nature communications

Article

https://doi.org/10.1038/s41467-025-66050-x

Complete structures of the YenTc holotoxin prepore and pore reveal the evolutionary basis for chitinase incorporation into ABC toxins

Received: 5 November 2024

Accepted: 27 October 2025

Yu Shang Low[1,2,11], **Solace G. Roche**[1,2,11], **Nadezhda A. Aleksandrova**[1,2,11], **Gabriel Foley**[1], **Jason KK Low**[3], **Joseph K. Box**[1], **Tristan I. Croll**[4,5], **Irene R. Chassagnon**[1,10], **J. Shaun Lott**[6], **Evelyne Deplazes**[1], **Mikael Bodén**[1], **Mark RH Hurst**[7], **Sarah J. Piper**[1,2,8,9,12] ✉ & **Michael J. Landsberg**[1,2,9,12] ✉

ABC toxins are toxin-translocating, pore-forming proteins found in a wide range of insecticidal bacteria and some mammalian pathogens. The *Yersinia entomopahaga* toxin complex (YenTc) belongs to a distinct subclass of ABC toxins, defined by a divergent molecular architecture. Structural details that define their mechanism of action remain to be elucidated. Here we determine structures of the YenTc holotoxin assembly in both prepore and pore-forming configurations using cryo-EM in conjunction with Alphafold2-assisted structural modelling of flexible domains. We define the structural mechanism via which enzymatically-active chitinase subunits are incorporated, and show using phylogenetic analyses that this subclass-defining feature has evolved relatively recently. Our structures point to the existence of distinct conformational states in YenTc, which may distinguish it from other structurally-characterised ABC toxins, or represent states on a shared mechanistic trajectory. Thus, our findings enhance our understanding of the structural diversity that defines distinct ABC toxin subclasses.

Yersinia entomophaga is an entomopathogenic bacterium that causes disease in susceptible insect hosts via targeted effects on midgut epithelial cells[1–3]. The major virulence factor associated with its pathogenicity is YenTc, a 2.4 MDa macromolecular assembly belonging to the ABC family of bacterial pore-forming toxins (also referred to as Tc toxins)[1]. ABC toxins are most commonly found in the genomes of insecticidal bacteria that are, for the most part, vectored by a nematode host[4–6]. Some mammalian pathogens also contain gene clusters that are either predicted or proven to encode for ABC toxins[7,8]. *Y. entomophaga* is unusual in that it is an insect pathogen that is able to both replicate and exert its toxic effects completely independently of any known nematode symbiont[1,9].

---

[1]School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD, Australia. [2]ARC Centre for Cryo-electron Microscopy of Membrane Proteins, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC, Australia. [3]School of Life and Environmental Sciences, University of Sydney, Camperdown, NSW, Australia. [4]Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. [5]Altos Labs, Cambridge, UK. [6]School of Biological Sciences, University of Auckland, Auckland, New Zealand. [7]Bioeconomy Science Institute, AgResearch, Lincoln Research Centre, Christchurch, New Zealand. [8]Drug Discovery Biology, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC, Australia. [9]Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, Australia. [10]Present address: Servatus Biopharmaceuticals, Coolum Beach, QLD, Australia. [11]These authors contributed equally: Yu Shang Low, Solace G. Roche, Nadezhda A. Aleksandrova. [12]These authors jointly supervised this work: Sarah J. Piper, Michael J. Landsberg. ✉e-mail: sarah.piper@monash.edu.au; m.landsberg@uq.edu.au

YenTc and other ABC toxins share some functional similarities with the wider family of binary toxins, the latter being two-component toxins comprising an active component and a binding component[10]. The "active" component of YenTc (TcBC) is a heterodimeric complex of two proteins (YenB and YenC) which cooperatively fold to form a protective cocoon that harbours a cytotoxic enzyme[11]. The pathogenicity island encoding for YenTc includes two paralogs of YenC (YenC1, YenC2) which encode different predicted cytotoxic enzymes at their C-terminus; either of these can fold together with YenB to form the active component[1]. A third YenC paralog (YenC3) at a distant genomic location is also able to recombine with YenB to form a functional, active component[12].

The "binding" component of YenTc (TcA) accounts for almost 90% of the complex by molecular mass. YenTc is a representative example of a distinct subtype of ABC toxins[13–15] characterised by a TcA component that is comprised of four, separately encoded, unique polypeptide chains. Each polypeptide chain is present in five copies, yielding a 2.1 MDa complex[13,16]. Two of the four separately encoded polypeptides are functional chitinases[17], while the remaining two polypeptides (YenA1 and YenA2) together form a predominantly α-helical structure that includes a pre-formed α-helical translocation channel, surrounded in its soluble state by an outer α-helical shell. Activation of YenTc occurs via an as yet undefined physiological trigger and leads to a large conformational change in the TcA subunit whereby the membrane-penetrating tip of the translocation channel becomes exposed[13]. Membrane engagement in vivo is likely to involve interaction with host-specific cell surface receptors that position the complex so that the tip is driven through an adjacent membrane bilayer[13,18]. In vitro studies suggest that YenTc engages with glycan motifs on the cell surface via the chitinase subunits[13]. The cell surface glycoprotein Vgr1 has been identified as a receptor for ABC toxins produced by *Photorhabdus luminescens,* but is not a receptor for YenTc[19]. Differences in the structure and architecture of the TcA component of YenTc, specifically the receptor-binding regions, offer the most likely explanation for these observed differences in specificity.

We previously reported structures of the separate TcA[13] and TcBC components of YenTc[11], but until now a structure of the combined, functional, holotoxin assembly has not been described. Here we report single particle cryo-EM structures of a native YenTc holotoxin in both the soluble prepore state and a lipid nanodisc-embedded configuration. The latter in particular represents a substantial improvement over previously reported structures, and therefore provides structural insights into the molecular transitions that accompany pore-formation. We identified structural differences between YenTc and other ABC toxins that may be representative of subtype-specific differences in structure or may alternatively represent different conformational states on a shared molecular trajectory. We were able to define the structural mechanism via which chitinase subunits (a defining feature of YenTc that distinguishes it from other ABC toxins) are incorporated into the holotoxin assembly. Phylogenetic analyses revealed that in a small, recently evolved subset of ABC toxins, chitinase-binding domains (CBDs) are inserted into the TcA α-helical shell at or near evolutionary hotspots which coincide with the location of putative receptor-binding domains (RBDs) in other ABC toxin subtypes. Unexpectedly, the CBDs share structural homology with bacterial adhesins, and the mechanism via which chitinases are incorporated shares similarities with the assembly mechanism seen in filamentous pili and fimbriae. Finally, we propose a refined classification system for bacterial ABC toxins and show that chitinase incorporation, as seen in the structure of YenTc, is a defining feature of a small subset of ABC toxins that has evolved relatively recently, suggesting this feature confers functional advantages resulting in enhanced fitness.

Results

Complete structures of the YenTc holotoxin assembly in two different conformational states

Structures of native, functional YenTc holotoxin assemblies in the soluble, prepore configuration and reconstituted in lipid nanodiscs were determined by single particle analysis following imaging by cryo-EM. Notably, all of the datasets used in the course of this work were collected manually using a side-entry cryo-holder on a Tecnai F30 FEG-TEM (Field Emission Gun-Transmission Electron Microscope) fitted with a K2 Summit direct electron detector. Given that financial and geographical barriers often prohibit access to cutting-edge, automated microscopes our results serve as an important example highlighting the capabilities of previous generation instruments and thus contribute to the "democratisation" of cryo-EM[20,21].

For the prepore structure, single particle images extracted from 616 micrographs were initially processed without any imposed symmetry, producing a structural map with an average resolution of 4.6 Å (Prepore map 1; Supplementary Fig. S1; Fig. 1A, B). Subsequent focused refinement of the TcA region with C5 symmetry imposed (Prepore map 2; Supplementary Fig. S1), and of the map density encompassing the TcBC subcomplex and the interfacing residues of TcA with no symmetry imposed (Prepore map 3; Supplementary Fig. S1), produced maps with resolutions of 4.0 Å and 5.7 Å for these regions, respectively.

To obtain the membrane-embedded pore form of the complex, we subjected complexes to mild chemical destabilisation by dialysis against guanidine hydrochloride at sub-denaturing (1.5 M) concentrations, prior to reconstitution into POPC lipid nanodiscs. Following size-exclusion chromatography, reconstituted pore-forms were identified in cryo-EM images, and single particle images extracted from a dataset of 1222 micrographs were processed with no symmetry imposed to produce a 3D reconstruction with a resolution of 6.4 Å (Pore map 1; Supplementary Fig. S1; Fig. 1C, D). Focused refinement on the TcA region, with imposed C5 symmetry, yielded a map with an average resolution of 5.9 Å (Pore map 2; Supplementary Fig. S2), substantially improving on the structure we reported previously at a resolution of 11 Å[13] and allowing detailed interpretation of the molecular transitions associated with pore-formation.

Analysis of the YenTc prepore and holotoxin assembly

Like other Tc toxins, and as suggested by our prior analysis of the individual components, the holotoxin structure of YenTc (Fig. 2A) features a central, pre-formed α-helical pore-forming domain that is surrounded by a predominantly α-helical outer shell that almost entirely encapsulates the pore. A funnel-shaped opening at the top of the pore-forming domain forms the interface with the TcBC component (TcB-binding domain). The asymmetric TcBC subcomplex breaks the otherwise fivefold symmetric structure. Model building was initiated by fitting our previously determined crystal structure of YenB-YenC2$^{\Delta CTD}$ into the prepore holotoxin map. An initial model of the TcB-binding domain of TcA was extracted from our previously solved cryo-EM structure of the TcA component, and the two models were simultaneously refined against the map density.

