Sun, P., Tropea, J. E., Austin, B. P., Cherry, S. & Waugh, D. S. Structural characterization of the Yersinia pestis type III secretion system needle protein YscF in complex with its heterodimeric chaperone YscE/YscG. *J. Mol. Biol.* **377**, 819–830 (2008).
- Quinaud, M. et al. Structure of the heterotrimeric complex that regulates type III secretion needle formation. *Proc. Natl. Acad. Sci. USA* **104**, 7803–7808 (2007).
- Troisfontaines, P. & Cornelis, G. R. Type III secretion: More systems than you think. *Physiology* **20**, 326–339 (2005).
- Iriarte, M. & Cornelis, G. R. Identification of SycN, YscX, and YscY, three new elements of the Yersinia Yop virulon. *J. Bacteriol.* **181**, 675–680 (1999).
- Day, J. B. & Plano, G. V. The Yersinia pestis YscY protein directly binds YscX, a secreted component of the type III secretion machinery. *J. Bacteriol.* **182**, 1834–1843 (2000).
- Diepold, A., Wiesand, U., Amstutz, M. & Cornelis, G. R. Assembly of the Yersinia injectisome: The missing pieces. *Mol. Microbiol.* **85**, 878–892 (2012).
- Vanden Bergh, P., Heller, M., Braga-Lagache, S. & Frey, J. The Aeromonas salmonicida subsp. salmonicida exoproteome: Determination of the complete

