

De novo design of mini-protein binders broadly neutralizing *Clostridioides difficile* toxin B variants

Received: 27 February 2024

Accepted: 12 September 2024

Published online: 02 October 2024

Xinchen Lv^{1,2,3,4,7}, Yuanyuan Zhang^{3,5,6,7}, Ke Sun^{3,4}, Qi Yang^{3,5},
Jianhua Luo^{1,2,3,5}, Liang Tao^{1,2,3,5}✉ & Peilong Lu^{1,2,3,4}✉

Clostridioides difficile toxin B (TcdB) is the key virulence factor accounting for *C. difficile* infection-associated symptoms. Effectively neutralizing different TcdB variants with a universal solution poses a significant challenge. Here we present the de novo design and characterization of pan-specific mini-protein binders against major TcdB subtypes. Our design successfully binds to the first receptor binding interface (RBI-1) of the varied TcdB subtypes, exhibiting affinities ranging from 20 pM to 10 nM. The cryo-electron microscopy (cryo-EM) structures of the mini protein binder in complex with TcdB1 and TcdB4 are consistent with the computational design models. The engineered and evolved variants of the mini-protein binder and chondroitin sulfate proteoglycan 4 (CSPG4), another natural receptor that binds to the second RBI (RBI-2) of TcdB, better neutralize major TcdB variants both in cells and in vivo, as demonstrated by the colon-loop assay using female mice. Our findings provide valuable starting points for the development of therapeutics targeting *C. difficile* infections (CDI).

Clostridioides difficile (also known as *Clostridium difficile*) is an opportunistic pathogen that mainly infects human colons, leading to the breakdown of the colonic epithelial barrier^{1,2}. This, in turn, results in diarrhea, colitis, and even gastroenteritis-associated death. CDI is responsible for approximately half a million cases (mostly nosocomial and community-acquired) and 15,000 deaths each year in the United States^{3,4}. The pathogenic *C. difficile* secretes one or more of the three exotoxins, including toxin A (TcdA), toxin B (TcdB), and *C. difficile* transferase (CDT), to cause intestinal injuries⁵. Of the three toxins, TcdB is required for the high virulence of pathogenic strains in animal models, and TcdB alone is sufficient to induce a wide range of diseases in humans^{6–9}.

TcdB is composed of multiple domains with a molecular weight of ~270 kDa. TcdB consists of an N-terminal glucosyltransferase domain

(GTD), a cysteine protease domain (CPD), a delivery and receptor-binding domain (DRBD), and a large C-terminal combined repetitive oligopeptides domain (CROPs) (Supplementary Fig. 1). TcdB utilizes the GTD to glucosylate small guanosine triphosphatase (GTPase) proteins, thereby disrupting the cytoskeleton and gradually causing cell death^{10–12}.

Natural variants of TcdB are commonly seen in the clinics^{13,14}. Based on recent studies, TcdB can be divided into at least eight subtypes, while four of them (known as TcdB1–4) are closely associated with human diseases^{15,16}. These TcdB variants recognize varied host receptors through two separate interfaces (Supplementary Fig. 1)^{17,18}. The first receptor-binding interface (RBI-1) is located at the convex edge of the DRBD: TcdB1 and TcdB3 bind Frizzled proteins (FZDs)^{19–22}, while TcdB2 and TcdB4, which are exclusively expressed by

¹Research Center for Industries of the Future, Westlake University, Hangzhou, Zhejiang 310024, China. ²Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang 310024, China. ³Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China. ⁴Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China. ⁵Center for Infectious Disease Research, School of Medicine, Westlake University, Hangzhou, Zhejiang 310024, China. ⁶Center for General Practice Medicine, Department of Infectious Diseases, Zhejiang Provincial People's Hospital, Hangzhou Medical College Affiliated People's Hospital, Hangzhou, Zhejiang 310014, China. ⁷These authors contributed equally: Xinchen Lv, Yuanyuan Zhang. ✉ e-mail: taoliang@westlake.edu.cn; lupeilong@westlake.edu.cn

hypervirulent *C. difficile* strains, recognize tissue factor pathway inhibitor (TFPI) as the host receptor^{18,23}. The second RBI (RBI-2) is a composite binding site formed by the CPD, DRBD, and CROPs domains, which recognizes chondroitin sulfate proteoglycan 4 (CSPG4)^{24,25}. TcdB1 and TcdB2 target CSPG4 via RBI-2 to impair the intestinal sub-epithelial myofibroblast layers during infection. Previous studies have demonstrated that interfering with the binding of CSPG4 can attenuate the toxicity of TcdB1^{19,25}.