Substantial remodelling of the TcA-TcBC interface was necessary, indicative of binding-induced structural changes in this part of the model. However, this was largely confined to the TcBC side of the interface (Fig. 2B–D and Supplementary Fig. S3C), with the TcA backbone remaining largely unchanged (Fig. 2E–G, RMSD = 3.33 Å across 750 residues). In the crystal structure of YenBC, the TcA-binding β-propeller domain adopts a distorted conformation (Fig. 2D), with two of the six blades closing the bottom of the cage. In the context of the holotoxin, the blades are repositioned, forming a less distorted, pseudo-symmetric structure (Fig. 2B), similar to what has been observed previously for the *P. luminescens* toxin complex 3 (PTC3)[22]. Thus, YenTc holotoxin formation involves an induced fit mechanism of

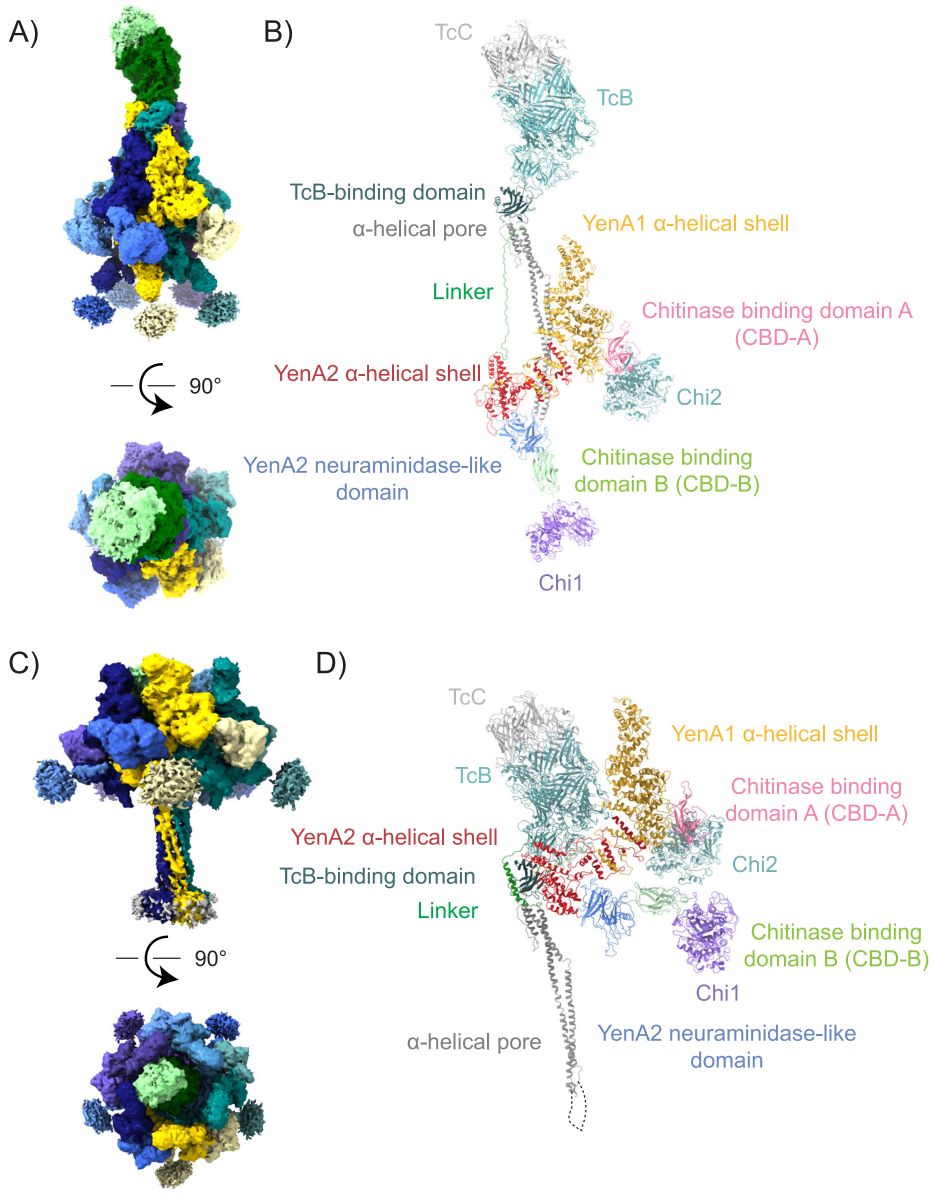

Fig. 1 | Cryo-EM maps and models of the YenTc holotoxin in both prepore and pore conformation. CryoEM map of the prepore (**A**) and pore (**C**) holotoxin structure showing the side (top) and top (bottom) views. Each protomer is shown in a different colour. **B**, **D** YenTc holotoxin protomer showing YenA1, YenA2, Chi2 and Chi1. The domains and subunits of the protomer are represented by different colours: Orange: YenA1 α-helical shell domain, red: YenA2 α-helical shell domain, dark green: TcB-binding domain, grey: α-helical pore domain, light green: linker, dark blue: neuraminidase-like domain, pink: chitinase-binding domain A (CBD-A), blue-green: Chi2 chitinase, pale green: chitinase-binding domain B (CBD-B), purple: Chi1 chitinase.

binding and recognition between the TcBC and TcA components, where the TcBC component undergoes structural adaptation to match the interface presented by TcA (Supplementary Movie S1).

Given the similarity to the PTC3 system, we conclude that the structural mechanism via which assembly of the two toxin components into a functional holotoxin occurs is likely replicated in other members of the ABC toxin family. The rearrangements lead to the formation of a single, continuous fenestration linking the interior of the TcBC cocoon with the lumen of the TcA pore-forming domain. Residual, non-continuous map density, which could not be accounted for in the subsequent model building steps, can be seen within the lumen of the TcBC cocoon and likely corresponds to the unfolded toxin domain of YenC (Supplementary Fig. S4). This density extends into the lumen of the TcA pore-forming domain (Supplementary Fig. S4), indicating that the interface between the TcA and TcBC components is indeed open.

Structural changes that define the prepore to pore transition

The transition of YenTc from the prepore to the pore conformation is characterised by substantial conformational changes that are almost exclusively confined to the TcA component. Whilst these changes are qualitatively similar to what was described in our previous manuscript that reported the structure of the TcA component reconstituted in liposomes[13], the resolution of this nanodisc-embedded holotoxin assembly is substantially improved, facilitating a more confident and complete positioning of the individual structural domains and allowing us to derive insights into aspects of the prepore-to-pore transition.

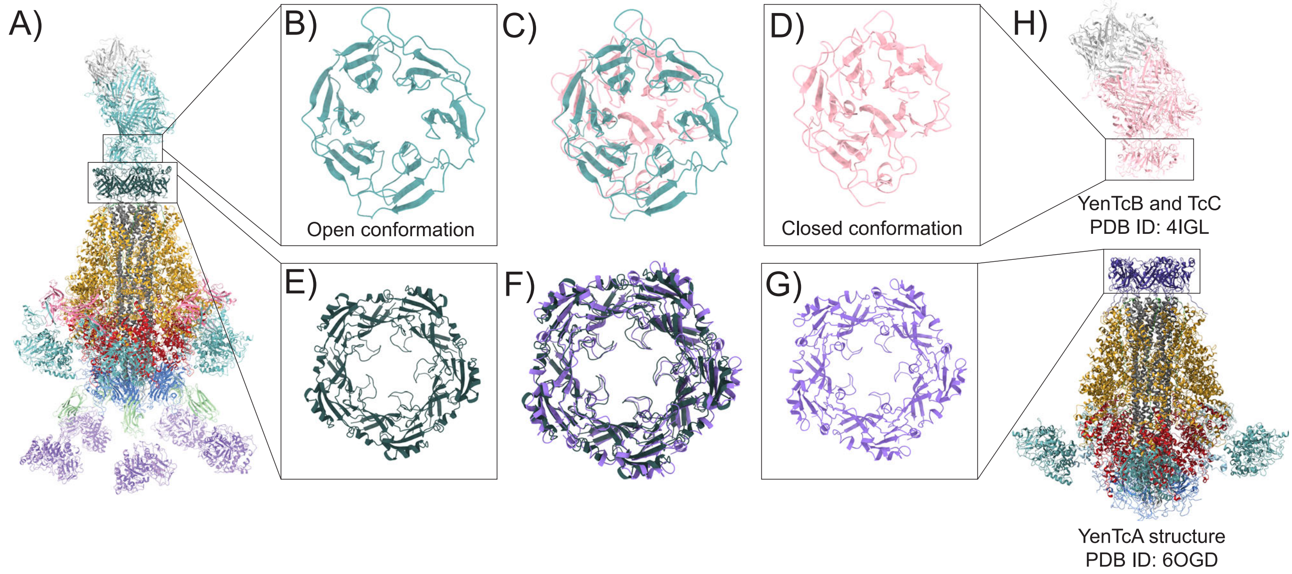

Fig. 2 | Conformational changes to YenTc subunits during complex formation. **A** Cryo-EM structure of the YenTc complex (this study). **B** β-propeller of TcB in the open conformation when in complex with TcA subunits. **C** Overlay of the β-propeller domain of TcB showing the conformational differences when TcBC forms a complex with TcA. **D** β-propeller region of TcB showing it in the closed conformation when isolated from the TcA subunit**.** **E** TcB-binding region in YenTc holotoxin structure. **F** Overlay of the TcB binding regions of the YenTcA (K9) structure and YenTc holotoxin structure showing little conformational changes as a result of TcB binding. **G** Structure of the TcB-binding region of the TcA subunits in the YenTcA (K9) structure. **H** Top panel: Crystal structure of *Y. entomophaga* TcB and TcC subunits (PDB ID: 4IGL). TcB and TcC are coloured pink and grey, respectively. Bottom panel: Structure of the YenTcA (K9) structure.

