

Structural dynamics of the CROPs domain control stability and toxicity of *Paeniclostridium sordellii* lethal toxin

Received: 15 June 2023

Accepted: 4 December 2023

Published online: 19 December 2023

 Check for updatesYao Zhou^{1,2,3,4,5}, Xiechao Zhan^{1,2,3,5} ✉, Jianhua Luo^{1,2,3,4}, Diyin Li^{1,2,3,4}, Ruoyu Zhou^{1,2,3}, Jiahao Zhang^{1,2,3}, Zhenrui Pan^{1,2,3}, Yuanyuan Zhang^{1,2,3}, Tianhui Jia^{1,2,3,4}, Xiaofeng Zhang^{1,2,3}, Yanyan Li^{1,2,3} & Liang Tao^{1,2,3,4} ✉

Paeniclostridium sordellii lethal toxin (TcsL) is a potent exotoxin that causes lethal toxic shock syndrome associated with fulminant bacterial infections. TcsL belongs to the large clostridial toxin (LCT) family. Here, we report that TcsL with varied lengths of combined repetitive oligopeptides (CROPs) deleted show increased autoproteolysis as well as higher cytotoxicity. We next present cryo-EM structures of full-length TcsL, at neutral (pH 7.4) and acidic (pH 5.0) conditions. The TcsL at neutral pH exhibits in the open conformation, which resembles reported TcdB structures. Low pH induces the conformational change of partial TcsL to the closed form. Two intracellular interfaces are observed in the closed conformation, which possibly locks the cysteine protease domain and hinders the binding of the host receptor. Our findings provide insights into the structure and function of TcsL and reveal mechanisms for CROPs-mediated modulation of autoproteolysis and cytotoxicity, which could be common across the LCT family.

P. aeniclostridium sordellii (also known as *Clostridium sordellii*) is a gram-positive anaerobic bacterium commonly found in the environment and causes fatal infections in animals like cattle, sheep, and horses¹. In humans, *P. sordellii* infections are not often but very severe, leading to life-threatening edema, gangrene, hypotension, and sepsis with overall mortality rates approaching ~70%^{2,3}. Approximately 3–4% of women in the United States carry *P. sordellii* in the vaginal and rectal tract; this number is further increased if a recent gynecologic procedure exists⁴. Clinical reports showed that during gynecological procedures, such as childbirth, miscarriage, and abortion, women were particularly at high risk for fulminant *P. sordellii* infections with a 100% mortality^{5–8}.

P. sordellii expresses lethal toxin TcsL (~270 kDa) and hemorrhagic toxin TcsH (~300 kDa). TcsL can cause major damage to endothelial and epithelial cells, particularly in the lungs, and is thought to be the primary cause of the high mortality for acute *P.*

sordellii infections^{9,10}. Both TcsL and TcsH belong to the LCT family, which also includes TcdA and TcdB in *Clostridioides difficile*, TpeL in *Clostridium perfringens*, and Tcn α in *Clostridium novyi*. Among LCTs, TcsL is most closely related to TcdB, with a sequence identity of ~75%¹¹. LCT family members, including TcsL consist of four functional domains: an N-terminal glucosyltransferase domain (GTD), a cysteine protease domain (CPD) that mediates the autocleavage, an intermingled domain responsible for both transmembrane delivery and receptor-binding (DRBD), and a C-terminal combined repetitive oligopeptides (CROPs) domain^{12,13}. Once the toxins enter target cells via receptor-mediated endocytosis, their GTD and CPD translocate from the lumen of the endosomes to the outside upon low pH^{14–16}. In the cytosol, the release of the GTD depends on the protease activity of the CPD triggered by the binding of inositol-hexakisphosphate (InsP6)¹⁷. The activation of CPD results in autocleavage between Leu543 and Gly544 in TcdB and TcsL^{18–20}. The released GTD then

¹Center for Infectious Disease Research, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang 310024, China. ²Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China. ³Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China. ⁴Research Center for Industries of the Future, Westlake University, Hangzhou, Zhejiang 310024, China. ⁵These authors contributed equally: Yao Zhou, Xiechao Zhan. ✉ e-mail: zhanxiechao@westlake.edu.cn; taoliang@westlake.edu.cn

glucosylates small GTPase proteins, leading to actin cytoskeleton disruption and cell death^{21–26}.

Great efforts have been made to understand the toxin action mechanisms of the LCTs through structural biological approaches^{21,27}. Recently, some high-resolution full-length or near-complete structures for TcdA^{28–30} and TcdB^{31–33} were reported. In those structures, the “core” domains (GTD, CPD, and DRBD) show a similar overall architecture, but the CROPs seem to be dynamic and have at least two major conformations: one curves upward around the GTD-CPD head and is referred to as the “open” conformation, the other curves downward alongside the DRBD and is referred to as the “closed” conformation. TcdA is usually displayed in the closed conformation^{29,30,34}. TcdB tends to display the open conformation in either the crystal or cryogenic electron microscopy (cryo-EM) structures^{31–33} while its closed form has been detected using SAX or XL-MS analyses^{31,34}. In previous studies, small regions located at the N-terminal boundary of TcdA and TcdB CROPs were defined as the “hinge”, which may be responsible for the swing of the CROPs domains^{29,31}. To our knowledge, no full-length structures for other LCTs have ever been reported.

The CROPs are unique structural modules found in almost all LCTs except for TpeL. These domains are composed of multiple repeating units: 19- to 24-amino acid short repeats (SRs) interspersed with 29- to 31-amino acid long repeats (LRs)²⁷. The LCT CROPs are variable in sequence and length, ranging from ~350 to ~900 amino acids, while TpeL is a natural CROP-less LCT³⁵. For TcsL, the CROPs domain is composed of four equivalent units (namely I–IV), each unit contains five SRs and one LR. The CROPs were thought to be the receptor-binding domains in LCTs mainly due to the early studies on TcdA. It was shown that the CROPs domain of TcdA was capable of binding the trisaccharide Gal- α 1,3-Gal- β 1,4-GlcNAc and thus considered as a domain responsible for binding carbohydrate moieties on cell surface^{36,37}. In recent years, multiple

protein receptors for LCTs were identified, including low-density lipoprotein receptor-related protein 1 (LRP1) for TpeL³⁸, low-density lipoprotein receptor (LDLR) family proteins for TcdA and Tcna^{39–42}, chondroitin sulfate proteoglycan 4 (CSPG4), poliovirus receptor-like 3 (PVRL3), Frizzled proteins (FZDs), and tissue factor pathway inhibitor (TFPI) for TcdB^{33,43–46}, Semaphorin 6 A and 6B (SEMA6A/6B) for TcsL^{47,48}, and transmembrane serine protease 2 (TMPRSS2) for TcsH⁴⁹. Many of them were structurally validated to bind LCTs independent of the CROPs^{33,48,50,51}, raising the speculation that the CROPs in LCTs are not solely for host recognition but have uncharacterized functions.

Here, we found that the CROPs completely or partially removed TcsL had remarkably increased autoproteolysis as well as cytotoxicity compared to the full-length toxin. To investigate the molecular basis of such CROPs-dependent regulations, we further determined the architecture of full-length TcsL at both neutral and acidic conditions by cryo-EM and captured two major conformations (open and closed). Both functional and structural analysis revealed two regions in the TcsL CROPs that are critical for the modulations: one closed to the DRBD (partly overlapped with the hinge region) and the other at the end of the CROPs domain. Lastly, we demonstrated that an integrated CROPs domain is important to retain the in vivo toxicity of TcsL pre-exposed to InsP6.