Developing neutralizing protein therapeutics as a strategy to prevent the entry of toxins into host cells holds promise for the treatment of CDI, and many attempts have been made^{2,26–32}. One notable example is the monoclonal antibody bezlotoxumab (ZINPLAVA), which received approval from the Food and Drug Administration for the treatment of recurrent CDI in 2016³³. Bezlotoxumab binds to the N-terminal half of the TcdB CROPs domain³⁴ (Supplementary Fig. 1), which possibly induces a conformational change in the RBI-2 and prevents the attachment of CSPG4²⁵. Complete neutralization of TcdB necessitates the blocking of both RBIs. Nanobodies and designed ankyrin repeat proteins (DARPs) have been developed to neutralize TcdB by interfering with the first RBI^{35–39}. However, due to residue variations in the DRBD site between different TcdB subtypes, these protein binders are specific to certain subtypes of TcdB toxins and have a narrow binding spectrum³⁷. There is an urgent need for a broad-spectrum TcdB neutralizer that is universally against various clinical TcdB variants.

Mini-protein binders are small globular proteins consisting of 37–65 amino acids^{40–45}. They have been specifically designed with surface residues that allow for the formation of shape- and chemical-complementary interactions with the structural epitope on the target molecule. The mini-protein scaffolds generally demonstrate high thermostability and exhibit low immunogenicity. Mini-protein binders that exhibit binding affinity ranging from the picomolar to the nanomolar range for a wide range of target proteins have been designed from scratch^{41,46}. Nevertheless, designing binders that can broadly and effectively bind different protein target subtypes presents a substantial challenge.

Here, we present the de novo design and characterization of pan-specific mini-protein binders against major TcdB subtypes. We employed de novo-designed mini-protein scaffolds, which were docked against the TcdB toxins to identify binding modes. The amino acid sequences of these scaffolds were then designed to optimize target binding, folding, and stability. Our designed mini-protein binder 43829 has demonstrated a high affinity for binding the RBI-1 of different TcdB subtypes. To validate the accuracy of our computer-generated models, we obtained cryo-electron microscopy (cryo-EM) structures of 43829 in complex with the TcdB. Remarkably, these experimental structures closely resemble the computational models. Next, we employed directed evolution techniques, and obtained a mini-protein binder variant, H3, with enhanced efficacy through directed evolution. Furthermore, we conducted evolutionary steps to enhance the functionality of CSPG4 and obtained 5N2 that better binds to the RBI-2 of TcdB. This evolved 5N2 binder proved to be highly effective in blocking CSPG4-mediated TcdB entry in cells. To maximize the neutralizing efficacy, we conjugated H3 and 5N2 to each arm of the bispecific IgG1-Fc and generated a powerful neutralizing agent. We demonstrated that our engineered molecule successfully neutralizes the major TcdB subtypes in vivo. These findings provide strategies for using de novo design to develop therapeutics tackling tough infectious diseases such as CDI.

Results

De novo design of broad-spectrum mini-protein binders against the RBI-1 of TcdB

We aimed to design pan-specific mini-protein binders that would target the RBI-1 of clinically dominant TcdBs. According to the previous study, the RBI-1 can be divided into two classes: the class I RBI-1 (found

in TcdB1 and TcdB3) binds FZD proteins while the class II RBI-1 (found in TcdB2 and TcdB4) recognizes TFPI (Supplementary Fig. 1a–c). We suggest that a designed binder targeting both classes of RBIs could potentially neutralize all major TcdB variants associated with human diseases.

The RBI-1s of both TcdB1 and TcdB4 exhibit similar structural organization of secondary structures and are located at nearly the same position on the TcdB DRBD domain (Supplementary Fig. 1c). The RBI-1 consists of a β -sheet (resi 1458–1511) serving as the base. A kinked α -helix (resi 1432–1450) and a short loop (resi 1594–1602) are situated on top of the two edges of this β -sheet, sandwiching the cognate receptors (Supplementary Fig. 1d). However, the residues on the RBI-1 are not conserved among TcdB1–4 (Supplementary Fig. 1a). Out of the 23 residues that interact with the receptors, 9 are not conserved. Importantly, the F1597 (in TcdB1) position is pivotal for promoting target specificity among different TcdB toxins, while the F1597S substitution in TcdB2/4 (cognate positions) prevents FZD from interacting and favors TFPI binding⁴⁷.