Overall, the α-helical shell of YenTc undergoes an iris-like expansion (Fig. 3) that results in the loss of all interchain contacts between YenA1 protomers, meaning the only points of contact remaining within the α-helical shell are between adjacent YenA2 protomers. Shell expansion accommodates a vertical displacement of the pore-forming domain and the attached YenBC cage. The pore-forming domain moves relative to the surrounding YenA shell by a distance of ~170 Å, resulting in the YenBC cage being almost completely enclosed within the expanded shell (Fig. 3A). The iris-like opening of the YenA shell is accompanied by a compound hinge-bending motion about two points in the structure, located on either side of the YenA1-YenA2 interface (the interface itself remains unchanged). The first hinge is located within the YenA2 region of the shell, and the bending motion results in an upward rotation of the lower part of the shell, the attached neuraminidase-like domain and CBD-B by ~15° (Fig. 3D). Movement about the first hinge point results in separation of the neuraminidase-like domains, making room for the emergence of the pore-forming domain from the base of the shell. The second hinge is within the YenA1 region of the shell, again close to the YenA1-YenA2 interface, and leads to a more subtle movement of the rest of the YenA1 shell and attached CBD-A. As a result, the opening of the shell is slightly more pronounced than what would occur if it rotated purely as a rigid body.

To our knowledge, the PTC3 toxin is the only other ABC toxin that has been structurally characterised in its pore-forming state. When comparing the YenTc pore with PTC3, the most striking difference is the degree to which the pore-forming domain is ejected out of the surrounding α-helical shell. Notably, only the TcB-binding domain and the attached TcBC cage remains inside the expanded shell of YenTc. This contrasts with PTC3, where the ejection of the pore-forming domain from the shell appears to be less extensive (Supplementary Fig. S5). In quantitative measures, the pore-forming domain moves relative to the surrounding shell by an additional 40 Å in YenTc compared to PTC3. The transition from prepore to pore in ABC toxins generally is accompanied by folding of an extended linker in the prepore state into a more compact α-helix in the pore form; this structural transition either drives or is a consequence of pore-formation. Interestingly, the folded linker domain in YenTc (Supplementary Fig. S5) is three turns longer than the equivalent helix in PTC3 (Supplementary Fig. S5C). Our observation of a more extensively folded linker helix in YenTc is likely a consequence of the more extensive pore ejection.

In the YenTc pore, the inner face of the linker helix packs against the TcB-binding domain, which represents an additional point of difference to the structural configuration seen in PTC3. To explore this further, we performed an in-depth analysis of the respective linker interfaces in PTC3 and YenTc (see Text S1) and concluded that poor sequence conservation in a structural element we termed the 'collar loop' may explain the differences in linker positioning. We also examined the corresponding sequences in other ABC toxins that have been structurally characterised in the prepore state[23] and found that this collar loop motif has a well-conserved sequence in all toxins we examined, with YenTc being the one exception (Supplementary Fig. S6 and Text S1). Notably, YenTc is the only example among this group that has a split TcA subunit architecture. Our analysis, therefore, suggests that the YenTc pore form could be representative of a distinct structural conformation that is potentially also adopted by other ABC toxins belonging to subtypes that have a split TcA subunit.

The structural basis of chitinase incorporation into YenTc

While our previous structural studies of YenTc identified the locations of the Chi1 and Chi2 domains within the TcA component of YenTc, poor local map resolution meant we had been unable to resolve the Chi2-YenA1 and Chi1-YenA2 interfaces[13]. Thus, the structural mechanism via which chitinases are incorporated into YenTc (and indeed, into any ABC toxin) has remained unclear. The improved maps reported here allow unequivocal identification of two previously unresolved CBDs inserted into the TcA shell (Fig. 4A). CBD-A (YenA1 residues 311–472) forms the major binding interface with the Chi2 subunit of TcA and could be modelled based on the map density alone. CBD-B (YenA2 residues 400–520) forms the major binding interface with the Chi1 subunit, but due to substantial conformational dynamics, required a hybrid modelling approach (detailed below).

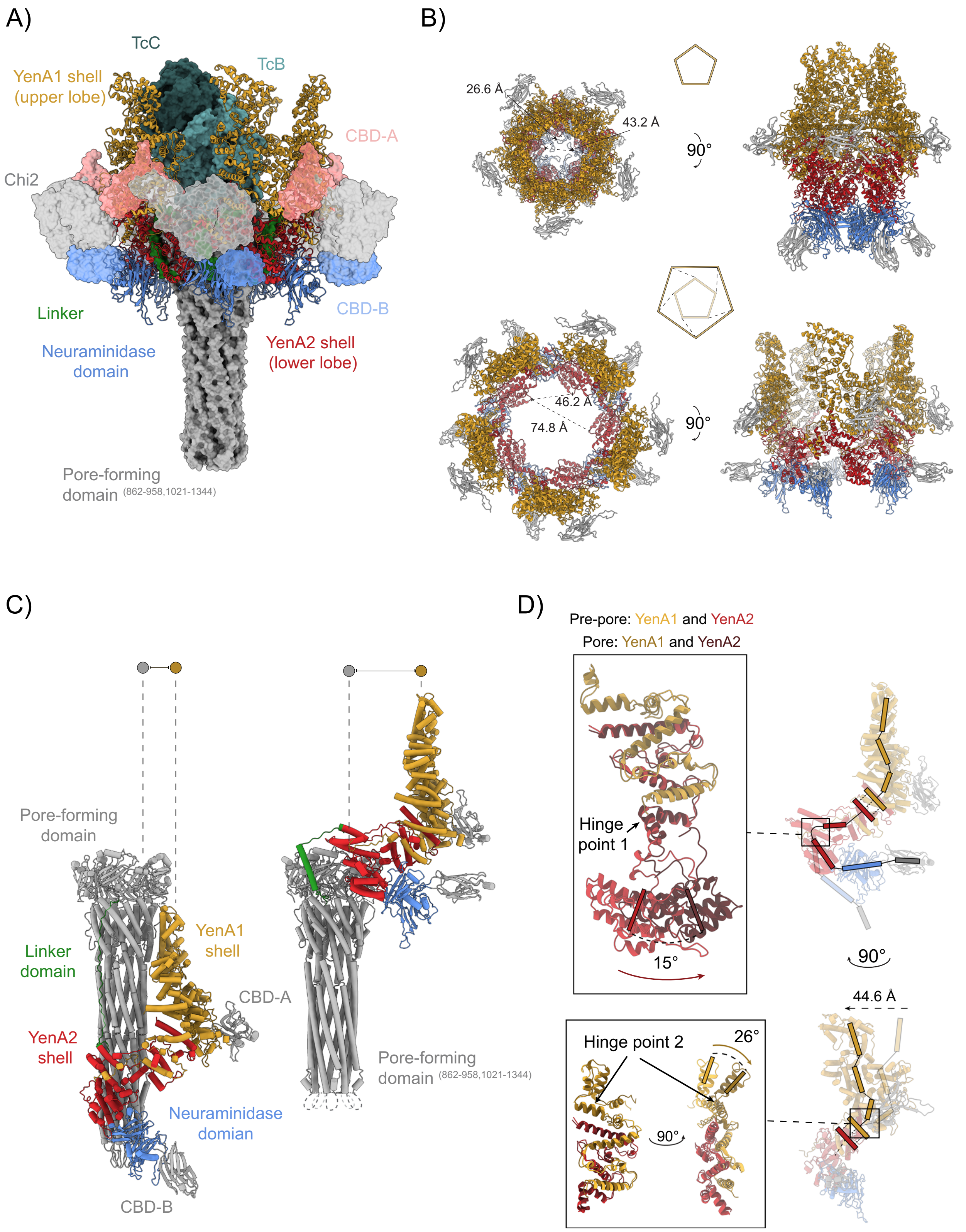

Fig. 3 | Shell expansion in the prepore to pore transition of YenTc. A In the pore-form, secondary structural features of the YenA1 (yellow) and YenA2 (red) shells are shown along with the neuraminidase-like domain (blue). Space-filling views of the linker domain (green), pore-forming domain (gray), TcB (cyan) and TcC (dark green) are shown. Peripheral domains (CBD-A, red; Chi2, gray; CBD-B, blue) are shown partially transparent. Chi1 has been hidden in this representation. **B** The structural transition from prepore (top) to pore (bottom) is marked by an iris-like expansion of the YenA1/A2 shell domain and neuraminidase-like domain. Distances showing shell expansion were measured between the Cα atoms of G142 between neighbouring and opposite YenA2 protomers. **C** The expansion of the shell accommodates the insertion of the pore-forming domain, and results in the diagonal translation of the shell interface from the bottom of the pore to the very top. **D** There are two major pivot points to accommodate the expansion of the shell. Relative to the YenA1/A2 interface, the YenA1 shell rotates by 26° resulting in a 44.6 Å displacement, and YenA2 rotates by 15° which displaces the neuraminidase-like domain and CBD-B.