Results

The CROPs robustly inhibit the autoproteolytic activity of CPD in TcsL

To investigate the potential roles of the CROPs domain in TcsL, we set out to generate the full-length TcsL and a TcsL truncation (TcsL^{1–1808}) with its entire CROPs deleted (Fig. 1a). These toxin proteins (and other LCTs thereafter) were C-terminal His-tagged and expressed in *Bacillus subtilis*. Unexpectedly, we found that TcsL^{1–1808} but not full-length TcsL

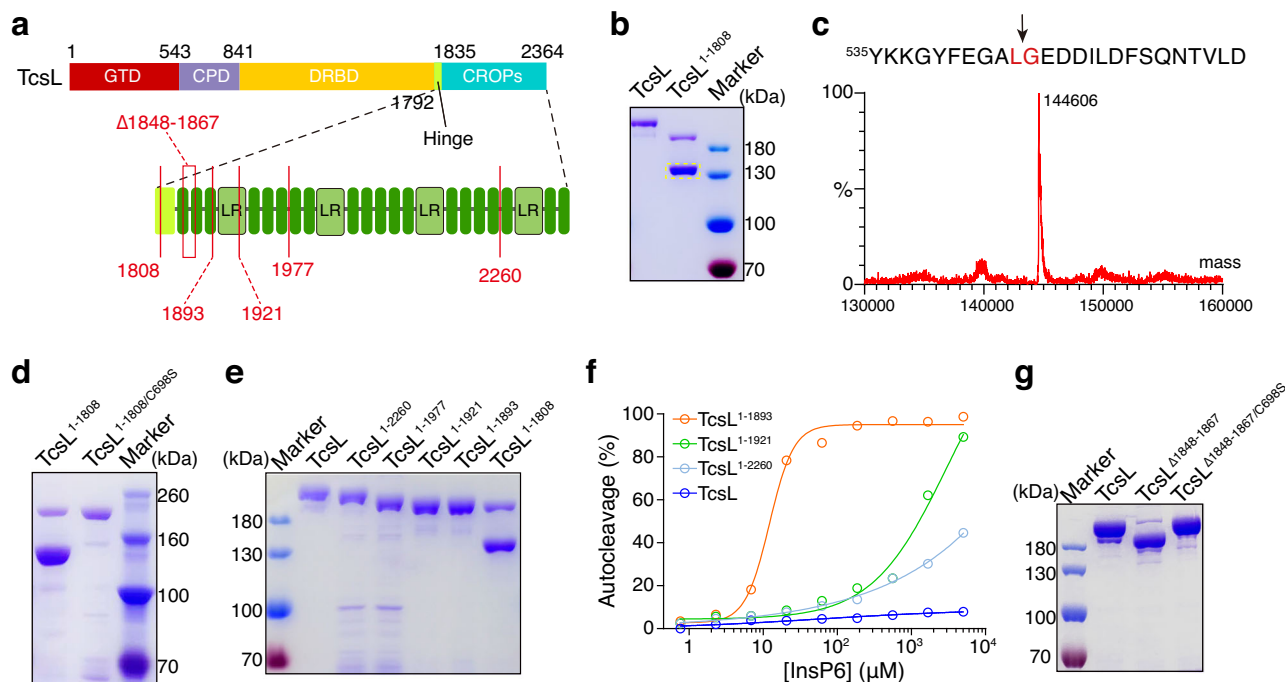


Fig. 1 | The CROPs domain protects TcsL from autoproteolysis. **a** A Schematic drawing of the TcsL truncations designed for this study. **b** Purified full-length TcsL and TcsL^{1–1808} were separated on a Coomassie-stained SDS-PAGE gel. TcsL^{1–1808} was largely fractured, leaving a ~140 kDa fragment (light dotted box). **c** The large fragment in (b) was analyzed by mass spectrometry. The black arrows indicate a breaking point between Leu543 and Gly544. **d** Purified TcsL^{1–1808} and TcsL^{1–1808/C6985} were separated on a Coomassie-stained SDS-PAGE gel. **e** Purified TcsL, TcsL^{1–2260},

TcsL^{1–1977}, TcsL^{1–1921}, TcsL^{1–1893}, and TcsL^{1–1808} were separated on a coomassie-stained SDS-PAGE gel, only TcsL^{1–1808} showed overt autoproteolysis. **f** The autocleavage of TcsL, TcsL^{1–2260}, TcsL^{1–1921}, and TcsL^{1–1893} was induced by the gradient concentrations of InsP6 for 3 hours. The percentage of cleavage was measured and plotted on the chart. **g** Purified full-length TcsL, TcsL ^{Δ 1848–1869}, and TcsL ^{Δ 1848–1869/C6985} were separated on a coomassie-stained SDS-PAGE gel. Source data are provided as a Source Data file.

was largely fractured after the purification, leaving a ~140 kDa C-terminal fragment (Fig. 1b).

Mass spectral (MS) analysis of the ~140 kDa fragment showed that the breakage happened between Leu543 and Gly544 (Fig. 1c), which is the reported cleavage site for CPD^{18,20}. We then mutated Cysteine at position 698 to Serine, which abolishes the proteolytic activity of the CPD^{52,53}, in TcsL¹⁻¹⁸⁰⁸. Unlike TcsL¹⁻¹⁸⁰⁸, TcsL^{1-1808/C698S} was purified as an intact polypeptide (Fig. 1d), demonstrating that the breakage in TcsL¹⁻¹⁸⁰⁸ is due to excessive autocleavage by its own CPD.

The CROPs suppress the autoproteolysis of TcsL in two manners

To determine how much the CROPs were required to prohibit the autoproteolysis of TcsL, we next generated several C-terminally truncated TcsL with partial CROPs remained, including TcsL¹⁻²²⁶⁰, TcsL¹⁻¹⁹⁷⁷, TcsL¹⁻¹⁹²¹, and TcsL¹⁻¹⁸⁹³. None of them showed obvious autoproteolysis, which is similar to the full-length TcsL but not TcsL¹⁻¹⁸⁰⁸ (Fig. 1e, f), indicating that the region between residues 1809 and 1893 tightly controls the autoproteolytic activity of the CPD. In addition, a TcsL mutant with 19 residues (residues 1848 to 1867) deleted also had substantial CPD-mediated autocleavage (Fig. 1g), further supporting that this small region is critical for inhibiting the autoproteolysis in TcsL.

We then quantitatively measured the rates of autoproteolysis for TcsL¹⁻²²⁶⁰, TcsL¹⁻¹⁹²¹, and TcsL¹⁻¹⁸⁹³ in the presence of exogenous InsP6. The full-length TcsL is resistant to InsP6-induced autoproteolysis, which is consistent with a previous report⁵⁴. Although these truncations are generally stable upon production, they showed varied autoproteolysis rates when exposed to gradient concentrations of InsP6: TcsL derivatives with shorter CROPs remaining could be induced by InsP6 more easily (Fig. 1f, Supplementary Fig. 1).

The C-terminus of CROPs affects the cytotoxicity of TcsL

To examine the cytotoxicity of these TcsL truncates, we performed the cytopathic cell-rounding assays on HeLa and A549 cells. To our surprise, the C-terminal truncated TcsL, including TcsL¹⁻²²⁶⁰, TcsL¹⁻¹⁹⁷⁷, TcsL¹⁻¹⁹²¹, and TcsL¹⁻¹⁸⁹³, were more potent than full-length TcsL in the tested cells (Fig. 2a). Albeit A549 cells are more susceptible to TcsL due to the differed expression of SEMA6A/6B^{47,48}, a similar pattern was observed in both HeLa and A549 cells: the truncated toxins showed close toxicity to each other and had ~30-fold increased cytotoxicity compared to full-length TcsL (Fig. 2b). Previous studies reported that CROPs-truncated TcdA, TcdB, and TcsH usually had similar or reduced (in varying degrees) cytotoxicity compared to the full-length toxins, due to the weakened binding to target cells^{42,45,49,55,56}.

Since TcsL is sequentially close to TcdB, including the CROPs (Supplementary Fig. 2), we then replaced the last 104 amino acids (residues 2261 to 2364) with the homologous sequence from TcdB (more precisely, TcdB1¹¹). This newly built chimeric toxin was named TcsL^{2261B} (Fig. 2c). TcsL^{2261B} also showed similar cytotoxicity to TcsL truncates and was more potent than the full-length TcsL (Fig. 2d). In addition, TcsL^{2261B} showed stronger InsP6-induced autocleavage (Fig. 2e, Supplementary Fig. 3), which resembled the TcsL truncates but not TcsL. These results indicate that the TcsL CROPs unit-IV can specifically suppress the cytotoxicity and InsP6-induced autoproteolysis of TcsL with sequence specificity to a certain extent.

Cryo-EM structure of the full-length TcsL

To characterize the potential molecular mechanisms of CROPs-mediated modulations on autoproteolysis and cytotoxicity of TcsL, we managed to determine the structure of the full-length TcsL at neutral pH using single particle cryo-EM. The two-dimensional class averages of TcsL at neutral pH exhibited clear structural features, and the final reconstruction of the full-length TcsL generated an EM map at 2.9 Å resolution (Supplementary Fig. 4).

The cryo-EM structure of TcsL reveals the four major domains of the toxin as previously described (Fig. 3a). The GTD and CPD domains constitute the central core modules of the toxin, interacting with the DRBD and CROPs domains. The DRBD extended and pointed away from the GTD and CPD. The CROPs domain emerges from the junction and curves upward around the core region like a hook, which forms a typical open conformation (Fig. 3a). The overall architecture of TcsL closely resembled that of previously reported structures of TcdB variants as determined by either crystallization or cryo-EM³¹⁻³³.

Low pH induces conformation change of TcsL

We next looked at the configuration of the full-length TcsL in a lowered pH (pH 5.0) using the single particle cryo-EM. Interestingly, we observed that TcsL at pH 5.0 simultaneously exhibits both open (~71.8%) and closed (~28.2%) conformations (Fig. 3b, Supplementary Fig. 5). The EM maps of the full-length TcsL at pH 5.0 were resolved at 2.9 Å (closed conformation) and 2.5 Å (open conformation) (Supplementary Fig. 6). The open conformation is close to the one obtained at neutral pH with a root-mean-square-deviation (RMSD) of 1.4 Å over 1856 Cα atoms (Fig. 3c, Supplementary Fig. 7a). In the closed conformation, the CROPs extend from the hinge region, curve around the section of the DRBD, and interact with the “tip” of the DRBD (Fig. 3d).

