

Mechanism of Tc toxin action revealed in molecular detail

Dominic Meusch^{1*}, Christos Gatsogiannis^{1*}, Rouslan G. Efremov^{1*}, Alexander E. Lang², Oliver Hofnagel¹, Ingrid R. Vetter³, Klaus Aktories^{2,4} & Stefan Raunser^{1,5}

Tripartite Tc toxin complexes of bacterial pathogens perforate the host membrane and translocate toxic enzymes into the host cell, including in humans. The underlying mechanism is complex but poorly understood. Here we report the first, to our knowledge, high-resolution structures of a TcA subunit in its prepore and pore state and of a complete 1.7 megadalton Tc complex. The structures reveal that, in addition to a translocation channel, TcA forms four receptor-binding sites and a neuraminidase-like region, which are important for its host specificity. pH-induced opening of the shell releases an entropic spring that drives the injection of the TcA channel into the membrane. Binding of TcB/TcC to TcA opens a gate formed by a six-bladed β -propeller and results in a continuous protein translocation channel, whose architecture and properties suggest a novel mode of protein unfolding and translocation. Our results allow us to understand key steps of infections involving Tc toxins at the molecular level.

Tc toxin complexes are virulence factors of many bacteria such as the plague pathogen *Yersinia pestis* and the insect pathogen *Photorhabdus luminescens*^{1–3}. Excreted by the bacteria as soluble proteins, Tcs bind to the cell surface, are endocytosed and perforate the host endosomal membrane by forming channels that translocate toxic enzymes into the host^{4,5}. This damages and ultimately kills the target cells. Tcs with specificity for insects are receiving great interest as potential biopesticides for expression in transgenic plants as alternatives to *Bacillus thuringiensis* toxins^{6,7}. Tc toxins are composed of TcA, TcB and TcC subunits^{4,8}. TcB and TcC together form a closed cage, in which the cytotoxic domain of TcC is autoproteolytically cleaved⁹. TcA forms a large bell-shaped pentameric structure^{5,10,11} and enters the membrane like a syringe, forming a translocation channel through which the cytotoxic domain is probably transported into the cytoplasm⁵. This unique mechanism differs from that of typical pore-forming toxins¹² or other toxins that form translocation pores, for example, diphtheria¹³ and anthrax toxin¹⁴. However, little is known about the force that drives this mechanism and how the three components interact with each other and with the host membrane to unfold and translocate the toxic domain. A greater understanding requires knowledge of the structures of the Tc complex and its subunits in different states, that is, before and after membrane insertion. We use a hybrid approach combining X-ray crystallography and electron cryomicroscopy (cryo-EM) to determine the crystal structures of the prepore TcA (TcdA1) subunit and TcB–TcC (TcdB2–TccC3) subunits, the cryo-EM structures of TcA in its pore state and the complete Tc toxin complex PTC3.

TcA has four receptor-binding domains

To understand how the pentameric TcA subunit (TcdA1 from *P. luminescens*) is built in molecular detail and how it functions, we solved its crystal structure—comprising 12,500 residues with a molecular weight of 1.41 MDa—at a resolution of 4.0 Å (Fig. 1a and Supplementary Video 1). The TcA protomer is composed of eight domains (Fig. 1b, c). Six domains form the shell and are connected by a ~42 amino acid linker to two domains forming the inner channel (Fig. 1b). The pentamer is

formed through the tight assembly of the protomers, which interact with all four other protomers (Extended Data Table 1a).

The large extended α -helical shell domain (amino acids 14–297, 434–1090, 1608–1632, 1762–1972) forms the inner scaffold of the shell (Fig. 1b). It contains over 60 irregularly arranged α -helices, including many short helices, and no β -strands, and can be separated into a larger and smaller lobe that are arranged perpendicular to each other (Fig. 1b). We could not identify any structural relative, suggesting that this structure represents a previously unknown fold type.

The α -helical shell domain contains three large insertions, forming a neuraminidase-like domain and four receptor-binding domains (Fig. 1b, c). The neuraminidase-like domain (1091–1307, 1581–1607), which is structurally homologous to neuraminidases¹⁵ (Extended Data Fig. 1), closes the bottom of the pentameric shell. The receptor-binding domains A–D (A, 298–433; B, 1308–1382, 1491–1580; C, 1383–1490; and D, 1633–1761) have an immunoglobulin (Ig)-like β -sandwich fold of two sheets with antiparallel β -strands (Extended Data Fig. 1). The domains are structurally reminiscent of the receptor-binding domains of the diphtheria¹⁶ and anthrax toxins¹⁷, suggesting that these domains act as receptor-binding domains. Thus, in contrast with other toxins that contain only one or two receptor-binding domains^{16,17}, TcA contains in total four putative receptor-binding domains.

Remarkably, the amino- and carboxy-terminal regions of the two insertions that form the receptor-binding domains are located less than 8 Å apart (distance between 273 and 433, and between 1089 and 1764), suggesting that the inserted domains were added to an originally simpler shell. Indeed, we found that several homologous toxins from *P. luminescens* and other species lack some of the domains (Extended Data Table 1b). This might also explain the different host specificity of Tc toxins.

The pore-forming entity of TcA consists of a funnel formed by the TcB-binding domain (2328–2516) and a 200-Å-long α -helical central channel (Fig. 2a). Like the receptor-binding domains, the TcB-binding domain has an Ig-like β -sandwich fold (Extended Data Fig. 1). Each protomer contributes an intertwined helical pair that bends in a left-handed

¹Department of Physical Biochemistry, Max Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. ²Institut für Experimentelle und Klinische Pharmakologie und Toxikologie, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany. ³Department of Mechanistic Cell Biology, Max Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. ⁴BIOSS Centre for Biological Signalling Studies, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany. ⁵Institute of Chemistry and Biochemistry, Freie Universität Berlin, Thielallee 63, 14195 Berlin, Germany.

*These authors contributed equally to this work.

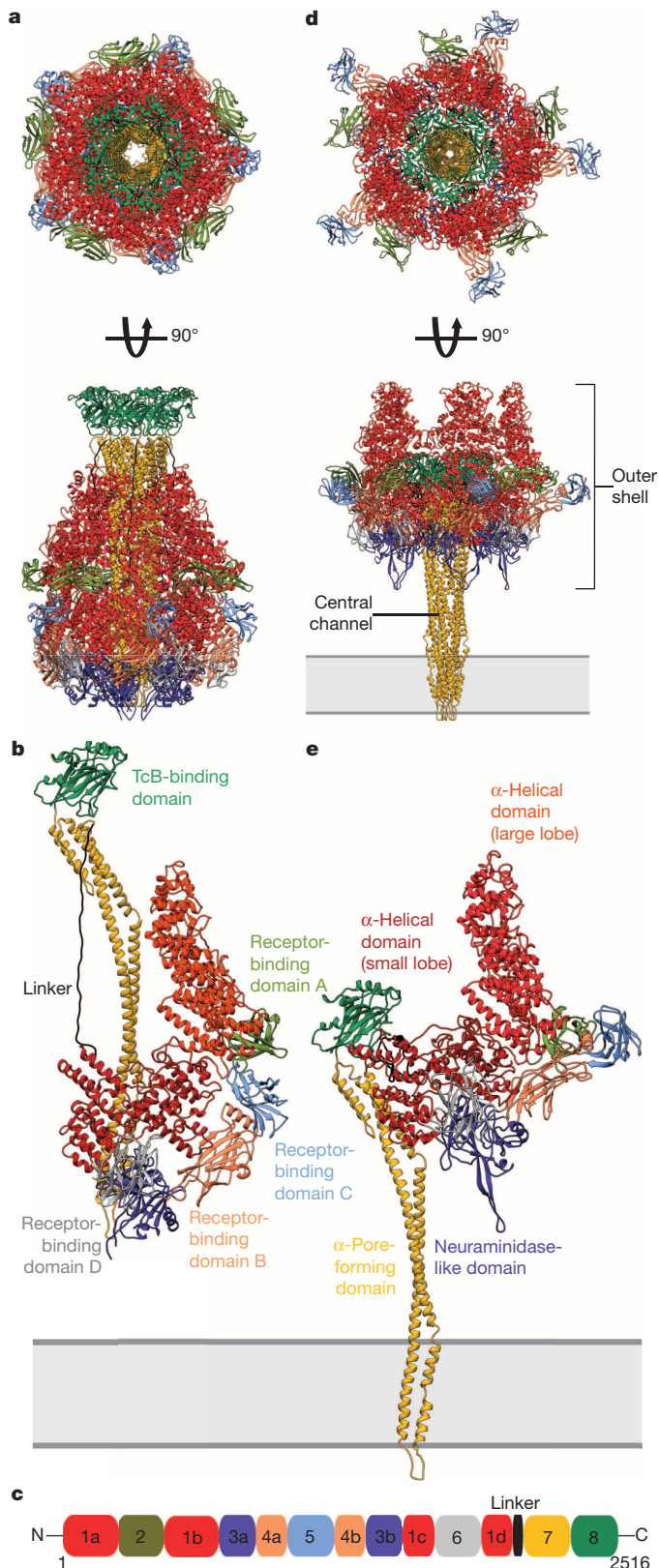


Figure 1 | Structures of the TcA prepore and pore complex. **a**, Top and side view of the TcA prepore pentamer. **b**, **c**, Side view of a TcA prepore protomer (**b**) and its domain organization (**c**). **d**, Top and side view of the TcA pore pentamer. **e**, Side view of a TcA pore protomer.

helical 180° turn over the length of the channel, forming an antiparallel coiled coil in the upper and middle part of the pore. At the top, the channel is further fortified by five helices (one per monomer) that

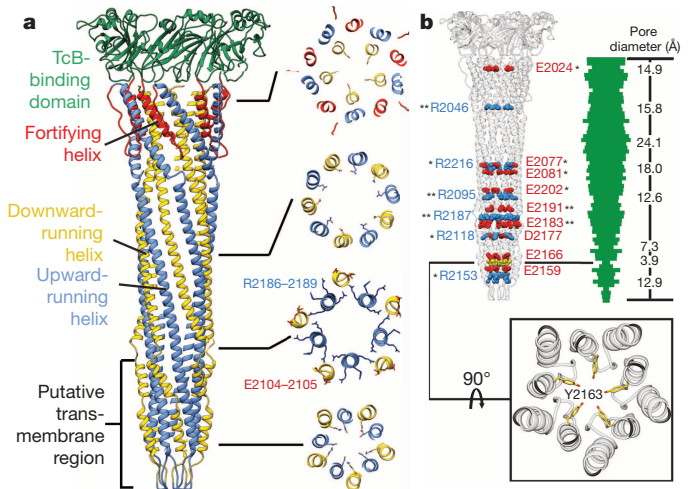


Figure 2 | The TcA prepore channel. **a**, The central channel is composed of ten long antiparallel α -helices that are interrupted by short unwound regions or loops. Helices running downwards and upwards are coloured yellow and blue, respectively. Helices that connect the channel with the funnel are depicted in red. **b**, Depiction of positively (blue), negatively (red) charged and aromatic (yellow) residues facing the interior of the channel. One and two asterisks depict highly conserved and conserved residues, respectively. The bar diagram (green) represents the inner diameter calculated using 3 \AA steps. The numbers indicate the channel diameters at the respective height. The narrowest position of the channel is at tyrosine 2163 (yellow), which minimizes the diameter of the channel to 3.9 \AA (see inset).

stabilize this region of the protein (Fig. 2a). At the bottom, the channel is closed by conserved β -turn loops formed by almost exclusively hydrophobic residues (2141–2153) (Fig. 2a and Extended Data Fig. 2a, b).

The outer surface of the lower end of the pore (pore surface residues between 2107–2158) is predominantly hydrophobic, confirming our previous prediction that this is the membrane-spanning part of the channel⁵ (Fig. 2a and Extended Data Fig. 2b). Notably, a ring of 15 arginines marks the upper end of the transmembrane region, indicating that this positively charged ring probably interacts with the negatively charged surface of the membrane and thereby stabilizes the pore within the lipid bilayer (Extended Data Fig. 2d).

Transition from the prepore to pore state

To reveal the mechanism of pore formation in molecular detail, we determined the structure of the TcA complex in its pore state by cryo-EM (Extended Data Fig. 3a–f) and built a molecular model derived from fitting the crystal structure of the TcA prepore into the cryo-EM structure (Fig. 1d, e and Extended Data Fig. 3g, h). Formation of the pore state is associated with marked conformational changes, resulting in the opening of the shell and a 12 nm shift of the central channel (Supplementary Video 2). Besides the linker between the shell and the channel, the conformations of the domains themselves are not considerably different from in the prepore state. However, their positions relative to each other change considerably, resulting in a completely different overall conformation (Fig. 1d, e). Comparing the prepore and pore structures, we identified three major hinge regions that are responsible for the opening of the shell (Extended Data Fig. 4a, b). Whereas two hinges involve receptor-binding domains, allowing their rearrangement (Extended Data Fig. 4c, d), a hinge between the two lobes of the α -helical shell domain is essential for the opening of the shell (Extended Data Fig. 4e). Notably, the arrangement of the four receptor-binding domains of TcA resembles the ectodomains of the interferon receptor IFNAR2 (ref. 18), suggesting that the different large-scale reorganization of these domains in the two TcA states represents an accommodation to different target geometries or an alteration of their binding affinity. The distance of receptor-binding domains to

the membrane is ~ 125 Å on average, suggesting that the receptor(s) on the cell surface probably have an elongated shape and protrude far from the membrane to bind to TcA, which is not unusual in eukaryotic cells. A good example of cell surface proteins that rise high above the membrane are integrins, which have an extracellular region of around 200 Å (ref. 19).

Entropic spring drives injection

The linker connecting the shell and the channel in the prepore state has a very unusual, stretched-out conformation, with a length of 113 Å for 48 residues (Fig. 3a). In the pore state, however, the linker appears to be contracted (Fig. 3b).

