Structure of a bacterial Rhs effector exported by the type VI secretion system

Patrick Günther[1,a], Dennis Quentin[1,a], Shehryar Ahmad[2], Kartik Sachar[2], Christos Gatsogiannis[1,3], John C. Whitney[2,4,5,b], Stefan Raunser[1,b]

[1]Department of Structural Biochemistry, Max Planck Institute of Molecular Physiology, Otto-Hahn-Str. 11, 44227 Dortmund, Germany

[2]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON, L8S 4K1, Canada

[3]Current address: Institute for Medical Physics and Biophysics and Center for Soft Nanoscience, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany

[4]Michael DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, ON, L8S 4K1, Canada

[5]David Braley Centre for Antibiotic Discovery, McMaster University, Hamilton, ON, L8S 4K1, Canada

[a]These authors contributed equally.

[b]Correspondence to: jwhitney@mcmaster.ca & stefan.raunser@mpi-dortmund.mpg.de

Abstract

The type VI secretion system (T6SS) is a widespread protein export apparatus found in Gram-negative bacteria. The majority of T6SSs deliver toxic effector proteins into competitor bacteria. Yet, the structure, function, and activation of many of these effectors remains poorly understood. Here, we present the structures of the T6SS effector RhsA from *Pseudomonas protegens* and its cognate T6SS spike protein, VgrG1, at 3.3 Å resolution. The structures reveal that the rearrangement hotspot (Rhs) repeats of RhsA assemble into a closed anticlockwise β-barrel spiral similar to that found in bacterial insecticidal Tc toxins and in metazoan teneurin proteins. We find that the C-terminal toxin domain of RhsA is autoproteolytically cleaved but remains inside the Rhs 'cocoon' where, with the exception of three ordered structural elements, most of the toxin is disordered. The N-terminal 'plug' domain is unique to T6SS Rhs proteins and resembles a champagne cork that seals the Rhs cocoon at one end while also mediating interactions with VgrG1. Interestingly, this domain is also autoproteolytically cleaved inside the cocoon but remains associated with it. We propose that mechanical force is required to remove the cleaved part of the plug, resulting in the release of the toxin domain as it is delivered into a susceptible bacterial cell by the T6SS.

Introduction

One way that bacteria interact with their environment is by secreting toxic molecules into their surroundings. Many species of Gram-negative bacteria have evolved specialized protein secretion systems for this purpose with arguably the best-characterized example being the T6SS[1,2]. T6SSs mediate bacterial antagonism by facilitating the injection of toxic effector proteins into competing bacterial cells in a contact-dependent manner[3].

A functional T6SS apparatus requires the concerted action of three protein subcomplexes: a cell envelope-spanning membrane complex, a cytoplasmic baseplate complex and a bacteriophage tail-like sheath-tube complex[4,5]. The tail-like complex consists of a contractile sheath that surrounds an inner tube comprised of many copies of stacked ring-shaped hexameric hemolysin co-regulated protein (Hcp)[6,7]. This Hcp tube is capped with a member of the homotrimeric valine-glycine repeat protein G (VgrG) spike protein family, which typically also interacts with a cone-shaped proline-alanine-alanine-arginine (PAAR) domain-containing protein to form a complete T6SS tail tube-spike complex[8]. During a T6SS

firing event, the Hcp-VgrG-PAAR tube-spike complex is rapidly assembled in the cytosol of the attacking bacterium, recruited to the membrane complex through its interaction with the baseplate, and subsequently surrounded by a contractile sheath that upon sheath contraction, propels the tail spike into the recipient cell[9]. Toxic effector proteins are recruited to the tube-spike complex and injected into recipient cells alongside Hcp and VgrG, typically resulting in death of the target cell. Self-protection from antibacterial T6SS effectors is accomplished via effector co-expression with cognate immunity proteins, which neutralize effector toxicity by occluding the effector active site or by the hydrolysis of effector generated toxic products[10,11].

Effectors are recruited to distinct regions of the T6SS tube-spike complex. Smaller effectors (< 50 kDa) are typically recognized and loaded within the lumen of the Hcp tube[12] whereas larger, multidomain effectors interact with VgrG proteins via an effector-encoded PAAR domain or by directly interacting with VgrG itself[13]. PAAR-containing effectors are widely distributed in Proteobacteria and can be further subdivided into subfamilies of proteins based on the presence of additional sequence motifs that may play a role in effector translocation into target cells. One family of PAAR effectors that was recently described in detail possess an additional N-terminal motif, called prePAAR, predicted to be involved in PAAR folding, as well as at least one N-terminal transmembrane domain (TMD) that is hypothesized to insert into the recipient cell inner membrane with one to three transmembrane helices per TMD that enable toxin translocation into the cytoplasm[14]. TMD-containing PAAR effectors require effector-associated gene (Eag) chaperones to bind and stabilize effector TMDs prior to export from the donor cell[14,15].

Previously, our group showed that the T6SS of the soil bacterium *Pseudomonas protegens* secretes the prePAAR and PAAR-containing effector RhsA[14]. The *rhsA* operon contains genes encoding its cognate VgrG spike and Eag chaperone proteins (Figure 1A), both of which are required for the delivery of this effector into susceptible competitors[14]. Previous negative stain electron microscopy experiments suggest that EagR1, RhsA and VgrG assemble to form a complex necessary for T6SS function (Figure 1B)[14]. The N-terminal domain of RhsA consists of a prePAAR motif as well as two predicted transmembrane helices. The former of these features is proposed to contribute sequence elements that 'complete' the PAAR domain and enable its interaction with the tip of VgrG1[14] whereas the latter likely plays a role in target cell penetration[16].

In addition to its N-terminal prePAAR, TMD and PAAR regions, RhsA possesses Rhs elements, which are characterized by the presence of repeating tyrosine-aspartate (YD) motifs. The chromosomal regions encoding Rhs repeats were originally described in *Escherichia coli*

genomes as sites where recombination frequently took place[17]. Later studies revealed that Rhs repeat-containing proteins are not unique to *E. coli* but are common across Proteobacteria[18,19]. Other more distant homologs, called teneurins, are involved in axon guidance in vertebrates[20–22]. Structures of bacterial Rhs proteins and human teneurin2 revealed that Rhs repeat regions form a large (~100 kDa) β-barrel cocoon-like structure spirally wound counterclockwise[23–26]. The C-terminus of the Rhs core is typically demarcated by a highly conserved PxxxxDPxGL motif, which is necessary for autoproteolytic cleavage[20]. Previous work has shown that C-terminal autoproteolysis of T6SS Rhs effectors and Tc toxins is required for toxin release from the Rhs cocoon[23,24,27].

While the Rhs repeat-containing region is highly conserved among Rhs proteins, their N- and C-terminal extensions can differ significantly. The C-terminus of bacterial Rhs proteins encode for a diverse range of toxins and is often referred to as the hypervariable region (HVR)[28]. In host cell targeting Tc toxins, the C-terminal toxin is encapsulated within the Rhs core[29], however, previous structural analyses of these proteins have been unable to resolve the structure of a HVR, suggesting that this region may be partially unfolded. The Rhs repeat-containing cocoon of Tc toxins (TcB and TcC) interacts through TcB with the A component (TcA) via a β-propeller domain. The propeller acts as a gate that opens upon A component binding, allowing the HVR to enter the translocation channel of TcA[29]. The penetration of the target cell membrane through the A component is triggered by a pH shift, which permits subsequent HVR release into the prey cell cytosol[23,30]. An similarly detailed toxin release mechanism has not been reported for the bacteria targeting T6SS Rhs effectors[3] and structures of T6SS-exported Rhs proteins have not been determined to date.

In this work, we present high-resolution structures of the T6SS effector RhsA and its cognate VgrG1 protein at resolutions of 3.3 Å. The structure of VgrG1 is nearly identical to our previously determined VgrG1 structure from *P. aeruginosa*[16], indicating a high structural conservation of VgrG proteins. The structure of RhsA reveals that the Rhs-repeats of RhsA form a β-barrel similar to that of the B and C components of Tc toxins and teneurins. However, unlike the cocoon of Tc toxins, which are capped by the aforementioned β-propeller, RhsA is closed at its N-terminus by a plug domain that connects the cocoon to the PAAR domain. The plug domain is held in place via hydrophobic interactions with the Rhs cocoon and its eventual opening is the most likely event for the toxin domain to be released from barrel. Our structural and biochemical findings show that the plug domain is autoproteolytically cleaved inside the cocoon. The cleaved 304 N-terminal residues, comprising the prePAAR domain, PAAR domain as well as the TMD also include an N-terminal seal of the plug and an anchor helix.

101 The helix interacts strongly with the inner surface of the cocoon and is likely responsible for
102 keeping the N-terminal 304 residues associated with the cocoon. In addition, we find that the
103 RhsA C-terminal toxin domain is also cleaved via a canonical Rhs aspartyl autoprotease and
104 that this cleavage event is required for RhsA-dependent bacterial killing. Taken together, our
105 results provide novel insights into the architecture and mechanism of action of Rhs effectors
106 of the T6SS.
107

Results and Discussion

RhsA forms a cocoon-like structure that undergoes a double autocleavage

We previously demonstrated that a fragment of RhsA lacking its N-terminal prePAAR motif and two predicted transmembrane helices of the TMD (Figure 1C, residues 74-1486), $RhsA_{\Delta NT}$, is stable in the absence of its EagR1 chaperone and can be purified to homogeneity as a soluble protein when over-expressed in *Escherichia coli*[14]. We therefore expressed and purified $RhsA_{\Delta NT}$ (Supplementary Figure 1A) and used it for cryo-EM and single particle analysis. The raw images and 2D class averages suggest that $RhsA_{\Delta NT}$ forms a stable dimer in solution, which facilitated structure determination due to the increased size of the particles. Subsequent image processing imposing C_2 symmetry resulted in a reconstruction with a resolution of 3.3 Å from 454,740 particles (Figure 1D, Table 1). The high quality of the map allowed us to build an atomic model of 75% of the protein, comprising residues 268-275, 289-302, 305-366, 371-386, 395-871, 891-1039 and 1051-1350 (Figure 1E). The 25% of remaining unresolved regions correspond to the PAAR domain (residues 75-267) and the toxin domain (residues 1351-1486), indicating that these regions likely exhibit a high degree of flexibility.

The two molecules in the $RhsA_{\Delta NT}$ dimer interact along their longitudinal axes and are slightly tilted with respect to one another. The dimer interface is comprised of several complementary hydrophilic surfaces indicating that it is mostly electrostatic in nature (Supplementary Figure 2A-B). The 76 antiparallel β-strands of $RhsA_{\Delta NT}$ spiral in an anticlockwise manner resulting in a large hollow cocoon-shaped structure with outer dimensions of 86 x 65 Å. The overall structure of the Rhs cocoon resembles that of other YD-repeat containing proteins, such as the BC component of Tc toxins[23,24] and human teneurin2[25]. We also identified the conserved catalytic center of an aspartyl autoprotease that was first identified in Tc toxins[23,24]. In line with functioning to self-cleave the C-terminal toxin domain, we observed that the density of the cryo-EM map stops abruptly after tryptophan 1350 (Figure 2A). This position is in agreement with cleavage sites in other Rhs-related toxins that share the same PxxxxDPxG W/L/F consensus sequence found in RhsA (Supplementary Figure 3C-E), indicating that the C-terminal toxin domain is likely autoproteolytically cleaved similar to the HVR found in Tc toxins. To experimentally test the proteolytic activity of this motif, we mutated either aspartate 1324 or aspartate 1346, which are the RhsA residues in the equivalent position to the catalytic dyad of Tc toxin aspartyl autoproteases, to asparagine. Consistent with their proposed role in autocleavage and toxin release, these RhsA variants did not undergo autoproteolysis at their C-terminus (Figure 2B) and were substantially less toxic when overexpressed in *E. coli* compared to the wild-type protein (Supplementary Figure 3E). We

next introduced these mutations into the chromosome of *P. protegens* to test whether C-terminal autoproteolytic cleavage is indeed required for interbacterial competition. In contrast to a strain expressing wild-type RhsA, strains expressing the D1324N or D1346N variants of the protein were unable to outcompete RhsA-sensitive recipient bacteria indicating that C-terminal autocleavage is required for T6SS-dependent killing by this effector (Figure 2C). These findings mirror what has been observed for T6SS-exported Rhs proteins in *Aeromonas dhakensis* and *Enterobacter cloacae* suggesting that toxin domain liberation is a universal property of Rhs effectors[18,31].

We also identified additional density inside the Rhs cocoon corresponding to the C-terminal toxin domain, which is predicted to function as a DNase based on its homology to several characterized endonucleases[32]. The density filling the cocoon is, for the most part, not well defined and is apparent only at a much lower density threshold compared to the rest of the map (Figure 3D), Nevertheless, we could build three β-strands of this domain, comprising residues 1408-1413, 1420-1423 and 1458-1461, respectively, all of which interact with the inner surface of the cocoon structure (Figure 3E). The interface between the β-strands and the cocoon is stabilized by both hydrophilic and hydrophobic interactions (Figure 3E, Supplementary Figure 4). In sum, these structural data indicate that most of the toxin domain is either only partially folded or very flexible.

Interestingly, the density inside the Rhs cocoon not only ends after tryptophan 1350 but also before proline 305 (Figure 2F), suggesting that RhsA is proteolytically cleaved at this position as well. However, the N-terminal region is presumably still associated with the cocoon structure because the protein was purified by affinity chromatography using an N-terminal His-tag. Therefore, we heat-denatured $RhsA_{WT}$ and $RhsA_{\Delta NT}$ prior to SDS-PAGE to determine if the N-terminal domain was indeed cleaved (Figure 2G). Consistent with our structural data, we observed bands that migrate at molecular weights consistent with the loss of ~304 and ~230 amino acids, respectively, corresponding to the 304 N-terminal residues of $RhsA_{WT}$ and 230 N-terminal residues of $RhsA_{\Delta NT}$. This demonstrates that the protein is indeed cleaved between residues 304 and 305. By contrast, a fragment of RhsA lacking this N-terminal region, $RhsA_{\Delta N}$, does not undergo N-terminal proteolysis. A similar observation was recently made for the T6SS effector TseI from *A. dhakensis*, which is also N-terminally cleaved even though it differs from RhsA at its N-terminus in that it lacks prePAAR, PAAR and a TMD[18]. This cleavage event was shown to be essential for the activity of TseI after its secretion by the T6SS. Two glutamate residues at the +7 and +8 position relative to the N-terminal cleavage site are responsible for

the autoproteolysis of this effector[18]. However, these glutamates are not conserved in RhsA and are replaced by alanine (A312) and lysine (K313), which cannot act as catalytic center (Supplementary Figure 3B-C). Therefore, we examined the direct vicinity of P305 in our structure and found a cysteine residue (C538) flanked by two histidines (H530 and H555) that protrude from the wall of the Rhs barrel and are located near proline 305 (Supplementary Figure 3G). Unlike the previously described glutamate residues, this site is conserved in class I prePAAR T6SS Rhs effectors, but does not exist in Tc toxins (Supplementary Figure 3F)[18]. In addition, we identified three consecutive aspartates (D318-D320) in close proximity to the cleavage site.