When superimposed the cryo-EM structures of TcsL at acidic pH in the open and closed conformations, with an RMSD of 5.1 Å over 1728 Cα atoms, the core domains, including GTD, CPD, and DRBD, are relatively identical, but a near-180° rotation of the entire CROPs domain were observed (Fig. 4a, Supplementary Fig. 7b). A major allosteric transition upon low pH induction was observed for the hinge region (residues 1792-1834) and the first SR swing for ~60 Å. As a result, the entire CROPs rotated for ~180° and the CROPs unit-IV touched the “tip” part of the DRBD (Fig. 4b, Supplementary Movie 1).

In the closed-form structure of TcsL, the CROPs domain potentially interacts with the DRBD through two intra-molecular interfaces (Fig. 4b, c). Interface I is formed between the Hinge region and the DRBD, while interface II is formed between the CROPs unit-IV and the middle part (residues 1176-1277) of the DRBD (Fig. 4b, c). Thus, low pH induces a conformational switch of TcsL from open to closed state, which is also in part supported by a previous XL-MS analysis for TcdB³¹.

CROPs retain the in vivo toxicity of TcsL upon exposure to InsP6

As an exotoxin, TcsL is naturally secreted by *P. sordellii* into its residing niches and exposed to environmental substances. Soils and feces, where *P. sordellii* normally resides, occasionally contain considerable amounts of InsP6⁵⁷. We pre-exposed TcsL and TcsL¹⁻¹⁸⁰⁸ with and without InsP6, followed by intraperitoneally (IP) injecting into mice. This procedure mimics the exposure of TcsL to extracellular InsP6 before entering the host cells. TcsL can induce toxic shock syndrome with similar vascular permeability and massive edema observed in lungs when intraperitoneally injected into mice^{9,47,48}. Both TcsL and TcsL¹⁻¹⁸⁰⁸ induced strong lung tissue damage at similar levels by evaluating both accumulated thoracic fluid (Fig. 5a, b) and hematoxylin and eosin (H&E) stained histological lung sections (Fig. 5c, d), suggesting that CROP-less TcsL is near-equally potent to TcsL once directly enter the host circulation system. Then we pre-mixed the toxins with InsP6 and then IP injected them into the mice. This serves as a simulation that the tested exotoxins are first exposed to environmental circumstances and then reach the host target. TcsL pre-incubated with or without InsP6 for 30 minutes induced similar thoracic fluid accumulation (Fig. 5a, b) and lung damage in mice (Fig. 5c, d), suggesting that the full-length TcsL is tolerant to environmental InsP6 to a certain extent. In contrast, TcsL¹⁻¹⁸⁰⁸ preincubated with InsP6 for only five minutes showed drastically reduced capability in inducing tissue damage, while TcsL¹⁻¹⁸⁰⁸ exposed to InsP6 for 30 minutes was nearly avirulent in vivo (Fig. 5a-d). These results demonstrated that

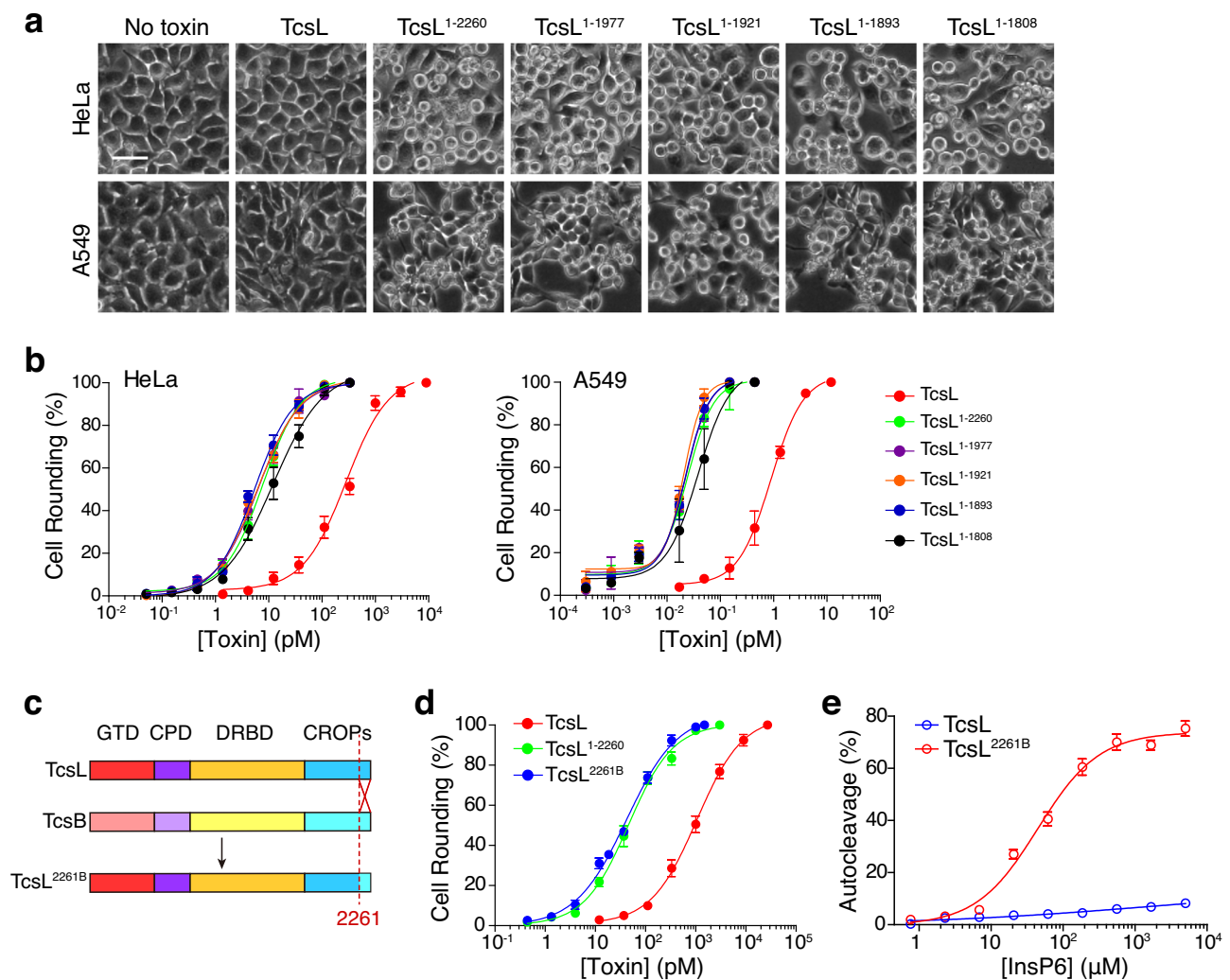


Fig. 2 | The C-terminus of CROPs affects the cytotoxicity of TcsL. **a** HeLa and A549 cells were incubated with TcsL, TcsL¹⁻²²⁶⁰, TcsL¹⁻¹⁹⁷⁷, TcsL¹⁻¹⁹²¹, TcsL¹⁻¹⁸⁹³, or TcsL¹⁻¹⁸⁰⁸ (37 pM for HeLa, 0.05 pM for A549) for 24 hours. Representative images are shown. The scale bar represents 50 μm. **b** The sensitivities of HeLa and A549 cells to TcsL, TcsL¹⁻²²⁶⁰, TcsL¹⁻¹⁹⁷⁷, TcsL¹⁻¹⁹²¹, TcsL¹⁻¹⁸⁹³, TcsL¹⁻¹⁸⁰⁸ were measured using cytopathic cell-rounding experiments. Error bars indicate mean ± SD, $n = 6$ biologically independent samples. **c** The schematic illustration of designed chimeric

toxin TcsL^{2261B} based on TcsL and TcdB. **d** The sensitivities of HeLa cells to TcsL, TcsL¹⁻²²⁶⁰, and TcsL^{2261B} were measured using cytopathic cell-rounding experiments. Error bars indicate mean ± SD, $n = 6$ biologically independent samples. **e** The autocleavage of TcsL^{2261B} was induced by the gradient concentrations of InsP6 for 3 hours. The percentage of cleavage was measured and plotted on the chart. Error bars indicate mean ± SD, $n = 3$ independent experiments. Source data are provided as a Source Data file.

CROPs-dependent autoproteolysis inhibition could be functionally important for maintaining the toxicity of TcsL before host invasion.