Steered molecular dynamics simulations of the transition between the two states indicate a gain in free energy of 20–66 kcal mol⁻¹, with a strong entropic contribution. This suggests that this process is exergonic and would therefore happen spontaneously, whether the collapsed state of the linker is folded or not (Fig. 3c). We therefore propose that the linker acts as an entropic spring. Comparable to a rubber band, it is stretched in the prepore state and can only achieve its preferred condensed conformation once the shell opens. This consequently pulls on the central channel and could therefore be the driving force for the previously described syringe-like injection mechanism of Tc toxins⁵ (Extended Data Fig. 4f and Supplementary Video 2).

The shell subunits from adjacent protomers overlap like an iris diaphragm, forming large interfaces in both TcA states. The interaction is mediated by electrostatic interactions involving different sets of residues in the prepore and pore state, respectively (Extended Data Fig. 2e, f). This indicates that both the prepore and pore state are stable conformations of the toxin complex. Interestingly, a shift to higher or lower pH values results in strong repulsions at the neuraminidase-like domains that close the shell at the bottom (Fig. 3d). We therefore propose that this domain acts as an electrostatic lock mainly responsible for the pH-induced opening of the shell⁵.

TcA translocation channel

The luminal surface of the channel is mainly negatively charged (Extended Data Fig. 2c), with many conserved charged residues protruding into the lumen of the channel (Fig. 2b), supporting our previous findings that the channel is cation selective²⁰. The cation selectivity of the anthrax translocation pore²¹ was shown to be a prerequisite for the unidirectional translocation of anthrax lethal factor (LF), making use of the pH gradient across the endosomal membrane²². It is thought that protonation before and deprotonation after the translocation of anionic proteins is the basis of a Brownian ratchet movement that drives the

translocation process in the direction of the compartment with the higher pH value, that is, the cytosol²³.

In contrast with LF, which is an anionic substrate, the translocated domains in Tc complexes from *P. luminescens*—that is, the hypervariable regions of TccC3 and TccC5—are both cationic substrates with isoelectric points of 9.68 and 8.65, respectively, and would therefore be translocated without additional pH gradients. This suggests a pH-independent mechanism of translocation in Tc toxins.

Interestingly, the diameter of the channel varies and opens at the height of a larger hydrophobic region, but generally narrows from the top to the bottom, where it is closed by the β -turn loops (Fig. 2b and Extended Data Fig. 2b). The narrowest diameter of the part of the channel that does not permeate the membrane is 12.6 Å, indicating that α -helices but not folded proteins would fit through. The putative transmembrane region, however, is narrower and almost closed at tyrosine 2163 (minimal diameter of 3.9 Å), indicating that membrane permeation of the toxin must induce conformational changes at this site, thus triggering the opening of the channel (Fig. 2b).

TcB–TcC forms a large cocoon

To understand better how TcB and TcC interact and to understand the detailed structural organization of these subunits, we solved the crystal structure of the TcB–TcC (TcdB2–TccC3 from *P. luminescens*) fusion protein at a resolution of 2.35 Å. TcB and TcC are built by large β -sheets that are wound in a continuous anticlockwise spiral, forming a large hollow cocoon which is closed by a distorted six-bladed non-symmetrical β -propeller that sits at a $\sim 45^\circ$ angle to the longitudinal axis of the cocoon and the N-terminal region of TcB (5–100) (Fig. 4a, b, Extended Data Figs 5, 6a–e and Supplementary Video 1). The structures are almost identical to the previously described structure of the YenB–YenC2 complex from the insect pathogen *Yersinia entomophaga*⁹, although the latter shares only 56.5% protein sequence identity with TcB and TcC (Supplementary Discussion and Extended Data Fig. 5e, f). We performed mutational studies and tested the proteolytic activity of TcC mutants *in vitro* (Extended Data Fig. 7a) and their toxicity by treating cells *in vivo* (Extended Data Fig. 7b), and found that TcC acts as an aspartyl autoprotease, similarly to YenC2, with aspartates 651 and 674 as the typical catalytic dyad (Extended Data Fig. 5a–d).

The inner surface of the TcB–TcC cocoon is mainly positively charged and contains large hydrophobic patches, which creates a hostile milieu for protein folding (Extended Data Fig. 5g–k). This indicates that the TcB–TcC cocoon provides an ideal microenvironment to harbour an unfolded protein and to shield it from the surroundings. It might even be able to directly unfold native proteins. In line with this, the cleaved

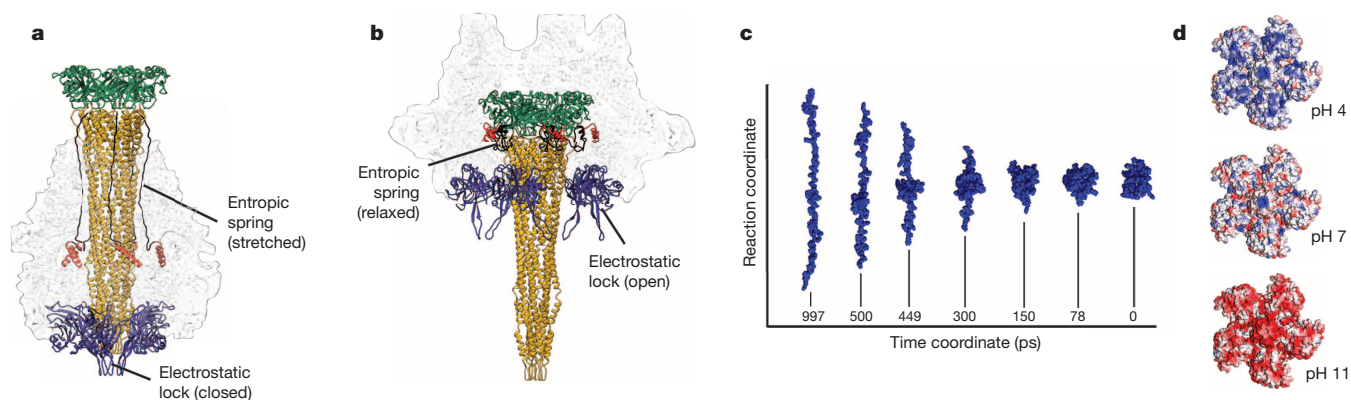
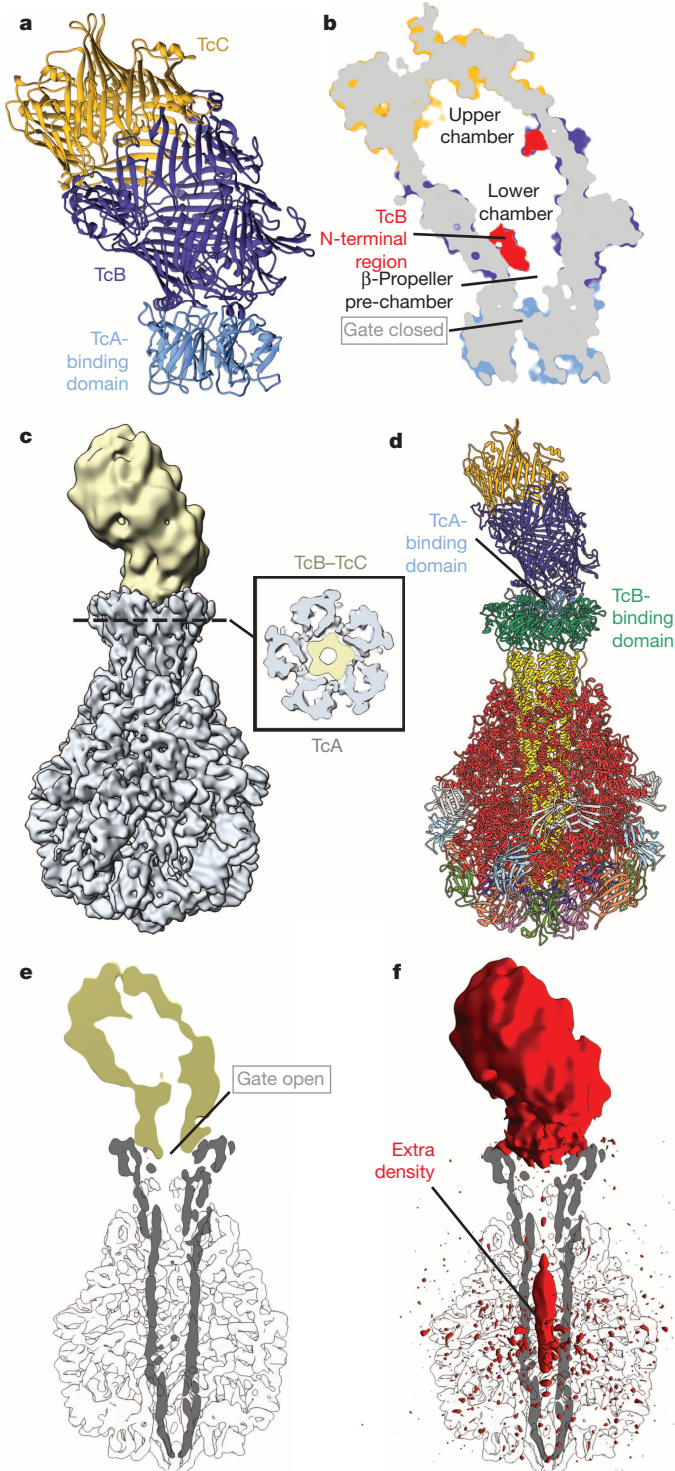


Figure 3 | An entropic spring drives the syringe-like injection mechanism of TcA. **a, b**, Structure of the TcA prepore (**a**) and pore complex (**b**), depicting the linker between the channel and shell domain (black) and the neuraminidase-like domain (blue). **c**, Steered molecular dynamics simulation, describing the contraction of the stretched linker, indicates that the linker is an entropic

spring. **d**, Surface electrostatic Coulomb potential of the neuraminidase-like domain at pH 4, 7 and 11. Strong repulsions at pH values of 4 and 11 but not at pH 7 indicate that the neuraminidase-like domain acts as an electrostatic lock that opens at high or low pH values.



off ADP-ribosyltransferase domain (679–960), which resides inside the cocoon before translocation (Extended Data Fig. 7c, d), is not resolved in our structure, indicating that the protein is either unfolded or in static disorder.

The TcA-binding domain acts as a gate

We then assembled the full holotoxin PTC3 prepore complex (TcdA1, TcdB2–TccC3) *in vitro* and determined its structure by cryo-EM (Fig. 4c and Extended Data Fig. 3i–l). The molecular model of PTC3 derived from fitting our crystal structures of the TcB–TcC and TcA complexes into the cryo-EM structure (Fig. 4d and Supplementary Videos 1, 3) clearly defines the interface between the two complexes

Figure 4 | Binding of TcB–TcC to TcA opens the β -propeller gate and the ADP-ribosyltransferase is secreted into the channel of TcA. **a**, TcB and TcC, depicted in blue and yellow, respectively, form a large cocoon. TcB also comprises the TcA-binding domain (light blue) that folds into a non-symmetrical six-bladed β -propeller. **b**, Central slice through TcB–TcC to visualize the upper and lower chamber as well as the β -propeller pre-chamber of the cocoon interior. The TcA-binding domain, TcB, TcC and the TcB N-terminal region are coloured in light blue, dark blue, yellow and red, respectively. **c**, Cryo-EM structure of the PTC3 (TcA/TcB–TcC) holotoxin complex. TcA and TcB–TcC are shown in grey and yellow, respectively. The inset shows a slice through the density at the indicated height of the interface between TcA and TcB–TcC, indicating that the β -propeller gate is open. **d**, Molecular model of the holotoxin complex, obtained by fitting the TcA crystal structure (Fig. 1a) and the TcB–TcC crystal structure (a) into the cryo-EM structure of the holotoxin complex shown in c. TcA domains are depicted in the same colours as in Fig. 1. TcB and TcC are coloured blue and yellow, respectively. The TcA-binding domain is shown in cyan. **e**, Slice through the density of the channel and TcB–TcC showing the open β -propeller gate and a continuous channel between the TcB–TcC cocoon and the channel of TcA. **f**, Difference map between the present cryo-EM structure of holotoxin and our previous cryo-EM structure of TcA alone (Electron Microscopy Data Bank accession EMD-2297). The difference map is shown in red overlaid with the central slice through the channel of the holotoxin shown in e depicted in grey.

(Extended Data Fig. 6f). The β -propeller of TcB and the domain of TcA that forms the highly conserved funnel of the central channel interact strongly with each other, identifying them as TcA- and TcB-binding domains, respectively (Fig. 4d).

Interestingly, the gate formed by the distorted β -propeller of TcB (Fig. 4b) is open in the holotoxin complex, resulting in a long continuous channel connecting the interior of the cocoon with that of the channel (Fig. 4e).

The density corresponding to the six-bladed β -propeller of TcB has a clear pseudo-five-fold symmetry (Fig. 4c and Extended Data Fig. 6g). This indicates that the conformational changes do not restore a six-fold symmetrical β -propeller, but that the β -propeller adapts to the symmetry mismatch by acquiring a pseudo-five-fold symmetry.

In the absence of complementary charges or matching hydrophobic patches at the interface (Extended Data Fig. 6f), shape complementarity and an induced fit are the basis for the specific interaction between TcB and TcA (Supplementary Video 3).

Most importantly, we find density inside the channel of TcA after holotoxin complex formation, indicating that the ADP-ribosyltransferase has passed through the β -propeller gate and entered the translocation pore before membrane permeation (Fig. 4f). The intensity of the density inside the pore is lower than the rest of the cryo-EM structure and fragmented, indicating that the protein is disordered inside the channel.

Thus, the binding of the distorted β -propeller of TcB to the pentameric funnel of TcA induces conformational changes that result in the rearrangement of the blades (Extended Data Fig. 6h and Supplementary Video 3) and thereby in the opening of the gate, allowing the passage of the toxic component into the target cell.

The major part of the open β -propeller pore through which the ADP-ribosyltransferase passes before it enters the TcA translocation channel is hydrophobic (Extended Data Fig. 6b). This region acts probably in a similar fashion to the Φ -clamp in anthrax toxin²⁴ and protects hydrophobic patches in the translocated protein. The narrow passage of the β -propeller also suggests that the ADP-ribosyltransferase must be unfolded before it passes through its gate, supporting our hypothesis that it is in an unfolded state inside the TcB–TcC cocoon.