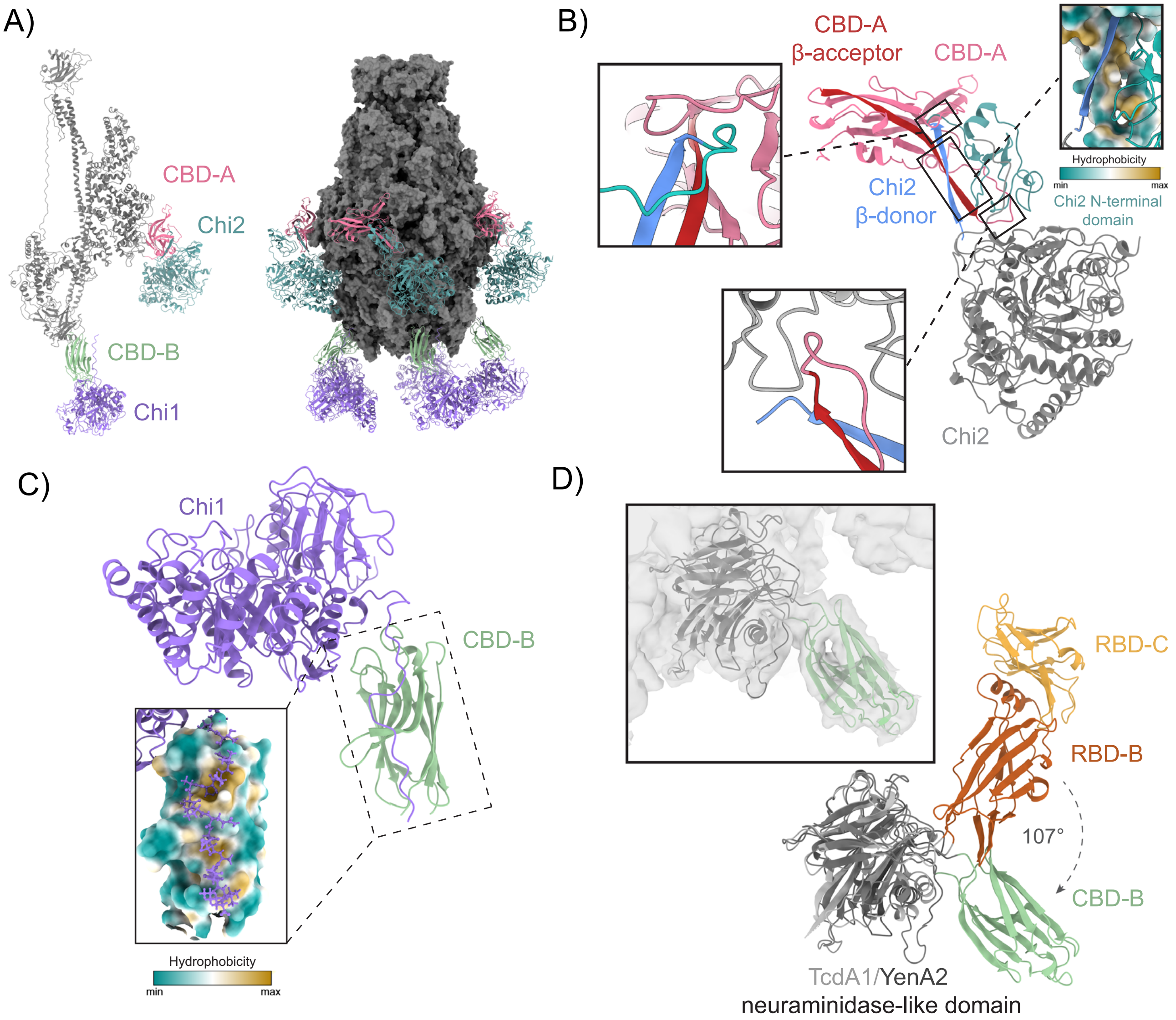

Fig. 4 | The structural basis of chitinase incorporation into the YenTc shell. **A** Architecture of the CBDs and chitinases from YenTc (right). **B** Structural overview of how CBD-A (red) interfaces with the Chi2^{NTD} (cyan) and Chi2 (grey). The interface is stabilised by interacting loops at each end of the complementary, antiparallel CBD-A β-acceptor (red) and Chi2 β-donor (blue) interface. The complementary antiparallel interface is stabilised further by a hydrophobic interaction, where the Chi2 β-donor occupies a hydrophobic cleft formed by the CBD-A β-acceptor. **C** The interface between Chi1 and CBD-B shows that the Chi2^{CTR} occupies a hydrophobic cleft formed by CBD-B, similar to the CBD-A interface. **D** Structural alignment of TcdA1 (light gray) and YenTc (dark gray) neuraminidase-like domains showing the conserved insertion point of the RBD-B/C (orange and red) daisy chain and CBD-B domain (green), respectively. When superimposed into the YenTc prepore map, the CBD-B domain is rotated 107° downward compared to the positioning of RBD-B/C in TcdA1.

Chi2 interacts with CBD-A through its N-terminal domain (Chi2^{NTD}; truncated in the Chi2 crystal structure (PDB: 4DWS) and absent from our previous model of the TcA subunit) which forms a relatively compact, mixed αβ fold with a single, long β-strand extending from its N-terminus. This β-strand appears to be an important component of the interface with CBD-A, interacting with a 10 residue stretch of YenA1 (residues 385–395) to form a two-stranded, antiparallel β-structure. The β-strand from Chi2 occupies a hydrophobic cleft formed by the complementary strand and surrounding structural elements of CBD-A, and the interface is likely further supported by several electrostatic interactions at both ends of the β-sheet (Fig. 4B and Supplementary Fig. S3E).

CBD-A adopts a classical β-sandwich fold and a DALI search[24] for structurally related proteins identified the adhesin domain of the *Escherichia coli* fimbriae protein CfaE (RMSD = 2.9 Å) as its nearest structural homologue (Supplementary Fig. 7B). Curiously, CfaE also forms heterodimeric interfaces involving β-donor and β-acceptor interactions that are remarkably similar to the interaction seen between CBD-A and Chi2^{NTD}. In the context of CfaE, the interaction seeds the formation of a filament of polymeric pilin proteins linked via a daisy chain of β-donor/β-acceptor interactions[25]. The nature of the β-strand interface between Chi2 and CBD-A differs slightly from that seen in CfaE, however, in that CBD-A has adapted a region that acts as a stabilising loop in the CfaE structure to act as the CBD-A β-donor (Supplementary Fig. 7C). This motif in turn recruits the N-terminal β-acceptor from Chi2^{NTD}, and an extension of the Chi2^{NTD} beyond the β-acceptor appears to act as a stabilising loop (Supplementary Fig. S7C). Thus, the mechanism of polymeric pilin assembly appears to have been co-opted here to facilitate recruitment of Chi2 into a stable complex with YenTc.

A 3D variability analysis using our cryo-EM data confirmed that substantial conformational dynamics (Supplementary Movie 2) are the most likely contributing factor to the poor local map resolution in the region of the Chi1:CBD-B interface, prohibiting high-resolution structure determination. To model this interface, we therefore employed a hybrid model-building approach using AlphaFold2 multimer (AF2m), cross-linking mass spectrometry (XL-MS), MD simulations and structural data derived from our previously determined crystal structure of the Chi1 subunit[17]. The AF2m-predicted structure of CBD-B (Supplementary Fig. S8A) showed a reasonable fit to the map density in the CBD-B region (Fig. 4D), and the predicted Chi1 structure was unsurprisingly consistent with the crystal structure. The absolute positioning of the attached Chi1 subunit was somewhat inconsistent with the residual map density attributed to Chi1 in the full complex, which we again expect is due to the extreme flexibility evident in this region. To validate the AF2m predicted interface, we performed XL-MS analysis on the holotoxin assembly. Three crosslinks between Chi1 and CBD-B ($\mathrm{Chi1^{E399}}$-$\mathrm{YenA2^{K447}}$, $\mathrm{Chi1^{K448}}$-$\mathrm{YenA2^{K449}}$, $\mathrm{Chi1^{E527}}$-$\mathrm{YenA2^{K449}}$) were identified (Supplementary Fig. S8B, C), two of which were satisfied in the AF2m structural model (Fig. S8B). To account for flexibility, as a final validation step, we performed triplicate, 1 μs MD simulations focused on this region of the holotoxin structure, and found that in addition to the predicted interface remaining stable throughout all simulations, all three contacts identified in our XL-MS analysis were satisfied throughout the simulations (Supplementary Fig. S8C). Thus, we concluded that our hybrid model provides a robust representation of the Chi1:CBD-B interface.

Chi1 lacks the 90 amino acid N-terminal domain found in Chi2 and is instead anchored to CBD-B via a single, unstructured strand located at its C-terminus ($\mathrm{Chi1^{CTS}}$) that is not resolved in the crystal structure[17]. $\mathrm{Chi1^{CTS}}$ integrates with an exposed hydrophobic cleft formed by CBD-B (Fig. 4C), and while the involvement of a hydrophobic cleft is somewhat similar to the interaction seen between CBD-A and $\mathrm{Chi2^{NTD}}$, the $\mathrm{Chi1^{CTS}}$ is predicted to adopt an extended strand topology in binding along the cleft, rather than the β-strand topology adopted by $\mathrm{Chi2^{NTD}}$. Notably, the *Streptococcal* factor H-binding protein (Fhb) and adhesin (SadP) are identified by DALI as structural homologues of CBD-B (Supplementary Fig. S9). Thus, CBD-A and CBD-B both appear to have features in common with bacterial adhesins or adhesin-like domains, leading us to speculate that this may represent a preferred structural fold for the anchoring of chitinases within the α-helical shell of ABC toxins.

Comparison of the receptor binding domain region of YenTc to other ABC toxins

While CBD-B shares structural homology with the adhesin proteins noted above, the closest structural homologues of CBD-B detected by DALI are putative receptor binding domains of other TcA proteins, in particular the RBD-A, RBD-B, and RBD-C domains of the TcA component from the *P. luminescens* PTC3 toxin, which are conserved to varying extents in other ABC toxins[23,26]. CBD-B shares a strikingly similar structural fold with all three RBDs identified in the PTC3 structure (RMSDs range from 2.6 to 3.4 Å), despite having low sequence identity (9–20%) (Supplementary Fig. S9). This low degree of sequence conservation provides a plausible explanation for why CBD-B (but apparently not the RBDs of other TcA proteins) facilitates incorporation of chitinase subunits into a holotoxin assembly.

Aside from the extent to which their sequences have diverged, a further difference exists with respect to the structural positioning of CBD-B relative to the RBDs found in other TcA proteins. When the prepore structures of YenTc and PTC3 are overlaid, CBD-B is rotated downwards through an angle of ~107° relative to the position occupied by RBD-B in PTC3 (Fig. 4D). Interestingly, independent experiments conducted on PTC3 during the course of our studies[27] provided evidence for the existence of structural intermediates between the prepore and pore conformations, one of which repositions the RBD-B/C region into a downward-facing orientation that more closely mirrors the positioning of CBD-B and the bound Chi1 subunit in YenTc (Supplementary Fig. S10). This led us to explore whether it was possible for the reverse motion to occur in YenTc. Repositioning of CBD-B, such that it mirrors the upward conformation seen in PTC3, was not possible without additional and substantial structural rearrangement of the rest of the complex. Moving CBD-B as a rigid body introduces steric clashes between the attached Chi1 subunit and the Chi2 subunit associated with the neighbouring protomer. Thus, binding of the Chi1 and Chi2 subunits appears to sterically preclude CBD-B from adopting a similar prepore configuration to the RBD domains seen in other ABC toxins that do not incorporate chitinases.