Our design process began by designing mini-protein binders for TcdB1 using RifDock⁴⁸. This involved generating the rotamer interaction field (RIF) by docking discrete amino acids against the conserved residues and important receptor-interacting residues but not F1597 on the RBI-1 of the TcdB1. F1597 is being excluded from consideration due to the aim of generating broadly neutralizing binders. Subsequently, 9606 mini-protein scaffolds⁴¹, across five different folds (HHH, HHHH, EHEE, HEEH, HEEHE, where H indicates an α -helix and E represents a β -strand), were docked against the RBI-1 of TcdB1, guided by RIF. The interface design was then performed, and more backbone geometries and interface compositions of the designer binders were sampled using the MotifGraft^{49,50} algorithm to optimize binding. A total of 69,752 mini-protein binders against the RBI-1 of TcdB1 were designed and filtered (Fig. 1).

From these designer binders, we computationally selected binders capable of binding the RBI-1 of TcdB2. TcdB1 was replaced with the aligned TcdB2 structure in each designer complex model, and the resulting binder-TcdB2 complex design model was relaxed using Rosetta. The design models were evaluated based on the score matrix, and 3434 mini-protein binder designs that interacted well with both TcdB1 and TcdB2 in silico were selected. These designs were named pan-specific designs (Fig. 1). In addition to these pan-specific designs, the top 15,289 mini-protein binder designs for TcdB1 evaluated by a scoring matrix were also selected (TcdB1-targeting designs) (Supplementary Fig. 2, Methods).

Characterization of TcdB mini-protein binders

Oligonucleotides encoding the designer mini-protein binders, including the pan-specific designs and TcdB1-targeting designs, were synthesized in an oligo pool. These oligonucleotides were then transformed into yeast cells along with an expression vector for surface display. Two screening approaches were employed: in the first approach, the yeast library was screened for binding to fluorescently tagged TcdB1 and TcdB2 variants alternatively; in the second approach, the yeast library was screened against TcdB1 (Supplementary Fig. 3a, b and Supplementary Fig. 13a). Through deep sequencing, designs that were enriched after fluorescence-activated cell sorting (FACS) for toxin binding were identified from both screen approaches. The top 15 designs enriched in the first approach, along with 3 designs mostly enriched from the second approach, were selected (Supplementary Fig. 3a, b). Among these 18 binders, designs 43866, 38778, 43707, and 43829 are from the pan-specific design library. Additionally, 16 out of the above 18 designs were successfully expressed in *Escherichia coli* and were purified as homogeneity proteins.

We then utilized bio-layer interferometry (BLI) to assess the binding affinities of these designs to TcdB1–4. Remarkably, seven binders demonstrated strong binding to all four TcdB variants