The structures of TcA and TcB–TcC and of the holotoxin PTC3 described here reveal that Tc toxins are exceptional and function differently from any known toxins. Our data allow us to describe their mechanism of action in an unprecedented level of molecular detail (Supplementary Discussion and Extended Data Fig. 8). They form distinguished compartments for protein unfolding and processing, and translocate their toxic cargo in a novel way. Hence, our data lay a strong

foundation for understanding the function of other pathogens, including human pathogenic bacteria such as *Y. pestis* and *Y. pseudotuberculosis*. Furthermore, our data serve as a strong foundation for the development of protein shuttles targeting specific cell types in medical therapy.

METHODS SUMMARY

TcdA1 and TcdB2–TccC3 from *P. luminescens* were purified as described previously⁵. Crystals of TcA and TcB–TcC grew either in buffer containing 0.1 M HEPES pH 7.0, Jeffamine ED-2001 and 1.1 M sodium malonate or in buffer containing 0.1 M trisodium citrate pH 5.5, 0.1 M MgCl₂, 0.1 M NaCl and 12% PEG 4,000, respectively. X-ray diffraction data were collected at the PXII-X10SA beamline at the Swiss Light Source, and processed using the CCP4 program suite²⁵. The structure of TcA was solved at 4 Å using the previous 6.3 Å cryo-EM structure of TcA⁵ as a search model using PHENIX²⁶. The 2.35 Å structure of TcB–TcC was solved by experimental phasing. Images of vitrified TcA pore complexes and of the holotoxin (TcA, TcB–TcC) were taken with a JEOL JEM 3200FSC electron microscope equipped with a field-emission gun at an operation voltage of 200 kV. An in-column omega energy filter was used with a slit width of 15 eV. Micrographs were recorded with an 8k × 8k TemCam-F816 CMOS camera (TVIPS) under minimal dose conditions (15–20 e[−] per Å²). SPARX software²⁷ was used for all image-processing steps. For visualization and analysis we used Chimera and implemented programs²⁸. Steered molecular dynamics simulations were calculated using GROMACS²⁹.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 October 2013; accepted 10 January 2014.

Published online 23 February; corrected online 2 April 2014 (see full-text HTML version for details).

- Waterfield, N. R., Bowen, D. J., Fetherston, J. D., Perry, R. D. & French-Constant, R. H. The *tc* genes of *Photorhabdus*: a growing family. *Trends Microbiol.* **9**, 185–191 (2001).
- French-Constant, R. H. & Bowen, D. J. Novel insecticidal toxins from nematode-symbiotic bacteria. *Cell. Mol. Life Sci.* **57**, 828–833 (2000).
- Hares, M. C. *et al.* The *Yersinia pseudotuberculosis* and *Yersinia pestis* toxin complex is active against cultured mammalian cells. *Microbiology* **154**, 3503–3517 (2008).
- Sheets, J. J. *et al.* Insecticidal toxin complex proteins from *Xenorhabdus nematophilus*: structure and pore formation. *J. Biol. Chem.* **286**, 22742–22749 (2011).
- Gatsogiannis, C. *et al.* A syringe-like injection mechanism in *Photorhabdus luminescens* toxins. *Nature* **495**, 520–523 (2013).
- Bravo, A. & Soberón, M. How to cope with insect resistance to Bt toxins? *Trends Biotechnol.* **26**, 573–579 (2008).
- French-Constant, R. H., Eleftherianos, I. & Reynolds, S. E. A nematode symbiont sheds light on invertebrate immunity. *Trends Parasitol.* **23**, 514–517 (2007).
- Lang, A. E. *et al.* *Photorhabdus luminescens* toxins ADP-ribosylate actin and RhoA to force actin clustering. *Science* **327**, 1139–1142 (2010).
- 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).
- 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).
- Lee, S. C. *et al.* Structural characterisation of the insecticidal toxin XptA1, reveals a 1.15 MDa tetramer with a cage-like structure. *J. Mol. Biol.* **366**, 1558–1568 (2007).
- Lesieur, C., Vécsey-Semjén, B., Abrami, L., Fivaz, M. & Gisou van der Goot, F. Membrane insertion: the strategies of toxins. *Mol. Membr. Biol.* **14**, 45–64 (1997).
- Murphy, J. R. Mechanism of diphtheria toxin catalytic domain delivery to the eukaryotic cell cytosol and the cellular factors that directly participate in the process. *Toxins* **3**, 294–308 (2011).
- Young, J. A. T. & Collier, R. J. Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu. Rev. Biochem.* **76**, 243–265 (2007).
- Varghese, J. N., Laver, W. G. & Colman, P. M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **303**, 35–40 (1983).
- Choe, S. *et al.* The crystal structure of diphtheria toxin. *Nature* **357**, 216–222 (1992).
- Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H. & Liddington, R. C. Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833–838 (1997).
- Li, Z., Strunk, J. J., Lamken, P., Piehler, J. & Walz, T. The EM structure of a type I interferon–receptor complex reveals a novel mechanism for cytokine signaling. *J. Mol. Biol.* **377**, 715–724 (2008).
- Humphries, M. J. Integrin structure. *Biochem. Soc. Trans.* **28**, 311–339 (2000).
- Lang, A. E., Konukiewitz, J., Aktories, K. & Benz, R. TcdA1 of *Photorhabdus luminescens*: electrophysiological analysis of pore formation and effector binding. *Biophys. J.* **105**, 376–384 (2013).
- Blaustein, R. O., Koehler, T. M., Collier, R. J. & Finkelstein, A. Anthrax toxin: channel-forming activity of protective antigen in planar phospholipid bilayers. *Proc. Natl Acad. Sci. USA* **86**, 2209–2213 (1989).
- Brown, M. J., Thoren, K. L. & Krantz, B. A. Charge requirements for proton gradient-driven translocation of anthrax toxin. *J. Biol. Chem.* **286**, 23189–23199 (2011).
- Feld, G. K., Brown, M. J. & Krantz, B. A. Ratcheting up protein translocation with anthrax toxin. *Protein Sci.* **21**, 606–624 (2012).
- Krantz, B. A. *et al.* A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science* **309**, 777–781 (2005).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Hohn, M. *et al.* SPARX, a new environment for Cryo-EM image processing. *J. Struct. Biol.* **157**, 47–55 (2007).
- Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Berendsen, H. & van der Spoel, D. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to R. S. Goody for continuous support and for useful comments on the manuscript. We thank V. Pfaumann for cloning, C. Schwan and S. Bergbrede for technical support and the Raunser laboratory for selecting particles. We thank the staff of beamline X10SA at the Paul Scherrer Institute, the X-ray communities at the Max Planck Institute (MPI) Dortmund and the MPI Heidelberg for help with data collection. This work was supported by the Deutsche Forschungsgemeinschaft, grants RA 1781/1-1 (to S.R.) and AK 6/22-1 (to K.A.), and by the Max Planck Society (to S.R.).

Author Contributions D.M. crystallized both protein complexes. D.M., R.G.E. and I.R.V. processed, refined and analysed X-ray data. C.G. and O.H. collected cryo-EM data; C.G. processed, refined and analysed cryo-EM data; A.E.L. purified proteins and performed mutational studies; I.R.V. calculated molecular dynamics simulations; K.A. and S.R. designed the study. D.M., C.G., R.G.E. and A.E.L. prepared figures. S.R. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The coordinates for the EM structures have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-2551 and EMD-2552. Coordinates of TcdB2–TccC3 and TcdA1 have been deposited in the Protein Data Bank under accession numbers 4O9X and 4O9Y, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.R. (stefan.raunser@mpi-dortmund.mpg.de).

METHODS

Protein expression and purification. TcdA1 and the TcdB2–TccC3 fusion protein of *P. luminescens* subsp. *akhurstii* were overexpressed and purified as described previously^{5,8}. After dialysis of TcdA1 against buffer containing 50 mM MES pH 5.0, 100 mM NaCl, 0.05% Tween-20 and 5% glycerol, and dialysis of TcdB2–TccC3 against buffer containing 50 mM Tris pH 8.0, 100 mM NaCl, 0.05% Tween-20 and 5% glycerol, size-exclusion chromatography was performed as the final purification step for each protein separately using a Superose 6 10/300 GL column (GE Healthcare Life Sciences). TcdA1 and TcdB2–TccC3 were concentrated to 15–20 mg ml⁻¹ and 10 mg ml⁻¹, respectively, using Amicon filter devices (Millipore), and crystallization experiments were performed directly. For the purification of the holotoxin complex, TcdB2–TccC3 was mixed in a 1.5 molar excess with TcdA1 and dialysed against a buffer containing 50 mM MES pH 5.0, 100 mM NaCl, 0.05% Tween-20 and 5% glycerol before separation using a Superose-6 10/300 GL column. The elution fractions containing protein with the highest molecular weight were further examined by negative-stain electron microscopy, and samples containing the holotoxin complex were finally used for plunge freezing and cryo-EM. The purification of the TcdB2–TccC3 mutants was carried out in the same way as described for the wild-type TcdB2–TccC3 complex.

Crystallization of TcdA1. Crystals of TcdA1 grew after 5–7 days in 6 µl hanging drops by vapour diffusion against reservoir solution (in a 2:1 protein-to-reservoir-solution ratio) containing 0.1 M HEPES pH 7.0, Jeffamine ED-2001 and 1.1 M sodium malonate, and reached a final size of approximately 300 × 200 × 200 µm. For cryo-protection, crystals were soaked step-wise in reservoir solution starting from 5% up to 25% glycerol and were flash-frozen in liquid nitrogen.

Crystallization of TcdB2–TccC3. The TcdB2–TccC3 fusion protein crystallized after 3–5 days in 3 µl hanging drops by vapour diffusion against reservoir solution in a 1:1 ratio containing 0.1 M trisodium citrate pH 5.5, 0.1 M MgCl₂, 0.1 M NaCl and 12% PEG 4,000 and reached a final size of approximately 100 × 60 × 60 µm. Crystals for phase determination were soaked in reservoir solutions containing either 1 mM K₂Pt(NO₂)₄ or ethyl mercury thiosalicylate (EMTS) for 4 h. For cryo-protection, crystals were soaked in reservoir solution containing 25% glycerol and flash frozen in liquid nitrogen.

Data collection and processing. X-ray diffraction data were collected at the PXII-X10SA beamline at the Swiss Light Source. For TcdA1 crystals, data were recorded using a 20 µm focused beam. An attenuated beam was used to centre and screen for highly ordered regions in the crystal. Frames were collected in 0.25° oscillation steps and exposure times of 0.5–1 s. One highly redundant low-resolution and four high-resolution data sets in 40° wedges were collected from a single crystal at different positions. Highly redundant data sets of heavy metal derivatized TcdB2–TccC3 crystals were collected with a 150 µm focused beam, an oscillation range of 0.25° and an exposure time of 0.25 s. All data subsets were indexed and integrated using XDS³⁰ and combined into a single complete high-quality data set.

Diffraction data were scaled and merged using SCALA³¹ and TRUNCATE³² from the CCP4 suite²⁵, respectively. Diffraction patterns of TcdA1 crystals displayed anisotropy and were therefore scaled and truncated using the diffraction anisotropy server³³ to a resolution of 3.7, 4.2 and 3.5 Å along the crystallographic axes *a*, *b* and *c*, respectively. The strong anisotropy is also the reason why the high *R* factors in the outer resolution shells (Supplementary Table 1) were disregarded.

TcdA1 structure solution and refinement. TcdA1 crystallized as a pentamer in space group P2₁2₁2₁, with two pentamers, a total number of 24,898 amino acids in the asymmetric unit (AU) and unit cell dimensions of 303 × 321 × 667 Å. Because of a very long *c* axis and the finite size of the PILATUS 2M detector at the Swiss Light Source, the final resolution was limited to 3.5 Å in one and 3.9 Å in the other direction.

We built a poly-alanine α-helical model into our previous cryo-EM TcdA1 structure (Electron Microscopy Data Bank accession EMD-2297⁵) using find_helices_strands implemented in PHENIX²⁶. This model was then used to obtain the correct phases by molecular replacement in PHASER³⁴. Phases of the best molecular replacement solution were truncated to 6.0 Å to avoid bias. Phase modification, extension and solvent flipping were performed in DMMULTI³⁵ using an initial mask based on the EM model. Because of tenfold NCS averaging and a solvent content of 78%, phase extension resulted in an excellent 3.5 Å electron density map (Extended Data Fig. 9) in which an initial poly-alanine model was built by iterative cycles of model building in O³⁶ and COOT³⁷. The register between amino acid sequence and the α-helical model was determined based on the connectivity of density and secondary structure prediction. The model was completed in PHENIX.refine²⁶, including NCS and secondary structure restraints, and was refined in several iterative manual building cycles to a final *R*_{free} and *R*_{work} of 24.3% and 29.5%, respectively. Finally, the model was validated using MOLPROBITY³⁸; data statistics are summarized in Supplementary Table 1.

Structure determination and refinement of TcdB2–TccC3. TcdB2–TccC3 crystals belong to the primitive hexagonal space group P3₂21 with unit cell dimensions of 232 × 232 × 142 Å and one molecule per AU. The structure was solved by experimental phasing. PHASER-EP³⁴ and AUTOSOL³⁹ implemented in PHENIX²⁶ identified and refined 41 Pt sites and produced a distinct initial experimentally phased electron density map (Extended Data Fig. 1). For crystals soaked in EMTS, the best model includes 9 Hg sites (Supplementary Table 1). The structure was further built using BUCCANEER⁴⁰ and manually completed in COOT³⁷ with a final *R*_{free} and *R*_{work} of 22.3% and 20.6%, respectively. The complete statistics are shown in Supplementary Table 1 and validated using MOLPROBITY³⁸.