CROPs-mediated regulation on autoproteolysis is common for other LCTs

Next, we managed to investigate whether the CROPs-mediated inhibition on autoproteolysis is general across the LCT family. For this purpose, we produced the full-length toxins of other LCTs, including TcdA, TcdB, TcsH, and Tcnα, and their CROP-less versions (TcdA¹⁻¹⁸³⁴, TcdB¹⁻¹⁸⁰⁵, TcsH¹⁻¹⁸³², and Tcnα¹⁻¹⁸⁰¹). These toxins were exposed to 100 μM InsP6 to test their autocleavage efficiency. Similar to TcsL, all tested LCTs in the full-length form showed weaker InsP6-induced autoproteolysis when compared to their CROP-less versions (Fig. 6a, Supplementary Fig. 8).

Finally, we simulated the situations in which TcdA and TcdB were pre-exposed to environmental InsP6 and then intoxicated the mice. TcdA, TcdA¹⁻¹⁸³⁴, TcdB, and TcdB¹⁻¹⁸⁰⁵ were pre-incubated with or without 100 μM InsP6 for 30 minutes and IP injected into the mice. InsP6-treated TcdA and TcdB could effectively kill the mice (Fig. 6b,

c), indicating that these full-length toxins are resistant to extracellular InsP6 with their toxicity largely retained. In contrast, InsP6-treated CROP-less TcdA and TcdB quickly lost their toxicity and became much less potent to mice (Fig. 6b, c), suggesting that the CROPs prevent the autoproteolysis of TcdA and TcdB induced by the extracellular InsP6.

Discussion

As exotoxins, LCTs, including TcsL are secreted from the bacteria into their living niches and required to be structurally stable before reaching their target cells. Notably, InsP6 is sometimes distributed in various environments such as soil⁵⁸ and human gastrointestinal tract⁵⁹ where pathogenic clostridial species including *P. sordellii* and *C. difficile* colonize. To maintain the toxin function, LCTs need to prevent environmental factors (such as InsP6)-induced autoproteolysis and degradation. Our findings indicate that a brief five-minute pre-incubation of TcsL¹⁻¹⁸⁰⁸ with InsP6 significantly reduced the majority of toxin activity in a mouse model. Conversely, the toxicity of full-length TcsL persisted unaltered 30 minutes after exposure to InsP6. Similar

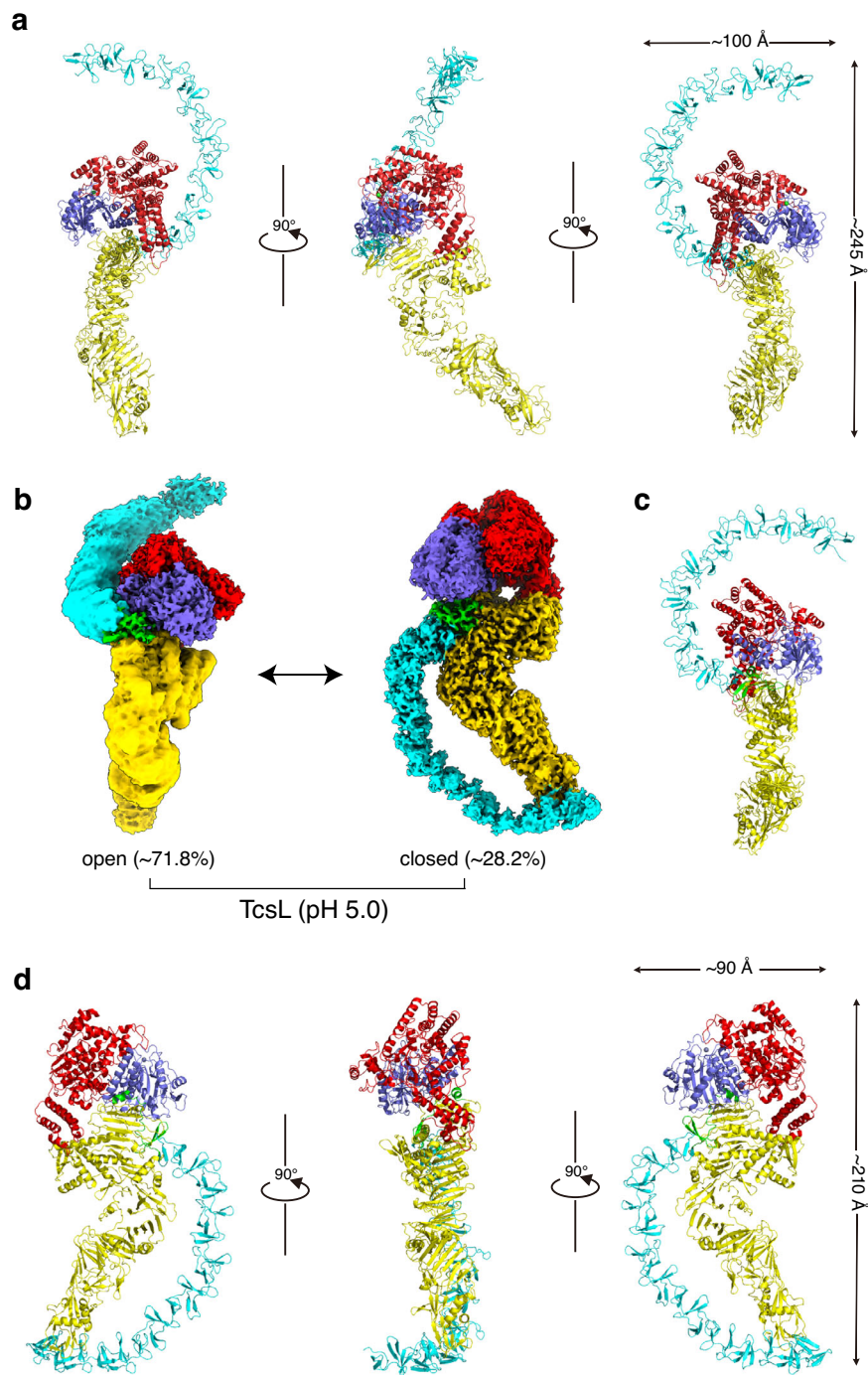


Fig. 3 | Overall structures of the full-length TcsL. **a** Cryo-EM structure of TcsL at neutral pH. **b** The EM-maps of TcsL at an acidic pH of 5.0 in both open (left) and closed (right) conformations. **c** Cryo-EM structure of TcsL in open conformation at

pH 5.0. **d** Cryo-EM structure of TcsL in closed conformation at pH 5.0. The GTD, CPD, DRBD, and CROPs are shown in red, purple, yellow, and cyan, respectively. The zinc atom is shown as a green sphere.

results were obtained for other LCTs like TcdA and TcdB as well. Thus, there is a compelling necessity for LCTs, such as TcsL, to develop an extensive CROPs domain to stabilize the toxin. The only exception is TpeL, which is naturally CROP-less. How TpeL stabilizes itself in the complex environment is unclear. On the other hand, the activity of TpeL CPD was reported to be low^{60,61} and TpeL is known to be least toxic in mice among LCTs²⁷.

It has been reported that the CROPs-truncated TcdA and TcdB underwent moderately increased InsP6-induced autocleavage compared to the full-length toxins^{31,62}. The deletion of amino acids 1769–1787 caused spontaneous autocleavage of subtype 2 TcdB in the

absence of InsP6⁶³. Here we demonstrate that the integrated CROPs domain forcibly protects TcsL from autoproteolysis by at least two mechanisms: the N-terminus of TcsL CROPs strongly restrains the autoproteolysis while longer CROPs further reduce the InsP6-induced autocleavage.

The cryo-EM structure of TcsL at neutral pH is an open conformation and similar to the previously reported structures of TcdB variants^{31–33}. Given the high sequential consensus between TcsL and TcdB, this result is somewhat expected. Intriguingly, cryo-EM structures of the full-length TcsL in both the open and closed conformations were captured in high resolution at pH 5.0.