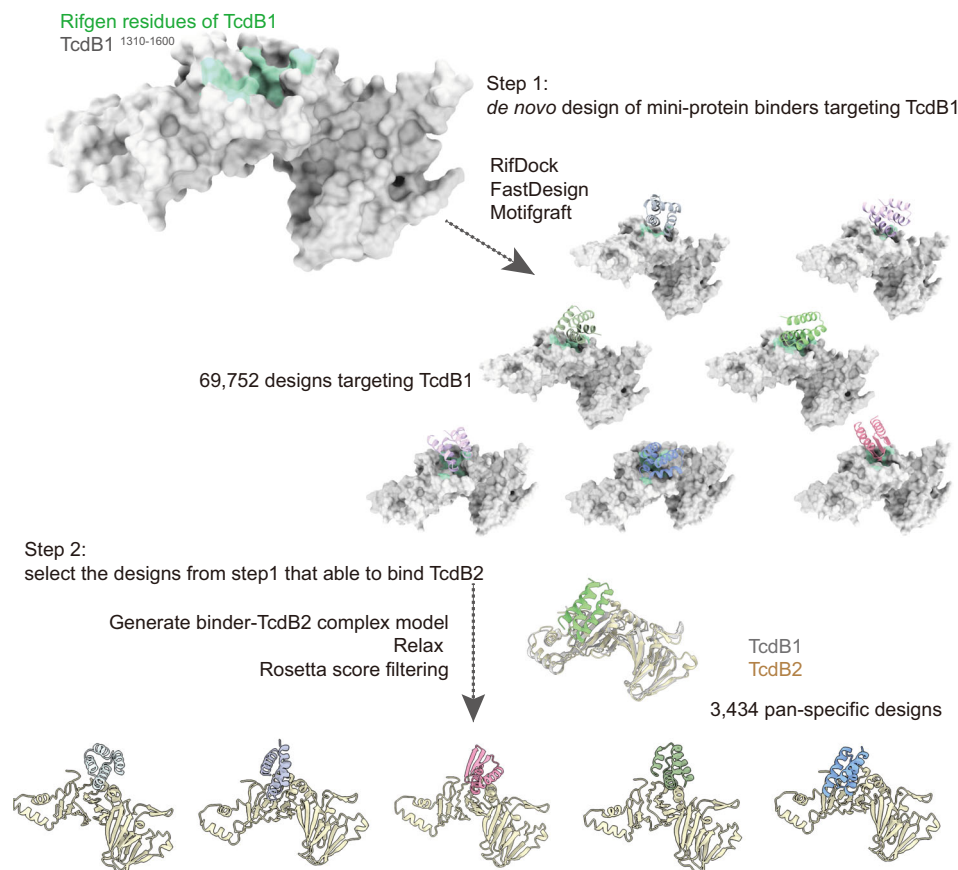


Fig. 1 | Overview of the computational design of broad-spectrum mini-protein binders against TcdB variants. Our computational design approach consists of two steps. In step 1, we performed the de novo design of mini-protein binders targeting TcdB1. Initially, we utilized RIFDock to generate the initial positions of the mini-protein binders relative to the RBI-1 of TcdB1. Interface design was conducted using Rosetta, along with the utilization of the MotifGraft algorithm for sampling additional backbone geometries and interface compositions. In total, 69,752 mini-protein binders against the RBI-1 of TcdB1 were designed and progressed to step 2.

In step 2, we selected binders with the capability to bind the RBI-1 of TcdB2 from the previously designed pool of mini-protein binders targeting TcdB1. For each design model, TcdB2 was superimposed onto the RBI-1 of TcdB1, and the resulting binder-TcdB2 coordinates were considered as the binder-TcdB2 complex design model. These models were further subjected to structure relaxation using Rosetta. Subsequently, the designed models underwent *in silico* evaluation, which led to the selection of 3,434 final mini-protein designs for targeting both TcdB1 and TcdB2 (Methods).

(Supplementary Fig. 3c, d). We further compared the variations in cytopathic toxicity generated by TcdB1 or TcdB4 in the presence of different mini-binders by using classical cell-rounding assays. Each minibinder exhibited a distinct level of cell-protection efficacy in HeLa *CSPG4*^{-/-} cells. Notably, mini-protein binder 43829 demonstrated the most potent performance among all our tested binders in the cell-protection assay (Fig. 2a, b). These data indicate that the designed binder can well compete with the native receptors, either FZDs or TFPI.

The mini-protein binder 43829 exhibited high thermal stability in a thermal-melting experiment and retained more than half of the mean residue ellipticity values. This suggests that the T_m value of the binder is greater than 95 °C (Fig. 2c). 43829 exhibited high affinities to all four TcdB variants ranging from 21.5 pM to 10.1 nM in BLI experiments (Fig. 2d). Specifically, 43829 showed affinities of approximately 21.5 pM and 49.1 pM for TcdB1 and TcdB3, respectively, with a notably slow dissociation rate (Supplementary Fig. 4a, c). Additionally, 43829 displayed affinities of approximately 10.1 nM and 5.1 nM for TcdB2 and TcdB4, respectively, approximately 2 orders of magnitude lower than those observed for class I TcdB toxins (Fig. 2d, Supplementary Fig. 4b, d). These data align with the calculated interface metrics, which indicate that 43829 has a higher affinity for TcdB1 when compared to TcdB2 (Supplementary Fig. 2a). When key interface residues were mutated (Fig. 2d and Supplementary Fig. 4e–h), the binding between 43829 and TcdB1 were greatly reduced or even nearly abolished, which is consistent with the design model.