To test the hypothesis of whether cysteine 538 or the other proximal residues are involved in N-terminal cleavage, we mutated these residues individually and performed a Western blot analysis against the N-terminus of the overexpressed constructs (Supplementary Figure 3H). Contrary to our assumption, we found that mutation of C538 to alanine did not abolish the N-terminal cleavage. Mutations of other conserved residues near the cleavage site had also little effect. Only the triple mutant (D318-D320 to alanine) led to a minimal reduction of N-terminal cleavage. However, we believe that this could be an indirect effect induced by the mutationally altered conformation in this region. Since there are no other obvious potential catalytically active residues besides the tested ones, it is likely that regions of the encapsulated toxin which are not resolved in our reconstruction catalyze the N-terminal cleavage by an unknown mechanism.

RhsA possesses a unique plug domain at its N-terminus

Our structural data show that the C-terminal aspartyl protease domain of RhsA seals the cocoon-shaped structure at one end, while the other end is capped by an N-terminal domain that adopts a smaller structure. This N-terminal domain of the barrel, formed by residues 268-386, caps the cocoon structure in a manner that is reminiscent of how a cork is used to plug a champagne bottle and thus we refer to it as the N-terminal plug domain ($RhsA_{plug}$) (Figure 3A). The interface between the plug domain and the Rhs core is mainly stabilized by hydrophobic interactions (Figure 3D-F) and a few hydrophilic interactions (Supplementary Figure 6A).

In addition to the ‘cork’ structure (residues 305-386), the plug domain contains an anchor helix (residues 289-302) and an N-terminal seal peptide (residues 268-275). The seal peptide fills a small opening in the ‘cork’ as it leaves the cocoon and connects the plug domain to the unmodelled PAAR domain (Figure 3A-B, Supplementary Figure 5A-B). In doing so, the seal peptide, together with the cork, closes off the cocoon entirely (Supplementary Figure 6B).

The anchor helix is amphipathic and stabilized by interactions with a hydrophobic patch of the Rhs repeats in this region (Figure 3B-C, Supplementary Figure 5C-E). The position of the anchor helix in our structure suggests its function is twofold. On the one hand it holds the N-terminal cleavage site in place. On the other hand, it ensures that the N-terminal plug domain remains stably attached to the cocoon, so that it remains sealed even after N-terminal cleavage.

The plug domain of RhsA possesses sequence and structural similarity to a domain found in the BC components of TcdB2-TccC3[23] and YenBC[24] (Figure 3C, Supplementary Figure 7 A-C) (18% sequence identity to TcdB2; 22% sequence identity to YenB). Interestingly, in Tc toxins this domain does not act as a plug that prevents the release of the toxin, but instead forms a negatively charged constriction through which the toxin domain is threaded prior to its translocation into a target cell[29]. A plug domain has also been described in teneurins (called fibronectin-plug, FN-plug)[25,26], however, it does not bear sequence or structural similarity to the plug domain of RhsA described herein. The FN-plug is narrower than the plug domains of RhsA and ABC toxins and extends further into the Rhs cocoon structure making numerous hydrophilic interactions with residues lining the inside of the YD-shell[25] (Figure 3C and Figure 4).

Our structure suggests that in contrast to Tc toxins the plug domain of RhsA tightly seals its Rhs cocoon. Because the plug domain appears to strongly interact with the Rhs core, mechanical removal of this entire domain after translocation of the cocoon into the target cell cytosol seems unlikely. Instead, we propose that only the N-terminal peptide seal of the RhsA plug is pulled out during T6SS-dependent delivery of RhsA into a susceptible bacterial cell. The seal peptide is part of the N-terminal region that is proteolytically cleaved, so we speculate that it would be more easily removed compared to the plug domain, which is held in place by the anchor helix and exists downstream of the N-terminal cleavage site. Penetration of the outer membrane and peptidoglycan layer as well as translocation of RhsA through the target cell inner membrane could provide the mechanical force to remove the plug. Since the plug domain resembles the constriction site in Tc toxins that effectors must pass during the initial translocation process, the same process is conceivable for the release of the C-terminal toxin domain of RhsA.

Comparison between Rhs proteins of known structure

A common feature of all Rhs proteins that have been structurally characterized to date is the cocoon structure formed by the Rhs repeats. While the Rhs cocoons from RhsA and teneurin2 are comprised of three β-helical turns of Rhs repeats, the cocoons of Tc toxins have

four turns and therefore have bigger overall dimensions and an internal cavity with larger volume. Consistent with this observation, the effector domain inside Tc toxin cocoons is larger (~30 kDa) than the toxin domain of RhsA (~15 kDa). All characterized Rhs proteins contain a hyperconserved region. For Tc toxins and RhsA this domain functions as aspartyl autoprotease whereas in teneurins it acts as a YD-shell plug that is not cleaved (Figure 4, blue). In all instances, the C-terminal hyperconserved region comprises 14 Rhs repeats that spiral into the inside of the cocoon and seal it at one side.

The plugs that close off the N-terminal side of the cocoon structures are more variable than their C-terminal counterparts (Figure 4, orange). While teneurins seal the barrel with an FN-plug that is mainly held in place by hydrophilic interactions with the barrel domain, TcdB2 from *Photorhabdus luminescens* and other BC components of Tc toxins close the cocoon using a β-propeller domain that acts as gatekeeper for toxin release. RhsA uses an N-terminal plug domain, which is homologous to the domain that forms the constriction site in Tc toxins. Both TcBC and RhsA encapsule a toxic effector, whereas teneurins act as scaffolding proteins with an empty cocoon (Figure 4, red). Nevertheless, teneurins encode a C-terminal ancient toxin component that sits outside of the barrel and is inactive (ABD Tox-GHH, Figure 4, red). While RhsA contains two autoprotease sites, namely an aspartyl protease that is responsible for the cleavage of the C-terminal toxic effector and an unknown protease site that cleaves the N-terminal 304 residues of the protein, Tc toxins only contain the former site and teneurins none at all.

In conclusion, based on its unique plug domain and its fusion to an N-terminal PAAR domain, we propose that our RhsA structure represents the founding member of a third structural class of Rhs repeat containing proteins (Figure 4).

Architecture of the pre-firing complex

To investigate how RhsA is mounted onto VgrG1 we set out to determine the structure of the secretion competent pre-firing complex (PFC) comprising VgrG1 in complex with full-length RhsA and EagR1. We purified the complex as described previously[14] and examined it by single particle cryo-EM. The 2D class averages enabled us to unequivocally characterize the arrangement of the subunits of the complex (Figure 5A). Expectedly, both EagR1 and RhsA are located at the tip of VgrG1. Interestingly, a large fraction of PFCs contained RhsA dimers instead of monomers (Supplementary Figure 8B), similar to what was observed in our analysis of $RhsA_{\Delta NT}$ alone. This demonstrates that the loading of two RhsA molecules onto one VgrG1

is sterically possible although it is unlikely to occur *in vivo* given that current data indicates a single VgrG homotrimer caps the Hcp tube[33].

As is the case for the previously characterized VgrG1-EagT6-Tse6 complex from *P. aeruginosa*[16], T6SS effectors can adopt multiple positions relative to their cognate VgrG spike protein. Unfortunately, this conformational heterogeneity prevented us from obtaining a high-resolution 3D reconstruction of the entire complex (Supplementary Movie 1). Instead, we applied C_3 symmetry during processing to determine the three-dimensional structure of *P. protegens* VgrG1 (Figure 5B). In applying this symmetry operator, the RhsA and EagR1 components of the complex, which do not adopt this symmetry, are averaged out during image processing. The cryo-EM map of VgrG1 reached a resolution of 3.3 Å and allowed us to build almost the complete atomic model of the protein comprising residues 8-643 (Figure 5C, Table 1). Not surprisingly given its high sequence homology, the structure of VgrG1 from *P. protegens* is nearly identical (71% sequence identity, 81% sequence similarity) to our previously determined VgrG1 structure from *P. aeruginosa* (r.m.s.d of 1.035 between 544 pruned Cα atoms; r.m.s.d of 1.428 across the complete structure, Supplementary Figure 8B-C)[16] even though their respective effectors, RhsA and Tse6, bear no sequence or structural similarity to one another beyond their N-terminal prePAAR and PAAR domains. Intriguingly, we identified two spherical densities in the center of the β-sheet prism of the VgrG1 trimer (Supplementary Figure 8E). Since we also observed these densities previously in the VgrG1 structure from *P. aeruginosa*[16], we speculate that this may be a common feature of VgrG1 proteins. Based on the exclusive clustering of positively charged residues around this density and its overall size, we hypothesize that it corresponds to either a phosphate or sulfate anion, which probably helps to stabilize the core of VgrG1 given that in the absence of this anion, the presence of positively charged residues in the core of the protein would likely be energetically unfavorable.

Model of RhsA firing events and toxin release

Based on the collective structural and functional data presented in this work combined with the findings of other recently published[18,31] work on T6SS-exported Rhs proteins, we propose a model for cytoplasmic delivery of RhsA and suggest a possible release mechanism for the toxin domain of the effector (Figure 6).

First, RhsA is expressed and autoproteolytically cleaved at an N-terminal and C-terminal position. After cleavage, the N-terminal domain comprising the prePAAR motif, the PAAR domain and the linker to the Rhs barrel interacts with the EagR1 chaperone and this complex remains associated with the RhsA cocoon. The C-terminal domain comprising the detached toxin domain is partially unfolded and remains inside the cocoon.

EagR1 then shields the transmembrane domains of RhsA from the aqueous milieu as the effector is loaded onto VgrG1 in the cytoplasm of the T6SS-containing cell resulting in a mature PFC. The exact location of T6SS effector delivery in the target cell remains unclear and may differ depending on the T6SS tail spike complex being exported[34,35]. Nonetheless, most characterized PAAR effectors exert their catalytic activity in the cytosol of the target cell, i.e. (p)ppApp synthetases, ADP-ribosyl transferases, DNases and NAD^+/$NADP^+$ hydrolases[11,36–38]. We propose that the RhsA-loaded VgrG1 tip is delivered to the periplasm as we have previously suggested for the Tse6 effector and that the PFC spontaneously enters the target cell inner membrane[16]. It remains unclear at which step the chaperone is stripped off; however, Coulthurst and colleagues detected a secreted T6SS Rhs effector in *Serratia* by mass spectrometry and did not detect its cognate Eag chaperone[39]. This finding supports a model whereby EagR1 dissociates from the spike complex during the loading event of the PFC onto an Hcp tube inside the lumen of the T6SS baseplate. But given the fact the TMD helices would be unstable without protective chaperones, it seems more likely that EagR1 is removed shortly before insertion into the membrane of the target cell. As proposed previously for Tse6, we suggest that the transmembrane helices of the TMD spontaneously enter the inner membrane of the target cell.

Given the rigid structure of RhsA, we propose that the barrel remains intact during the firing and translocation events. Two different scenarios are conceivable for the release of the toxin domain from the cocoon and its delivery into the cytosol of the target cell if the spike only protrudes into the periplasm. In one scenario, translocation of the linker between the PAAR domain and the plug domain would lead to the removal of the seal peptide. This would result in the opening of a passageway through which the toxin domain could be threaded in an unfolded or partially unfolded state and subsequently translocated, with the assistance of the N-terminal transmembrane helices, into the cytosol. The RhsA barrel would remain in the periplasm in this case but would interact directly with the inner membrane of the target cell. In a second scenario, the TMD region would facilitate the translocation of the entire RhsA protein. Mechanical forces during translocation in combination with hydrophobic interactions in the membrane could lead to ‘tugging’ of the N-terminal domain and a release of the toxin domain

342 from the cocoon. However, since the barrel is large (86 x 65 Å), its transport across the
343 membrane is relatively unlikely. In both cases, the energy required for release and translocation
344 of the toxin domain would be driven by its spontaneous refolding in the target cell cytosol[40].
345 An alternative route would be the direct delivery of the barrel into the cytosol. This
346 would require that the PFC not only penetrates the outer membrane, but also the inner
347 membrane to advance into the cytosol of the target cell. In this case, the cocoon would be
348 located inside the cytosol of the target cell, and the toxin domain would be released analogously
349 to above-described scenario B. However, it is unclear what the role of the TMD would be in
350 this alternative delivery mechanism.
351 This study provides detailed molecular insights into the autoproteolytic processing of
352 Rhs effectors and its importance for toxin release inside the target cell. It not only enhances
353 our knowledge about Rhs effector function, but also lays a foundation for a mechanistic
354 understanding of how the T6SS machinery functions. The unique ability of the T6SS to mediate
355 contact-dependent killing of a wide range of bacteria may enable development of novel
356 therapeutics for selective depletion of drug-resistant pathogens.
357

358 **Materials and methods**
359 Bacterial strains and culture conditions
360 *Pseudomonas protegens* Pf-5 (Table 2) was grown in Lysogeny Broth (LB) (10 g L^{-1}
361 NaCl, 10 g L^{-1} tryptone, and 5 g L^{-1} yeast extract) at 30°C or on solid LB containing 1.5% or
362 3% agar. Media were supplemented with gentamicin (30 µg mL^{-1}) and irgasan (25 µg mL^{-1})
363 as needed.
364 *Escherichia coli* strains XL-1 Blue, SM10 and CodonPlus (DE3) were used for plasmid
365 maintenance and toxicity experiments, conjugative transfer and protein overexpression,
366 respectively (Table 2). All *E. coli* strains were grown at 37°C in LB medium. Unless otherwise
367 noted, media was supplemented with 150 µg mL^{-1} carbenicillin, 50 µg mL^{-1} kanamycin, 200
368 µg mL^{-1} trimethoprim, 15 µg mL^{-1} gentamicin, 0.25-1.0 mM isopropyl β-D-1-
369 thiogalactopyranoside (IPTG), 0.1% (w/v) rhamnose or 40 µg mL^{-1} X-gal.
370

371 DNA manipulation and plasmid construction
372 All primers used in this study were synthesized by Integrated DNA Technologies
373 (IDT). Molecular biology reagents (Phusion polymerase, restriction enzymes and T4 DNA

ligase) were obtained from New England Biolabs (NEB). Sanger sequencing was performed by Genewiz Incorporated.