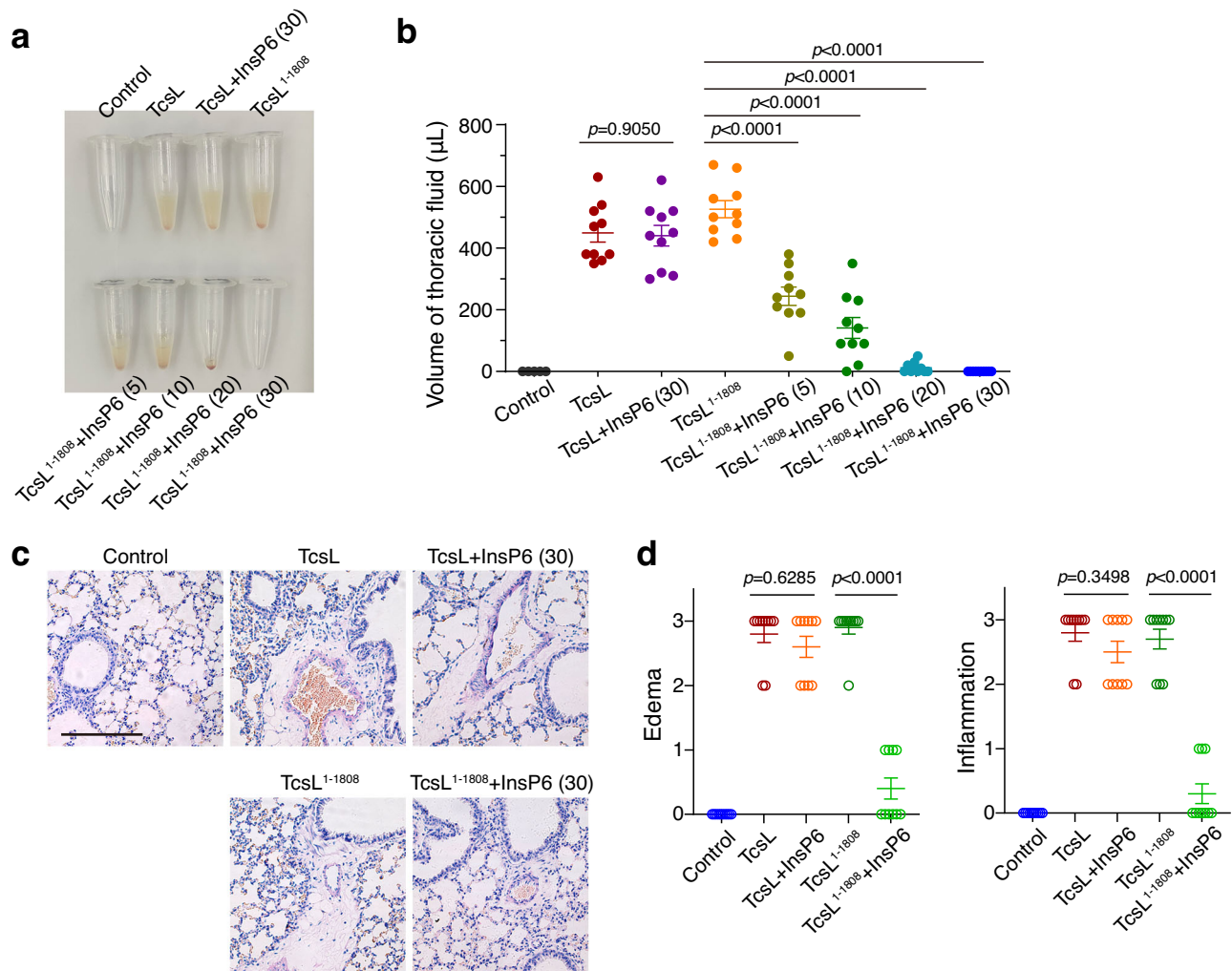


Fig. 5 | CROPs retain toxicity of TcsL on mouse lung tissues. **a** TcsL or TcsL¹⁻¹⁸⁰⁸ was pre-incubated with InsP6 for the indicated time (in parentheses) and then IP-injected into the mice. After four hours, fluid was collected from the mice's thoracic cavity. The representative fluid accumulations are shown. **b** The volume of fluid collected from each mouse was measured and plotted. Error bars indicate mean \pm SEM, $n = 10$ mice per group, two-tailed Mann-Whitney test. **c** H&E staining

showing the lung tissue sections from mice injected with 0.8 $\mu\text{g}/\text{kg}$ TcsL or 0.6 $\mu\text{g}/\text{kg}$ TcsL¹⁻¹⁸⁰⁸ pre-incubated with or without 100 μM InsP6. The scale bar represents 200 μm . **d** Histopathological scores for (c) were assessed based on indicated pathological features. Error bars indicate mean \pm SEM, $n = 10$ mice per group, two-tailed Mann-Whitney test. Source data are provided as a Source Data file.

In the closed conformation, the CROPs domain of TcsL may sterically hinder the binding of the toxin receptor SEMA6A/6B (Supplementary Fig. 9). This could in part explain why TcsL has reduced cytotoxicity than C-terminal truncated TcsL in A549 cells. However, TcsL¹⁻²²⁶⁰ and TcsL^{2261B} are also more potent than the full-length TcsL to HeLa cells, which express little to no SEMA6A/6B⁴⁸, implying TcsL might have additional uncharacterized receptor(s).

Based on these findings, we propose a CROPs-mediated self-stabilizing strategy for TcsL and likely other LCTs in complex extracellular environments. The closed conformation toxin could be tighter and more resistant to autoproteolysis and degradation, as additional interactions between the DRBD and CROPs were observed. The open conformation TcsL is more suitable to attack the host cells as its receptor binding region for SEMA6A/B is fully exposed. After endocytosed by the target cells, the CPD is translocated to the cytosol while the CROPs stay in the lumen of the acidic endosome. In this case, the CROPs and CPD are located on either side of the endosomal membrane; and InsP6-induced and CPD-mediated autoproteolysis would conduct normally (Fig. 7). Taken together, our study renovates the current understanding of the intrinsic functions of the CROPs domain and intramolecular regulatory mechanisms of TcsL. The structural and

functional insights into the CROPs-dependent autoproteolysis inhibition also expose vulnerabilities of TcsL (and possibly other LCTs), which can be utilized to develop potential therapeutic avenues against LCT-related diseases.

Methods

Ethics statement

All procedures were conducted following the guidelines approved by the Institutional Animal Care and Use Committee at Westlake University (IACUC Protocols #19-010-TL and #22-018-TL-5). To minimize the pain and distress, mice were monitored every hour after the toxin injection. Animals with signs such as labored breathing, inability to move after gentle stimulation, or disorientation were euthanized immediately. This method was approved by the IACUC and monitored by a qualified veterinarian.

Materials

HeLa (HI, CRL-1958) and A549 (CCL-185) cells were originally obtained from ATCC. They were tested negative for mycoplasma contamination and authenticated via STR profiling (Shanghai Biowing Biotechnology Co. LTD, Shanghai, China). Cells were cultured in DMEM medium plus

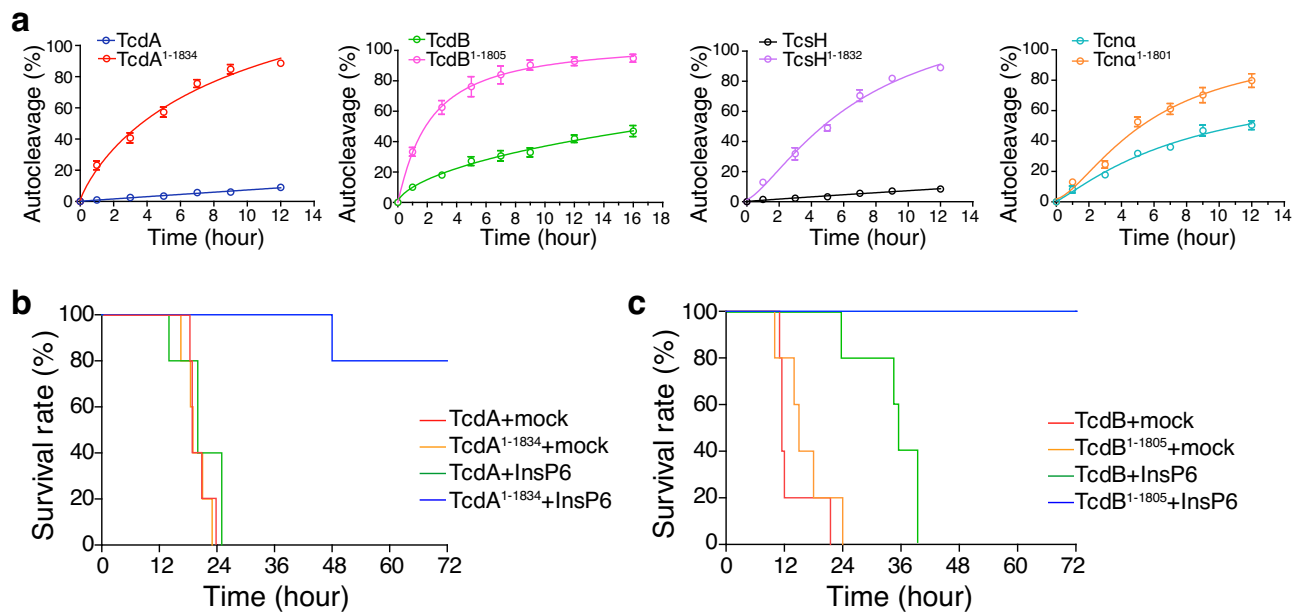


Fig. 6 | CROPs-mediated inhibition on autoproteolysis of TcdA and TcdB. a In the presence of 100 μ M InsP6, the time course cleavage revealed increased autoproteolysis in the CROP-less LCTs (TcdA¹⁻¹⁸³⁴, TcdB¹⁻¹⁸⁰⁵, TcsH¹⁻¹⁸³², and Tcna¹⁻¹⁸⁰¹) compared to the full-length toxins. Error bars indicate mean \pm SD, $n = 3$ independent experiments. **b** 2 μ g/kg TcdA and 1.35 μ g/kg TcdA¹⁻¹⁸³⁴ were pre-incubated with or without 100 μ M InsP6 for 30 minutes and then IP injected into

the mice. The survival of the mice is shown by the Kaplan-Meier curves, $n = 5$ mice per group. **c** 1 μ g/kg TcdB and 0.75 μ g/kg TcdB¹⁻¹⁸⁰⁵ were pre-incubated with or without 100 μ M InsP6 for 30 minutes and then IP injected into the mice. The survival of the mice is shown by the Kaplan-Meier curves, $n = 5$ mice per group. Source data are provided as a Source Data file.