Cryo-EM structures

To structurally validate the design and elucidate how minibinder 43829 achieves broad neutralization to both classes of TcdB, we solved the structures of TcdB1 and TcdB4 in complex with minibinder 43829 by cryo-EM at overall resolutions of 3.13 and 3.04 Å, respectively. The mini-binders bind to the RBI-1 of the TcdB DRBD domain at a 1:1 ratio, competing with FZD- or TFPI-binding (Fig. 3 and Supplementary Figs. 5 and 6). The minibinder region exhibits relatively lower resolution than the rest of the structure, with calculated map-to-model resolutions of 4.05 Å and 4.45 Å in the TcdB1-43829 map and the TcdB4-43829 map, respectively (Fig. 3c, h). These EM maps resolved the interactions between minibinder 43829 and both classes of TcdB (Fig. 3 and Supplementary Fig. 8). We superimposed our designed models of the 43829/TcdB complex onto the corresponding cryo-EM structures and found that the 43829 structure and the overall binding modes closely matched the design models. When aligning the structures of 43829 alone from the design model and the cryo-EM structure, the backbone C α RMSDs (root mean square deviations) of 43829 are within 1 Å. Moreover, by aligning the TcdB structures in the design model and cryo-EM structure, and subsequently calculating the backbone C α RMSDs of the resulting 43829 poses, the measurements yielded values of 1.7 Å and 2.4 Å for the 43829/TcdB1 and 43829/TcdB4 structures, respectively (Fig. 3b, g).

The 43829 minibinder consists of two α -helices and three β -sheets with a HEEHE topology, where H indicates an α -helix and E represents a

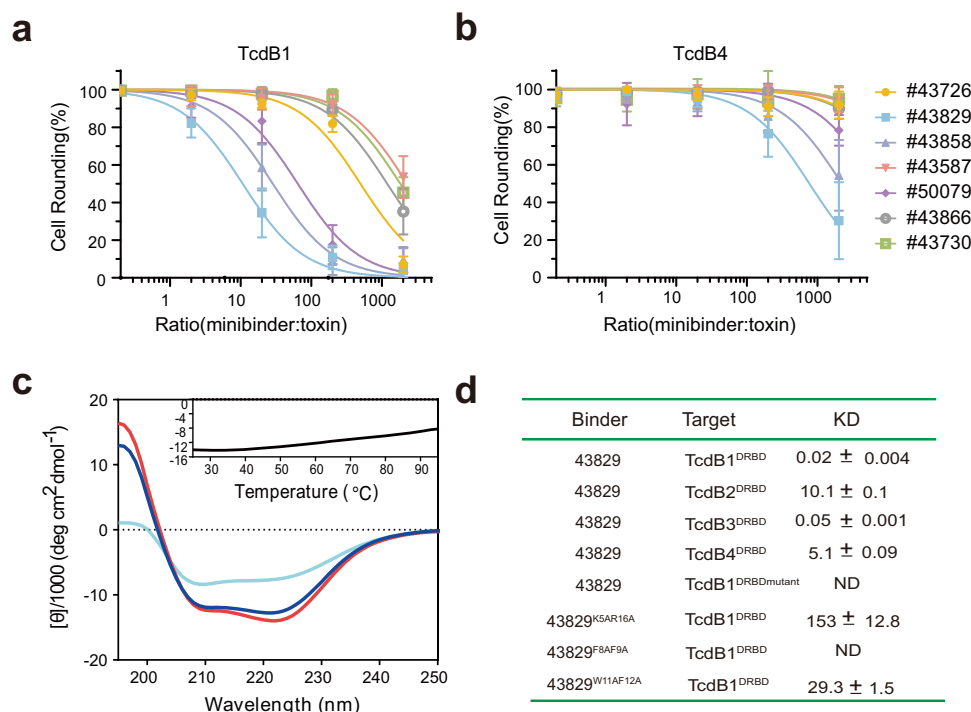


Fig. 2 | Characterization of pan-specific mini-protein binders. a, b Neutralization of TcdB toxin variants by designer mini-protein binders. The cytopathic effects of HeLa *CSPG4*^{-/-} cells induced by TcdB1 or TcdB4 at a concentration of 100 pM were quantified using the cell-rounding assay. Different mini-binders were present at various concentrations. The concentration fold of the minibinder to TcdB is depicted on the X axis. Data from 6 experiments are presented as mean ± SD. **c** Circular dichroism spectra of 43829. The circular dichroism spectra of the 43829

mini-protein binder were measured at different temperatures: 25 °C (red), 95 °C (light blue), and cooling back to 25 °C (blue). The inset shows the thermo-melting curve, with the circular dichroism signals measured at 222 nm from 25 °C to 95 °C. **d** Binding affinity of the de novo mini-protein binder 43829 and its variants for TcdB toxin variants, was assessed by BLI, with the resulting affinity values reported in nanomolar (nM) units. TcdB1^{DRBdmutant} indicates the double mutant V1491E/V1493E of TcdB1. Source data are provided as a Source Data file.