Cryo-EM and image processing. PTC3 (0.02 mg ml⁻¹) was applied to glow-discharged holey carbon grids (C-flats, Protochips) with a thin layer of continuous carbon and vitrified by plunging into liquid ethane using a Cp3 (Gatan). Digital micrographs were taken with a JEOL JEM 3200FSC electron microscope equipped with a FEG at an operation voltage of 200 kV, and a defocus range of 0.4–3.2 µm. An omega in-column energy filter was used with a slit width of 15 eV. The images were recorded with an 8k × 8k CMOS camera F816 (TVIPS) at a pixel size of 1.25 Å per pixel. A total of 53,128 single particles were manually selected from 1,250 images using boxer⁴¹ or e2boxer2.py⁴² and windowed into 512 × 512 pixel boxes. The defocus was estimated using an in-house defocus-determination program (unpublished). Single particles were then aligned and classified using reference-free alignment and *k*-means classification procedures in SPARX²⁷. The class averages were then manually inspected and those showing no density corresponding to TcdB2–TccC3 were selected and corresponding class members were removed from the data set.

The three-dimensional reconstruction of PTC3 was created by a multi-symmetry projection matching procedure using the SPARX software package, as previously described^{5,43} with a final data set of 43,000 particles. Our previously obtained negative-stain reconstruction of PTC3 (Electron Microscopy Data Bank accession EMD-2299) was down-filtered to 40 Å and used as a starting reference for refinements without imposing symmetry. After each refinement cycle, the density of TcdA1 was masked out and symmetrized using C5 symmetry. The remaining density corresponding to TcdB2–TccC3 and background was scaled to match the threshold of the density of TcdA1. Finally, we combined the two densities into one, masked it to remove background noise and used it as a reference for the subsequent refinement cycle. After obtaining global parameters, local refinements were performed without symmetrization of TcdA1 after each refinement cycle. The final density map was then sharpened using a negative *b* factor of –200 Å² as calculated by the program embfactor⁴⁴ and filtered to its nominal resolution of 9.1 Å according to its Fourier shell correlation (FSC) curve at 0.5. The final density of TcdB2–TccC3 was obtained after subtracting the previously obtained cryo-EM structure of TcdA1 (Electron Microscopy Data Bank accession EMD-2297) from the present volume of PTC3, after proper scaling of both density maps using the tool ‘vop subtract’ from the Chimera software package²⁸. Densities showing no connection to TcdB2–TccC3 were manually removed using ‘Volume Eraser’ from the Chimera software package. The resulting density of TcdB2–TccC3 was then down-filtered to 11 Å according to the local resolution in this region of the density. Local resolution calculations were performed as described previously⁴⁵ using a 20 Å sphere for windowed FSC calculations at 0.5.

To obtain a higher resolution reconstruction for the five-fold symmetric TcdA1 component of PTC3, we performed an independent reconstruction of PTC3 using five-fold symmetry. In the resulting reconstructions, the additional density corresponding to the TcdB2–TccC3 subcomplex appeared as a featureless blob, which was removed after each refinement cycle. The final reconstruction of the TcdA1 was then sharpened using a negative *b* factor of –140 Å² and filtered to its nominal resolution of 7.4 Å according to its FSC curve at 0.5.

The final volume of PTC3 shown in Fig. 4 was obtained after adding the 7.4 Å reconstruction of TcdA1 obtained by refinements with C5 symmetry imposed and the asymmetric 11 Å reconstruction of TcdB2–TccC3 after proper scaling of their densities.

For the analysis of the TcdA1 pore complex, 605 images were taken at a magnification of ×124.472 and a defocus range of 0.4–2.9 µm. Images were reduced by 2 × 2 pixel averaging to a final pixel size of 2.5 Å per pixel. Particles were selected manually using boxer⁴¹ and windowed in 128 × 128 pixel boxes. After stack creation, the data set was aligned and classified using reference-free alignment and *k*-means classification procedures implemented in SPARX²⁷. Resulting two-dimensional class averages were manually inspected and classes representing noise or distorted particles were rejected. The final data set of 13,300 particles was refined against the 50 Å filtered structure of the TcdA1 pore complex (Electron Microscopy Data Bank accession EMD-2301), using the projection matching protocol implemented in SPARX. During the refinement, C5 symmetry was imposed. The final reconstruction of the TcdA1 pore complex was then sharpened using a negative *b* factor of –220 Å² as calculated by the program

embfactor⁴⁴ and filtered to its average resolution of 9 Å according to its FSC curve at 0.5. For visualization, analysis and preparation of the figures and animations, we used Chimera²⁸. All three-dimensional reconstructions were filtered to their nominal resolutions. For the difference map between the holotoxin PTC3 and TcA, the contour level of the difference map was set to the estimated volume of the TcB–TcC complex, namely $277 \text{ kDa} \times 1.21 \text{ \AA}^3/\text{kDa} = 335 \text{ \AA}^3$. This estimate uses 1.21 \AA^3 per kDa (ref. 46). However, because the C terminus of the TcB–TcC complex is only partially resolved, this contour level is probably slightly overestimated and therefore the map also includes some background noise.

Molecular model of the TcA pore state. The experimental density map of the TcdA1 pore complex was segmented into two components: the central channel and the outer shell. The crystal structure of the channel was fitted as a rigid body in the density corresponding to the channel in the TcA pore complex using Chimera. Because the central channel was not as well resolved as the outer shell (Extended Data Fig. 3d), we avoided refining this molecular model against the density map in this region of the complex. Although the putative transmembrane region of the TcA pore complex is not resolved in our cryo-EM structure (Extended Data Fig. 3f, g), we did not remove this region in the resulting model of the pore state to provide a more complete comparison between both conformers. However, major conformational changes during pore formation are expected in this region⁷.

For the fitting of the shell domains we split the crystal structure of the TcA prepore complex into its respective domains and fitted them as rigid bodies into the electron density corresponding to the TcA shell in its pore state. The densities corresponding to the linker and receptor-binding domain 3 were not well defined, and we therefore excluded these domains from the flexible fitting procedure using DireX⁴⁷ (Extended Data Fig. 3g, h). The resulting model of the shell in the pore state was combined with the molecular model of the central channel. To complete the model we added the structures of receptor-binding domain 3 and also the simulated model of the linker in its relaxed state, according to their covalent bonding in the prepore state, to obtain the resulting complete molecular model of TcA in the pore state as shown in Fig. 1. Atom clashes were removed by energy minimization using PHENIX²⁶. Morphing between the two conformers was calculated using the tool ‘Morph Conformations’ in Chimera and recorded as a movie (Supplementary Video 2). Detailed analysis of the prepore–pore domain motion and identification of hinge residues was performed with the DynDom Server⁴⁸.

Structure analysis. The structures of TcdA1 and TcdB2–TccC3 were first separated in putative domains, which were then used as a search model for structural homologues using DALI⁴⁹. Topology diagrams of verified domains were created with ProFunc⁵⁰ and compared with the structural model and homologues aligned in Chimera²⁸. The atomic models of the TcdA1 protomers, TcdA1 prepore channel, linker and shell, as well as TcdB2 and TccC3, were protonated using H⁺ (ref. 51) at pH 7.5 and the electrostatic Coloumb potential of the protein surface was calculated ranging from -12 to 12 kcal mol^{-1} in Chimera²⁸. The degree of hydrophobicity per amino acid residue was calculated using scores generated as described previously⁵². All sequence alignments were prepared in ClustalW2⁵³, clustalw and pdb files were supplied to ConSurf⁵⁴ to calculate conserved amino acid residues, and the final alignment and model files were visualized in Chimera. The interaction interfaces between protomers and domains were generated using the PISA tools⁵⁵ and summarized in table format. The pore diameter of the TcdA1 prepore channel (amino acids 2016–2327) was defined and calculated by PoreWalker⁵⁶ and the transmembrane region (amino acids 2089–2140) of the TcdA1 prepore channel was solvated and embedded into a dioleoyl-phosphatidylcholine bilayer via simulation in CHARMM⁵⁷. Coiled coil regions were identified with the help of the program Socket⁵⁸. All structural images were prepared using Chimera²⁸.

Molecular dynamics simulations and estimation of free energy. In an attempt to estimate the free energy that could be gained by releasing the tensioned entropic spring, we performed steered molecular dynamics simulations of the 48-residue peptide comprising the linker according to the umbrella sampling method⁵⁹ using the GROMACS package²⁹.

Although there was additional density in our cryo-EM structure of the pore conformation corresponding to a condensed linker domain, it was not defined well enough to identify secondary structures. We therefore built several models using different algorithms that consistently and with high confidence comprised a domain with one α -helix and two adjacent short β -sheets. The nFOLD⁶⁰ model based on template 2bh1 (chain X of the *Vibrio cholerae* secretion pathway protein E) featured conserved hydrophobic residues (mostly leucines) in the core, and hydrophilic residues pointing outside. This starting structure was placed into a cubic box of water containing 100 mM NaCl.

After initial minimization, 100 ps of equilibration were performed, and then the N and C termini were pulled apart with constant velocity, using a pull rate of 20 nm ns^{-1} and a pull force constant of $20,000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. From these

trajectories, snapshots were taken to generate the starting configurations for the 70 umbrella sampling windows. In each window, a 10 ns dynamics simulation was run. The results were analysed with the weighted histogram analysis method (WHAM) as implemented in GROMACS⁶¹. This method allows the calculation of the potential of mean force for each configuration along the reaction trajectory and, if summed up, an estimate of the free energy of unfolding for the peptide. This value is approximately 66 kcal mol^{-1} for the observed distance of 113.4 \AA , corresponding to 75.6% of the maximum stretch length. As it is well known that steered molecular dynamics simulations overestimate the forces needed for peptide pulling owing to the large pull speeds necessary to achieve manageable calculation times⁶², this value represents an upper limit. We therefore applied a statistical model of polymers to get another estimate of the entropy gained. The entropic effect is supposed to be important as the chain is stretched to about 50–80% of its maximum extension⁶³. If we calculate the gain in free energy according to the ‘freely jointed chain’ model⁶⁴, we gain 25 kcal mol^{-1} for a change of the end-to-end distance from the observed 113.4 \AA to the folded state. As the freely jointed chain model is somewhat artificial because it does not account for the stiffness of the protein chain, we also used the ‘worm-like chain’ standard model⁶⁵, using a persistence length of 4 \AA (a typical value for proteins)⁶⁶, and find a value of 20 kcal mol^{-1} . Because these values do not account for any enthalpic interactions, if the collapsed state has a defined fold as assumed above, the free energy for folding of the domain, usually around 5 – 15 kcal mol^{-1} , would have to be added to these values, leading to the range of 25 – 40 kcal mol^{-1} . Still, the entropic spring would contribute the major part of the free energy.

Generation of point mutations within TccC3. To engineer the TccC3 mutants D651A, D657A, D674A, D676A, D651A/D657A and D674A/D676A, site-directed mutagenesis of wild-type TcdB2–TccC3 (TcdB2–TccC3-pET-28a) was performed according to the QuikChange protocol (Stratagene) as described by the supplier. The following primers were used: D651A sense, GTGGTTAAGCTCCGCTCCG GCAGGAAC; D651A antisense, GTTCCTGCCGGAGCGGAGCTTAACCAC; D657A sense, CCGGCAGGAACAATCGCTGGGCTGAATTTATATC; D657A antisense, GATATAAATTCAGCCAGCGATGTTCTCTGCCGG; D674A sense, CCAGTTACCCCTCTGCTCCTGATGGATTAATGC; D674A antisense, GCA TTAATCCATCAGGAGCAAGGAGGGTAAGTGG; D676A sense, ACCCTCCT TGATCCTGCTGGATTAATGCCAACAA; D676A antisense, TTGTTGGCATT AATCCAGCAGGATCAAGGAGGGT; D674A/D676A sense, CAGTTACCCCTC CTGCTCCTGCTGGATTAATGCCAACAA; D674A/D676A antisense, TGTTG GCATTAATCCAGCAGGAGCAAGGAGGGTAAGTGG.

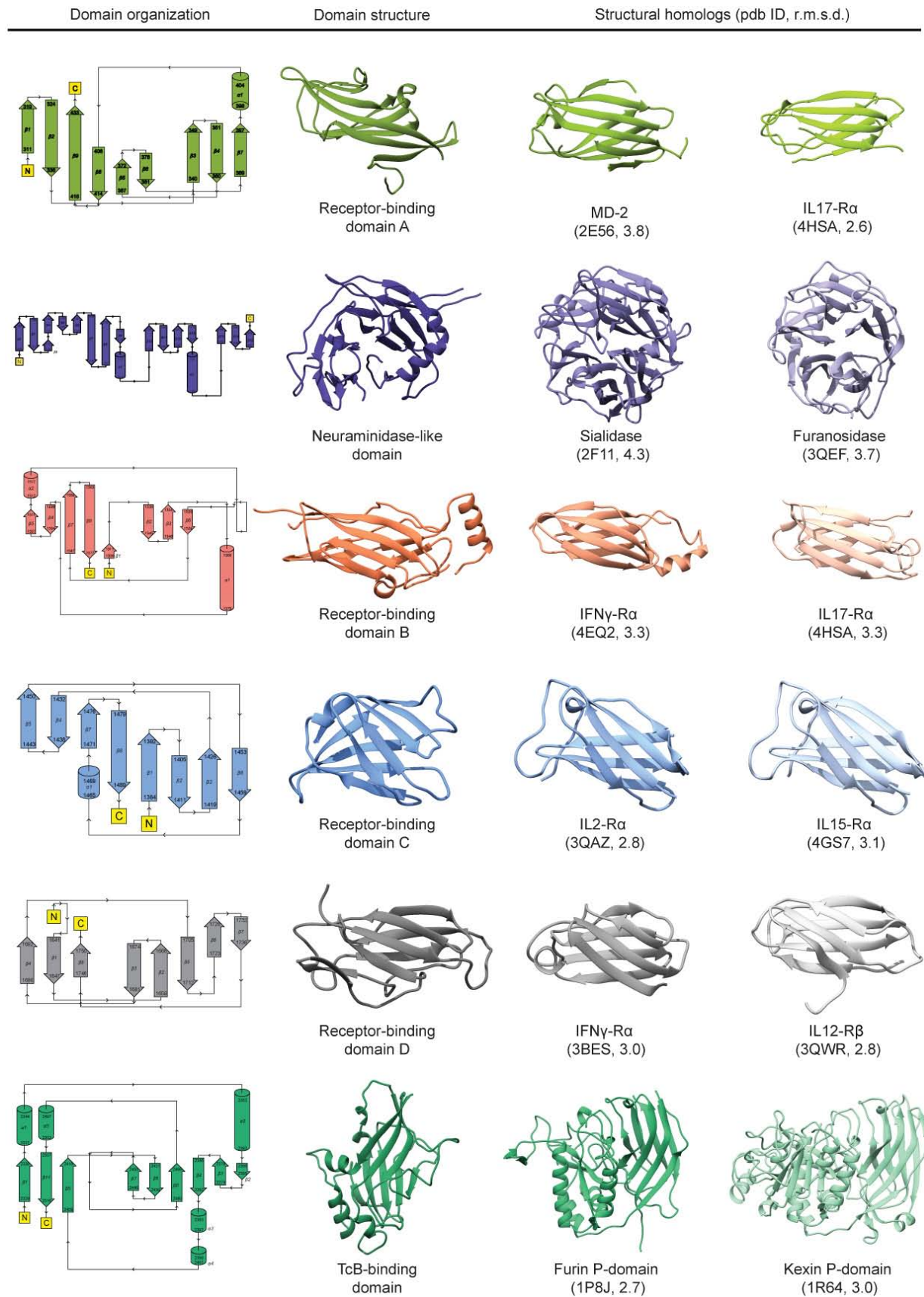
Note, for engineering D651A/D657A, the mutant D651A was used as a template together with the primers D657A sense and D657A antisense. After PCR, the methylated DNA was digested by the addition of $1 \mu\text{l}$ of DpnI (10 units per μl) for 1 h at 37°C and transformed into chemical competent *Escherichia coli* TG1 cells. Plasmid DNA was prepared using standard procedures and sequenced (GATC Biotech GmbH) to confirm gene sequences.