We also noted that the CBD-B domain of YenA2 not only shares structural homology with the putative RBDs of other TcA proteins, but that CBD-B is inserted into the same structural loop of the conserved neuraminidase-like domain. Thus, it appears likely that CBD-B may have been adapted from an RBD (although the converse may also be true, i.e. that other RBDs are adapted from CBD-B). Alternatively, this specific loop in the neuraminidase-like domain may represent an evolutionary hot spot, within which (at least) two different domains implicated in receptor recognition have been inserted via separate evolutionary events.

We investigated whether a similar, evolutionary hot spot exists within the α-helical shell of TcA, where CBD-A is inserted into YenTc and RBD-A is inserted into PTC3. Surprisingly, however, we found that RBD-A and CBD-A are not inserted into the same location, with CBD-A inserted ~200 residues closer to the N-terminus of YenA1 relative to the corresponding insertion site of RBD-A (Supplementary Fig. S7A). Attempts to superimpose the two domains also revealed little structural similarity between CBD-A and RBD-A, further supporting the conclusion that incorporation of these domains likely occurred via independent evolutionary events. Additionally, YenTc appears to have retained a small domain with no regular secondary structure in place of the canonical RBD-A fold. This 'pseudo RBD-A' domain is not found in the shell of other, structurally characterised ABC toxins that lack an RBD-A (for example, the *P. luminescens* TcdA4 protein). This suggests that, from an evolutionary perspective, toxins containing the RBD-A domain likely emerged prior to those containing the CBD-A domain (typified by YenTc), and that the RBD-A domain became the pseudo RBD-A domain that remains in YenTc, either soon before or subsequent to CBD-A being acquired.

CBDs are exclusively found in a clade of ABC toxins that directly associate with chitinases

To provide an enhanced context for understanding the evolutionary events that led to insertion of the CBD-A and CBD-B domains, facilitating the direct structural association of chitinases with ABC toxins, we constructed phylogenetic trees based on a curated database of ABC toxin loci assembled from published genomes. A detailed description of how this database was assembled is provided (see “Methods” and Text S2). The dataset was initially curated so that only complete ABC toxin loci were included (a toxin locus was considered complete if it contained a TcA component predicted to include a TcB-binding domain, and at least one copy of a TcB and a TcC component). To facilitate unambiguous association between components, a toxin locus was only included in the analysis if there were no other identifiable toxin components within the genome from which it was derived, or if unequivocal association between components could be inferred from other forms of evidence (e.g. from structural studies, as is the case with several *Photorhabdus* and *Xenorhabdus* toxins).

When the filtered database of toxin loci was subject to phylogenetic analyses, four clearly separated and well-defined clusters emerged when trees were constructed based on the alignment of TcA subunits (Fig. 5). Two of these clusters represent ABC toxins that have a split TcA gene. In a previous study, we defined these split TcA toxins as Type 2 toxins[13,14] and it was evident in our analysis here that these

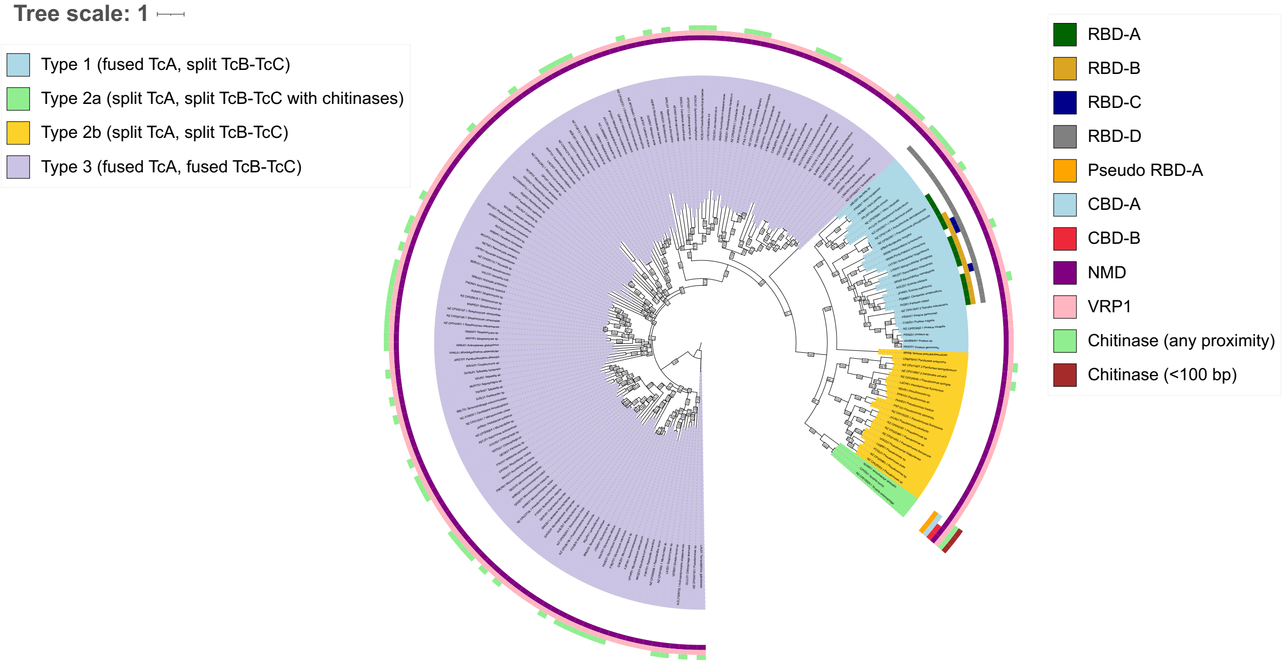

Fig. 5 | Phylogenetic tree of TcA components belonging to genomes with unambiguously associated pathogenicity islands. Colours on the tree represent the different subtypes as defined by genome architecture types. Dark green, gold, dark blue, grey boxes indicate presence of receptor binding domains in Type 1 components only. Orange, red, and light blue boxes indicate presence of pseudo RBD-A, CBD-A, and CBD-B in Type 2A components only. The presence of a neuraminidase-like domain (NMD - purple) was a requirement for sequences to be retained in the analysis and all sequences also contain an additional domain, annotated in Pfam as a *Salmonella* virulence plasmid (VRP1) domain (pink), found in ABC toxins and in plasmid encoded virulence factors from *Salmonella typhimurium*. Chitinases (light green) are found across many genomes, but are only found in close proximity to a pathogenicity island (<100 bp - brown) in genomes containing a Type 2A component. The closest chitinases to non-Type 2A components are over 85 kb. Internal labels show bootstrap values generated from 1000 ultrafast bootstraps. Figure generated with Interactive Tree of Life v5[56].

separate further into two distinct phylogenetic clades. One clade includes toxins like YenTc, that incorporate chitinases into their TcA component while the other has a split TcA gene but no predicted chitinase genes within 100 bp of the toxin locus. For the sake of consistency with our previously defined nomenclature, we refer to these two clades hereafter as Type 2A (split TcA with chitinases) and Type 2B (split TcA without chitinases).

We searched the entire sequence database for evidence of CBD conservation within the various TcA sequences. Using hidden Markov model sequence profiles, we were unable to find sequences matching the profiles constructed for CBD-A and CBD-B in any other TcA proteins, outside the limited number of Type 2A toxins present in our dataset. This was somewhat unsurprising given that Type 2A toxins were also the only toxins for which chitinases appear to be co-localised with other toxin components, and supports their role as bona fide CBDs. We subsequently searched the entire UniProtKB database for CBD-A and CBD-B domains and detected these in 23 genomes, predominantly from the *Xenorhabdus* sp., but also *Photorhabdus asymbiotica*, *Photorhabdus laumondii*, and *Yersinia nurmii* (Supplementary Fig. S11). The additional sequences detected in this way are all derived from genomes that include multiple TcA components and were therefore not included in the phylogenetic analysis. However, only TcA components that included hits for CBD-A or CBD-B were flanked by putative chitinases. Thus, we conclude that either a CBD-A or a CBD-B domain is a requirement for chitinase incorporation.

Evolutionary relationships between clades

The phylogenetic analysis undertaken here additionally revealed that at least two separate clades of single TcA toxins exist. The smaller, more recently evolved clade includes the *P. luminescens* PTC3 as well as toxins incorporating the *P. luminescens* TcdA4 subunit, *Xenorhabdus nematophila* XptA1 subunit, and *Morganella morganii* TcdA4 subunit[23]. The larger clade appears to be the most phylogenetically ancient clade, and to date no toxins belonging to this clade have been structurally characterised.

Using the same hidden Markov model sequence profile analysis invoked to search for CBD-A and CBD-B, we built profiles representing the RBDs previously identified within solved structures[23] or defined within InterPro (see Supplementary Text 2) and found that RBDs were exclusively conserved (albeit to varying extents) within TcA sequences belonging to the more recently evolved of the two, single TcA clades identified above. RBD-D was the most strongly conserved, found in 77.8% of the sequences belonging to this clade, with RBD-A, -B and -C less well conserved. In a previous analysis conducted by Song et al. it was independently found that the prevalence of RBD-D domains in Type 1 ABC toxins was 83.4%[15]. We therefore conclude that Type 1 ABC toxins are characterised by the presence of a single TcA component, with a high prevalence of RBD-D domains and variable rates of incorporation of PTC3-like (and potentially other) RBDs.