β-strand. The first α-helix comprises the grafted motif shared by 8 binders among the final 18 binders selected in FACS (Supplementary Fig. 7). It perches on the RBI-1 with high shape and charge complementarity. Five hydrophobic residues (F8, F9, W11, F12, and I15) on the first α-helix of 43829 lie in the hydrophobic pockets of RBI-1, thereby making extensive van der Waals interactions with M1437, L1493, Y1509, L1433, and V1497 in TcdB1, as well as with the cognate conserved residues in TcdB4 (Fig. 3e, j). Additionally, two positively charged residues (K5 and R16), located near the two ends of the first α-helix of 43829, contribute to polar interactions with both TcdB1 and TcdB4 (Fig. 3d, i). Specifically, K5 interacts with D1501 (in TcdB1) and N1502 (in TcdB4), while R16 inserts into acidic pockets forming charge-charge interactions with E1468 and E1471 (in TcdB1), as well as E1469 and E1472 (in TcdB4). Interestingly, in the superimposed FZD-TcdB1 and 43829-TcdB1 structures, K5 of 43829 is located close to K127 of FZD, which also interacts with D1501 in TcdB1 (Supplementary Fig. 8a). Similarly, in the superimposed TFPI-TcdB4 and 43829-TcdB4 structures, R16 of 43829 is positioned near R135 of TFPI, which inserts into the acidic pocket formed by E1433, D1467, and E1469 in TcdB4 (Supplementary Fig. 8c). Our 43829 design binds both TcdB1 and TcdB4 by capturing interacting features of both FZD-TcdB1 and TFPI-TcdB4. This achievement is particularly noteworthy because these interactions were designed from scratch. Apart from the first α-helix, E48 on the second α-helix of 43829 also interacts with TcdB by forming salt bridges with K1434 (TcdB1) and K1435 (TcdB4) (Fig. 3d, i).

There are residue variants between TcdB1 and TcdB4 on the 43829 interfaces. L1438, S1495, S1505, and F1597 in TcdB1 were substituted by A1439, I1496, P1506, and S1598 in TcdB4 (Supplementary Fig. 1d). These residue substitutions are situated on the periphery of the 43829 interfaces and can be accommodated when interacting with 43829. Notably, the F1597 (TcdB1) and S1598 (TcdB4) residues are

crucial for receptor-binding specificities for FZD and TFPI. The substitution of F1597 in TcdB1 to S1598 in TcdB4 completely abolishes FZD binding by disrupting the interaction between the buried F1597 residue and the co-receptor palmitoleic acid (PAM) molecule bound in FZD. However, in the 43829-TcdB1 complex, F1597 of TcdB1 adopts a different rotamer compared to that in the FZD-TcdB1 complex (Supplementary Fig. 8b). The sidechain of F1597 is partially solvent accessible and packs with the aromatic rings of F8 and W11 of the 43829 minibinder in the EM structure (Fig. 3e). Substituting F1597 of TcdB1 with S1598 of TcdB4 may slightly weaken the hydrophobic interaction, but it would not abolish TcdB4 binding. This structural analysis is consistent with the relatively lower affinity of 43829 for TcdB4 compared to that for TcdB1 (Fig. 2d).

Directed evolution of 43829

To further improve the binding affinity of 43829 to the class II RBI-1, we employed directed evolution methods. We first generated a site saturation mutagenesis (SSM) library where each residue of 43829 was substituted with the 20 amino acids. The SSM library was sorted against TcdB2¹²⁸⁵⁻¹⁸⁰⁴ and TcdB4¹²⁸⁵⁻¹⁸³⁴, respectively (Supplementary Fig. 13b, c). By comparing the yeast library before and after FACS sorting, we found that residues at the binding interface and protein core were mostly conserved in both TcdB2 and TcdB4 samples (Supplementary Fig. 9a, b). Then, we constructed a combinatorial library by incorporating mutations enriched in FACS sorting and sorted this library for binding to TcdB4 and TcdB1 (Supplementary Fig. 9c, d and Supplementary Fig. 13d). By sorting this library, we identified two closely related sequences, namely H1 and H3, which were shown to bind TcdB4 with a slow disassociation rate (Supplementary Fig. 9e–g). It is evident from the cell-rounding assay that both H1 and H3 were more effective than 43829 in reducing TcdB-induced cytopathic toxicity (Fig. 4a, b). Specifically, H1