Immunoblot analysis. Samples were subjected to SDS–PAGE and transferred onto a PVDF membrane. The C-terminal ADP-ribosyltransferase domain hvr (hypervariable region) of TccC3 was detected with a specific antibody (anti-TccC3hvr, gift from J. J. Sheets). Binding of the second horseradish-peroxidase-coupled antibody was detected using enhanced chemoluminescent detection reagent (100 mM Tris-HCl, pH 8.0, 1 mM luminol (Fluka), 0.2 mM p-coumaric acid and 3 mM H₂O₂).

Cytotoxicity of TcdB2–TccC3 mutants and fluorescence microscopy. HeLa cells, seeded onto glass coverslips, were washed with PBS and the medium was exchanged to DMEM containing 0.5% fetal calf serum. Afterwards, all samples were incubated for 2 h with TcdA1 ($2 \mu\text{g ml}^{-1}$) and wild-type or mutant TccC3, fused to TcdB2 ($1 \mu\text{g ml}^{-1}$). Then, cells were fixed with 4% paraformaldehyde for 15 min, washed with PBS, and permeabilized with 0.15% (v/v) Triton X-100 in PBS for 10 min. Subsequently, cells were incubated for 2 h with phalloidin-TRITC and washed again with PBS. Coverslips were dried and embedded with Mowiol supplemented with 1,4-diazabicyclo[2.2.2]octane (DABCO) and DAPI for nuclear staining. Samples were cured overnight and analysed by fluorescence microscopy, using an Axiophot system (Zeiss) processed by using Metamorph 7.0 software (Universal Imaging).

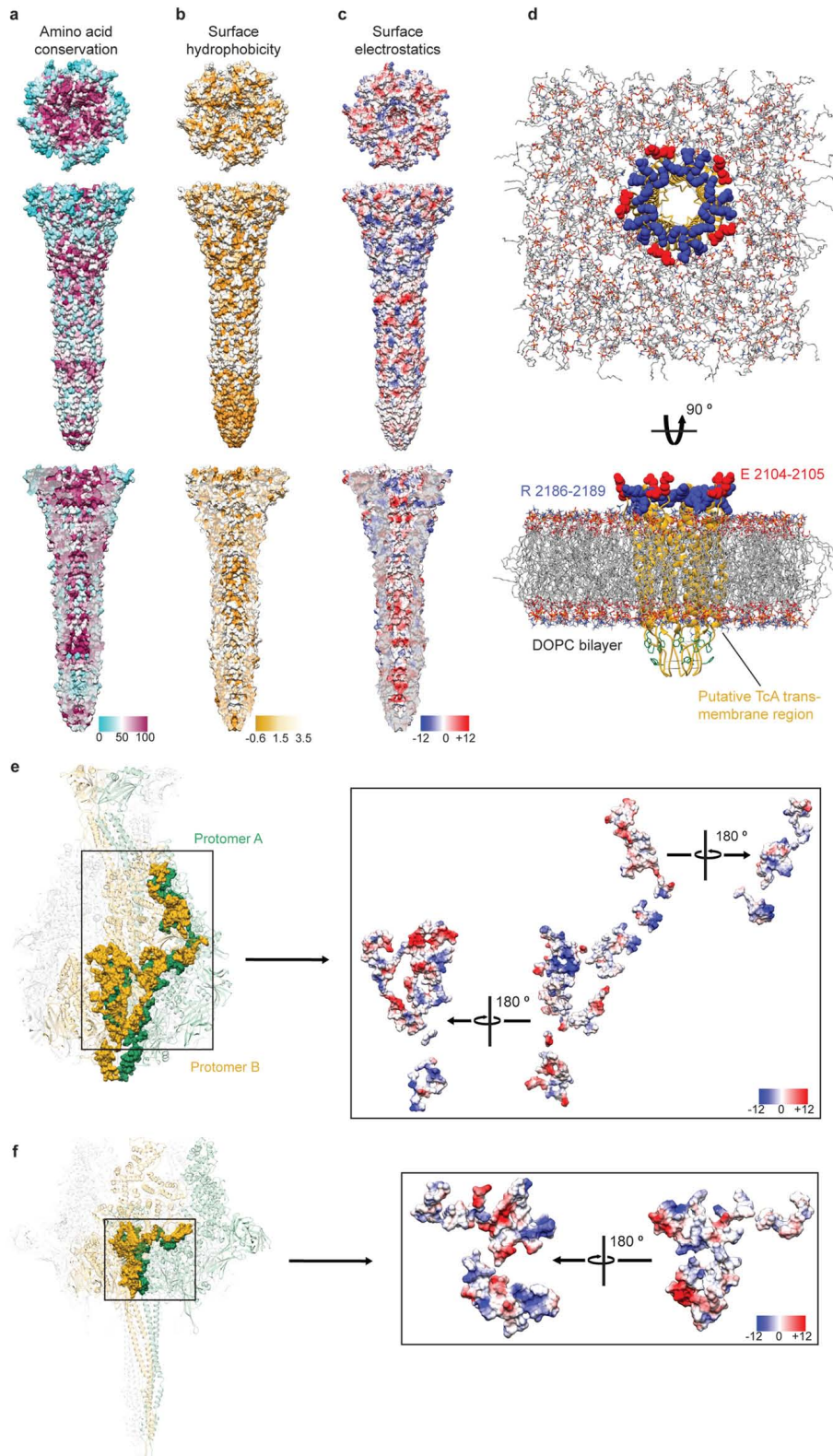
Mass spectrometry of wild-type TcdB2–TccC3. Gel slices were cut into small pieces, dehydrated and rehydrated in acetonitrile and distilled H₂O, respectively. After digestion of wild-type TcdB2–TccC3 overnight with 2 mM trypsin solution, peptides of the protein were evaporated and dissolved in 0.1% formic acid before HPLC–MS separation. Peptides were analysed and identified using MASCOT⁶⁷. For ESI-MS measurement, $20 \mu\text{l}$ of a 1 mg ml^{-1} wild-type TcdB2–TccC3 solution was applied to a desalting column after electro-spray ionization. Results were analysed using the Magtran software package⁶⁸ and visualized using Excel from the Microsoft office suite.

30. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
31. Evans, P. R. in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* Vol. 33, 22–24 (European Science Foundation, 1997).
32. French, S. & Wilson, K. On the treatment of negative intensity observations. *Acta Crystallogr. A* **34**, 517–525 (1978).
33. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).
34. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
35. Baker, D., Bystroff, C., Fletterick, R. J. & Agard, D. A. PRISM: topologically constrained phased refinement for macromolecular crystallography. *Acta Crystallogr. D* **49**, 429–439 (1993).
36. Jones, T. A. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272 (1978).
37. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
38. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
39. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65**, 582–601 (2009).
40. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
41. Ludtke, S. J. 3-D structures of macromolecules using single-particle analysis in EMAN. *Methods Mol. Biol.* **673**, 157–173 (2010).
42. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
43. Raunser, S. *et al.* Rubisco in complex with Rubisco large subunit methyltransferase. *Proc. Natl Acad. Sci. USA* **106**, 3160–3165 (2009).
44. Fernández, J. J., Luque, D., Castón, J. R. & Carrascosa, J. L. Sharpening high resolution information in single particle electron cryomicroscopy. *J. Struct. Biol.* **164**, 170–175 (2008).
45. Anger, A. M. *et al.* Structures of the human and *Drosophila* 80S ribosome. *Nature* **497**, 80–85 (2013).
46. Harpaz, Y., Gerstein, M. & Chothia, C. Volume changes on protein folding. *Structure* **2**, 641–649 (1994).
47. Wang, Z. & Schröder, G. F. Real-space refinement with DireX: from global fitting to side-chain improvements. *Biopolymers* **97**, 687–697 (2012).
48. Hayward, S. & Lee, R. A. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* **21**, 181–183 (2002).
49. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
50. Laskowski, R. A., Watson, J. D. & Thornton, J. M. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res.* **33**, W89–W93 (2005).
51. Pérez, A., Marchán, I., Svoboda, D. & Spöner, J. Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.* (2007).
52. Hessa, T. *et al.* Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381 (2005).
53. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
54. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
55. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
56. Pellegrini-Calace, M., Maiwald, T. & Thornton, J. M. PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure. *PLoS Comput. Biol.* **5**, e1000440 (2009).
57. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **97**, 50–58 (2009).
58. Walshaw, J. & Woolfson, D. N. Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307**, 1427–1450 (2001).
59. Lemkul, J. A. & Bevan, D. R. Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics. *J. Phys. Chem. B* **114**, 1652–1660 (2010).
60. Jones, D. T. *et al.* Prediction of novel and analogous folds using fragment assembly and fold recognition. *Proteins* **61** (suppl. 7), 143–151 (2005).
61. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
62. Sotomayor, M. & Schulten, K. Single-molecule experiments *in vitro* and *in silico*. *Science* **316**, 1144–1148 (2007).
63. Erickson, H. P. Reversible unfolding of fibronectin type III and immunoglobulin domains provides the structural basis for stretch and elasticity of titin and fibronectin. *Proc. Natl Acad. Sci. USA* **91**, 10114–10118 (1994).
64. Flory, P. J. Theory of elastic mechanisms in fibrous proteins. *J. Am. Chem. Soc.* **78**, 5222–5235 (1956).
65. Zhang, B., Xu, G. & Evans, J. S. A kinetic molecular model of the reversible unfolding and refolding of titin under force extension. *Biophys. J.* **77**, 1306–1315 (1999).
66. Noy, A. *Handbook of Molecular Force Spectroscopy* (Springer, 2007).
67. Koenig, T. *et al.* Robust prediction of the MASCOT score for an improved quality assessment in mass spectrometric proteomics. *J. Proteome Res.* **7**, 3708–3717 (2008).
68. Zhang, Z. & Marshall, A. G. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass Spectrom.* **9**, 225–233 (1998).



Extended Data Figure 1 | Topology diagram and structural homologues of selected TcA domains. Receptor-binding domains A to D, neuraminidase-like domain and TcB-binding domain of TcA are coloured according to the colour