Heterologous expression plasmids: pETDuet-1, pET29b and pSCrhaB2-CV. Splicing by overlap-extension PCR was used to make mutant constructs. Standard restriction enzyme-based cloning procedures were subsequently used to ligate wild-type or mutant PCR products into the plasmid of interest.

Generation of *P. protegens* mutants

In-frame chromosomal deletion mutants in *P. protegens* were made using the pEXG2 plasmid as described previously for *Pseudomonas aeruginosa*[41]. Briefly, ~500 bp upstream and downstream of target gene were amplified by standard PCR and spliced together by overlap-extension PCR. The resulting DNA fragment was ligated into the pEXG2 allelic exchange vector using standard cloning procedures (Table 3). Deletion constructs were transformed into *E. coli* SM10 and subsequently introduced into *P. protegens* via conjugal transfer. Merodiploids were directly plated on LB (lacking NaCl) containing 5% (w/v) sucrose for *sacB*-based counter-selection. Deletions were confirmed by colony PCR in strains that were resistant to sucrose, but sensitive to gentamicin. Chromosomal point mutations or epitope tags were constructed similarly with the constructs harboring the mutation or tag cloned into pEXG2. Sucrose-resistant and gentamicin-sensitive colonies were confirmed to have the mutations of interest by Sanger sequencing of appropriate PCR amplicons.

Pseudomonas growth competition assays

Recipient *P. protegens* strains contained a Δ*pppA* mutation to stimulate T6SS effector secretion and induce a 'counterattack' from *P. protegens* donor strains[42]. Recipient strains were also marked with a tetracycline-resistant, *lacZ*-expression cassette at a neutral phage site (*attB*) to differentiate from unlabeled donor strains.

Stationary-phase overnight cultures of *P. protegenes* donors and recipients were mixed in a 1:1 (v/v) ratio and relative abundance of donor:recipient was determined by plating part of the competition mixture on LB plates with 40 µg mL-1 X-gal. Ten microlitres of each competition mixture was then spotted on a 0.45 µm nitrocellulose membrane that was overlaid on a 3% LB agar plate and incubated face up at 30 °C. Competition spots were harvested after 20-25 hours by resuspending in LB and counting CFU by plating on LB agar with 40 µg mL^{-1} X-gal. The final ratio of donor:recipient colony forming units were then normalized to the initial ratios of donor and recipient strains and reported as a competitive index.

408

409 Toxicity assays

410 Wild-type RhsA and the various RhsA truncations and site-specific mutants used in this
411 study were cloned into the rhamnose-inducible pSCrhaB2-CV vector[43]. RhsI was cloned into
412 the IPTG-inducible vector pPSV39[12]. The various RhsA expressing pSCrhaB2-CV plasmids
413 were co-transformed into *E. coli* XL-1 Blue with pPSV39::*rhsI*. Stationary-phase overnight
414 cultures containing these plasmids were serially diluted 10^{-6} in 10-fold increments and each
415 dilution was spotted onto LB agar plates containing 0.1% (w/v) L-rhamnose, 500 µM IPTG,
416 trimethoprim 250 µg mL^{-1} and 15 µg mL^{-1} gentamicin. Photographs were taken after overnight
417 growth at 37°C.

418

419 Protein expression and purification

420 Wild-type RhsA, its N-terminal (C538A, D322A, H530A, D318-320A) and C-terminal
421 autoproteolysis mutants (D1346A, D1364N) were cloned into MCS-1 of pETDuet-1 and co-
422 expressed with RhsI, which was cloned into MCS2. Plasmids were co-transformed into *E. coli*
423 BL21 Codon Plus alongside a pET29b vector expressing EagR1.

424

425 Purification of $RhsA_{\Delta NT}$ and VgrG1 for cryo-EM

426 $RhsA_{\Delta NT}$ was co-expressed with RhsI using pETDuet-1 (see Table 3 for details). VgrG1
427 was expressed in isolation using pET29b. Plasmids were individually expressed in *E. coli* BL21
428 Codon Plus. Strains harboring pETDuet-1 expressing $RhsA_{\Delta NT}$-RhsI or pET29b expressing
429 VgrG1 were inoculated in separate flasks containing 100 mL LB with selection and incubated
430 overnight in a shaking incubator at 37°C. Following 14-18hr incubation, the culture was
431 subinoculated (1/50 dilution) into four flasks each with 1 litre of LB and appropriate antibiotic
432 selection. Cultures were initially incubated at 37°C until the culture reached an OD of ~0.3.
433 The incubator was subsequently cooled to 18°C and each culture induced with 1 mM IPTG
434 upon reaching an OD of 0.6-0.7. Cultures were harvested by centrifugation at 9,800 *g* for 10
435 minutes. Pellets were resuspended in 30 mL lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 10
436 mM imidazole) and lysed by sonication (6 x 30 second pulses, amplitude 30%) and then spun
437 at 39,000 *g*. Cleared lysates were applied to a Ni-NTA gravity flow column equilibrated using
438 lysis buffer. The column was washed with the lysis buffer three times and the samples were
439 eluted in 3 mL of elution buffer (lysis buffer with 400 mM imidazole). The samples were

applied to a HiLoad 16/600 Superdex 200 column equilibrated in 20 mM Tris-HCl pH 8.0 150 mM NaCl and collected in the same buffer.

Negative stain electron microscopy

Four microliters of sample at a concentration of 0.005 mg/ml were applied to freshly glow-discharged carbon-coated copper grids. After 90s incubation time, excess sample was blotted away with Whatman No. 4, then washed twice with four microliters purification buffer and once with 0.75% (w/v) uranyl formate. A second batch of staining solution was incubated on the grid for 90s before excess was again blotted away. Grids were air-dried and imaged on a JEOL JEM-1400 microscope, equipped with a LaB_6 cathode and 4k × 4k CMOS detector F416 (TVIPS), operating at 120 kV.

Sample vitrification and data collection

Three microliters of the EagR1-RhsA-VgrG1 complex, at a concentration of 0.1 mg/ml, were applied to a freshly glow-discharged holey carbon grid (QF 2/1 200-mesh). The grid was blotted for 3s (blot force -5, drain time 0.5 s, 8 °C, 100 % humidity) and immediately plunged into nitrogen cooled liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). Data collection was performed on a C_S-corrected Titan Krios (Thermo Fisher Scientific) operating at 300 kV in an automated fashion using EPU (Thermo Fisher Scientific). Movies were recorded on a Falcon 3 detector in linear mode at a nominal magnification of 59,000x with a calibrated pixel size of 1.1 Å/pixel. Image stacks were acquired in a defocus range from -1.2 to -2.2 µm with an accumulated dose of 90 $e^-/Å^2$ fractionated over 40 frames with a total exposure time of 1.5 s.

For $RhsA_{\Delta NT}$, three microliters of sample, at a concentration of 4 mg/ml, were applied to a freshly glow-discharged holey carbon grid (QF 1.2/1.3 200-mesh). The grid was blotted for 3s (blot force -5, drain time 0.5 s, 8 °C, 100 % humidity) and immediately plunged into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). The grid was transferred to a Titan Krios (Thermo Fisher Scientific) operating at 300 kV equipped with a GIF BioQuantum energy filter (Gatan), set to a slit width of 20 eV, and K3 Summit Detector (Gatan). Movies were recorded in counting mode at a nominal magnification of 105,000x with a calibrated pixel size of 0.91 Å/pixel in an automated fashion using EPU (Thermo Fisher Scientific). Image stacks were acquired in a defocus range from -0.8 to -2.2 µm with an accumulated dose of 61 $e^-/Å^2$ fractionated over 20 frames with a total exposure time of 2 s.

Both datasets were monitored live with TranSPHIRE[44] to evaluate i.e. the defocus range and astigmatism. Pre-processing was performed on-the-fly in TranSPHIRE including drift-correction and dose-weighting with MotionCor2[45], CTF estimation on dose-weighted micrographs with CTFFIND4[46] and picking using the general model of crYOLO[47].

Cryo-EM image processing

After preprocessing in TranSPHIRE all processing steps were carried out in the SPHIRE[48] software package unless otherwise stated. Images with a resolution limit less than 6 Å were unselected using the graphical CTF assessment tool in SPHIRE for both datasets.

In case of the PFC, particles were extracted with box size of 408 x 408 pixels from 1250 good micrographs. Reference-free 2D classification and cleaning of the dataset was performed with the iterative stable alignment and clustering approach ISAC[49] implemented in SPHIRE. ISAC was performed at pixel size of 6.29 Å/pix. Using the Beautifier tool, the original pixel size was restored creating sharpened 2D class averages showing high-resolution details. A subset of particles showing clear high-resolution details were selected for structure refinement. 3D refinement was performed in MERIDIEN imposing C_3 symmetry with a 25 Å lowpass-filtered reference of our previously determined VgrG1 structure (EMD-0136). The two half-maps were combined with the PostRefiner tool in SPHIRE using a soft mask and automatic estimation of B-factors. Details of the processing workflow are shown in Supplementary Figure 8.

For $RhsA_{\Delta NT}$, particles were extracted from 13,090 good micrographs with a final box size of 288 x 288 pixels. Particles were subjected to ISAC which was performed at pixel size of 3.3 Å/pix. A subset of particles showing high-resolution features were selected from beautified class averages. An initial reference for refinement was generated from the beautified averages using RVPER. All refinements and classifications were performed with implied C_2 symmetry. Subsequent 3D refinement of the good particles with MERIDIEN yielded a map of overall 3.6 Å resolution but showed clear resolution anisotropy in the peripheries of the barrels. We imported the particle stack into RELION 3.1.0[50] with projection parameters obtained from MERIDIEN. Particles were classified into 4 classes without image alignment (T=4, C_2 symmetry, soft mask) using the map from MERIDIEN as reference low-pass filtered to 30 Å. The particles belonging to the class displaying the highest resolution were selected for another round of MERIDIEN after removal of duplicated particles (minimum inter-particle

distance threshold 100 Å). Particles were further CTF-refined and polished in Relion. The final refinement was performed with MERIDIEN. The final map was evaluated using the 3D FSC tool[51]. Map for Figure 1 and was postprocessed with the DeepEMhancer[52] using the high-resolution model. Local resolution was estimated with a normal postprocessed map in SPHIRE. Details of the processing workflow are shown in Supplementary Figure 1.

Model building, refinement and validation

The previously obtained VgrG1 structure from *P. aeruginosa* (PDB ID: 6H3L) was docked into the map as rigid body in UCSF Chimera[53]. The *P. protegens* VgrG1 sequence was manually adjusted in Coot[54] and iteratively refined in Phenix[55] and ISOLDE[56]. Model validity was assessed in Phenix with MolProbity[57]. Final model statistics are given in Table 1.

For $RhsA_{\Delta NT}$, initial backbone traces were identified with automated model building software Buccaneer[58]. The model was manually adjusted to completion in Coot[54] using both the DeepEMHanced and a normal postprocessed map to exclude any bias. Stabilized β-strands of the toxin domain were identified via secondary structure prediction and manually placed inside the densities. The single interpretable α-helix was not predicted by secondary structure predictions. Instead, we manually check the rest of the sequence to determine the sequence register. We placed residues 289-302 in this density guided by AlphaFold[59] predictions. This was moreover guided by the matched hydrophobicity towards the interacting Rhs repeats. To identify residues corresponding to the linker density which complements the plug domain, we searched in the vicinity of the helix sequence for residues encoding bulky side chains. F271 served as anchor point to build residues 268-275 guided by AlphaFold[59] predictions.
Iterative refinement of the model in Phenix[55] and ISOLDE[56] was performed until convergence. Model validity was assessed in Phenix with MolProbity[57]. Final model statistics are given in Table 1.

Figures were prepared in Chimera X[60]. Multiple sequence alignments and secondary structure predictions were calculated using the MPI Bioinformatics toolkit[61,62] and visualized for creation of figures in Jalview[63].

Supplementary Material

Table 1: Statistics of cryo-EM data collection, image processing and model validation.