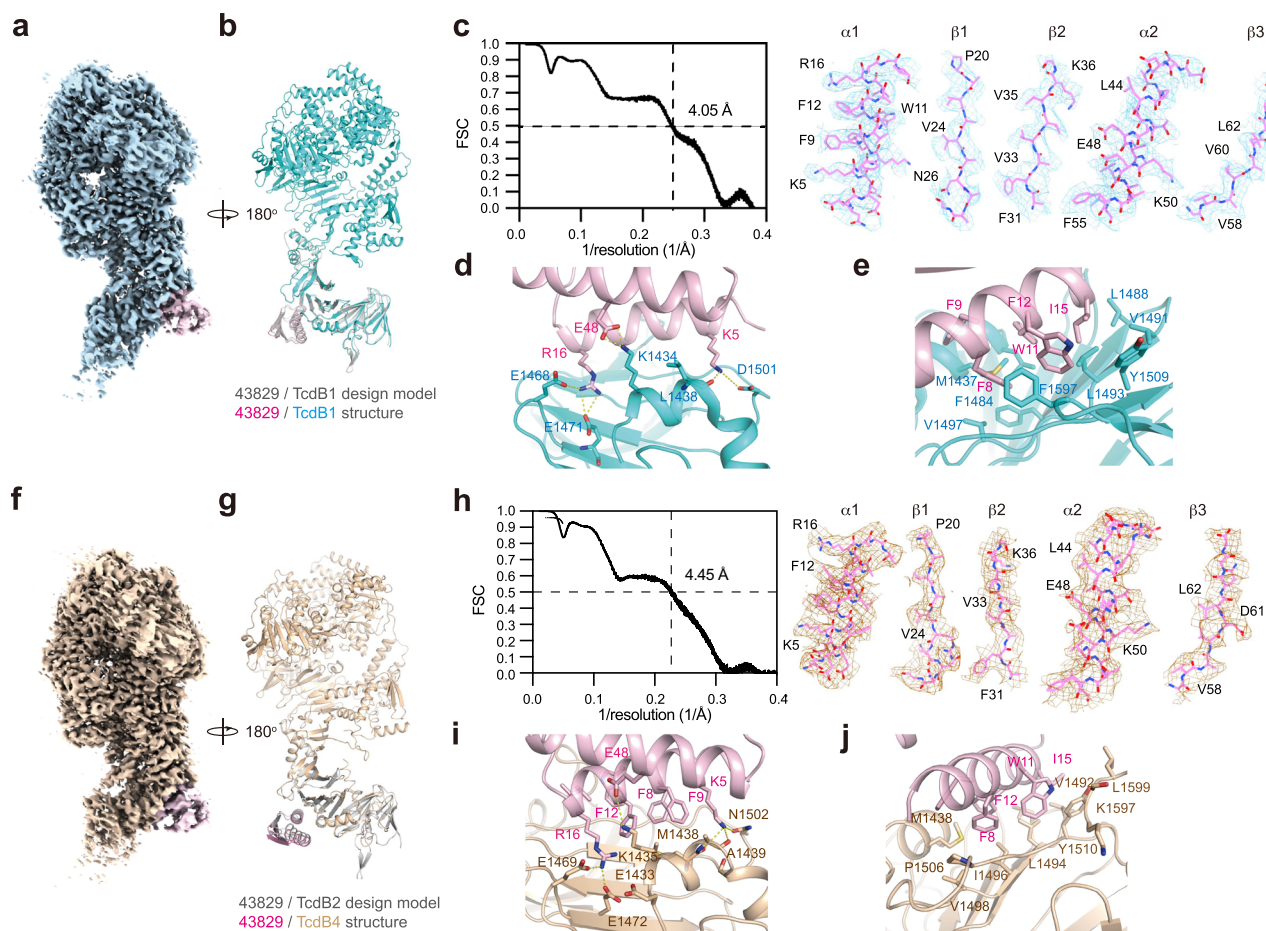


Fig. 3 | Cryo-EM characterization of TcdB1 and TcdB4 in complex with minibinder 43829. **a** A cryo-EM map of the TcdB1-43829 complex, with an overall resolution of 3.13 Å and a contour level set at 0.5. **b** Superimposition of the computational design model (gray) on the experimentally determined cryo-EM structure. The minibinder 43829 is depicted in pink, while TcdB1 is shown in cyan. **c** The cryo-EM density of the 43829 binder in the TcdB1-43829 complex. The FSC curve was generated using Mtriage by comparing the structural model of 43829 to the map of 43829 complexed with TcdB1. The map-to-model resolution for minibinder 43829 alone in the TcdB1-43829 map is 4.05 Å. **d–e** Zoomed views of the cryo-EM structures. Possible hydrogen bonds are highlighted by yellow dashed lines. **f** A

3.04 Å cryo-EM map of TcdB4-43829 complex with a contour level set at 0.5. **g** Superimposition of the computational design model (gray) on the experimentally determined cryo-EM structure, with the minibinder 43829 depicted in pink and TcdB4 in tan. **h** The cryo-EM density of the 43829 binder in the TcdB4-43829 complex. The FSC curve was generated using Mtriage by comparing the structural model of 43829 to the map of TcdB4-43829. The map-to-model resolution for minibinder 43829 alone in the TcdB4-43829 map is 4.45 Å. **i, j** Zoomed views of the cryo-EM structures. Possible hydrogen bonds are highlighted by yellow dashed lines. Interactions between 43829 (pink) and the RBI-1 of TcdB4 (tan).

and H3 showed comparable effectiveness in protecting HeLa *CSPG4*^{−/−} cells against TcdB1 intoxication compared to 43829, while both H1 and H3 showed substantially improved efficacy in reducing TcdB4 intoxication compared to 43829. In the BLI assay, H3 demonstrated a remarkable improvement in binding affinity of over 10-fold towards TcdB2 (−965 pM) and TcdB4 (−278 pM) compared to 43829, with no overt reduction in binding to TcdB1 (−99.3 pM) and TcdB3 (−41.8 pM) (Fig. 4d and Supplementary Fig. 9h–k).

Directed evolution of protein binders against the RBI-2 of TcdB