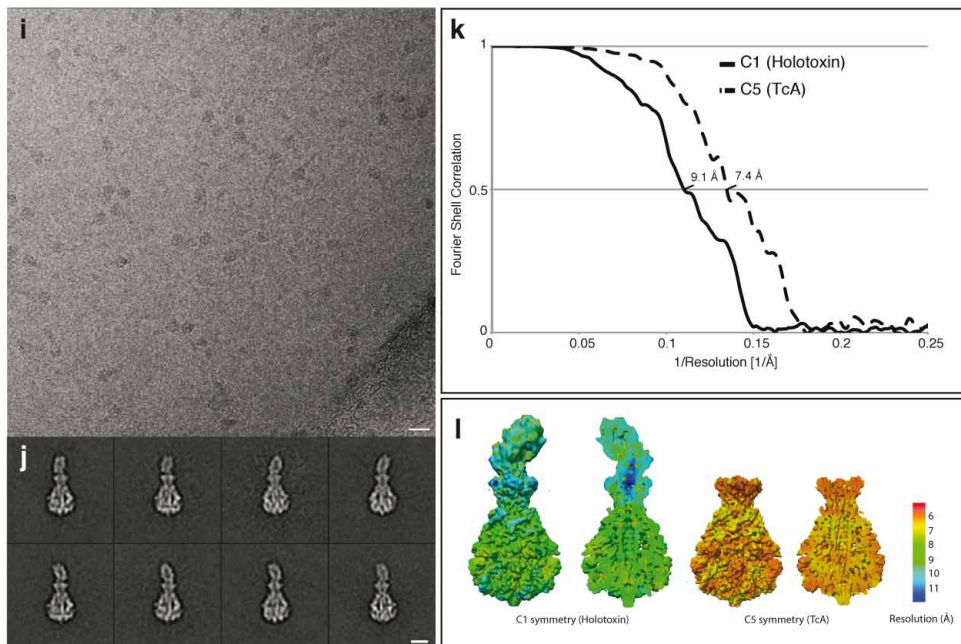
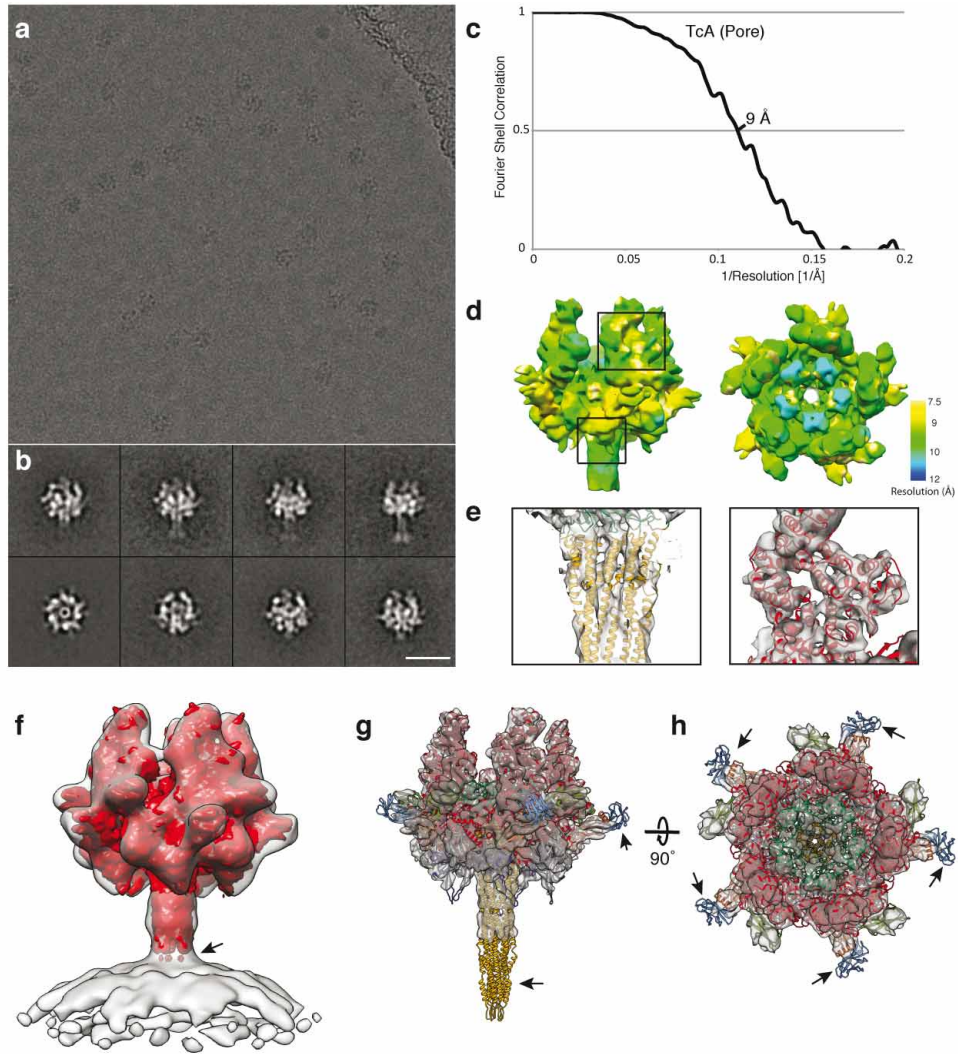
code used in Fig. 1. Two closest structural homologues (Protein Data Bank (PDB) accession and corresponding root mean squared deviation (r.m.s.d.)) were aligned to the query structure.



Extended Data Figure 2 | Biophysical properties of the TcA channel and the interface between two shell subunits in the pre-pore and pore state of TcA.

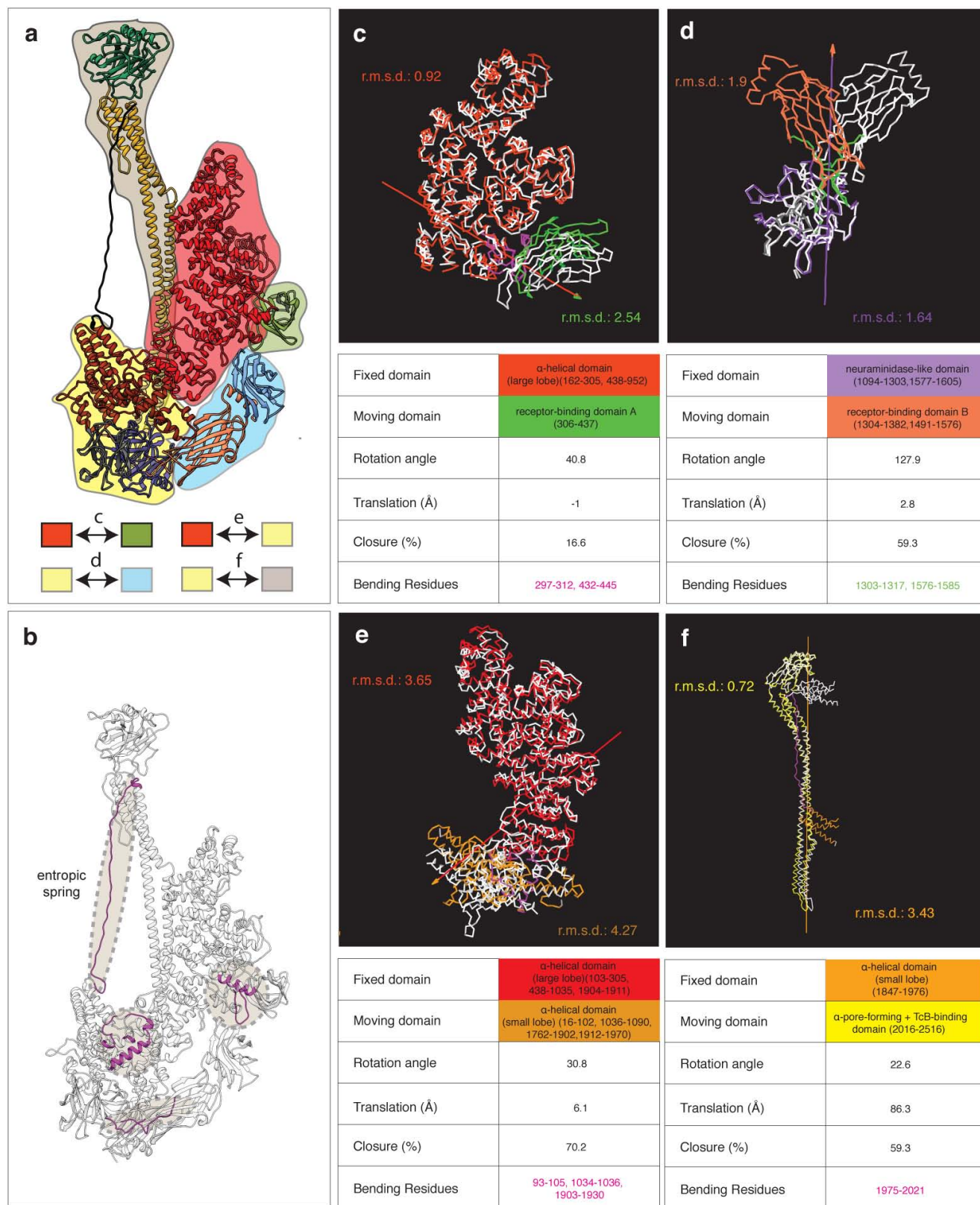
a–c, Top and side views on the surface and in the lumen of the channel, demonstrating the conservation of residues (**a**), the surface hydrophobic potential (**b**) and the surface electrostatic Coulomb potential at pH 7.5 (**c**). Conserved residues are depicted in purplish-red, hydrophobic patches are shown in orange, positive-charge and negative-charge density are coloured in blue and red, respectively. **d**, The putative transmembrane region (yellow) of

the TcA prepore channel (amino acids 2104–2190) was integrated into a homogeneous dioleoyl-phosphatidylcholine lipid bilayer ($115 \times 115 \times 40 \text{ \AA}$) using molecular dynamics simulation. Spheres of D2186–2189 and E2104/2105 are highlighted in blue and red, respectively. **e**, Residues of the large interface between two shell subunits are depicted in yellow and green on the crystal structure of the TcA prepore state (left) and flipped open to demonstrate the complementary charges at the surface (electrostatic Coulomb potential at pH 7.5) (right). **f**, Same as in **e**, but for the TcA pore state.



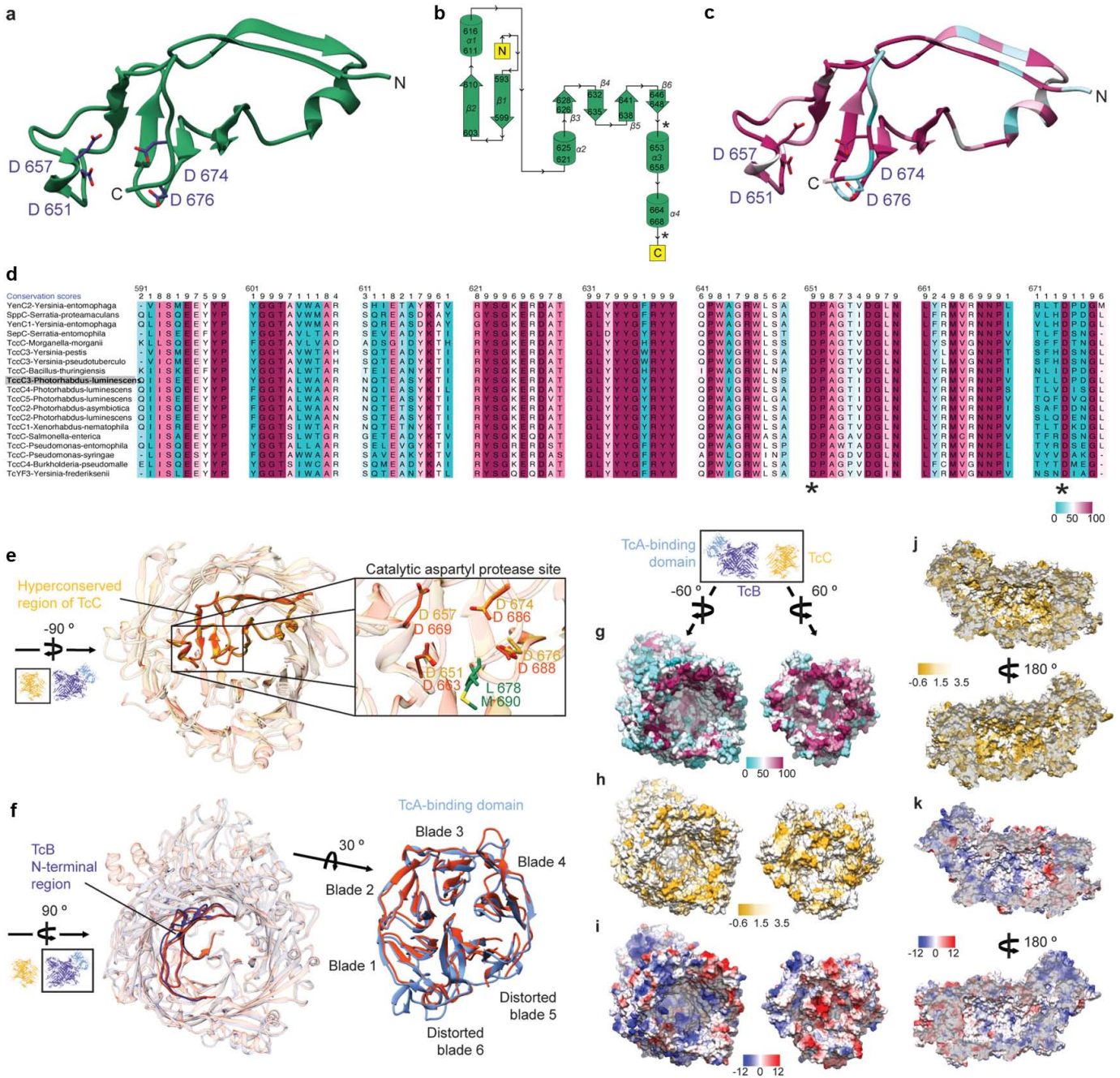
Extended Data Figure 3 | Cryo-EM of the TcA pore complex and the holotoxin complex PTC3 (TcA/TcB-TcC). **a**, Representative digital micrograph area of vitrified TcA pore complexes. **b**, Corresponding representative class averages, each containing 300–500 single particles. Scale bars, 20 nm. **c**, Fourier shell correlation (FSC). The 0.5 FSC cut-off criterion indicates that the cryo-EM map has an average resolution of 9 Å. **d**, Surface renderings of the TcA-pore reconstruction in top (left) and side (right) view, coloured according to localized resolution. **e**, Fitting of the TcA prepore crystal structure in the electron density of the TcA pore. **f**, Fitting of the present cryo-EM reconstruction of the TcA pore complex (shown in red) in our previous low-resolution cryo-EM structure of the TcA pore complex reconstituted in liposomes (Electron Microscopy Data Bank accession EMD-2298) (grey). The arrow indicates that the transmembrane part of the channel is not resolved in the present structure of the pore. **g**, **h**, Top and side view of the molecular model

of the TcA pore, obtained by flexible fitting in the cryo-EM density. The different domains are coloured according to the colour code used in Fig. 1. Parts of the model that were not resolved in the cryo-EM density map (receptor-binding domain A, transmembrane part of the central channel) are indicated by black arrows. **i**, A typical digital micrograph area of PTC3 holotoxin complexes in vitrified ice. Scale bar, 20 nm. **j**, Characteristic reference-free class averages of PTC3 complexes, each containing 300–400 single particles. Scale bar, 10 nm. **k**, The 0.5 FSC cut-off criterion indicates that the cryo-EM maps of PTC3 (TcA/TcB-TcC, C1 symmetry—that is, no symmetry) and TcA (TcB-TcC was masked during image processing, C5 symmetry) have an average resolution of 9.1 Å and 7.4 Å, respectively. **l**, Surface and cross-sections of PTC3 and TcA coloured according to the local resolution. For better clarity the respective maps are filtered to their average resolutions. See Methods for more detail.