Importantly, we could not find any evidence supporting the conservation of the currently defined RBDs or CBDs in the more ancient clade of single TcA toxins. Sequences from this clade are vastly under-represented in the independent analysis conducted by Song et al[15]. Thus, we conclude that this subset of TcA sequences is indicative of a previously undefined clade of ABC toxins, which we defined as Type 3 ABC toxins. In addition to the absence of any CBDs or RBDs, we also noted that members of this clade were characterised by a fused TcBC protein, rather than having separate TcB and TcC components as is typical of the other clades. We constructed additional phylogenetic trees based on the TcB and TcC components (for the purpose of this analysis, TcBC fusions were split into two sequences and classified separately) and found that broadly, the clustering of the TcB and TcC sequences mirrored that seen for the TcA sequences (Supplementary Fig. S12).

Our phylogenetic analyses therefore suggest that – as typified by the Type 3 architecture—ABC toxins were originally two component toxins comprised of TcA and TcBC components. Given the high bootstrap values on the branches between Type 1, 2 and 3 toxins, and on the branch between Type 2A and 2B toxins (Fig. 5), we can be confident that Type 1 and 2 toxins emerged subsequently, following the fission of TcBC into two separate components, and that a second fission event in the TcA component produced Type 2A and Type 2B toxins. Our combined structural and phylogenetic analyses further suggest a plausible evolutionary scenario giving rise to the distribution of RBDs and CBDs and delineation of the more recently evolved Type 1 and Type 2 clades. Because the *P. luminescens* PTC3 (Type 1) RBD-B and YenTc (Type 2A) CBD-B domains are similar in their structure and insertion point (Fig. 4D), we hypothesise they are orthologous domains derived from a common RBD-like β-sandwich domain acquired by the ancestors of Type 1 and 2 toxins after splitting from Type 3 toxins. Following another split within the Type 2 toxins, the 2A clade adapted these domains for recruitment of chitinases (which we infer as a recent evolutionary event from the phylogenetic tree), while the 2B clade appears to have lost these domains. Type 2A toxins likely acquired CBD-A in a separate evolutionary event given it has no structural similarity—nor does it share an insertion point—with any RBDs of known Type 1 toxin structures (Supplementary Fig. S7A). The identification of a pseudo RBD-A domain in YenTc and other Type 2A toxins (Fig. 5) indicates that this domain likely descended from RBD-A of Type 1 toxins and was lost via genetic drift, and further supports the conclusion that Type 2 toxins evolved most recently.

Discussion

In summary, this study presents what are to our knowledge the first complete structures of YenTc in prepore and pore configurations, providing the first example of an ABC holotoxin structure featuring a split TcA component that incorporates chitinases, and only the second structurally characterised example of any ABC holotoxin assembly. Some of the earliest functional studies of ABC toxin components established that it was possible to recombine genes encoding for different toxin components into functional holotoxins[23,28–30]. Our observation that holotoxin formation involves an induced-fit mechanism of interface formation that appears to be conserved across ABC toxin subtypes provides a structural explanation for these early observations.

The Chi1 and Chi2 subunits are unusual components of YenTc, in that they bear no sequence or structural similarity to the prototypical TcA, TcB or TcC proteins that are a defining feature of ABC toxin components. Our studies revealed that chitinase incorporation requires the presence of CBDs in the TcA subunit, inserted in place of (or potentially adapted from) the classical RBDs characterised previously in *P. luminescens*, *X. nematophila* and *M. morganii* toxins[23,26,31]. We found that YenTc and other chitinase-incorporating ABC toxins belong to a distinct phylogenetic clade that has evolved most recently, suggesting that these evolutionary changes may be associated with enhanced fitness. Previous studies demonstrating enhanced oral toxicity of YenTc at lower doses than for ABC toxins belonging to other clades[1] align with this conclusion.

We observed important differences in both the prepore and pore structures when comparing the overall conformation of the YenTc holotoxin to the PTC3 holotoxin from *P. luminescens*. The unique, downward-facing conformation of CBD-B seen in the prepore structure of YenTc resembles a recently described, putative structural intermediate of PTC3[27] and given that CBD-B and RBD-B are inserted into the exact same evolutionary hotspot within the conserved neuraminidase domain of their respective TcA subunits, it invites speculation that the YenTc prepore reflects a pre-primed, intermediate state. Given that chitinase binding sterically precludes YenTc from adopting a prepore ground state similar to PTC3 (Supplementary Fig. S10), we can conclude that in YenTc at least, this conformation exists as a direct consequence of chitinase binding.

The pore-form structure of YenTc also differs from the conformation described previously for PTC3[31,32], primarily with respect to the ejection of the pore-forming domain from the surrounding TcA shell. Previously, we provided evidence that the Chi1 and Chi2 subunits of YenTc may be implicated in receptor recognition[13], and when these are bound to CBD-B and CBD-A, respectively, the receptor-binding regions of YenTc extend further from the TcA shell than the corresponding, putative RBDs of PTC3. Thus, if bound to equivalently-sized receptors, the TcA shell of YenTc would be positioned further away from the target cell membrane, which may necessitate or require the greater displacement of the pore-forming domain relative to the TcA shell that is observed in YenTc. Based on conservation of the collar loop sequence, which appears to control the positioning of the pore-forming domain, we predict that other Type 1 ABC toxins are likely to adopt similar pore structures to PTC3 and that the YenTc conformation represents a distinct structural conformation that may be representative of Type 2A toxins. However, the recent discovery of intermediate structural states in PTC3[27] means we cannot rule out the possibility that the YenTc pore structure reported here reflects a distinct structural state on a molecular trajectory shared by ABC toxins from both clades. Characterisation of more ABC toxins from different subtypes and in different conformational states may help to address these questions.

Finally, our phylogenetic analyses suggest that ABC toxins can be classified into three major clades, one of which has not been defined previously and is characterised by the presence of a single TcA component and a fused TcBC component. Given that members of this clade do not appear to contain RBDs or CBDs, structure-function studies of representative members from this clade will be essential to understand what further diversity exists with regard to potential mechanisms of cell surface recognition.

Methods

Bacterial strains used

The *Y. entomophaga* toxin complex (YenTc) was purified from wild-type *Y. entomophaga* (strain MH96). *E. coli* BL21 DE3 was used for expression of membrane scaffold proteins (MSP1D1) for preparing lipid nanodiscs.

Expression and purification of YenTc

Full-length YenTc was expressed and purified as described previously[16,33]. In brief, *Y. entomophaga* was cultured in antibiotic-free LB medium at 25 °C overnight, shaking at 180 rpm. Cultures were incubated at 25 °C for 1 h without shaking to allow for YenTc secretion, then harvested by centrifugation at 4800 × *g* to pellet the cells, and the media supernatant collected. Ammonium sulphate was added to the supernatants to a final concentration of 70% (w/v) and the precipitated protein collected by centrifugation at 13,000 × *g*, following which the remaining supernatant was discarded. Precipitated protein was resuspended in TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.6), filtered, and purified further by Superose 6 (Cytiva) size-exclusion chromatography. Fractions corresponding to YenTc were pooled and concentrated for further experiments.

MSP1D1 nanodisc scaffolding protein expression and purification

A plasmid vector containing the scaffolding protein MSP1D1 gene (provided by Dr Carus Lau) was transformed into *E. coli* BL21DE3 for expression. Expressed protein was purified by gravity flow immobilised metal affinity chromatography as described[34] and dialysed into TBS buffer. Nanodiscs (NDs) were prepared as described previously[34].

Formation and reconstitution of pore-form YenTc in nanodiscs

To induce pore formation, YenTc was first subjected to dialysis against buffer containing guanidine hydrochloride (GnHCl) and Tween-20 (50 mM Tris, 150 mM NaCl, 1.5 M GnHCl, 0.05% Tween-20, pH 7.6). A second dialysis step was then performed using buffer without GnHCl. This pre-solubilised, pore-form YenTc sample was then mixed with POPC lipids solubilised in buffer containing sodium cholate (50 mM Tris, 150 mM NaCl, 23.2 mM sodium cholate, pH 7.6) and incubated for 30–45 min at 4 °C. Subsequently, MSP1D1 was added, and the sample was incubated for a further 1 h at 4 °C, achieving a final cholate concentration of 16–18 mM and molar ratios of 1:16:440 for YenTc:MSP1D1:POPC. Detergent removal was achieved in one of two ways—either through overnight incubation with Biobeads-SM2 (Biorad), or through dialysis against detergent-free TBS buffer—following which the sample was subjected to Superose 6 (Cytiva) size-exclusion chromatography.

Cryo-EM grid preparation and data collection

R2/2 holey carbon grids (Quantifoil) or lacey carbon grids with graphene oxide support film (ProSciTech) were washed using ethylacetate prior to glow discharge. Vitrification of YenTc prepore and pore-form samples was performed using a Vitrobot Mark IV (ThermoFisher Scientific) operating at 4 C and 100% humidity. 4 µL of each sample (0.2–1.1 mg/mL in TBS) was applied onto a grid and allowed to absorb for 3–10 s. The grid was then blotted for 3–4.5 s before plunge freezing in liquid ethane in preparation for cryo-EM. Data collection for both YenTc prepore and pore-form samples was performed on a Tecnai F30 FEG-TEM (ThermoFisher Scientific) operating at 300 kV and fitted with a K2 Summit direct electron detector (Gatan). A nominal magnification of 23,000×, corresponding to a pixel size of 1.72 Å at the specimen level, was used for data collection. Datasets were collected using SerialEM software with dose fractionation in electron counting mode, with 40 subframes collected per image and a total dose of 45–49.2 e/Å^2 (dose rate 8–9 e/pix/s). The defocus range used was from 1.8 to 3 µm.