10% fetal bovine serum (FBS) and 1% penicillin/streptomycin in a humidified atmosphere of 95% air and 5% CO₂ at 37°C. InsP6 (#S3793) was purchased from Selleck.

Mice

BALB/c mice (6–8 weeks, male, specific-pathogen-free) were purchased from the Laboratory Animal Resources Center at Westlake University (Hangzhou, China). Mice were housed with food and water without limitation and monitored under the care of full-time staff.

Cloning, expression, and purification of recombinant proteins

DNA fragments encoding TcsL (reference sequence: *P. sordellii* 9048), TcdB1 (reference sequence: *C. difficile* 630), Tcna (reference sequence: *C. novyi* GD211209) were codon-optimized, synthesized by Genscript (Nanjing, China). DNA fragments encoding the full-length and truncated toxins were PCR amplified and cloned into a modified pHT01 vector. Point mutations and deletion were performed using a Quick-change II Site-Directed Mutagenesis Kit (#200523, Agilent Technologies) following the manufacturer's instructions. C-terminal His-tagged recombinant proteins, including TcsL, TcsL¹⁻¹⁸⁰⁸, TcsL¹⁻¹⁸⁹³, TcsL¹⁻¹⁹²¹, TcsL¹⁻¹⁹⁷⁷, TcsL ^{Δ 1848-1867}, TcsL^{1-1808/C698S}, TcsL ^{Δ 1848-1867/C698S}, TcsL^{2261B}, TcdB, TcdB¹⁻¹⁸⁰⁵, Tcna, and Tcna¹⁻¹⁸⁰¹ were expressed and purified from *Bacillus subtilis* strain SL401. *B. subtilis* cells were cultured at 37°C till OD₆₀₀ reached 0.6 and then induced with 1 mM isopropyl- β -D-thiogalactoside at 25°C for 20 hours. All recombinant proteins were purified by Ni-affinity chromatography, followed by size-exclusion chromatography (GE Healthcare). The purified TcsL was further applied to the size-exclusion chromatography at a neutral pH using the buffer (20 mM Tris-Cl pH 7.4, 150 mM NaCl) or at an acidic pH using the buffer (20 mM sodium acetate pH 5.0, 150 mM NaCl) and the peak fractions were applied to the cryo-EM sample preparation.

Cryo-sample preparation and EM data collection

Each four μ L aliquots of the purified TcsL at pH 7.4 and pH 5.0 were applied onto Holey carbon grids (Quantifoil, Au, 300-mesh, R1.2/1.3),

which were glow-discharged in the Plasma Cleaner (HARRICK PLASMA Company). The grids were then blotted for 3.5 seconds and quickly plunged into liquid ethane cooled by liquid nitrogen using Vitrobot Plunge IV (Thermo Fisher) at 8°C and 100% humidity.

The cryo-grids were transferred to a 300-kV Titan Krios electron microscope (Thermo Fisher Scientific) equipped with a Gatan K3 detector and GIF Quantum energy filter (slit width 20 eV). The micrographs with the preset defocus range from -1.5 to -2.0 μ m were collected at a normal magnification of 81,000 \times in super-resolution mode. Each stack with 32 frames was exposed for 2.56 s with a total dose of ~ 50 e-/ \AA^2 using EPU (Thermo Fisher Scientific). The movies are aligned and summed using MotionCor2 with a binning factor of 2, resulting in a pixel size of 1.087 \AA ⁶⁴. Dose weighting was performed concurrently and the defocus value for each image was determined by Gctf⁶⁵ (Supplementary Table S1).

Cryo-EM data processing

The cryo-EM data processing procedures for the TcsL at pH 7.4 were mainly carried out in RELION 3.0⁶⁶ except that is specially mentioned. First, 1,217,203 particles were auto-picked from 1,862 images using Gautomatch (developed by Kai Zhang, <https://www2.mrc-lmb.cam.ac.uk/download/gautomatch-053/>). We then performed the 2D classifications on the total particles and used part of the good particles to generate the initial models. After 2D selection, 637,210 particles were further applied to 3D classifications. Finally, we generated the reconstruction of TcsL at pH 7.4 at an average resolution of 2.9 \AA .

The cryo-EM data processing procedures for the TcsL at pH 5.0 were mainly carried out in cryoSPARC⁶⁷. A total of 12,265 micrographs were collected using the same condition as that of TcsL at neutral pH. About 12 million particles were generated from 11,464 manually selected micrographs. After 2D classifications, about 3.3 million particles were selected and two distinct features were observed. Following the initial model generation, the Heterogeneous refinement was applied to the remaining particles. The two conformations of the

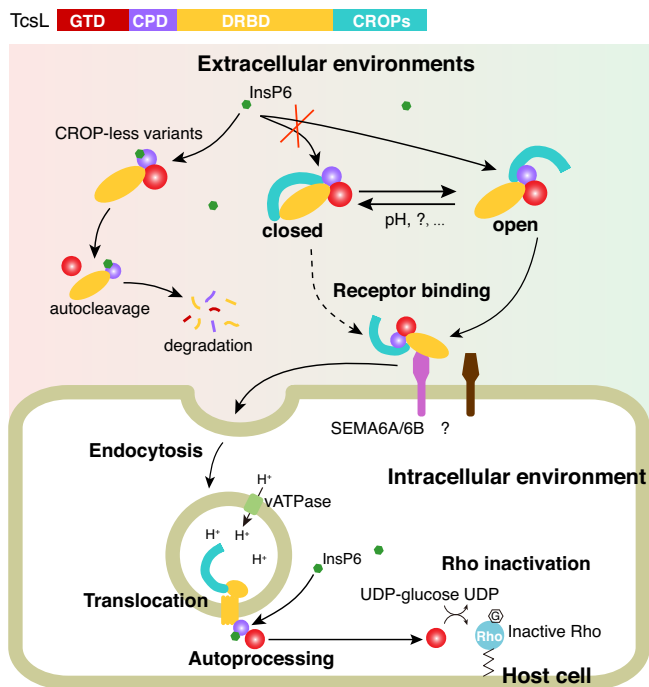


Fig. 7 | Proposed model for CROPs-mediated self-stabilization of TcsL. In the complex extracellular environments, TcsL may exist in either open or closed conformation. TcsL in the open conformation better attacks the host cells, while the closed-form TcsL might be more resistant to autoproteolysis and degradation. The CROP-less toxin variants are sensitive to InsP6 and will easily undergo autoproteolysis and degradation. Once enter the host cell via endocytosis, the toxins undergo conformational changes in the acidic endosomes and translocate the GTD and CPD into the cytosol. Thus, the CROPs would no longer inhibit the InsP6-induced autocleavage.

closed and open TcsL at an acidic pH were finally reconstructed at an average resolution of 2.9 Å and 2.5 Å using 410,049 and 1,045,888 particles, respectively.

The reported resolutions are calculated based on the FSC value of 0.143⁶⁸. The angular distributions of the particles used for the final reconstructions are reasonable. Local resolution variations for the TcsL at pH 7.4 and pH 5.0 were estimated using Resmap⁶⁹ and cryoSPARC, respectively. The EM density maps display clear features that allow detailed structural examination of protein sidechains (Supplementary Table S1).

Model building and refinement

The atomic models were mainly generated by de novo modeling facilitated by AlphaFold⁷⁰. The predicted structure was fitted into the individual EM density map for the TcsL at pH 7.4 or 5.0 using Chimera⁷¹ and manually adjusted using Coot⁷². The final models of the TcsL at pH 7.4 and pH 5.0 were refined using the phenix.real_space_refine program in PHENIX with secondary structure restraints⁷³. The structures were further validated through examination of the Molprobit scores and statistics of the Ramachandran plots. Molprobit scores were calculated as described⁷⁴ (Supplementary Table S1).