	VgrG1	**RhsA$_{\Delta NT}$**
Data collection		
Microscope	Titan Krios (X-FEG, Cs-corrected)	Titan Krios (X-FEG, Cs 2.7 mm)
Magnification	59,000	105,000
Voltage (kV)	300	300
Defocus range (μm)	-1.2 to -2.2	-0.8 to -2.2
Camera	F3 linear	K3 counting
Pixel size (Å/pixel)	1.1	0.91
Total electron dose (e^{-}/Å^2)	90	62
Exposure time (s)	1.5	2
Number of images	1250	13,090
Refinement		
Number of final particles	423,980	454,740
Final resolution (Å)	3.3	3.3
Symmetry	C_3	C_2
Map sharpening B factor (Å^2)	-174.5	DeepEMhancer
Model composition		
Non-hydrogen atoms	14868	17070
Protein residues	1881	2094
RMSD bond	0.007	0.006
RMSD angles	0.722	0.570
Model-to-map fit, CC Mask	0.77	0.77
Validation		
MolProbity	1.85	1.42
Clashscore	6.86	4.15
EMRinger score	2.77	4.70
Poor rotamers (%)	0.00	0.45
Ramachandran		
Favored (%)	92.46	96.59
Allowed (%)	7.22	3.31
Outliers (%)	0.32	0.10

542 Table 2: Bacterial strains used in this study.

Organism	**Genotype**	**Description**	**Reference**
P. protegens Pf-5	wild-type		Paulsen et al., Nature Biotechnology, 2005[64]
	PFL_6096_D1324N	Expresses RhsA D1324N	This study
	PFL_6096_D1346N	Expresses RhsA D1346N	This study
	ΔPFL_6079 attB::*lacZ*, TetR	*pppA* deletion, constitutive *lacZ* expression, Tet^R	Ahmad et al., eLife, 2018[14]
	ΔPFL_6079 ΔPFL_6096 ΔPFL_6097 attB::lacZ, TetR	*pppA rhsA rhsI* deletion, constitutive *lacZ* expression, Tet^R	Ahmad et al., eLife, 2018[14]
E. coli SM10 λpir	*thi thr leu tonA lac Y supE recA*::RP4-2-Tc::Mu	Conjugation strain	BioMedal LifeScience
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB* $lacI^q$ ZΔ*M15* Tn*10* (Tet^R)]	Cloning strain	Novagen
E. coli BL21 (DE3) CodonPlus	F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3) pLysS(cm^R)	Protein expression strain	Novagen

543

544 Table 3: Plasmids used in this study.

Plasmid	Relevant features	Reference
pETDuet-1	Co-expression vector with *lacI*, T7 promoter, N-terminal His_6 tag in MCS-1, Amp^R	Novagen
pET29b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Kan^R	Novagen
pEXG2	Allelic exchange vector containing *sacB*, Gm^R	Hmelo et al., Nature Protocols, 2015[41]
pSCrhaB2-CV	Expression vector with *PrhaB*, Tmp^R	Cardona et al., Plasmid, 2005[43]
pPSV39-CV	Expression vector with *lacI*, *lacUV5* promoter, C-terminal VSV-G tag, Gm^R	Silverman, Cell, 2013[12]
pETDuet-1::PFL_6096 ::PFL_6097	Co-expression vector for RhsA and RhsI	Ahmad et al., eLife, 2018[14]
pETDuet-1::PFL_6096_74-CT ::PFL_6097	Co-expression vector for $RhsA_{\Delta NT}$ and RhsI	Ahmad et al., eLife, 2018[14]
pET29b::PFL_6094	Expression vector for VgrG1	This study
pETDuet-1::PFL_6096_74-CT_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged $RhsA_{\Delta NT}$ and RhsI	This study
pETDuet-1::PFL_6096_74-CT_D1324N_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged $RhsA_{\Delta NT}$ D1324N and RhsI	This study
pETDuet-1::PFL_6096_74-CT_D1346N_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged $RhsA_{\Delta NT}$ D1346N and RhsI	This study
pETDuet-1::PFL_6096_290-CT ::PFL_6097	Co-expression vector for $RhsA_{\Delta N}$ and RhsI	This study
pET29b::PFL_6095	Expression vector for EagR1	Ahmad et al., eLife, 2018[14]
pEXG2::PFL_6096_D1324N	Allelic exchange construct for generating *rhsA*_D1324N	This study
pEXG2::PFL_6096_D1346N	Allelic exchange construct for generating *rhsA*_D1346N	This study
pSCrhaB2-CV:: PFL_6096_74-CT	Expression vector for $RhsA_{\Delta NT}$	This study
pSCrhaB2-CV:: PFL_6096_74-CT_D1324N	Expression vector for $RhsA_{\Delta NT}$ D1324N	This study
pSCrhaB2-CV:: PFL_6096_74-CT_D1346N	Expression vector for $RhsA_{\Delta NT}$ D1346N	This study

pPSV39-CV::PFL_6097	Expression vector for RhsI	Tang et al, JBC, 2018[38]

545

Figures

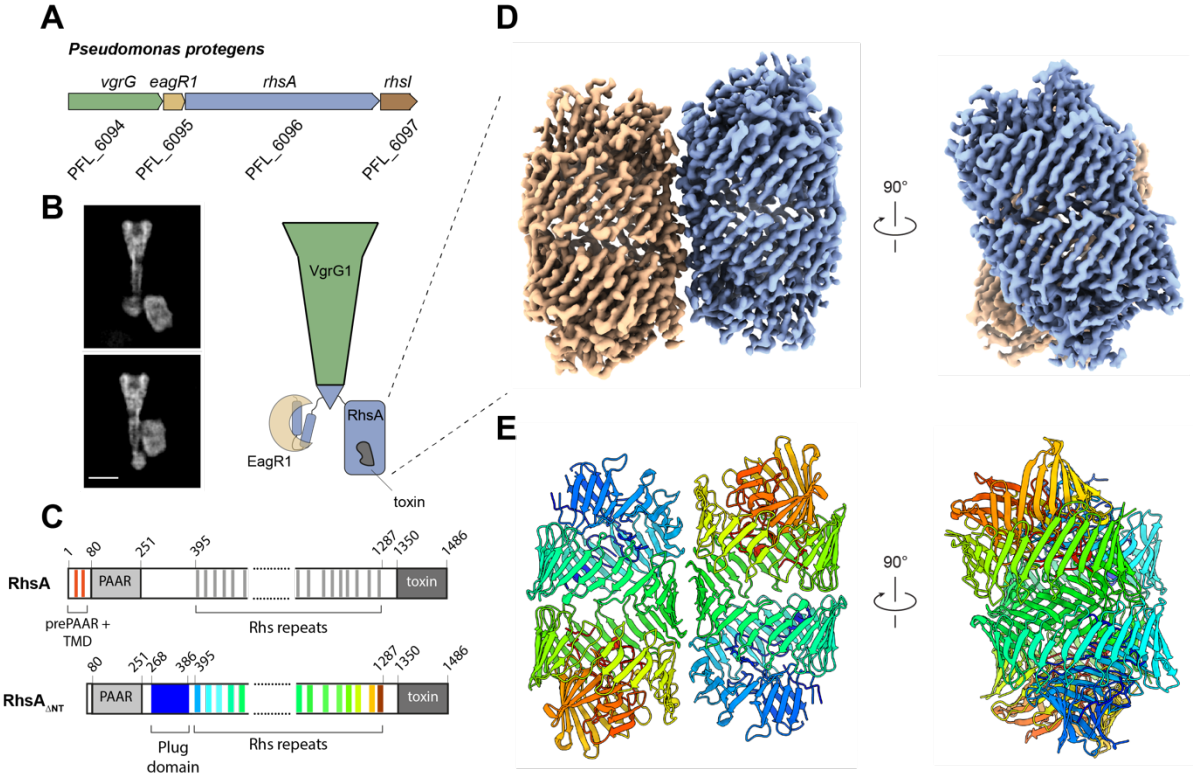


Figure 1: High-resolution structure of the T6SS effector RhsA.

(A) Genomic context of *rhsA* (PFL_6096, blue) in *Pseudomonas protegens* Pf-5. Upstream genes encoding the cognate VgrG protein, *vgrG1* (PFL_6094, green) and RhsA-specific chaperone, *eagR1* (PFL_6095, light brown) are shown. Self-protection against RhsA is conferred via expression of the downstream immunity-encoding gene, *rhsI* (dark brown). (B) Representative cryo-EM 2-D class averages of the assembled pre-firing complex composed of VgrG1, RhsA and EagR1 (left) and a schematic representation of each of the components that comprise this complex (right). Scale bar, 10 nm. (C) Full-length RhsA contains a prePAAR motif and a TMD comprising two transmembrane helices upstream of its PAAR domain. The $RhsA_{\Delta NT}$ truncation of RhsA was used in this study to determine the high-resolution structure of the RhsA cocoon. (D) Cryo-EM density of $RhsA_{\Delta NT}$ displayed perpendicular to the central symmetry axis of the barrel and rotated 90° clockwise (map postprocessed with DeepEMhancer). (E) Cartoon representation of the atomic model of RhsA colored in rainbow from N-terminus (blue) to C-terminus (red).

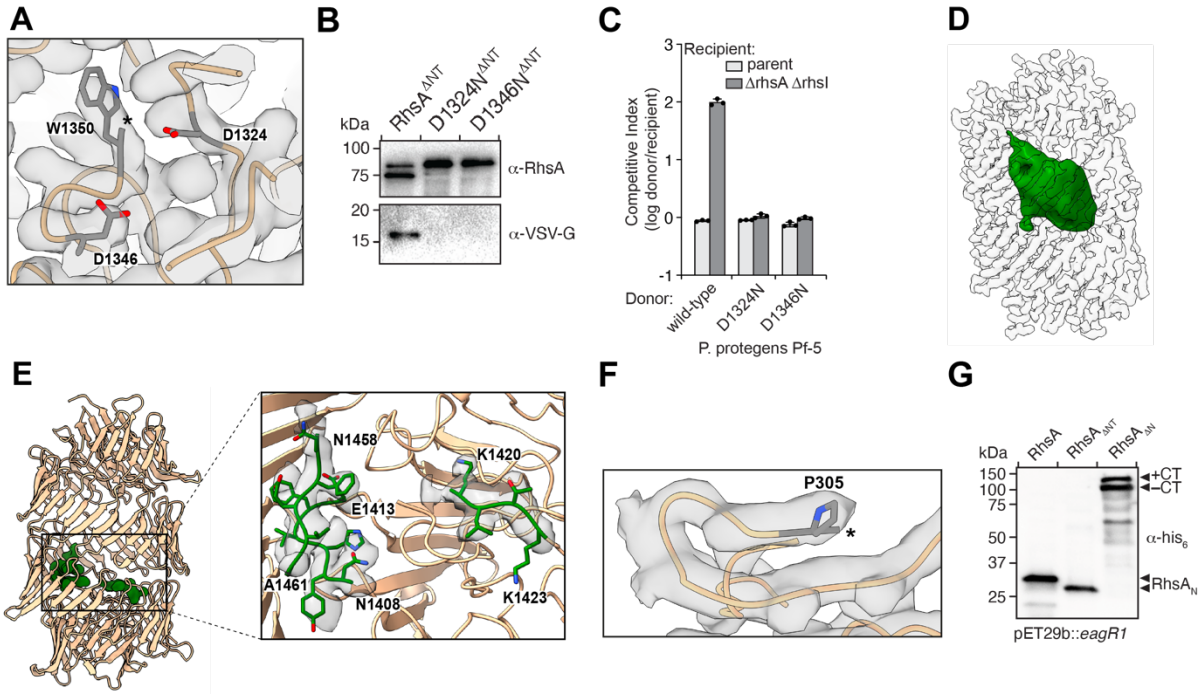


Figure 2: Autoproteolysis of RhsA occurs at its N- and C-terminus.

(A) RhsA is autoproteolytically cleaved at its C-terminus at position W1350. The end of the connected density is indicated with an asterisk. Catalytic aspartates D1324 and D1346, are shown in stick representation. (B) Western blot analysis shows proteolytic cleavage of the C-terminal toxin domain. Mutation of either D1324 or D1346 to asparagine prevents autoproteolytic cleavage of the C-terminal RhsA toxin domain. Blots were performed against the Rhs barrel (α-RhsA) and against a C-terminal VSV-G epitope tag (α-VSV-G). (C) Outcome of intraspecific growth competition assays between the indicated *P. protegens* donor and recipient strains. Donor strains were competed against recipient strains that either contain (light grey) or lack the *rhsA-rhsI* effector-immunity pair (dark grey). The recipient strains also lack *pppA* to stimulate type VI secretion in donors[14]. Data are mean ± s.d. for $n = 3$ biological replicates and are representative of two independent experiments. (D) The toxin domain of RhsA is encapsulated by its cocoon-shaped Rhs repeat-containing domain. Difference map of the encapsulated toxin at extremely high-density threshold and low pass-filtered to 20 Å (green) is shown. The Rhs cocoon is depicted using a transparent space-filling representation at normal map threshold. (E) Regions of the toxin domain are stabilized inside the cocoon through interactions with the C-terminal autoprotease domain. The densities appear at the same density threshold as the rest of the map. The atomic models are shown in stick representation. (F) RhsA undergoes N-terminal cleavage at proline 305. The end of the density is indicated with an asterisk. (G) N-terminal cleavage occurs in RhsA and a mutant lacking the prePAAR-TMD ($RhsA_{\Delta NT}$) but not in a mutant lacking the entire N-terminal region ($RhsA_{\Delta N}$). The indicated RhsA constructs were purified from *E. coli* and subject to Western blot and detected using an N-terminal His_6-tag antibody (α-His).

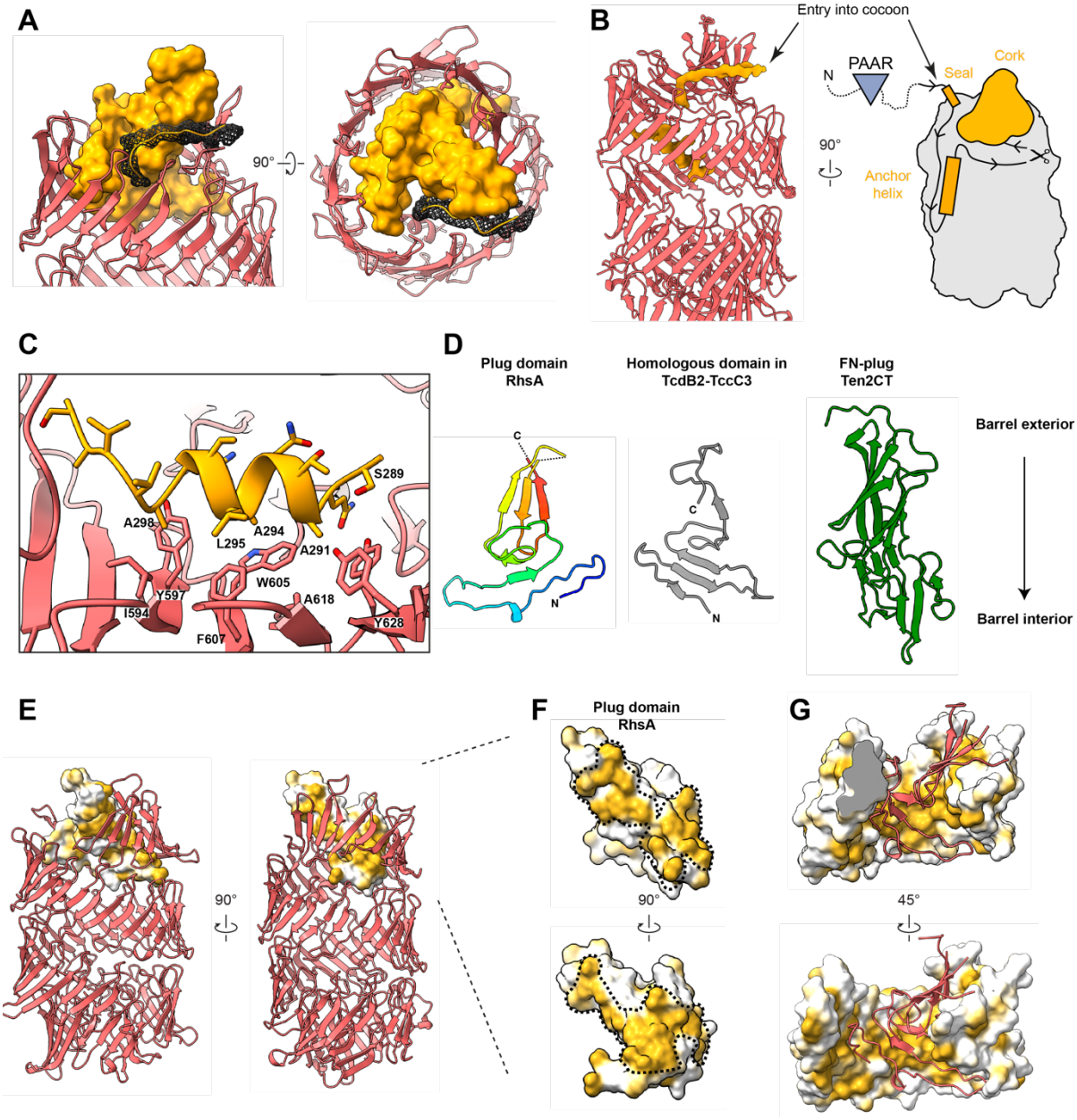


Figure 3: A unique plug domain seals the Rhs barrel of RhsA.