Single-particle processing of YenTc prepore and pore images

For the YenTc prepore, a total of 616 movies from three data collection sessions were merged and processed using cryoSPARC (version 3.2.1)[35]. Patch-motion correction and patch CTF estimation were performed, and 137,309 particles were picked. Multiple rounds of 2D classification filtered the dataset to 72,265 particles, and these particles were used to generate an *ab intio* model. The best ab *initio* model (60,315 particles) was subjected to homogenous refinement using C1 and C5 symmetries, resulting in maps with average, estimated resolutions (FSC = 0.143) at 4.6 Å and 4.0 Å, respectively. Focused refinement of the BC domain was performed, yielding a map at 5.7 Å. The final data collection and modelling parameters are shown in Table S1 and the processing strategy is summarised in Supplementary Fig. S1.

For the YenTc pore, 1222 movies collected across multiple sessions were patch-motion corrected with dose-weighted averaging using cryoSP ARC (version 3.2.1)[35]. CTF parameters were estimated with patch CTF, and 116,252 particles were picked using the template picker. Following particle curation and 2D classification, 97,198 particles remained. To overcome issues with preferred particle orientation, the particles were re-distributed in six new superclasses using 2D class rebalancing, which reduced the particle dataset to a total of 40,973 particles. These were used to generate two ab initio models, one of which exhibited clear density for the pore domain of YenA. This ab initio model was used as input for homogeneous, non-uniform and heterogeneous refinements. The final models were refined using C5 and C1 symmetry, producing final maps with average, estimated resolutions (FSC = 0.143) of 5.9 Å and 6.4 Å, respectively. The final data collection and modelling parameters are shown in Table S1, and the processing strategy is summarised in Supplementary Fig. S2.

Model building and refinement of YenTc prepore and pore structures

Model building into the YenTc prepore and pore maps was initiated using previously published models for the TcA (6OGD) and TcBC (4IGL) subunits. Newly-resolved structural domains that were missing from these models were built using COOT 0.8.9.3[36]. AlphaFold2 (AF2) multimer[37] was used to generate an initial structure of the CBD-B-Chi1 interface, using residues 400–520 of YenA2 and the complete sequence of Chi1 as inputs. The source coordinates are summarised graphically in Supplementary Fig. S13. The entire structure was then subject to iterative refinement cycles using ISOLDE 1.6.0[38] and Phenix v1.21.5207[39]. For the YenTc pore form, within most domains there were few structural changes between the prepore and pore form, and these were able to be built initially by splitting the prepore structure into domains and transferring the corresponding coordinates one domain at a time from the prepore model into the pore map density. For the remaining regions, the Cα backbone was traced in COOT, side chains added, and then the entire assembly was refined using ISOLDE and Phenix. Side chains were used to assist model refinement, but were stubbed in the final deposited model as the resolution of this map was considered insufficient to reliably infer side chain conformations.

Protein cross-linking, digestion mass spectrometry and peptide identification

YenTc used for crosslinking experiments was prepared in essentially the same way as for structural studies, with the exception that it was necessary to purify in HEPES-buffered saline. For each crosslinking experiment, ~50 µg of the complex was used, and experiments were performed essentially as described elsewhere[40]. Azido-benzoic-acid-succinimide (ABAS) crosslinking experiments were performed in the same manner as for the sulfosuccinimidyl 4,4′-azipentanoate (sulfo-SDA) crosslinker, with the exception that 254 nm UV was used for ABAS.

Liquid chromatography and mass analysis parameters were essentially as described elsewhere[41] with the following exceptions: (1) Mass analyses were performed using an Eclipse Tribrid mass spectrometer (ThermoFisher Scientific); (2) For sulfo-SDA, ABAS and DMTMM, following each full-scan MS1 at 60,000 resolution at 200 *m/z* (300–1800 *m/z*, AGC = 4×10^5, 50 ms max injection time), precursor ions were selected for sequential HCD-EThcD-MS2 acquisitions in a data-dependent manner (HCD-MS2: $R = 15{,}000$, 30 NCE; EThcD-MS2: $R = 15{,}000$, calibrated charged dependent ETD parameters enabled, supplemental HCD at 30 NCE; both: 22 ms injection time, 1.4 *m/z* isolation window, 5×10^4 intensity threshold, charge states +4–10, dynamic exclusion of 20 s).

DSSO and DHSO cross-linked peptides were identified using the XlinkX 2.0[42] nodes as implemented in Proteome Discover v2.3 (ThermoFisher Scientific). Sulfo-SDA, ABAS and DMTMM datasets were processed with MS Studio (v2.4.0.3545)[43]. Generally, the following key parameters were used: peptide mass between 300 and 10,000 Da, minimum peptide length of 5–6 residues, precursor mass tolerance ±10 ppm, product-ion mass tolerance of ±20 ppm for Orbitrap data and ±0.5 Da for ion trap data. Allowable variable modification = oxidation (M), allowable static modification = carbamidomethyl (C), enzyme specificity of Trypsin with up to two (DSSO, DHSO) or four (Sulfo-SDA, ABAS, DMTMM) missed cleavages (excluding the site of cross-linking), and FDR control at 1%. The search database used consisted of YenTc proteins. The S20D4 default search setting was initially used for DSSO and DHSO datasets but filtered to S80D20 on manual spectral verification. Cross-linked spectral matches from MS Studio were also manually visually assessed. All mass spectrometry data and XL-MS search results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[44] with dataset identifier PXD068275.

Molecular dynamics structure preparation and simulations

Simulations to validate both the AF2m-predicted interface and the identified Cα-Cα distances from crosslinking mass spectrometry experiments between CBD-B and Chi1 were performed using the GROMACS 2022[45] simulation package in conjunction with the united-atom GROMOS54a7 forcefield[46]. To prepare the system for simulations, the CBD-B:Chi1 structure was placed in a triclinic box with dimensions 13 nm × 9 nm × 9 nm and solvated with 33,385 water molecules. A single point charge[47] model for water was used. Na^+ and Cl^- ions were added to neutralise the system charge and reach a final ionic strength of 150 mM. The final system, containing 107,112 atoms, was energy minimised using a steepest descent algorithm with a tolerance of 250 $kJ.mol^{-1}$.

The system was equilibrated for 5 ns in an NPT ensemble using a 2 fs timestep and position restraints with a force constant of 1000 $kJ.mol^{-1}.nm^{-2}$ applied to all atoms. Subsequently, a series of 5 ns NPT simulations were performed with decreasing positions restraints of 500 $kJ.mol^{-1}.nm^{-2}$, 250 $kJ.mol^{-1}.nm^{-2}$, 100 $kJ.mol^{-1}.nm^{-2}$ and 20 $kJ.mol^{-1}.nm^{-2}$.

Production runs were carried out in triplicate for 1 µs using a 4 fs timestep. Non-bonded interactions had a cutoff of 1.4 nm with the neighbour list updated every 5 steps and were modelled using the Verlet scheme. Electrostatic interactions beyond the cutoff were modelled using the reaction-field method[48]. All system components were temperature coupled using a Berendsen thermostat[49] with a temperature of 300 K and a time constant of 0.1 ps. All components used a Berendsen barostat with isotropic pressure coupling, with a compressibility of 4.5×10^{-5} bar^{-1} and a time constant of 1 ps. The periodic boundary conditions were applied in the x, y, and z dimensions.

Analysis of molecular dynamics data

Cluster analysis was performed on backbone atoms using the linkage method with a 0.15 nm cutoff in GROMACS 2022 on both the individual simulated replicates and a concatenated trajectory containing all three replicates, to ensure the conformations of the protein were independent from the starting structure (Supplementary Fig. S14).

From each simulation, the complete trajectory (1250 frames) was used to track the distances between residues identified in XL-MS experiments between CBD-B and Chi1. MDAnalysis[50,51] was used to calculate distances between crosslinked atoms over the course of the simulation. Distance over time plots were used to ensure distances had converged for the duration of the production simulations (Supplementary Fig. S14). For all residue pairs identified in the crosslinks, the Cα-Cα and reaction centre distances between the residues were tracked. For the DSSO crosslink identified between K447 of CBD-B and K448 of Chi1, the reaction centres were defined as the NZ atoms of each sidechain. For the DMTMM crosslinks identified between K449 of CBD-B and E527 and E399 of Chi1, the reaction centres were defined as the NZ atom of lysine and the OE1 atom of glutamic acid.

Bioinformatic analysis and phylogenetic tree building

Refer to Supplementary Information.

Structure visualisation and analysis

Structural data were visualised using UCSF ChimeraX-1.3. Multiple sequence alignments generated for the purpose of structural comparisons (i.e. excluding the phylogenetic analysis detailed above) were produced using the Molecular Evolutionary Genetics Analysis software[52] using ClustalW with default parameters[53]. Structural homologues of the CBD-A, CBD-B and $Chi2^{NTD}$ were identified using DALI[24] and TM-align[54]. Interface analyses were performed using PDBePISA to identify residues involved in electrostatic interactions [55].

Reporting summary

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

Data availability

All data are available in the main text or the supplementary materials. Structures presented in this study have been deposited in the Electron Microscopy Data Bank and Protein Data Bank under the following accession codes: YenTc holotoxin prepore EMDB-45190 [https://www.ebi.ac.uk/emdb/EMD-45190], PDB-9C4K; YenTcA prepore EMDB-45405 [https://www.ebi.ac.uk/emdb/EMD-45405]; YenTcBC prepore EMDB-45409 [https://www.ebi.ac.uk/emdb/EMD-45409]; YenTc holotoxin pore EMDB-45422 [https://www.ebi.ac.uk/emdb/EMD-45422], PDB-9CBC, YenTc pore EMDB-45423 [https://www.ebi.ac.uk/emdb/EMD-45423]. The mass spectrometry data used in this study are available in the ProteomeExchange via PRIDE under accession code PXD068275. Molecular dynamics trajectories and data have been deposited to Zenodo [https://doi.org/10.5281/zenodo.17438278]. The following PDB structures were referred to in the text: 4O9Y; 6YRO; 5BOA; 6HBO; 5UD7; 6DFQ; 4IGL; 6OGD; 6RW8; 6RWA; 6RW9; 6RWB.