Extended Data Figure 4 | Flexible regions and hinges of TcA. **a**, Structure of the TcA prepore protomer, with the different domains coloured according to the colour code used in Fig. 1. Four groups of domains were identified as ‘quasi-rigid bodies’ during motion analysis and are highlighted in different colours (the large lobe of the α -helical domain, red; receptor-binding domain D and the neuraminidase-like domain, yellow; receptor-binding domain B and receptor-binding domain C, cyan; the pore-forming domain and the

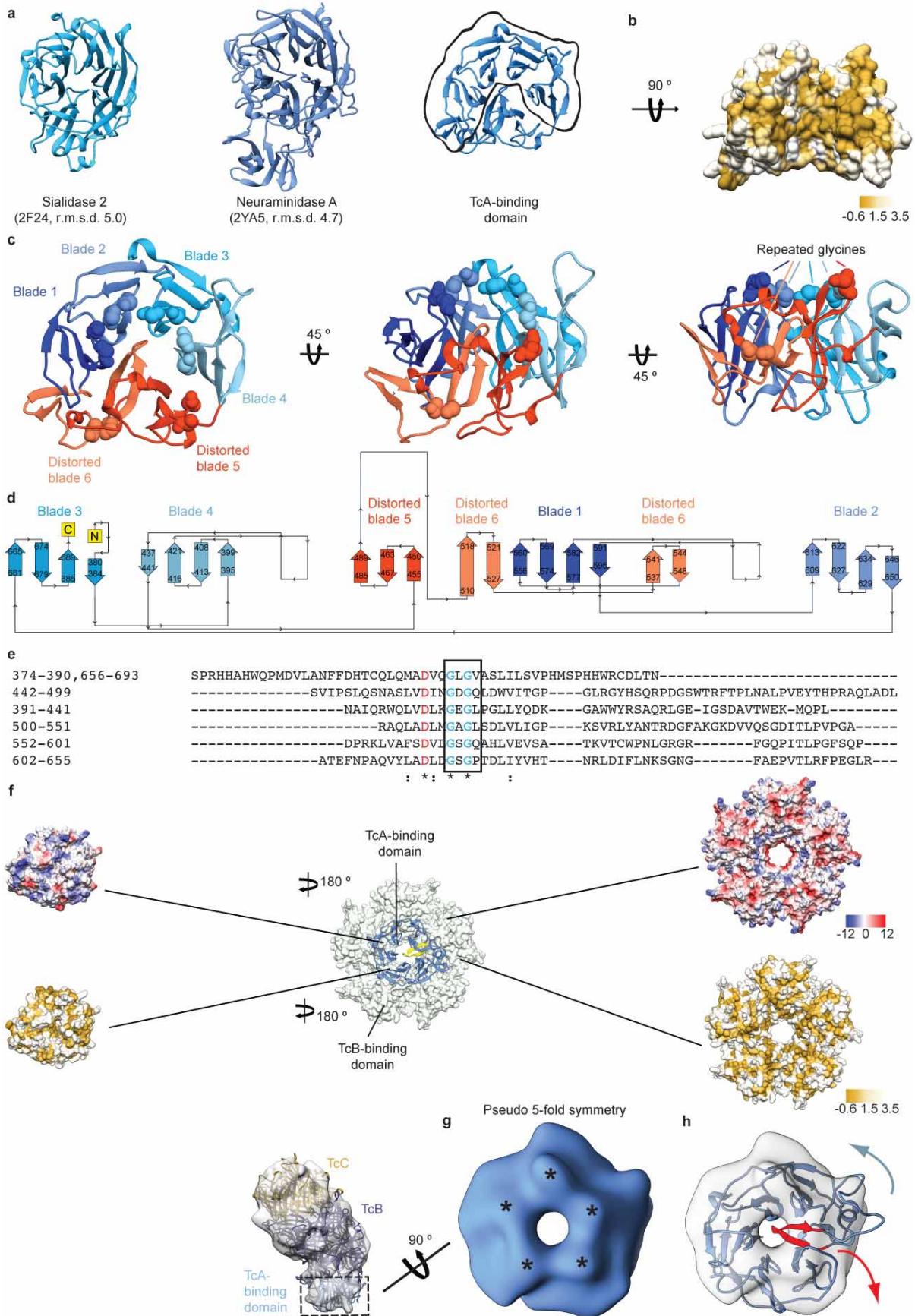
TcB-binding domain, grey). **b**, Structure of the TcA prepore protomer, with the four bending regions as identified by the domain motion analysis highlighted in magenta. **c–f**, Superimposition of the quasi-rigid bodies of TcA in the prepore (coloured) and pore state (white) for the four different pairs of covalently linked quasi-rigid bodies, as highlighted in **a**. Details of the interdomain motion are given for each pair.



Extended Data Figure 5 | The hyperconserved region of TcC, the interior of the TcB–TcC cocoon and the interaction surface between TcB and TcC.

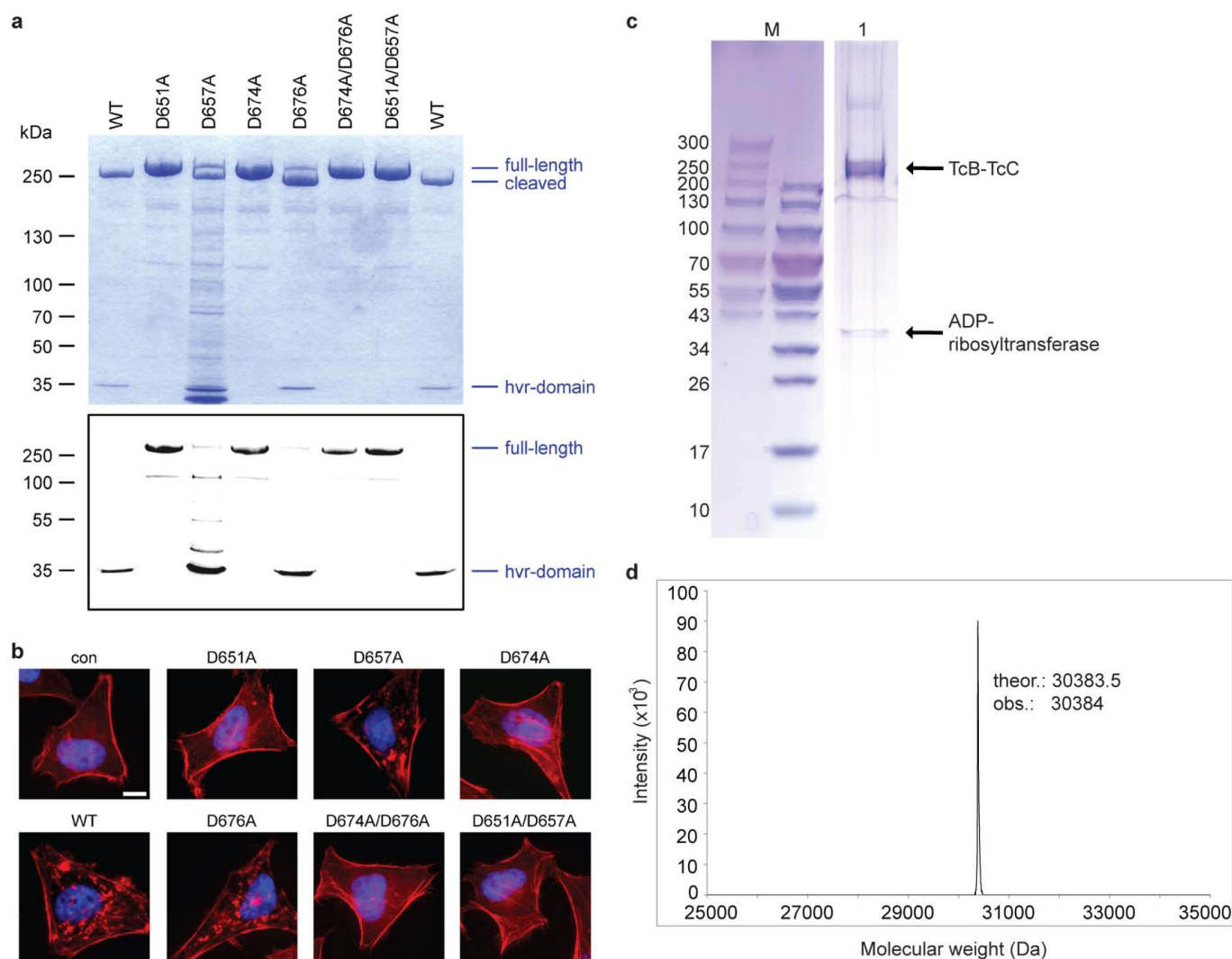
a, Ribbon representation of the hyperconserved region of TcC (green). Aspartic acid residues at the intrinsic aspartyl protease site are coloured in dark blue. D651 and D674 built the actual catalytic dyad, whereas D657 and D676 are not directly involved in proteolysis. **b**, Topology diagram of the hyperconserved region of TcC. Positions of the two highly conserved catalytic aspartate residues (D651 and D674) are marked with an asterisk. **c, d**, Amino acid sequences of the hyperconserved region of TcC and 18 homologues were aligned. The conservation score corresponds to the percentage of conserved amino acid residues and is coloured from cyan for no conservation to magenta for complete conservation. The two conserved catalytic aspartic acid residues (D651 and D674) of the intrinsic aspartyl protease domain are marked with an asterisk.

e, f, Hyperconserved region of TcC (yellow) (**e**) as well as the TcB N-terminal region (dark blue) and TcA-binding domain (light blue) (**f**). The crystal structure of YenB and YenC2 from *Yersinia entomophaga* (coloured in red) (PDB accession 4IGL) are overlaid for comparison. The depicted domains match each other with an overall r.m.s.d. of 0.88 Å. **g**, Conservation of amino acid residues at the interface between TcB and TcC. Conservation score corresponds to the percentage of conserved amino acid residues and is coloured from cyan for no (0%) conservation to magenta for complete (100%) conservation. **h–k**, Sliced views displaying the surface hydrophobic potential (**h, j**) and the surface electrostatic Coulomb potential at pH 7.5 (**i, k**) of the inner surface of TcB–TcC. Hydrophobic patches are shown in orange, positively and negatively charged surfaces are coloured in blue and red, respectively.



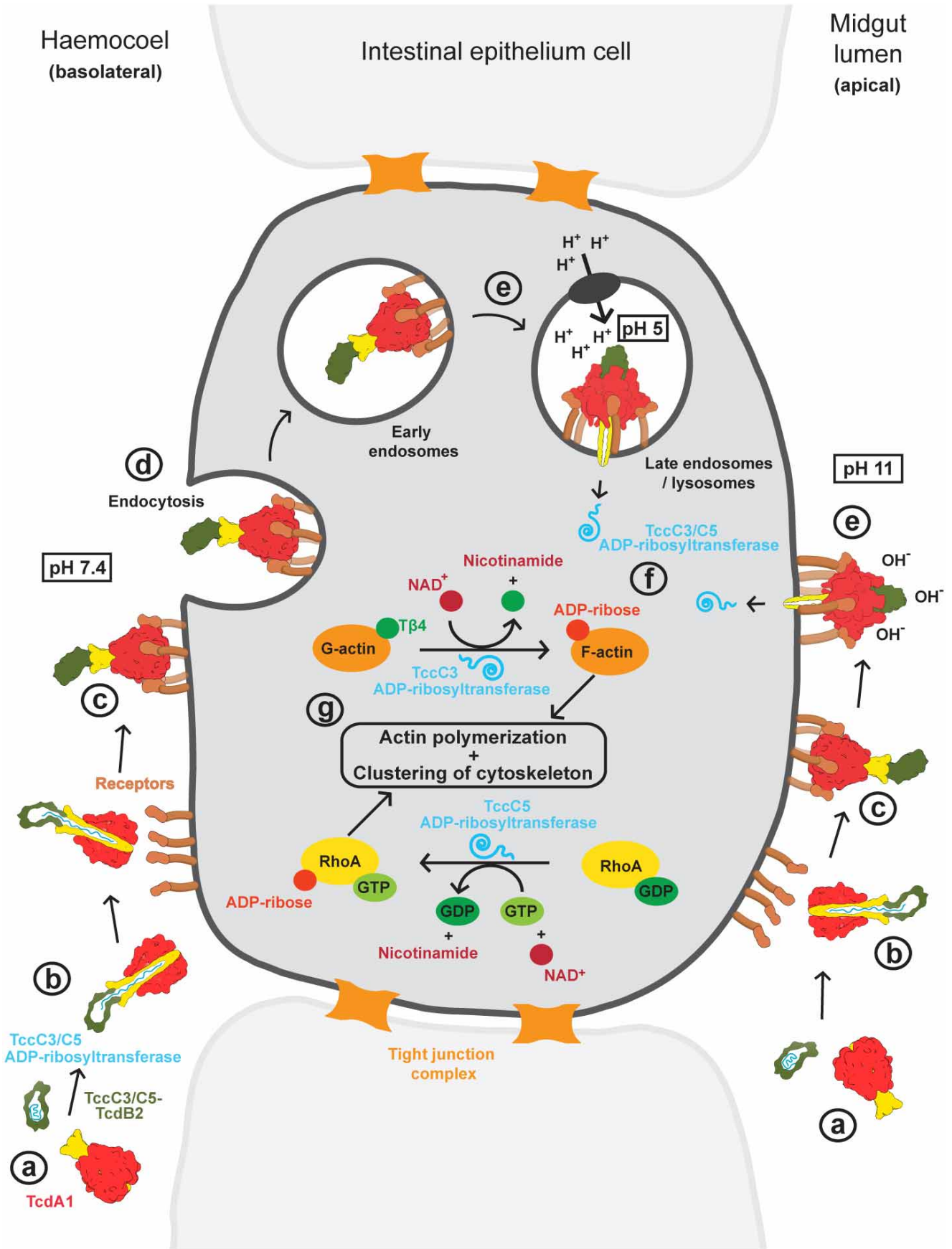
Extended Data Figure 6 | TcA-binding domain of TcB and the binding interface between TcA and TcB. **a**, Ribbon presentation of TcA-binding domain and its homologues. Two closest structural homologues (PDB accession and corresponding r.m.s.d.) were aligned to the query structure. **b**, Surface hydrophobicity of the four symmetric blades of the TcA-binding domain. Note the hydrophobic surface in the central passage of the β -propeller of the TcA-binding domain through which the ADP-ribosyltransferase passes before it enters the TcA translocation channel. **c, d**, Ribbon presentation (**c**) and topology diagram (**d**) of the TcA-binding domain (amino acids 374–693) of TcB. Blades 1 to 4 of the six-bladed β -propeller are coloured from dark to light blue whereas distorted blades 5 and 6 are coloured in red and orange, respectively. **e**, Sequence alignment of the six β -propeller blade regions. The two glycines of the DxxGxGxxxL(W)V(I) repeat (blue) are highlighted in the sequence with an asterisk and displayed in spheres in the structure depicted in

a. Note the different orientation of both glycines in distorted blade 6 compared with the glycines in the other blades. **f**, Specific interaction between TcA (TcB-binding domain) depicted as surface representation and TcB–TcC (β -propeller of the TcA-binding domain) depicted as ribbon representation. The β -hairpin that closes the central passage of the β -propeller before its interaction with TcA is depicted in yellow. Positively and negatively charged surfaces are coloured in blue and red, respectively. Hydrophobic patches are depicted in orange. **g**, The electron density of the TcB– β -propeller in the holotoxin complex has pseudo-five-fold symmetry. **h**, Rigid-body fit of the TcA-binding domain into the corresponding density. Arrows indicate putative conformational changes that are necessary to open the central passage of the β -propeller of the TcA-binding domain. The β -hairpin that closes the central passage of the β -propeller before its interaction with TcdA1 is depicted in red.