(A) The cocoon structure of RhsA is closed off by an N-terminal plug comprised of a 'cork' domain (orange, density representation), a 'seal' peptide (mesh), and an anchor helix (B, C). The seal and the cork density together form a cap structure. (B) The seal peptide, which also functions as the linker to the unmodelled PAAR domain, not only complements the shape of the cork but is also the entry point of the N-terminal part of the protein into the inside of the cocoon. The cocoon remains stably bound to the cleaved N-terminal region, including the PAAR domain, due to the anchor helix inside the cocoon. (C) The anchor helix is stabilized by hydrophobic interactions with the inner wall of the Rhs repeats. (D) Cartoon representation of the cork region of the RhsA plug domain and a comparison with the homologous N-terminal plug domain of *Photorhabdus luminescens* TcdB2 (grey, PDB ID: 6H6G) and the unique FN-plug domain of human teneurin2 (green, PDB ID: 6FB3). The lower part of each depicted plug domain inserts into each of their respective Rhs barrels. The cork domain of RhsA is colored in rainbow. (E) Surface representation of the cork domain of RhsA. The molecular surface is colored according to hydrophobicity where ochre and white indicate hydrophobic and hydrophilic regions, respectively. The Rhs barrel

is shown as a cartoon. (F) Enlarged view of the cork domain of RhsA. The hydrophobic surface spirals around the domain as indicated by the black dashed line. (G) The upper Rhs repeats of RhsA possess complementary hydrophobic patches to those found on its plug domain (shown in cartoon representation, red).

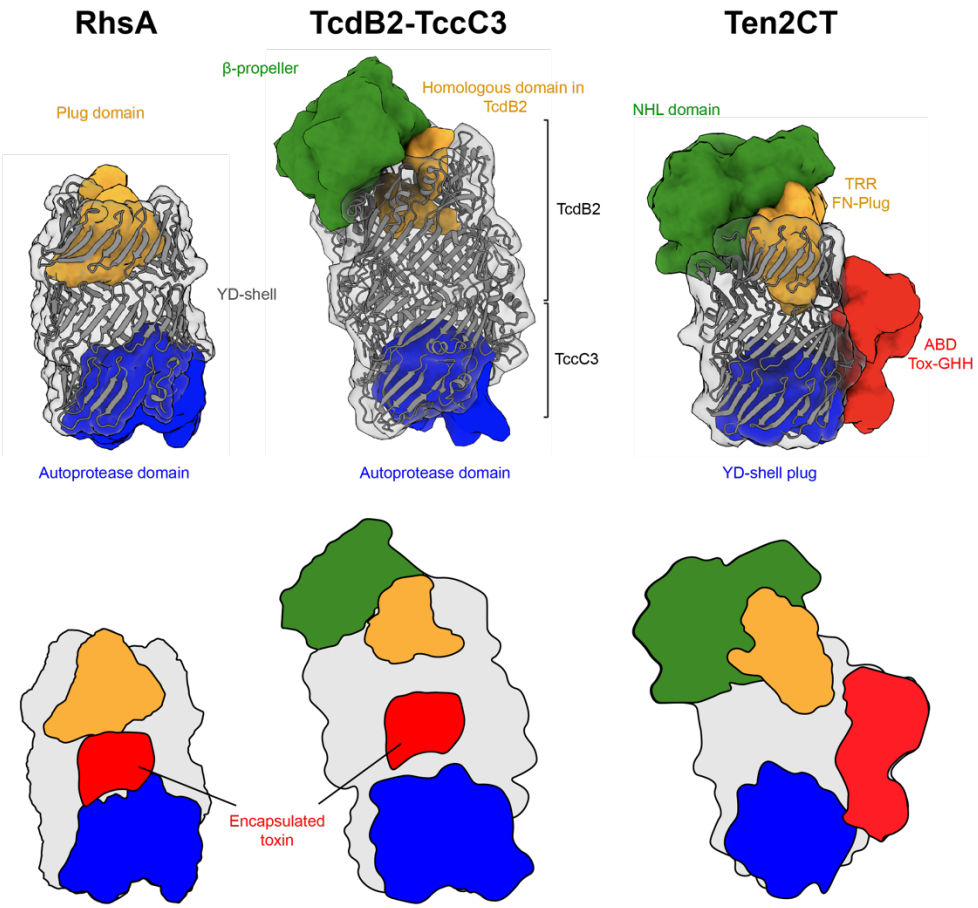


Figure 4: Structural comparison of Rhs repeat containing proteins.

The C-terminal autoproteolysis domain (left and middle) or YD-shell plug (right) is structurally conserved among diverse Rhs proteins (blue). RhsA and BC components of Tc toxin complexes encapsulate their toxic effector domains (red) whereas in teneurin proteins the toxin domain is appended to the outside of the Rhs barrel (red). A distinguishing feature of these Rhs proteins is the unique N-terminal plug domain for each protein family (orange). RhsA is capped by a cork-like plug domain that seals the Rhs barrel (orange). In Tc toxins, the homologous plug domain acts as constriction site and the cocoon is sealed off by a β-propeller domain (green). Teneurin proteins are capped with a non-homologous FN-plug (orange) that is stabilized by an NHL domain (green). The lower row shows schematic representations of the domain organizations.

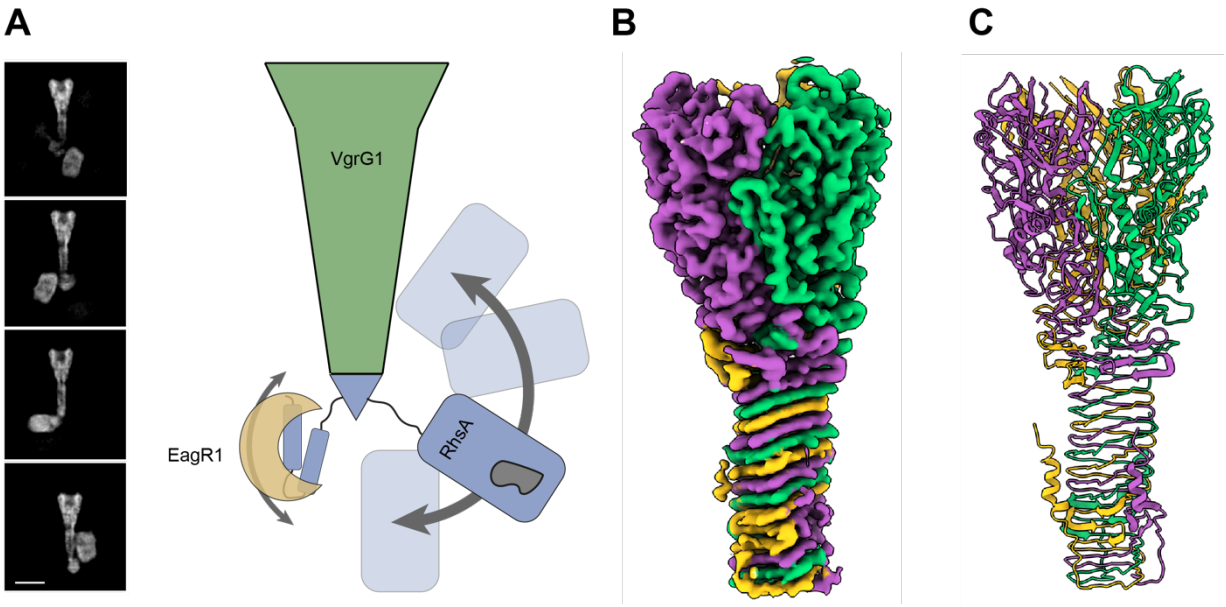


Figure 5: High-resolution structure of *P. protegens* VgrG1.

(A) The RhsA barrel of the assembled pre-firing complex (PFC) displays high positional flexibility relative to VgrG1. Representative cryo-EM 2-D class averages depicting flexibility are shown. Scale bar, 10 nm. (B) Cryo-EM density and (C) ribbon representation of the molecular model of the *P. protegens* VgrG1 trimer viewed along the long axis of the protein. Each protomer is colored differently to highlight their positions within the homotrimeric VgrG1 spike.

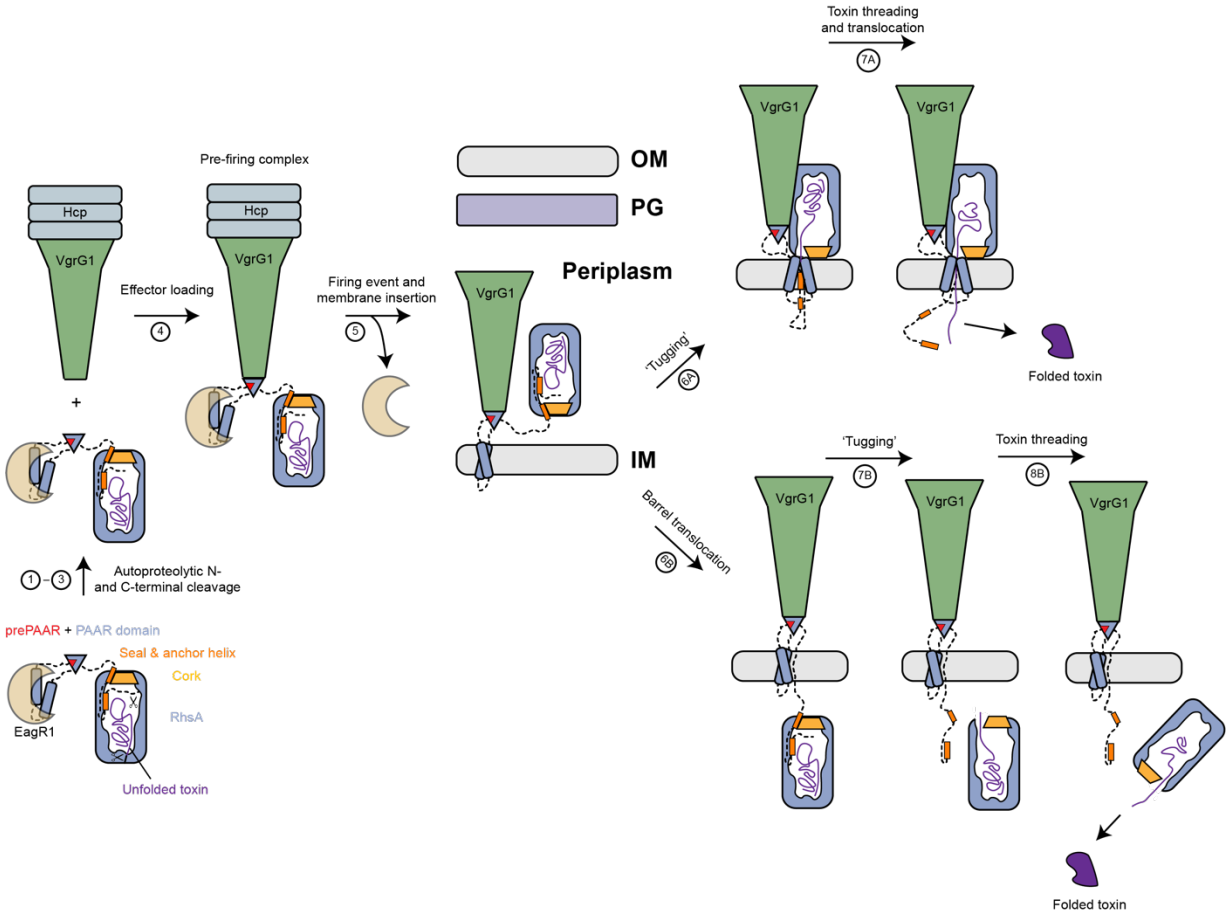
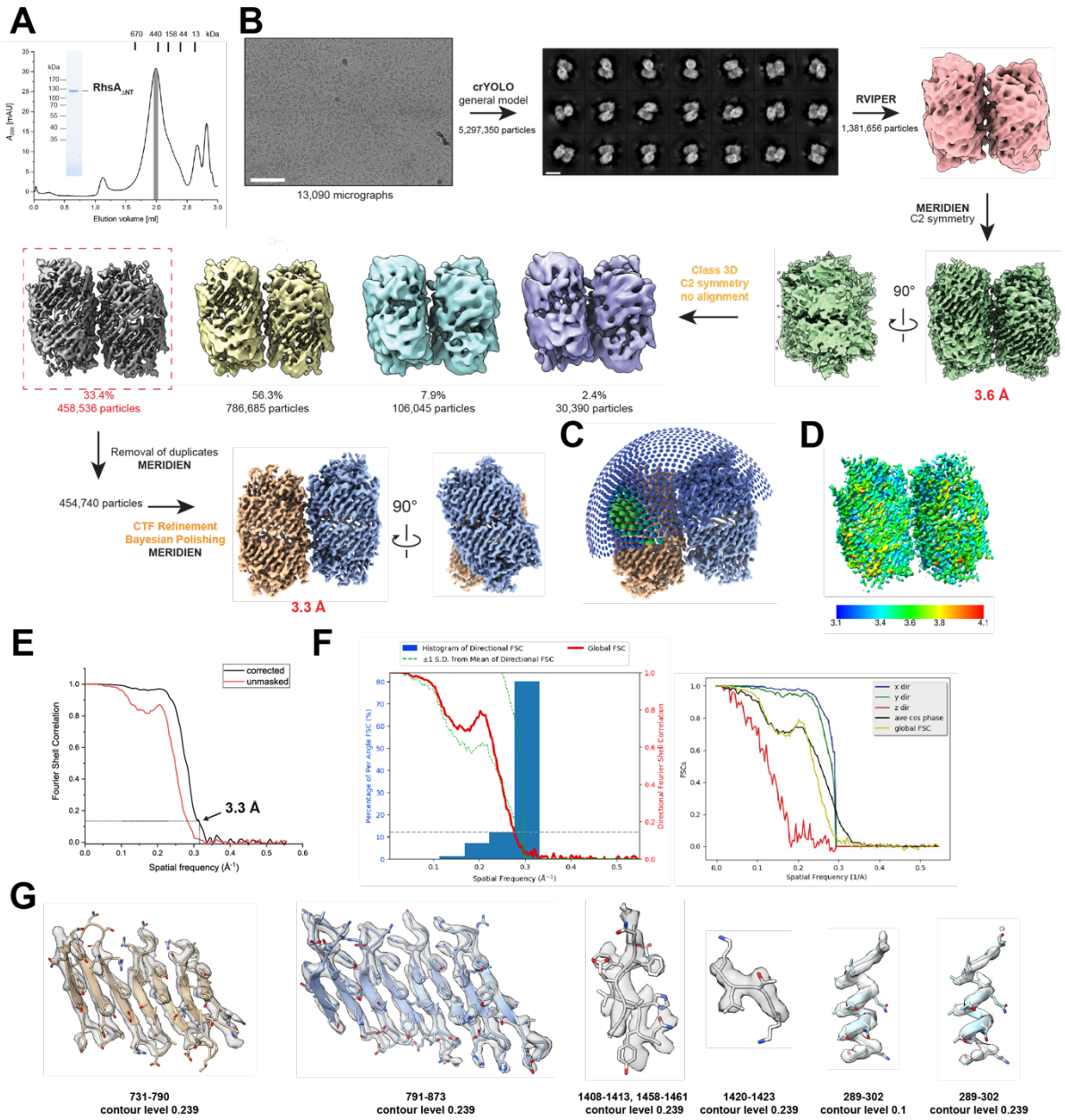