Cytopathic cell rounding assay

Cells were seeded at a density of 3×10^5 cells per well in 24-well plates (Corning) and cultivated at 37 °C and 5% CO₂ overnight. Toxins were then added into the culture medium in a serial dilution of 1/3 and incubated at 37 °C for 24 hours. The phase-contrast images of cells were then captured (Olympus IX73, 10× objectives). A zone of 200 × 200 μm was selected randomly, each containing ~20–50 cells. Round-shaped and normal-shaped cells were counted manually. All experiments were

performed in three independent biological replicates. Statistical analysis was performed using OriginPro v8.5 (OriginLab).

Toxin challenge assay in mice

0.8 μg/kg TcsL, 0.6 μg/kg TcsL¹⁻¹⁸⁰⁸, 2 μg/kg TcdA, 1.35 μg/kg TcdA¹⁻¹⁸³⁴, 1 μg/kg TcdB, or 0.75 μg/kg TcdB¹⁻¹⁸⁰⁵ was pre-incubated with or without 100 μM InsP6 in 200–500 μL of saline (0.9 % NaCl, pH 7.4) for 5, 10, 20, or 30 minutes. The mixtures were then IP-injected into the mice. For TcsL, all mice were euthanized after four hours. The fluid present in the thoracic cavity was collected for volume measurement and the lung segments of these mice were dissected for histopathological analysis. For TcdA and TcdB, the survival of the mice was illustrated by the Kaplan-Meier curves (monitored for 72 hours).

H&E staining and histopathological analysis

Lung specimens were fixed in formalin for 12 hours before being dehydrated with an alcohol gradient, cleared with xylene, and then embedded in paraffin. Paraffin blocks were cut into 5 mm thick sections. The sections were stained with H&E. The H&E staining sections were scored blinded by two pathologists based on edema and inflammatory cell infiltration on a scale of 0 to 3 (mild to severe). The average scores were plotted on the charts.

Statistics and reproducibility

Data are presented as mean ± standard deviation (SD) for biochemical experiments and mean ± standard error of the mean (SEM) for pathological experiments. The number of the sample size (n) and statistical hypothesis testing method are described in the legends of the corresponding figures. Statistical analyses of data were performed with GraphPad Prism v9.3 or OriginPro v8.5. All experiments in Figs. 1b, 1d–e, and 1g have been repeated for at least three times with similar results.

Reporting summary

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

Data availability

The atomic coordinates for the TcsL at pH 7.4 and at the pH 5.0 in its open and closed conformations have been deposited in the Protein Data Bank (PDB) with the accession codes 8JB5, 8X2H, and 8X2I, respectively. The EM maps have been deposited in the EMDB with the accession codes EMD-36141, EMD-38010, and EMD-38011 for the TcsL at pH 7.4 and 5.0 in its open and closed conformations, respectively. The other data that support the findings of this study are available in the publicly accessible repository. Source data are provided with the paper. Source data are provided with this paper.

References

- Vidor, C., Awad, M. & Lyras, D. Antibiotic resistance, virulence factors and genetics of *Clostridium sordellii*. *Res. Microbiol.* **166**, 368–374 (2015).
- Aldape, M. J., Bryant, A. E. & Stevens, D. L. *Clostridium sordellii* infection: epidemiology, clinical findings, and current perspectives on diagnosis and treatment. *Clin. Infect. Dis.* **43**, 1436–1446 (2006).
- Aronoff, D. M. & Ballard, J. D. *Clostridium sordellii* toxic shock syndrome. *Lancet Infect. Dis.* **9**, 725–726 (2009).
- Chong, E. et al. Vaginal and Rectal *Clostridium sordellii* and *Clostridium perfringens* Presence Among Women in the United States. *Obstet. Gynecol.* **127**, 360–368 (2016).
- Fischer, M. et al. Fatal toxic shock syndrome associated with *Clostridium sordellii* after medical abortion. *N. Engl. J. Med.* **353**, 2352–2360 (2005).
- Meites, E., Zane, S. & Gould, C. Fatal *Clostridium sordellii* Infections after Medical Abortions. *N. Engl. J. Med.* **363**, 1382–1383 (2010).

7. Agrawal, P. & Garg, R. Fulminant leukemoid reaction due to postpartum *Clostridium sordellii* infection. *J. Glob. Infect. Dis.* **4**, 209–211 (2012).
8. Cohen, A. L. et al. Toxic shock associated with *Clostridium sordellii* and *Clostridium perfringens* after medical and spontaneous abortion. *Obstet. Gynecol.* **110**, 1027–1033 (2007).
9. Geny, B. et al. *Clostridium sordellii* lethal toxin kills mice by inducing a major increase in lung vascular permeability. *Am. J. Pathol.* **170**, 1003–1017 (2007).
10. Carter, G. P. et al. TcsL is an essential virulence factor in *Clostridium sordellii* ATCC 9714. *Infect. Immun.* **79**, 1025–1032 (2011).
11. Shen, E. et al. Subtyping analysis reveals new variants and accelerated evolution of *Clostridioides difficile* toxin B. *Commun. Biol.* **3**, 347 (2020).
12. Jank, T. & Aktories, K. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends Microbiol.* **16**, 222–229 (2008).
13. Chandrasekaran, R. & Lacy, D. B. The role of toxins in *Clostridium difficile* infection. *FEMS Microbiol. Rev.* **41**, 723–750 (2017).
14. Florin, I. & Thelestam, M. Lysosomal involvement in cellular intoxication with *Clostridium difficile* toxin B. *Microb. Pathog.* **1**, 373–385 (1986).
15. Qa'Dan, M., Spyres, L. M. & Ballard, J. D. pH-induced conformational changes in *Clostridium difficile* toxin B. *Infect. Immun.* **68**, 2470–2474 (2000).
16. Pfeifer, G. et al. Cellular uptake of *Clostridium difficile* toxin B. Translocation of the N-terminal catalytic domain into the cytosol of eukaryotic cells. *J. Biol. Chem.* **278**, 44535–44541 (2003).
17. Reineke, J. et al. Autocatalytic cleavage of *Clostridium difficile* toxin B. *Nature* **446**, 415–419 (2007).
18. Rupnik, M., Pabst, S., von Eichel-Streiber, C., Urlaub, H. & Soling, H. D. Characterization of the cleavage site and function of resulting cleavage fragments after limited proteolysis of *Clostridium difficile* toxin B (TcdB) by host cells. *Microbiol. (Read.)* **151**, 199–208 (2005).
19. Giesemann, T., Egerer, M., Jank, T. & Aktories, K. Processing of *Clostridium difficile* toxins. *J. Med. Microbiol.* **57**, 690–696 (2008).
20. Egerer, M. & Satchell, K. J. Inositol hexakisphosphate-induced autoprocessing of large bacterial protein toxins. *PLoS Pathog.* **6**, e1000942 (2010).
21. Aktories, K., Schwan, C. & Jank, T. *Clostridium difficile* Toxin Biology. *Annu. Rev. Microbiol.* **71**, 281–307 (2017).
22. Papatheodorou, P., Barth, H., Minton, N. & Aktories, K. Cellular Uptake and Mode-of-Action of *Clostridium difficile* Toxins. *Adv. Exp. Med. Biol.* **1050**, 77–96 (2018).
23. Genth, H. et al. Haemorrhagic toxin and lethal toxin from *Clostridium sordellii* strain vpi9048: molecular characterization and comparative analysis of substrate specificity of the large clostridial glucosylating toxins. *Cell Microbiol.* **16**, 1706–1721 (2014).
24. Just, I., Selzer, J., Hofmann, F., Green, G. A. & Aktories, K. Inactivation of Ras by *Clostridium sordellii* lethal toxin-catalyzed glucosylation. *J. Biol. Chem.* **271**, 10149–10153 (1996).
25. Popoff, M. R. et al. Ras, Rap, and Rac small GTP-binding proteins are targets for *Clostridium sordellii* lethal toxin glucosylation. *J. Biol. Chem.* **271**, 10217–10224 (1996).
26. Hofmann, F., Rex, G., Aktories, K. & Just, I. The ras-related protein Ral is monoglucosylated by *Clostridium sordellii* lethal toxin. *Biochem. Biophys. Res. Commun.* **227**, 77–81 (1996).
27. Pruitt, R. N. & Lacy, D. B. Toward a structural understanding of *Clostridium difficile* toxins A and B. *Front. Cell. Infect. Mi* **2**, 28 (2012).
28. Chumbler, N. M. et al. Crystal structure of *Clostridium difficile* toxin A. *Nat. Microbiol.* **1**, 15002 (2016).
29. Aminzadeh, A., Larsen, C. E., Boesen, T. & Jorgensen, R. High-resolution structure of native toxin A from *Clostridioides difficile*. *EMBO Rep.* **23**, e53597 (2021).
30. Chen, B. et al. Structure and conformational dynamics of *Clostridioides difficile* toxin A. *Life Sci. Alliance* **5**, e202201383 (2022).
31. Chen, P. et al. Structure of the full-length *Clostridium difficile* toxin B. *Nat. Struct. Mol. Biol.* **26**, 712–719 (2019).
32. Jiang, M. et al. Structural dynamics of receptor recognition and pH-induced dissociation of full-length *Clostridioides difficile* Toxin B. *PLoS Biol.* **20**, e3001589 (2022).
33. Luo, J. et al. TFPI is a colonic crypt receptor for TcdB from hypervirulent clade 2 *C. difficile*. *Cell* **185**, 980–994 (2022).
34. Pruitt, R. N., Chambers, M. G., Ng, K. K., Ohi, M. D. & Lacy, D. B. Structural organization of the functional domains of *Clostridium difficile* toxins A and B. *Proc. Natl Acad. Sci. Usa.* **107**, 13467–13472 (2010).
35. Orrell, K. E. & Melnyk, R. A. Large Clostridial Toxins: Mechanisms and Roles in Disease. *Microbiol. Mol. Biol. Rev.* **85**, e0006421 (2021).
36. Krivan, H. C., Clark, G. F., Smith, D. F. & Wilkins, T. D. Cell surface binding site for *Clostridium difficile* enterotoxin: evidence for a glycoconjugate containing the sequence Gal alpha 1-3Gal beta 1-4GlcNAc. *Infect. Immun.* **53**, 573–581 (1986).
37. Greco, A. et al. Carbohydrate recognition by *Clostridium difficile* toxin A. *Nat. Struct. Mol. Biol.* **13**, 460–461 (2006).
38. Schorch, B. et al. LRP1 is a receptor for *Clostridium perfringens* TpeL toxin indicating a two-receptor model of clostridial glucosylating toxins. *Proc. Natl Acad. Sci. Usa.* **111**, 6431–6436 (2014).
39. Zhou, Y. et al. Sulfated glycosaminoglycans and low-density lipoprotein receptor mediate the cellular entry of *Clostridium novyi* alpha-toxin. *Cell Res.* **31**, 935–938 (2021).
40. Zhou, Y. et al. LDLR, LRP1, and Megalin redundantly participate in the uptake of *Clostridium novyi* alpha-toxin. *Commun. Biol.* **5**, 906 (2022).
41. Schottelndreier, D., Langejürgen, A., Lindner, R. & Genth, H. Low Density Lipoprotein Receptor-Related Protein-1 (LRP1) Is Involved in the Uptake of *Clostridioides difficile* Toxin A and Serves as an Internalizing Receptor. *Front. Cell. Infect. Mi* **10**, 565465 (2020).
42. Tao, L. et al. Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to *Clostridium difficile* toxin A entry into cells. *Nat. Microbiol.* **4**, 1760–1769 (2019).
43. Yuan, P. et al. Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B. *Cell Res.* **25**, 157–168 (2015).
44. LaFrance, M. E. et al. Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity. *Proc. Natl Acad. Sci. Usa.* **112**, 7073–7078 (2015).
45. Tao, L. et al. Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. *Nature* **538**, 350–355 (2016).
46. Tian, S. et al. Identification of TFPI as a receptor reveals recombination-driven receptor switching in *Clostridioides difficile* toxin B variants. *Nat. Commun.* **13**, 6786 (2022).
47. Tian, S. et al. Genome-Wide CRISPR Screen Identifies Semaphorin 6A and 6B as Receptors for *Paenoclostridium sordellii* Toxin TcsL. *Cell Host Microbe* **27**, 782–792 (2020).
48. Lee, H. et al. Recognition of Semaphorin Proteins by *P. sordellii* Lethal Toxin Reveals Principles of Receptor Specificity in Clostridial Toxins. *Cell* **182**, 345–356 (2020).
49. Li, X. et al. *Paenoclostridium sordellii* hemorrhagic toxin targets TMPRSS2 to induce colonic epithelial lesions. *Nat. Commun.* **13**, 4331 (2022).
50. Chen, P. et al. Structural basis for recognition of frizzled proteins by *Clostridium difficile* toxin B. *Science* **360**, 664–669 (2018).
51. Chen, P. et al. Structural basis for CSPG4 as a receptor for TcdB and a therapeutic target in *Clostridioides difficile* infection. *Nat. Commun.* **12**, 3748 (2021).
52. Egerer, M., Giesemann, T., Jank, T., Satchell, K. J. & Aktories, K. Autocatalytic cleavage of *Clostridium difficile* toxins A and B depends on cysteine protease activity. *J. Biol. Chem.* **282**, 25314–25321 (2007).