References

1. Hurst, M. R. et al. The main virulence determinant of Yersinia entomophaga MH96 is a broad-host-range toxin complex active against insects. *J. Bacteriol.* **193**, 1966–1980 (2011).
2. Hurst, M. R. H., van Koten, C. & Jackson, T. A. Pathology of Yersinia entomophaga MH96 towards Costelytra zealandica (Coleoptera; Scarabaeidae) larvae. *J. Invertebr. Pathol.* **115**, 102–107 (2014).
3. Hurst, M. R., Becher, S. A., Young, S. D., Nelson, T. L. & Glare, T. R. Yersinia entomophaga sp. nov., isolated from the New Zealand grass grub Costelytra zealandica. *Int. J. Syst. Evolut. Microbiol.* **61**, 844–849 (2011).
4. ffrench-Constant, R. & Bowen, D. Novel insecticidal toxins from nematode-symbiotic bacteria. *Cell. Mol. Life Sci.* **57**, 828–833 (2000).
5. Bowen, D. et al. Insecticidal toxins from the bacterium Photorhabdus luminescens. *Science* **280**, 2129–2132 (1998).
6. Blackburn, M., Golubeva, E., Bowen, D. & Ffrench-Constant, R. H. A novel insecticidal toxin from Photorhabdus luminescens, toxin complex a (Tca), and its histopathological effects on the midgut of Manduca sexta. *Appl. Environ. Microbiol.* **64**, 3036–3041 (1998).
7. Hinchliffe, S. J. et al. Application of DNA microarrays to study the evolutionary genomics of Yersinia pestis and Yersinia pseudotuberculosis. *Genome Res.* **13**, 2018–2029 (2003).
8. Hares, M. C. et al. The Yersinia pseudotuberculosis and Yersinia pestis toxin complex is active against cultured mammalian cells. *Microbiology* **154**, 3503–3517 (2008).
9. Marshall, S. D. et al. Histopathological effects of the Yen-Tc toxin complex from Yersinia entomophaga MH96 (Enterobacteriaceae) on the Costelytra zealandica (Coleoptera: Scarabaeidae) larval midgut. *Appl. Environ. Microbiol.* **78**, 4835–4847 (2012).
10. Barth, H., Aktories, K., Popoff, M. R. & Stiles, B. G. Binary bacterial toxins: biochemistry, biology, and applications of common Clostridium and Bacillus proteins. *Microbiol. Mol. Biol. Rev.* **68**, 373–402 (2004).
11. Busby, J. N., Panjikar, S., Landsberg, M. J., Hurst, M. R. & Lott, J. S. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature* **501**, 547–550 (2013).
12. Busby, J. N. et al. The ABC toxin complex from Yersinia entomophaga can package three different cytotoxic components expressed from distinct genetic loci in an unfolded state: the structures of both shell and cargo. *IUCrJ* **11**, 299–308 (2024).
13. Piper, S. J. et al. Cryo-EM structures of the pore-forming A subunit from the Yersinia entomophaga ABC toxin. *Nat. Commun.* **10**, 1952 (2019).

14. Chassagnon, I. R., Piper, S. J. & Landsberg, M. J. in *Comprehensive Nanoscience and Nanotechnology* 279–298 (Elsevier, 2019).
15. Song, N. et al. Genome-wide dissection reveals diverse pathogenic roles of bacterial Tc toxins. *PLoS Pathog.* **17**, e1009102 (2021).
16. Landsberg, M. J. et al. 3D structure of the Yersinia entomophaga toxin complex and implications for insecticidal activity. *Proc. Natl Acad. Sci. USA* **108**, 20544–20549 (2011).
17. Busby, J. N. et al. Structural analysis of Chi1 chitinase from Yen-Tc: the multisubunit insecticidal ABC toxin complex of Yersinia entomophaga. *J. Mol. Biol.* **415**, 359–371 (2012).
18. Roderer, D., Hofnagel, O., Benz, R. & Raunser, S. Structure of a Tc holotoxin pore provides insights into the translocation mechanism. *Proc. Natl Acad. Sci. USA* **116**, 23083–23090 (2019).
19. Xu, Y. et al. CRISPR screens in Drosophila cells identify Vsg as a Tc toxin receptor. *Nature* **610**, 349–355 (2022).
20. Stuart, D. I., Subramaniam, S. & Abrescia, N. G. A. The democratization of cryo-EM. *Nat. Methods* **13**, 607–608 (2016).
21. Venugopal, H. et al. High-resolution cryo-EM using a common LaB6 120-keV electron microscope equipped with a sub–200-keV direct electron detector. *Sci. Adv.* **11**, eadr0438 (2025).
22. Gatsogiannis, C. et al. Tc toxin activation requires unfolding and refolding of a β-propeller. *Nature* **563**, 209–213 (2018).
23. Leidreiter, F. et al. Common architecture of Tc toxins from human and insect pathogenic bacteria. *Sci. Adv.* **5**, eaax6497 (2019).
24. Holm, L., Laiho, A., Törönen, P. & Salgado, M. DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci.* **32**, e4519 (2023).
25. Li, Y.-F. et al. Structure of CFA/I fimbriae from enterotoxigenic Escherichia coli. *Proc. Natl Acad. Sci. USA* **106**, 10793–10798 (2009).
26. Martin, C. L. et al. Structures of the insecticidal toxin complex subunit XptA2 highlight roles for flexible domains. *Int. J. Mol. Sci.* **24**, 13221 (2023).
27. Ng'ang'a, P. N. et al. Multistate kinetics of the syringe-like injection mechanism of Tc toxins. *Sci. Adv.* **11**, eadr2019 (2025).
28. Waterfield, N. An ABC guide to the bacterial toxin complexes. *Adv. Appl. Microbiol.* **58**, 169–183 (2005).
29. Sergeant, M., Jarrett, P., Ousley, M. & Morgan, J. A. W. Interactions of insecticidal toxin gene products from Xenorhabdus nematophilus PMFI296. *Appl. Environ. Microbiol.* **69**, 3344–3349 (2003).
30. Sheets, J. J. et al. Insecticidal toxin complex proteins from Xenorhabdus nematophilus: structure and pore formation*. *J. Biol. Chem.* **286**, 22742–22749 (2011).
31. Meusch, D. et al. Mechanism of Tc toxin action revealed in molecular detail. *Nature* **508**, 61–65 (2014).
32. Gatsogiannis, C. et al. Membrane insertion of a Tc toxin in near-atomic detail. *Nat. Struct. Mol. Biol.* **23**, 884–890 (2016).
33. Jones, S. A. & Hurst, M. R. Purification of the Yersinia entomophaga Yen-TC toxin complex using size exclusion chromatography. in *Microbial-Based Biopesticides: Methods and Protocols*, 39–48 https://doi.org/10.1007/978-1-4939-6367-6_4 (Springer-Nature, 2016).
34. Bayburt, T. H. & Sligar, S. G. Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Sci.* **12**, 2476–2481 (2003).
35. Punjani, A., Rubinstein, J., Fleet, D. & Brubaker, M. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017)
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
37. Bryant, P., Pozzati, G. & Elofsson, A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* **13**, 1265 (2022).
38. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).
39. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
40. Vergara-Cruces, Á et al. Structure of the plant plastid-encoded RNA polymerase. *Cell* **187**, 1145–1159.e1121 (2024).
41. Bartolec, T. K. et al. Cross-linking mass spectrometry discovers, evaluates, and corroborates structures and protein-protein interactions in the human cell. *Proc. Natl Acad. Sci. USA* **120**, e2219418120 (2023).
42. Liu, F., Lössl, P., Scheltema, R., Viner, R. & Heck, A. J. R. Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* **8**, 15473 (2017).
43. Crowder, D. A. et al. High-sensitivity proteome-scale searches for crosslinked peptides using CRIMP 2.0. *Anal. Chem.* **95**, 6425–6432 (2023).
44. Vizcaíno, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2015).
45. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
46. Schmid, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856 (2011).
47. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
48. Hünenberger, P. H. & van Gunsteren, W. F. Alternative schemes for the inclusion of a reaction-field correction into molecular dynamics simulations: influence on the simulated energetic, structural, and dielectric properties of liquid water. *J. Chem. Phys.* **108**, 6117–6134 (1998).
49. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
50. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput Chem.* **32**, 2319–2327 (2011).
51. Gowers, R. et al. *Proc. 15th Python in Science Conference* (SciPy, 2016).
52. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
53. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
54. Zhang, Y. & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).
55. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
56. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

Acknowledgements

The authors acknowledge the facilities and the scientific and technical assistance of the Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy and Microanalysis, The University of Queensland. We are grateful to Matthias Floetenmeyer and Lou Brillault (University of Queensland) for their advice and input. We wish to acknowledge The University of Queensland's Research Computing Centre (RCC) for its support in this research. This work was supported by the Australasian Leadership Computing Grants scheme, with computational resources provided by NCI Australia, an NCRIS enabled capability supported by the Australian Government. Australian Research Council

(DP220101681, DP170104484) – MJL; Royal Society of New Zealand Marsden Fund (14-UOA-146) – JSL; New Zealand Foundation of Research, Science and Technology (C10X0804) – MRHH.

Author contributions

Experimental design: S.J.P., M.J.L., M.B., E.D., J.S.L., J.K.K.L., M.R.H.H. Protein expression and purification: S.J.P., I.R.C. Electron microscopy data collection: S.J.P. Electron microscopy data processing, model building and model validation: Y.S.L., N.A.A., S.G.R., T.I.K., S.J.P. and J.K.B. Structural analysis: Y.S.L., S.G.R., N.A.A. MD Simulations: S.G.R. Phylogenetic analysis: G.F. Cross-linking mass-spectrometry experiments and data analysis: J.K.K.L. Writing: Y.S.L., S.G.R., N.A.A., S.J.P., M.J.L. All authors discussed the results, read and edited the manuscript.

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/s41467-025-66050-x.

Correspondence and requests for materials should be addressed to Sarah J. Piper or Michael J. Landsberg.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025