Figure 6: Model of T6SS-dependent delivery of RhsA into the cytoplasm of a susceptible bacterial cell.

(1-3) RhsA undergoes N- and C-terminal autoproteolytic processing. The prePAAR motif 'completes' the PAAR domain fold. The EagR1 chaperone solubilizes the two transmembrane helices of RhsA so that the effector can be loaded onto cytoplasmic VgrG1 via its prePAAR + PAAR domain. (4) The secretion competent RhsA effector is loaded onto the VgrG1 spike. (5) During a firing event the EagR1 chaperone is dissociated from the complex and the T6SS injects the PFC into target cells where it crosses the peptidoglycan (PG) layer and inserts into the inner membrane (IM). (6-8) Two different scenarios (6A or 6B) are proposed as possible mechanisms for toxin domain release from the cocoon into the cytosol of the target cell. Either the toxin domain alone (mechanism 6A) or the entire RhsA barrel is translocated across the inner membrane (mechanism 6B). In both cases, the seal of the cocoon is likely removed by translocation-induced pulling and the energy required for release and translocation of the toxin domain out of the Rhs cage is probably driven by its spontaneous refolding in the prey bacterium's cytosol.

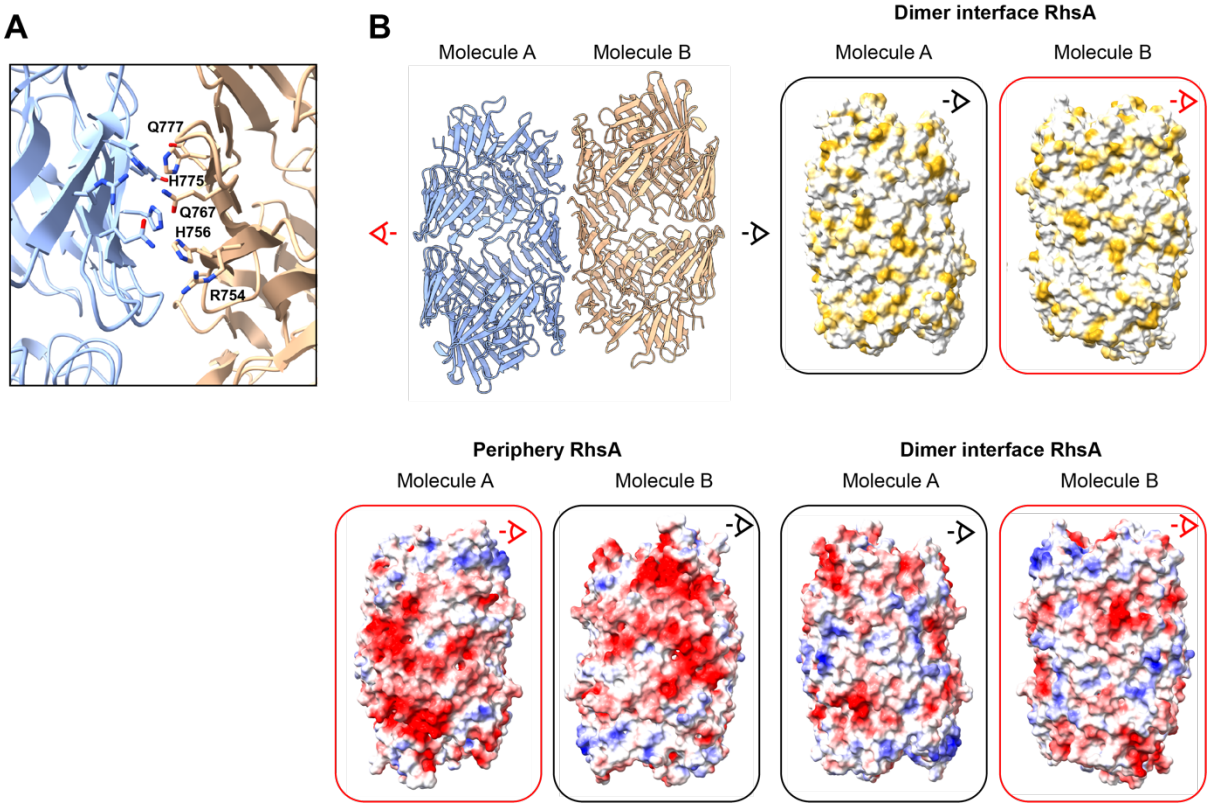
Supplementary Material



Supplementary Figure 1: Cryo-EM processing workflow used to obtain the structure of RhsA$_{\Delta NT}$.

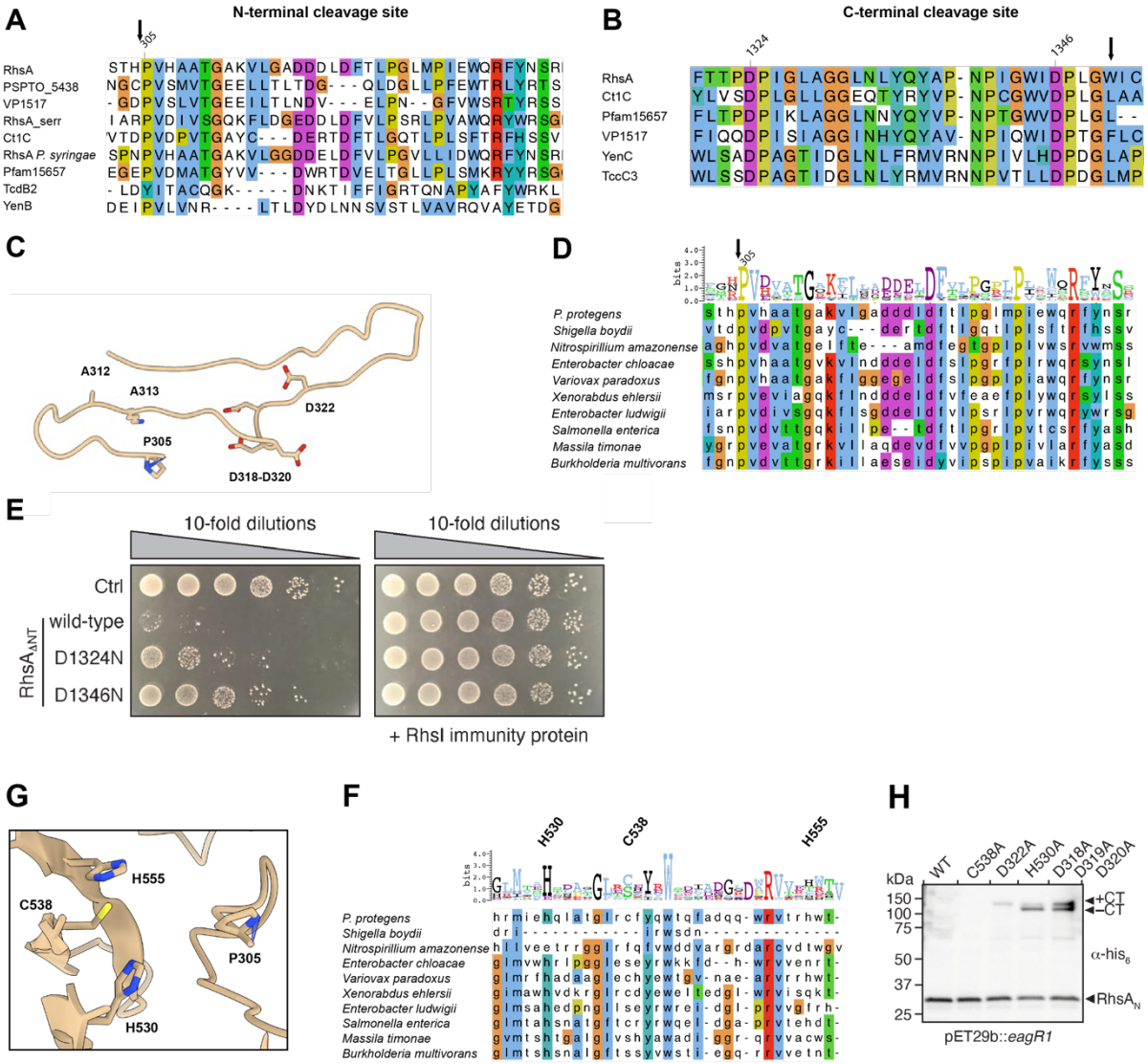
(A) Final purification step of RhsA$_{\Delta NT}$ via size-exclusion chromatography using a Superose 6 5/150 increase column. Molecular weight standards are indicated at their respective elution volumes. The grey bar denotes pooled and concentrated fractions used for cryo-EM analysis. The same material was analyzed for purity via semi-denaturing SDS-PAGE imaged with a stain-free filter. (B) Representative cryo-EM micrograph of RhsA$_{\Delta NT}$ used for structural determination. Scale bar EM micrograph, 100 nm. Particles were picked with the general model of crYOLO. The rest of the processing workflow is indicated and summarized in the methods section. Scale bar 2D class averages, 10 nm. Used software packages are highlighted. Orange font depict steps carried out in Relion. The final map was calculated using MERIDIEN and postprocessed with DeepEMhancer. (C) Angular distribution plot of the final reconstruction. (D) Local resolution estimates visualized on a map postprocessed in SPHIRE. (E)

Fourier shell correlation plot calculated from two independently processed maps. Resolution estimation is reported at the gold standard cutoff of 0.143. (F) Resolution anisotropy was assessed with the 3DFSC online server tool. (G) Selected regions of the map are shown as transparent surface with the built atomic models as stick representations.



Supplementary Figure 2: Structural analysis of the dimer interface of RhsA.

(A) Close up of the dimer interface of the two RhsA molecules at the central symmetry axis. Potential candidate residues engaging in stabilizing interactions are labelled and shown in stick representation. (B) Surface properties of the dimer interface and the periphery of RhsA. The surface of the interface is colored according to its Coulomb potential indicating positively (red, -20) and negatively (blue, +20) charged areas. The second representation shows the same interface but colored according to hydrophobicity. Ochre indicates hydrophobic and white indicates hydrophilic regions. The area facing the outer peripheries of both barrels shows that these regions would electrostatically repel each other and explain why only two barrels can interact at the same time.

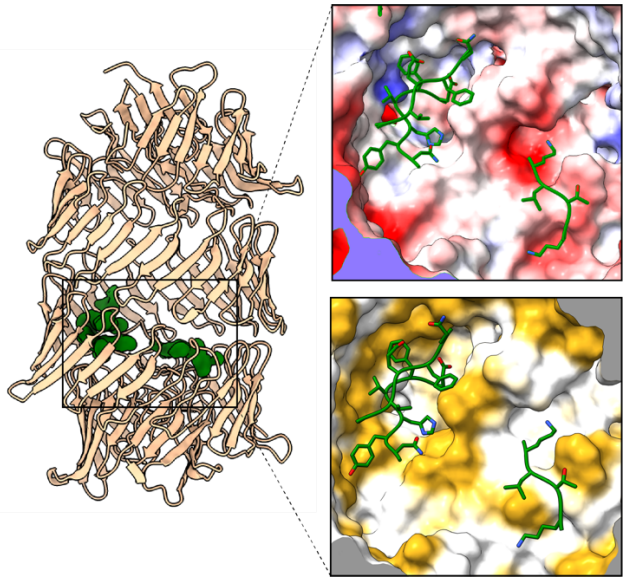


Supplementary Figure 3: Identification and characterization of the autoproteolysis sites of RhsA.