- repertoire of Type-Three Secretion System effectors and identification of other virulence factors. *Proteome Sci.* **11**, 1–20 (2013).
20. Pallen, M. J., Francis, M. S. & Fütterer, K. Tetratricopeptide-like repeats in type-III-secretion chaperones and regulators. *FEMS Microbiol. Lett.* **223**, 53–60 (2003).
  21. Gurung, J. M. et al. Heterologous complementation studies with the YscX and YscY protein families reveals a specificity for Yersinia pseudotuberculosis type III secretion. *Front. Cell. Infect. Microbiol.* **8**, 80 (2018).
  22. Francis, M. S., Lloyd, S. A. & Wolf-Watz, H. The type III secretion chaperone LcrH co-operates with YopD to establish a negative, regulatory loop for control of Yop synthesis in Yersinia pseudotuberculosis. *Mol. Microbiol.* **42**, 1075–1093 (2001).
  23. Bröms, J. E., Edqvist, P. J., Carlsson, K. E., Forsberg, A. & Francis, M. S. Mapping of a YscY binding domain within the LcrH chaperone that is required for regulation of Yersinia type III secretion. *J. Bacteriol.* **187**, 7738–7752 (2005).
  24. Lara-Tejero, M., Kato, J., Wagner, S., Liu, X. & Galán, J. E. A sorting platform determines the order of protein secretion in bacterial type III systems. *Science* **331**, 1188–1191 (2011).
  25. Diepold, A., Kudryashev, M., Delalez, N. J., Berry, R. M. & Armitage, J. P. Composition, formation, and regulation of the cytosolic C-ring, a dynamic component of the type III secretion injectisome. *PLoS Biol.* **13**, e1002039 (2015).
  26. Saijo-Hamano, Y. et al. Structure of the cytoplasmic domain of FlhA and implication for flagellar type III protein export. *Mol. Microbiol.* **76**, 260–268 (2010).
  27. Moore, S. A. & Jia, Y. Structure of the cytoplasmic domain of the flagellar secretion apparatus component FlhA from *Helicobacter pylori*. *J. Biol. Chem.* **285**, 21060–21069 (2010).
  28. Bange, G. et al. FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. *Proc. Natl. Acad. Sci. USA* **107**, 11295–11300 (2010).
  29. Worrall, L. J., Vuckovic, M. & Strynadka, N. C. J. Crystal structure of the C-terminal domain of the Salmonella type III secretion system export apparatus protein InvA. *Protein Sci.* **19**, 1091–1096 (2010).
  30. Abrusci, P. et al. Architecture of the major component of the type III secretion system export apparatus. *Nat. Struct. Mol. Biol.* **20**, 99–104 (2013).
  31. Majewski, D. D., Lyons, B. J. E., Atkinson, C. E. & Strynadka, N. C. J. Cryo-EM analysis of the SctV cytosolic domain from the Enteropathogenic E.coli T3SS injectisome. *J. Struct. Biol.* **212**, 107660 (2020).
  32. Jensen, J. L., Yamini, S., Rietsch, A. & Spiller, B. W. The structure of the Type III secretion system export gate with CdsO, an ATPase lever arm. *PLoS Pathog.* **16**, e1008923 (2020).
  33. Matthews-Palmer, T. R. S. et al. Structure of the cytoplasmic domain of SctV (SsaV) from the Salmonella SPI-2 injectisome and implications for a pH sensing mechanism. *J. Struct. Biol.* **213**, 107729 (2021).
  34. Xu, J., Wang, J., Liu, A., Zhang, Y. & Gao, X. Structural and functional analysis of SsaV cytoplasmic domain and variable linker states in the context of the InvA-SsaV chimeric protein. *Microbiol. Spectr.* **9**, e01251-21 (2021).
  35. Kuhlen, L., Johnson, S., Cao, J., Deme, J. C. & Lea, S. M. Nonameric structures of the cytoplasmic domain of FlhA and SctV in the context of the full-length protein. *PLoS One* **16**, e0252800 (2021).
  36. Diepold, A. et al. A dynamic and adaptive network of cytosolic interactions governs protein export by the T3SS injectisome. *Nat. Commun.* **8**, 15940 (2017).
  37. Diepold, A., Wiesand, U. & Cornelis, G. R. The assembly of the export apparatus (YscR,S,T,U,V) of the Yersinia type III secretion apparatus occurs independently of other structural components and involves the formation of an YscV oligomer. *Mol. Microbiol.* **82**, 502–514 (2011).
  38. Xing, Q. et al. Structures of chaperone-substrate complexes docked onto the export gate in a type III secretion system. *Nat. Commun.* **9**, 1773 (2018).
  39. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
  40. Walter, T. S. et al. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617–1622 (2006).
  41. Yang, J. et al. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2014).
  42. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
  43. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
  44. Souza, C. D. A. et al. The YscE/YscG chaperone and YscF N-terminal sequences target YscF to the Yersinia pestis type III secretion apparatus. *Microbiology* **164**, 338–348 (2018).
  45. Schubot, F. D. et al. Three-dimensional structure of a macromolecular assembly that regulates type III secretion in Yersinia pestis. *J. Mol. Biol.* **346**, 1147–1161 (2005).
  46. Job, V., Mattei, P. J., Lemaire, D., Attree, I. & Dessen, A. Structural basis of chaperone recognition of type III secretion system minor translocator proteins. *J. Biol. Chem.* **285**, 23224–23232 (2010).
  47. Portaliou, A. G. et al. Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic *Escherichia coli*. *EMBO J.* **36**, 3517–3531 (2017).
  48. Inoue, Y. et al. Structural insights into the substrate specificity switch mechanism of the type III protein export apparatus. *Structure* **27**, 965–976 (2019).
  49. Terahara, N. et al. Insight into structural remodeling of the FlhA ring responsible for bacterial flagellar type III protein export. *Sci. Adv.* **4**, eao7054 (2018).
  50. Mueller, U. et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J.* **130**, 1–10 (2015).
  51. Gabadinho, J. et al. MxCuBE: A synchrotron beamline control environment customized for macromolecular crystallography experiments. *J. Synchrotron. Radiat.* **17**, 700–707 (2010).
  52. Cianci, M. et al. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J. Synchrotron. Radiat.* **24**, 323–332 (2017).
  53. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
  54. Sparta, K. M., Heinemann, U., Mueller, U. & Weiss, M. S. XDSAPP2.0. *J. Appl. Crystallogr.* **49**, 1085–1092 (2016).
  55. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
  56. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
  57. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
  58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
  59. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
  60. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
  61. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
  62. Malý, M., Diederichs, K., Dohnalek, J. & Kolenko, P. Paired refinement under the control of PAIREF. *IUCrJ* **7**, 681–692 (2020).
  63. Malý, M., Diederichs, K., Dohnalek, J. & Kolenko, P. PAIREF: Paired refinement also for Phenix users. *Acta Crystallogr. F Struct. Biol. Commun.* **77**, 226–229 (2021).
  64. Morin, A. et al. Cutting Edge: Collaboration gets the most out of software. *Elife* **2**, e01456 (2013).
  65. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125–142 (1984).
  66. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
  67. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).

## Acknowledgements

We thank the Helmholtz-Zentrum für Materialien und Energie for the allocation of synchrotron radiation beamtime. We acknowledge the Paul Scherrer Institut, Villigen, Switzerland for provision of synchrotron radiation beamtime at beamline X06SA (PX1) of the SLS and would like to thank Tomizaki Takashi for assistance. The synchrotron data were collected at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank Isabel Bento for the assistance in using the beamline. We acknowledge the help of Dr. Jens Sproß (Bielefeld University) in preparing, collecting, and analyzing mass spectrometry data as well as Dr. Gregor Hagelueken (University of Bonn) for generating the AlphaFold 2 models. We thank Martin Malý (Czech Technical University) for his help in running PAIREF. D.G. acknowledges funding from the Bielefelder Nachwuchsfonds. H.H.N. thankfully acknowledges the financial support by HZB (travel grant “BESSY 2 MX: 17”).

## Author contributions

H.H.N., D.G. and M.S. conceived the project. D.G. and M.S. performed experiments. D.G., M.S. and H.H.N. collected and analyzed data. D.G. prepared the figures. D.G. and H.H.N. wrote the manuscript. H.H.N. supervised the project.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30487-1>.

**Correspondence** and requests for materials should be addressed to Hartmut H. Niemann.

**Peer review information** *Nature Communications* thanks Gregory Plano, Benjamin Spiller, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022