53. Chumbler, N. M. et al. *Clostridium difficile* Toxin B causes epithelial cell necrosis through an autoprocessing-independent mechanism. *PLoS Pathog.* **8**, e1003072 (2012).
54. Guttenberg, G. et al. Inositol hexakisphosphate-dependent processing of *Clostridium sordellii* lethal toxin and *Clostridium novyi* alpha-toxin. *J. Biol. Chem.* **286**, 14779–14786 (2011).
55. Olling, A. et al. The repetitive oligopeptide sequences modulate cytopathic potency but are not crucial for cellular uptake of *Clostridium difficile* toxin A. *PLoS One* **6**, e17623 (2011).
56. Gupta, P. et al. Functional defects in *Clostridium difficile* TcdB toxin uptake identify CSPG4 receptor-binding determinants. *J. Biol. Chem.* **292**, 17290–17301 (2017).
57. Liu, X., Han, R., Cao, Y., Turner, B. L. & Ma, L. Q. Enhancing Phytate Availability in Soils and Phytate-P Acquisition by Plants: A Review. *Environ. Sci. Technol.* **56**, 9196–9219 (2022).
58. Gerke, J. Phytate (Inositol Hexakisphosphate) in Soil and Phosphate Acquisition from Inositol Phosphates by Higher Plants. A Review. *Plants (Basel)* **4**, 253–266 (2015).
59. Owen, R. W., Weisgerber, U. M., Spiegelhalter, B. & Bartsch, H. Faecal phytic acid and its relation to other putative markers of risk for colorectal cancer. *Gut* **38**, 591–597 (1996).
60. Pruitt, R. N. et al. Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in *Clostridium difficile* toxin A. *J. Biol. Chem.* **284**, 21934–21940 (2009).
61. Guttenberg, G. et al. Molecular characteristics of *Clostridium perfringens* TpeL toxin and consequences of mono-O-GlcNAcylation of Ras in living cells. *J. Biol. Chem.* **287**, 24929–24940 (2012).
62. Olling, A. et al. The combined repetitive oligopeptides of *Clostridium difficile* toxin A counteract premature cleavage of the glucosyl-transferase domain by stabilizing protein conformation. *Toxins (Basel)* **6**, 2162–2176 (2014).
63. Bland, S. J., Larabee, J. L., Shadid, T. M., Lang, M. L. & Ballard, J. D. Deletion of a 19-Amino-Acid Region in *Clostridioides difficile* TcdB2 Results in Spontaneous Autoprocessing and Reduced Cell Binding and Provides a Nontoxic Immunogen for Vaccination. *Infect. Immun.* **87** (2019).
64. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
65. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
66. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
67. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
68. Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
69. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. methods* **11**, 63–65 (2014).
70. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
71. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
72. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
73. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* **74**, 531–544 (2018).
74. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

Acknowledgements

We thank Drs. Min Dong and Peng Chen for the suggestion and discussion. We also appreciate the technical support from the Biomedical Research Core Facility, High-Performance Computing Center, and Laboratory Animal Resources Center at Westlake University. This study was partially supported by the National Key R&D Program of China (Grant no. 2023YFC2308400 to L.T.), the National Natural Science Foundation of China (Grant no. 31970129 to L.T.), the Zhejiang Provincial Natural Science Foundation of China (Grant no. LR20C010001 to L.T. and Grant no. LQ23C010001 to J.L.), Postdoctoral Innovation Talents Support Program (BX20230325 to Y.Z.), and the Westlake Center for Genome Editing (Program no. 21200000A992210 to L.T.). L.T. also acknowledges support from the Westlake Laboratory of Life Sciences and Biomedicine and the Westlake Education Foundation.

Author contributions

L.T. conceived the project and designed the experiments. Y.Zhou, D.L., Z.P., and Y.Zhang performed microbiological, biochemical, and cell biological assays. X.Zhan, R.Z., J.Z., X.Zhang, and Y.L. performed protein purification, cryo-EM structure determination, and structural data analysis. Y.Zhou, J.L., and T.J. performed animal experiments and pathological analyses. X.Zhan and L.T. wrote the manuscript with input from all co-authors.

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/s41467-023-44169-z>.

Correspondence and requests for materials should be addressed to Xiechao Zhan or Liang Tao.

Peer review information *Nature Communications* thanks Hideaki Tsuge, Steven Mileto and the other, anonymous, reviewer(s) 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 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. 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/4.0/>.

© The Author(s) 2023