Multiple sequence alignments of the N- and C-terminal cleavage sites. The cleavage sites are indicated by the black arrows. Critical residues are highlighted by the residue numbering. Coloring is according to the ClustalW color code. (A) and (B) Rhs proteins from different species and including T6SS effectors and BC-components of Tc toxins were aligned to show conserved cleavage sites. (C) Cartoon representation of the N-terminal cleavage site in RhsA. Residues A312 and K313 as well as conserved and potential catalytically active residues are show in stick representation. Residues A312 and K313 correspond to the catalytically active glutamates in TseI[18]. (D) Sequence alignment of the T6SS class I prePAAR effectors highlighting the N-terminal cleavage site. Residue conservation is depicted as Weblogo. (E) In *E. coli* toxicity assays show reduced toxicity of $RhsA_{\Delta NT}$ harboring mutations of the catalytic aspartates D1324N and D1346N, respectively. Toxicity could by reversed by overexpression of the immune protein RhsI. (G) Hypothesized cysteine protease motif which is near the N-terminal cleavage site P305. (F) Multiple sequence alignment highlighting the conservation of the hypothesized cysteine protease motif among class I prePAAR effectors. The Weblogo represents residue conservation. (H) Western blot analysis of potential residues involved in N-terminal cleavage and generation of the cleavage product $RhsA_N$. Impaired cleavage was assessed by appearance of the full-length RhsA chain which is a mixture of both species, the C-terminally cleaved fragment (-CT) and the C-terminally uncleaved fragment (+CT). RhsA was

679 coexpressed with its cognate chaperone EagR1. The blot was performed against N-terminal His_6-tagged proteins
680 (α-His).

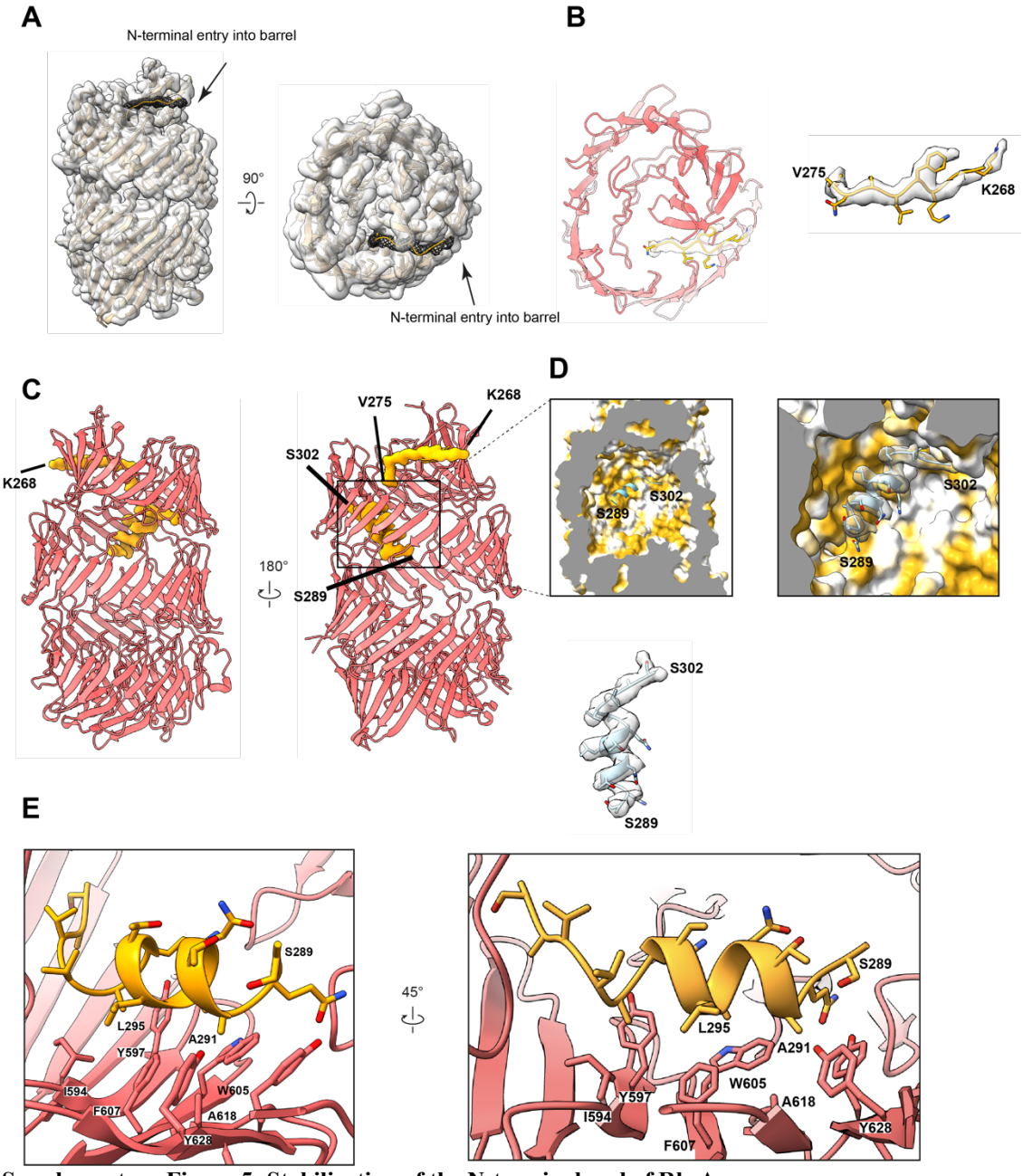
681
682



683
684 **Supplementary Figure 4: Local environment of the ordered toxin fragments inside the Rhs cocoon.**

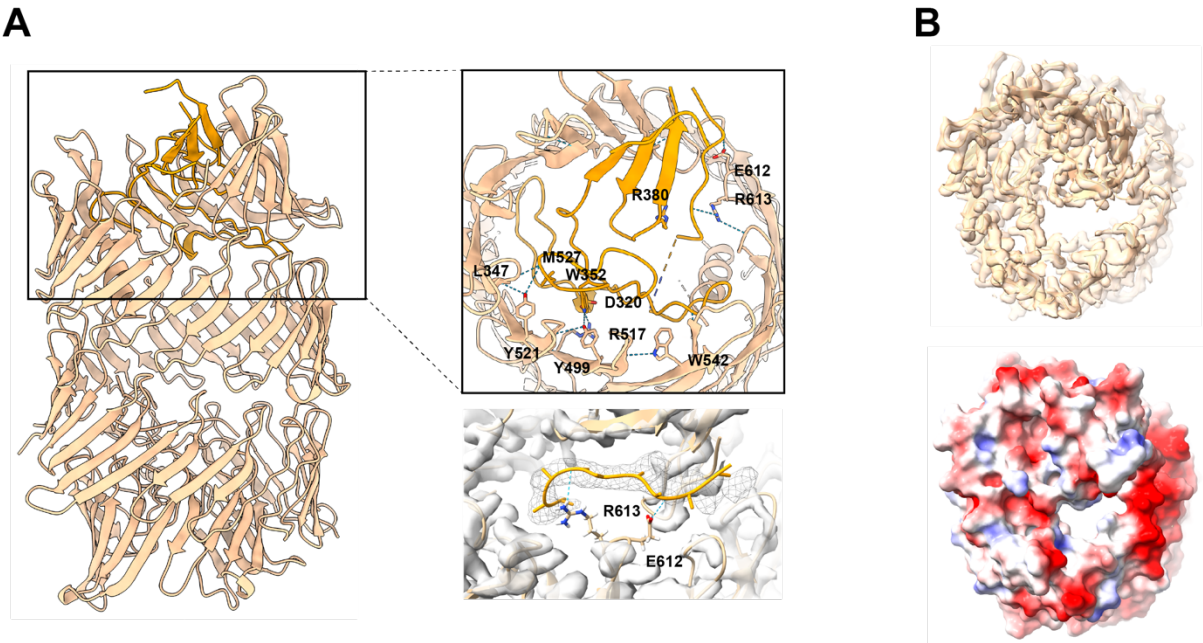
685 Interaction of three β-strands belonging to the toxin domain of RhsA with the Rhs core. The toxin fragments are
686 stabilized by the interaction with hydrophobic and hydrophilic surfaces inside the Rhs barrel. The molecular
687 surface is colored according to hydrophobicity with ochre and white indicating hydrophobic and hydrophilic
688 regions, respectively. The electrostatic representation is colored according to Coulomb potential, which depicts
689 positively (red, -20) and negatively (blue, +20) charged areas.

690



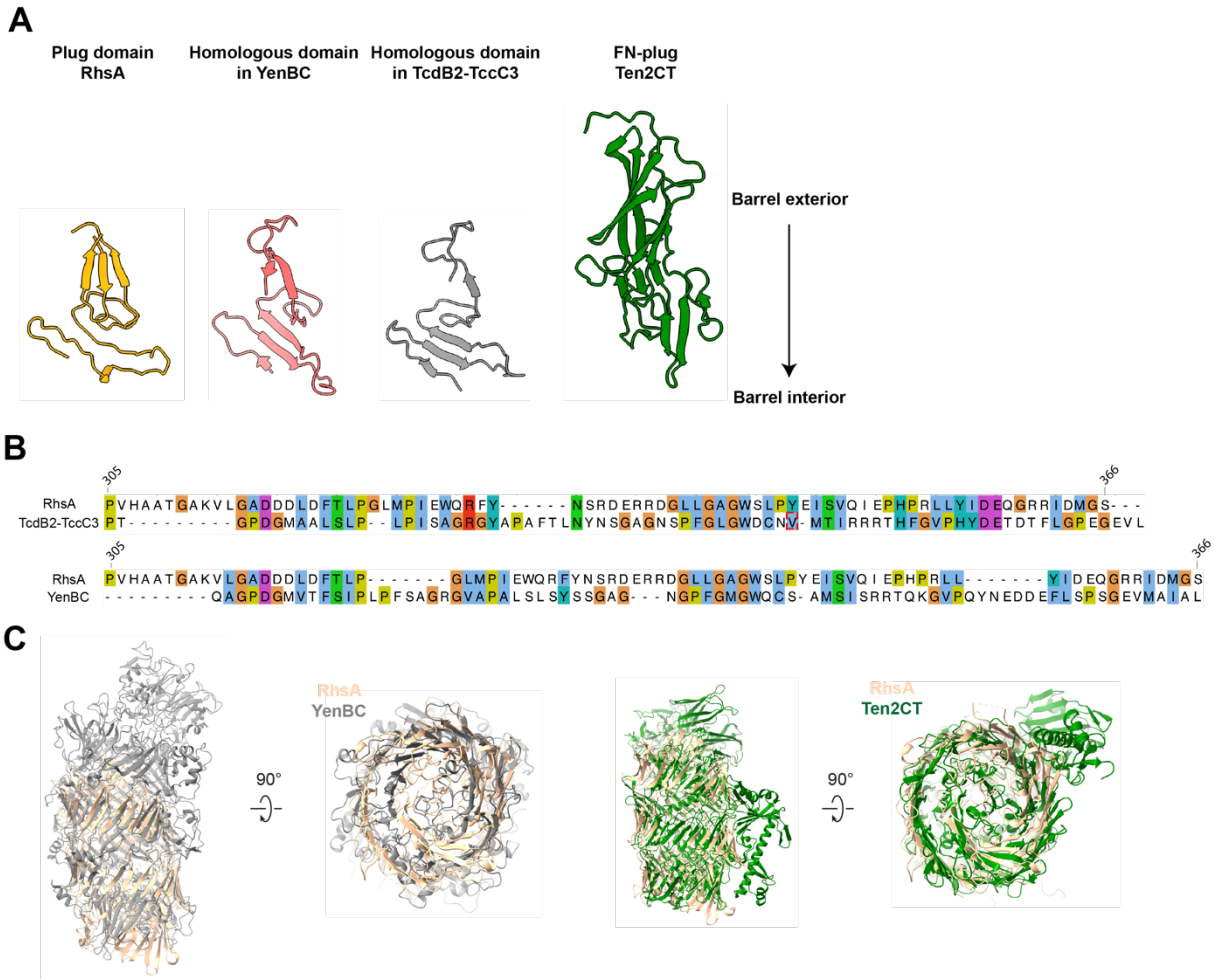
Supplementary Figure 5: Stabilization of the N-terminal seal of RhsA.

(A-B) A small density, corresponding to seal (residues K268-V275) of RhsA enters the barrel from the top (mesh). This results in a complete sealing of the cocoon. (C–E) The seal leads further down into an amphipathic helix which strongly interacts with the inner surface of the cocoon and thus serves as an anchor point for the N-terminal domain. The amphipathic helix is stabilized by hydrophobic interactions with Rhs repeats.