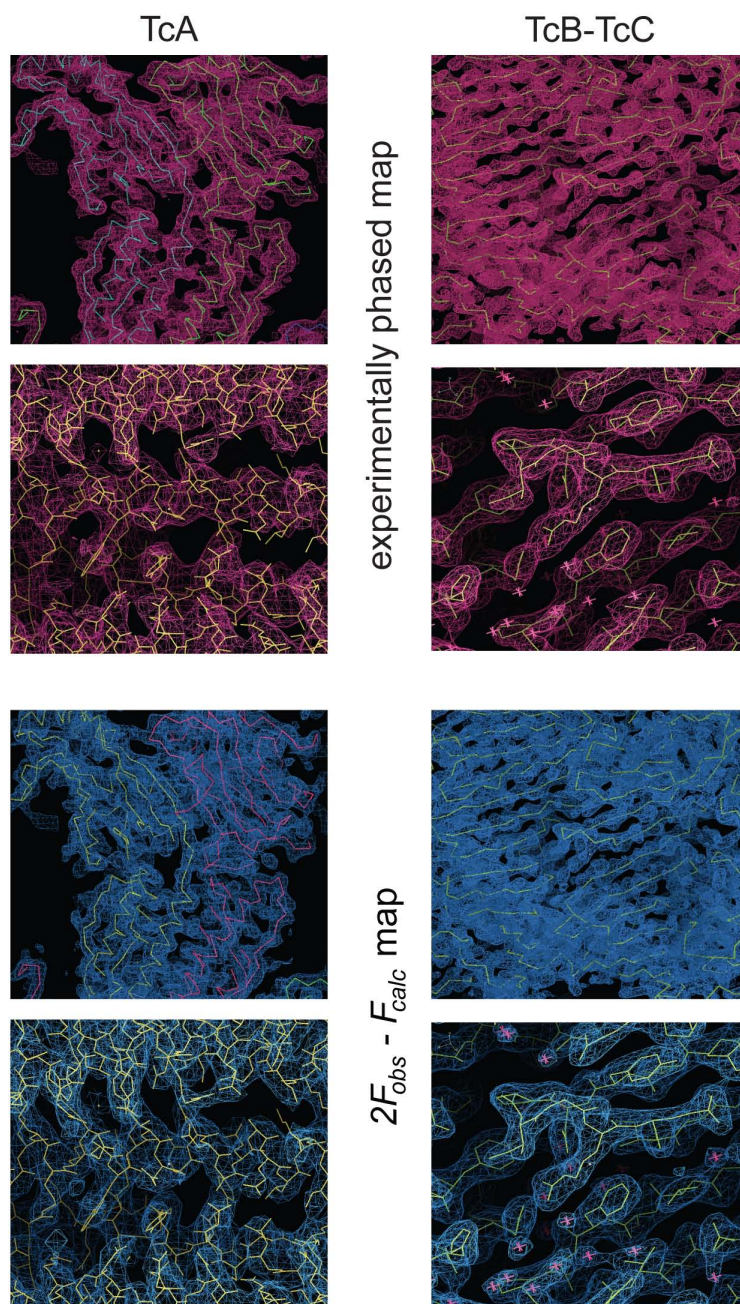
Extended Data Figure 7 | Cleavage of the C-terminal ADP-ribosyltransferase domain and *in vivo* intoxication with wild-type and mutant TcC. **a**, Two micrograms of recombinant wild-type (WT) and mutant TcC (TccC3), fused to TcB, were loaded onto a gradient SDS gel (3–12% SDS) and stained with Coomassie blue (top). Wild-type and mutant TcB–TcC were subjected to SDS–PAGE (12.5% SDS) and the C-terminal hvr domain was detected by immunoblotting (bottom). Note, in uncleaved TcB–TcC, the hvr domain is detected at ~280 kDa. The data are representative of four repetitions. **b**, ADP-ribosylation of actin induces actin aggregation that can be detected by fluorescence microscopy. Intoxication of HeLa cells with wild-type and mutant TcC. All samples were treated for 2 h with TcA ($2 \mu\text{g ml}^{-1}$) and the indicated wild-type or mutant TcB–TcC complexes ($1 \mu\text{g ml}^{-1}$). Untreated cells were used as controls (con). Then, cells were fixed and stained with

tetramethyl rhodamine isothiocyanate (TRITC)-conjugated phalloidin and 4',6-diamidino-2-phenylindole (DAPI). For each construct, ten different cells were imaged. The experiment was replicated three times. Scale bar, 10 μm . **c**, SDS–PAGE of 15–20 TcB–TcC crystals. The gel represents one out of four crystals that were screened, yielding the same result. Protein bands marked with an arrow were digested with trypsin before high-performance liquid chromatography–mass spectrometry (HPLC–MS) analysis. HPLC–MS analysis clearly identified this band as the ADP-ribosyltransferase domain, showing that this domain is located in our crystals but cleaved off from the rest of the protein. **d**, Electrospray ionization (ESI)–MS measurements of TcB–TcC showed that the ADP-ribosyltransferase domain is trapped inside the TcB–TcC cocoon, as the measured molecular weight of the sample corresponds to the full-length complex, including the ADP-ribosyltransferase domain.



Extended Data Figure 8 | Mechanism of action of Tc toxin complexes in insect hosts. **a**, *P. luminescens* secretes Tc toxin complexes that attack insect midgut cells. We propose that this can happen from the haemocoel, that is, the basolateral membrane, or the midgut lumen, that is, the apical membrane. The C-terminal hypervariable region of TccC3 or TccC5, which are both ADP-ribosyltransferases, is trapped in a closed cocoon, formed by the large β -sheet structures of TcdB2–TccC3 or TcdB2–TccC5, respectively. It is probably unfolded inside the hostile environment of the cocoon and autoproteolytically cleaved by the intrinsic aspartate protease activity of TcC. **b**, With the assembly of the holotoxin, that is, complex formation between TcB–TcC and TcA, the closed gate inside the connecting β -propeller opens owing to an induced fit mechanism, resulting in a long continuous channel. The ADP-ribosyltransferase penetrates into the channel of TcA. However, the channel is blocked until TcA permeates the membrane. **c**, The PTC3 holotoxin complex

binds with its receptor-binding domains to receptors on the surface of host cells. **d**, This triggers its endocytosis on the basolateral membrane. **e**, High pH values in the midgut lumen or low pH values in endosomes open the electrostatic lock formed by the neuraminidase-like domain of TcA, which results in an opening of the shell. The linker between the shell and channel of TcA, which acts as an entropic spring, is released and drives the syringe-like injection of the TcA translocation channel into the membrane. **f**, The ADP-ribosyltransferase is translocated through the TcA pore into the host cytoplasm. **g**, The ADP-ribosyltransferase corresponding to the C-terminal hypervariable region of TccC3 modifies G-actin such that it cannot bind to β -thymosin-4 anymore. The ADP-ribosyltransferase corresponding to the C-terminal hypervariable region of TccC5 fixes RhoA in its active state. Both modifications result in uncontrolled actin polymerization, collapse of the cytoskeleton and ultimately to cell death.















































Extended Data Figure 9 | Representative model-based and experimentally phased electron density maps of TcdA1 and TcdB2-TccC3. Visualization of the experimentally phased 4 Å and 2.35 Å resolution electron density maps (purple) of the TcB-binding and α -helical pore-forming domain of TcdA1 (top left) and the β -sheet-rich cocoon region of TcdB2-TccC3 (top right) contoured

at 1.5σ , respectively. The $2F_{\text{obs}} - F_{\text{calc}}$ Fourier maps (blue) of TcdA1 at 3.5 Å resolution (bottom left) show the interaction interface between two protomers, whereas the Fourier maps of TcdB2-TccC3 at 2.5 Å resolution (bottom right) represent distinct amino acid side chains in the β -sheet cocoon contoured at 1.5σ , respectively.

Extended Data Table 1 | TcA interfaces and incidence and distribution of TcA domains in Tc toxin complexes

a

TcdA1 Prepore								TcdA1 Pore							
Interface	Residues		Solvent access area (Å ²)	Hydrogen bonds	Salt bridges	Solvation free energy gain (kcal/mol)		Interface	Residues		Solvent access area (Å ²)	Hydrogen bonds	Salt bridges	Solvation free energy gain (kcal/mol)	
<i>Protomer</i>	#1	#2	<i>Interface</i>			<i>Protomer</i>	#1	#2	<i>Interface</i>						
A B	327	351	11301	108	40	-88.3		A B	271	262	9648	100	34	-70.4	
A C	49	50	1417	13	2	-5.9		A C	14	12	241	3	2	-1.2	
A D	50	49	1374	12	2	-6.3		A D	13	16	249	3	2	-1.2	
A E	352	330	11448	109	44	-87.9		A E	259	272	9591	102	34	-70.4	
<i>Channel</i>	#1	#2	<i>Interface</i>			<i>Channel</i>	#1	#2	<i>Interface</i>						
A B	137	134	5131	55	22	-58.4		A B	137	134	5131	55	22	-58.4	
<i>Shell</i>	#1	#2	<i>Interface</i>			<i>Shell</i>	#1	#2	<i>Interface</i>						
A B	150	168	4906	47	19	-19.0		A B	91	81	3039	30	10	-14.6	
<i>Protomer A</i>	#1	#2	<i>Interface</i>			<i>Protomer A</i>	#1	#2	<i>Interface</i>						
1  2 	33	28	1035	17	5	-0.8		1  2 	28	21	831	11	0	-5.8	
1  3 	40	40	1456	15	1	-13.2		1  3 	52	47	1709	20	1	-1.9	
1  4 	16	18	431	8	0	1.3		1  4 	22	26	690	5	0	-3.4	
1  5 	17	17	518	4	0	-3.2		1  6 	46	35	1527	23	3	-11.1	
1  6 	19	14	558	3	0	-3.7		1  7 	11	14	411	3	0	0.1	
1  7 	19	15	464	7	2	-1.9		1  8 	8	10	253	1	0	-2.1	
2  5 	3	4	66	0	0	-1.5		2  4 	4	3	73	1	0	1.1	
3  4 	18	17	600	6	5	0.9		3  4 	29	19	917	10	2	-2.7	
3  6 	30	28	1003	10	3	-3.9		3  6 	18	20	661	6	5	-1.9	
4  5 	23	26	895	15	2	-8.8		4  5 	24	28	889	11	1	-7.4	
7  8 	14	24	817	7	4	-7.4		7  8 	12	20	615	3	1	-7.1	

b

<i>Protein name</i>	<i>Organism</i>	Domain							
		#1	#2	#3	#4	#5	#6	#7	#8
TcdA1	<i>Photobacterium luminescens</i>	+	+	+	+	+	+	+	+
TcdA2	<i>Photobacterium luminescens</i>	+	-	+	+	+	+	+	+
TcdA4	<i>Photobacterium luminescens</i>	+	-	+	+	+	+	+	+
TcdA2	<i>Photobacterium asymbiotica</i>	+	-	+	+	+	+	+	+
XptA1	<i>Xenorhabdus nematophila</i>	+	+	+	+	+	+	+	+
XptA2	<i>Xenorhabdus nematophila</i>	+	+	+	+	+	+	+	+
YenA	<i>Yersinia entomophaga</i>	+	+	+	-	-	-	+	+
TcYF1	<i>Yersinia frederiksenii</i>	+	+	+	+	+	-	+	+
TcA	<i>Yersinia pseudotuberculosis</i>	+	+	+	-	-	-	+	+
TcA	<i>Yersinia pestis</i>	+	+	+	-	-	-	+	+
SppA	<i>Serratia proteamaculans</i>	+	+	+	+	+	-	+	+
SepA	<i>Serratia entomophila</i>	+	+	+	+	+	-	+	+
TcA	<i>Burkholderia pseudomallei</i>	+	+	+	+	+	+	+	+
TcA	<i>Morganella morganii</i>	+	+	+	+	+	+	+	+
TcA	<i>Pseudomonas syringae</i>	+	+	+	+	+	+	+	+

a, Protomer, shell and domain interfaces in TcA. Each interface between protomers, shells and domains was calculated using PISA tools. #1 and #2 refer to values in column 1 and 2, respectively. **b**, Incidence and distribution of TcA domains in Tc toxin complexes from different bacterial organisms. The sequences were aligned against the sequences of full-length TcA and of each TcA domain of *P. luminescens*. The incidence is summarized (+: domain exists; -: domain absent).