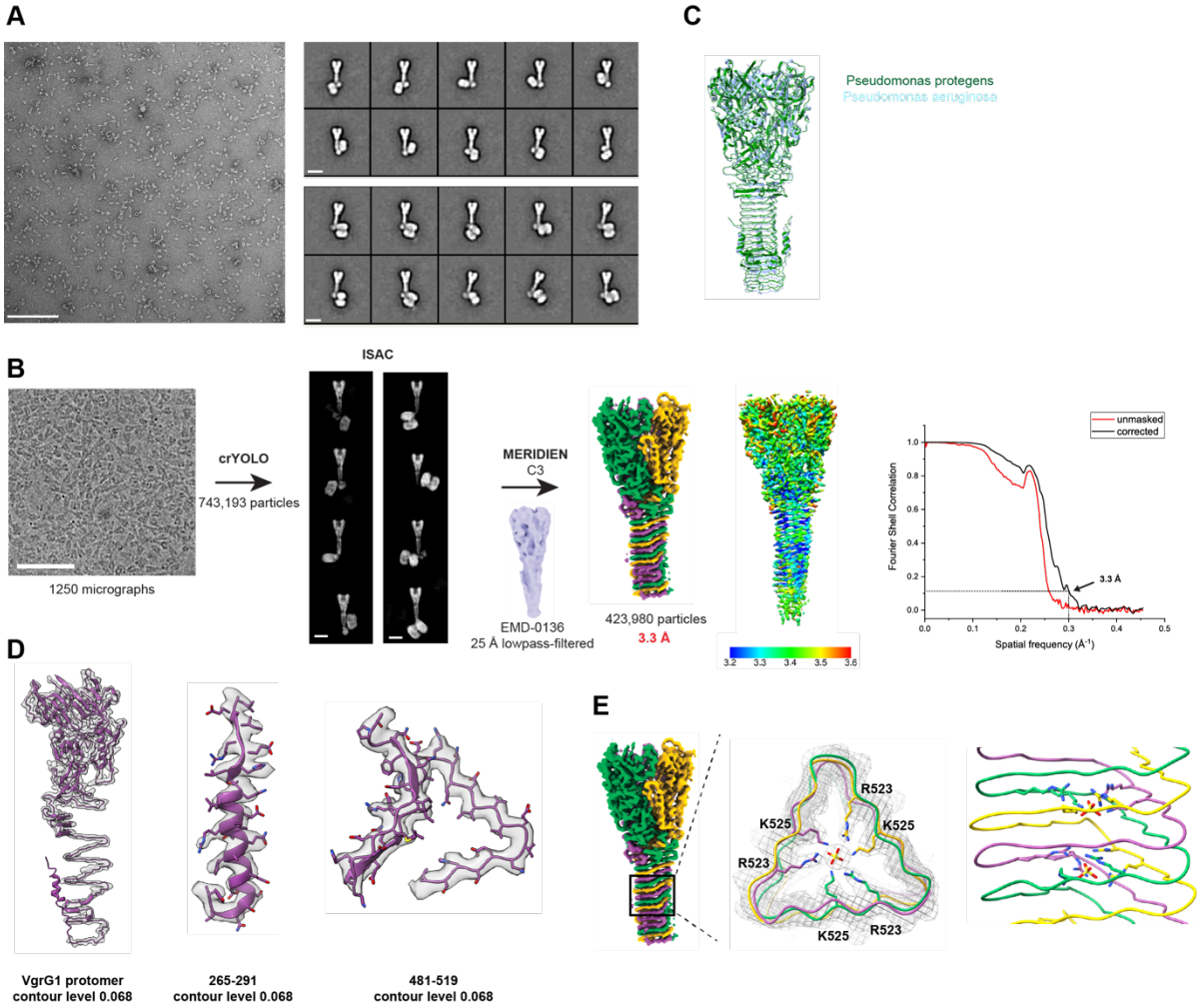
Supplementary Figure 6: Hydrophilic interactions of the plug domain with the Rhs core.

(A) Only a few hydrophilic interactions stabilize the plug domain. Residues participating in hydrogen bonds are labeled and shown in stick representation. The cork and the seal are colored in orange whereas the Rhs core is colored in beige. (B) The barrel would not be closed without the observed density corresponding to the seal. The model for the seal was manually removed to visualize the opening though which toxin is threaded into the target cell after removal of the seal and the anchor helix.



Supplementary Figure 7: Comparison of Rhs repeat containing proteins of known structure.

(A) Comparison of the plug domains of RhsA (orange) with YenBC (PDB ID: 4IGL), TcdB2-TccC3 (PDB ID: 6H6G) and Ten2CT (PDB ID: 6FB3). (B) Sequence alignment of the plug domain of RhsA with the sequences of the homologous domains found in TcdB2-TccC3 (top) and YenBC (bottom) are shown. Residues are colored according to the ClustalW color code. (C) Structural overlay of RhsA with YenBC (left) and Ten2CT (right).



Supplementary Figure 8: Cryo-EM of the PFC.

(A) Representative negative stain micrograph and 2D class averages of the intact PFC consisting of VgrG1, RhsA and EagR1. Scale bar micrograph, 100 nm. Scale bar 2D class averages, 10 nm. (B) Cryo-EM micrograph of the PFC and image processing workflow. Particles were picked with the general model of crYOLO and classified with ISAC. Refinement in MERIDIEN led to a reconstruction of 3.3 Å (FSC=0.143 criterion). The final reconstruction was either colored individually for each protomer or according to local resolution. Fourier shell correlation was plotted according to two independent refined maps. Dashed line indicates the gold standard FSC criterion of 0.143. Scale bar micrograph, 100 nm. Scale bar 2D class averages, 10 nm. (C) Overlay of VgrG1 from *P. protegens* (green) with VgrG1 from *P. aeruginosa* (blue, PDB ID: 6H3N). (D) Map quality of selected parts of the structure (transparent surface) with built atomic models (cartoon representation). (E) Potential ion binding site in the middle part of the β-prism of *P. protegens* VgrG1. Sulfate ions were modelled into the spherical densities. Coordinating residues are displayed in stick representation.

Supplementary Movie 1: Flexibility of the VgrG1-RhsA-EagR1 complex.

The video highlights the positional flexibility of RhsA relative to the VgrG1 spike protein.

Acknowledgments

We thank D. Prumbaum and O. Hofnagel for assistance in electron microscopy data collection and maintaining the EM facility. This work was supported by the Max Planck Society (to S.R.) and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2017–05350 to J.C.W.). J.C.W. is the Canada Research Chair in Molecular Microbiology and holds an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund.

Author Contributions

S.R. and J.C.W. designed the project. J.C.W. and S.A. provided protein complexes. D.Q. prepared specimens and collected data. D.Q., C.G. and P.G. analyzed and processed the EM data. P.G. built the atomic models, analyzed the structures and prepared figures. S.A., K.S. and J.C.W. performed cellular and biochemical assays. P.G. and S.R. wrote the manuscript with input from all authors.

Data Availability

The cryo-EM maps of VgrG1 and RhsA have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes of XXX and YYY, respectively. The refined models for VgrG1 and RhsA were uploaded in the PDB and the entries have the IDs XXXX and YYYY, respectively.

Ethics declaration

The authors declare no competing interests.

References

1. Mougous, J. D. *et al.* A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science (80-.).* **312**, 1526–1530 (2006).
2. Pukatzki, S. *et al.* Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1528–1533 (2006).
3. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells.

Nature **475**, 343–349 (2011).

4. 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).
5. Silverman, J. M., Brunet, Y. R., Cascales, E. & Mougous, J. D. Structure and regulation of the type VI secretion system. *Annu. Rev. Microbiol.* **66**, 453–472 (2012).
6. Cianfanelli, F. R., Monlezun, L. & Coulthurst, S. J. Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends Microbiol.* **24**, 51–62 (2016).
7. Ho, B. T., Dong, T. G. & Mekalanos, J. J. A view to a kill: The bacterial type VI secretion system. *Cell Host Microbe* **15**, 9–21 (2014).
8. Shneider, M. M. *et al.* PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* **500**, 350–353 (2013).
9. Lossi, N. S. *et al.* The HsiB1C1 (TssB-TssC) complex of the Pseudomonas aeruginosa type VI secretion system forms a bacteriophage tail sheathlike structure. *J. Biol. Chem.* **288**, 7536–7548 (2013).
10. English, G. *et al.* New secreted toxins and immunity proteins encoded within the type VI secretion system gene cluster of Serratia marcescens. *Mol. Microbiol.* **86**, 921–936 (2012).
11. Ting, S. Y. *et al.* Bifunctional Immunity Proteins Protect Bacteria against FtsZ-Targeting ADP-Ribosylating Toxins. *Cell* **175**, 1380-1392.e14 (2018).
12. Silverman, J. M. *et al.* Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. *Mol. Cell* **51**, 584–593 (2013).
13. Flaugnatti, N. *et al.* A phospholipase A1 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).
14. Ahmad, S. *et al.* Structural basis for effector transmembrane domain recognition by type VI secretion system chaperones. *Elife* **9**, 1–29 (2020).
15. Whitney, J. C. *et al.* An Interbacterial NAD(P)+ Glycohydrolase Toxin Requires Elongation Factor Tu for Delivery to Target Cells. *Cell* **163**, 607–619 (2015).
16. Quentin, D. *et al.* Mechanism of loading and translocation of type VI secretion system effector Tse6. *Nat. Microbiol.* **3**, 1142–1152 (2018).
17. Lin, R. J., Capage, M. & Hill, C. W. A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome. *J. Mol. Biol.* **177**, 1–18 (1984).

18. Pei, T. T. *et al.* Intramolecular chaperone-mediated secretion of an Rhs effector toxin by a type VI secretion system. *Nat. Commun.* **11**, (2020).
19. Jurėnas, D. & Journet, L. Activity, delivery, and diversity of Type VI secretion effectors. *Mol. Microbiol.* **115**, 383–394 (2021).
20. Jackson, A. P., Thomas, G. H., Parkhill, J. & Thomson, N. R. Evolutionary diversification of an ancient gene family (rhs) through C-terminal displacement. *BMC Genomics* **10**, 1–16 (2009).
21. Tucker, R. P., Beckmann, J., Leachman, N. T., Schöler, J. & Chiquet-Ehrismann, R. Phylogenetic analysis of the teneurins: Conserved features and premetazoan ancestry. *Mol. Biol. Evol.* **29**, 1019–1029 (2012).
22. Minet, A. D., Rubin, B. P., Tucker, R. P., Baumgartner, S. & Chiquet-Ehrismann, R. Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene ten-m, is a neuronal protein with a novel type of heparin-binding domain. *J. Cell Sci.* **112**, 2019–2032 (1999).
23. Meusch, D. *et al.* Mechanism of Tc toxin action revealed in molecular detail. *Nature* **508**, 61–65 (2014).
24. Busby, J. N., Panjikar, S., Landsberg, M. J., Hurst, M. R. H. & Lott, J. S. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature* **501**, 547–550 (2013).
25. Jackson, V. A. *et al.* Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. *Nat. Commun.* **9**, 1–9 (2018).
26. del Toro, D. *et al.* Structural Basis of Teneurin-Latrophilin Interaction in Repulsive Guidance of Migrating Neurons. *Cell* **180**, 323-339.e19 (2020).
27. Lovejoy, D. A., Chawaf, A. Al & Cadinouche, M. Z. A. Teneurin C-terminal associated peptides: An enigmatic family of neuropeptides with structural similarity to the corticotropin-releasing factor and calcitonin families of peptides. *Gen. Comp. Endocrinol.* **148**, 299–305 (2006).
28. Lang, A. E. *et al.* Photorhabdus luminescens ToxinsADP-Ribosylate Actin and RhoA to ForceActin Clustering. *Science (80-.).* **327**, 1139–1143 (2010).
29. Gatsogiannis, C. *et al.* Tc toxin activation requires unfolding and refolding of a β-propeller. *Nature* **563**, 209–233 (2018).
30. Gatsogiannis, C. *et al.* A syringe-like injection mechanism in Photorhabdus luminescens toxins. *Nature* **495**, 520–523 (2013).
31. Donato, S. L. *et al.* The β-encapsulation cage of rearrangement hotspot (Rhs) effectors

is required for type VI secretion. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33540–33548 (2020).

32. Tang, Y., Romano, F. B., Breña, M. & Heuck, A. P. The Pseudomonas aeruginosa type III secretion translocator PopB assists the insertion of the PopD translocator into host cell membranes. *J. Biol. Chem.* **293**, 8982–8993 (2018).
33. Renault, M. G. *et al.* The gp27-like Hub of VgrG Serves as Adaptor to Promote Hcp Tube Assembly. *J. Mol. Biol.* **430**, 3143–3156 (2018).
34. Ho, B. T., Fu, Y., Dong, T. G. & Mekalanos, J. J. Vibrio cholerae type 6 secretion system effector trafficking in target bacterial cells. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9427–9432 (2017).
35. Vettiger, A. & Basler, M. Type VI Secretion System Substrates Are Transferred and Reused among Sister Cells. *Cell* **167**, 99-110.e12 (2016).
36. Ahmad, S. *et al.* An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. *Nature* **575**, 674–678 (2019).
37. Mok, B. Y. *et al.* A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* **583**, 631–637 (2020).
38. Tang, J. Y., Bullen, N. P., Ahmad, S. & Whitney, J. C. Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *J. Biol. Chem.* **293**, 1504–1514 (2018).
39. Cianfanelli, F. R. *et al.* VgrG and PAAR Proteins Define Distinct Versions of a Functional Type VI Secretion System. *PLoS Pathog.* **12**, 1–27 (2016).
40. Roderer, D., Schubert, E., Sitsel, O. & Raunser, S. Towards the application of Tc toxins as a universal protein translocation system. *Nat. Commun.* **10**, 1–11 (2019).
41. Hmelo, L. R. *et al.* Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. *Nat. Protoc.* **10**, 1820–1841 (2015).
42. Basler, M., Ho, B. T. & Mekalanos, J. J. Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152**, 884–894 (2013).
43. Cardona, S. T. & Valvano, M. A. An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in Burkholderia cenocepacia. *Plasmid* **54**, 219–228 (2005).
44. Stabrin, M. *et al.* TranSPHIRE: automated and feedback-optimized on-the-fly processing for cryo-EM. *Nat. Commun.* **11**, 1–14 (2020).
45. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

46. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
47. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 1–13 (2019).
48. Moriya, T. *et al.* High-resolution single particle analysis from electron cryo-microscopy images using SPHIRE. *J. Vis. Exp.* **2017**, 1–11 (2017).
49. Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* **20**, 237–247 (2012).
50. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, 1–22 (2018).
51. Zi Tan, Y. *et al.* Addressing preferred specimen orientation in single-particle cryo-EMthrough tilting. *Nat. Methods* **14**, 793–796 (2017).
52. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* 1–8 (2020) doi:10.1101/2020.06.12.148296.
53. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
55. 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).
56. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 519–530 (2018).
57. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
58. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
59. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
60. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
61. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit

905 with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
906 62. Gabler, F. *et al.* Protein Sequence Analysis Using the MPI Bioinformatics Toolkit.
907 *Curr. Protoc. Bioinforma.* **72**, 1–30 (2020).
908 63. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J.
909 Jalview Version 2-A multiple sequence alignment editor and analysis workbench.
910 *Bioinformatics* **25**, 1189–1191 (2009).
911 64. Paulsen, I. T. *et al.* Complete genome sequence of the plant commensal Pseudomonas
912 fluorescens Pf-5. *Nat. Biotechnol.* **23**, 873–878 (2005).
913