As TcdB variants recognize varied host receptors through two separate interfaces (Supplementary Fig. 1)^{17,18}, it is imperative to target both RBI-1 and RBI-2 simultaneously to achieve better protection. We next developed a protein binder against the RBI-2 of TcdB, which recognizes CSPG4 as the cellular receptor. The first CSPG4 repeat (residues 410–551), referred to as CSPG4-R1²⁵, has been demonstrated as a natural decoy for TcdB2 but exhibits relatively weaker affinity towards TcdB1. To render CSPG4-R1 the universal binding ability to TcdB, we conducted directed evolution by constructing an error-prone library of R1 displayed on the yeast surface. The library was then sorted against TcdB1, and this mutation-selection cycle was repeated for four rounds (Supplementary

Fig. 14a, b). Through sequencing of the sorted yeast population, we identified three mutants (5N1, 5N2, 5N3) (Supplementary Fig. 10a). The levels of expression of Fc_5N2 and Fc_5N3 secreted by mammalian cells were found to be higher compared to Fc_5N1 (Supplementary Fig. 10e). In BLI assays, all three variants exhibited enhanced binding to TcdB1 when compared to Fc_R1 (Supplementary Fig. 10f). Among the mutations, an interface mutation (E462K) was identified (Supplementary Fig. 10a, b), which potentially forms a salt bridge with E1805 of TcdB1 that is conserved across all TcdB1–4 variants.

Notably, the single point E462K mutation in R1 led to an approximately three-fold improvement in binding affinity towards TcdB1, resulting in a reduction from 30.3 nM to 11.8 nM (Supplementary Fig. 10c, d). In the cell-rounding assay, Fc_5N2 exhibited a more pronounced inhibitory effect on TcdB1-induced cytotoxicity in HeLa *FZD1/2/7*^{−/−} cells compared to other R1 variants (Fig. 5a, b). Besides E462K, 5N2 also carries two additional surface mutations, namely Q517K and D484N (Fig. 5c). Measurement of the binding affinities of Fc_5N2 for TcdB1–4 using BLI assay revealed values ranging from 4 nM to 6.5 nM (Fig. 5d and Supplementary Fig. 10g–j). The Fc_5N2 variant exhibited a significant enhancement of approximately ten-fold in its affinity towards TcdB1 when compared to Fc_R1.

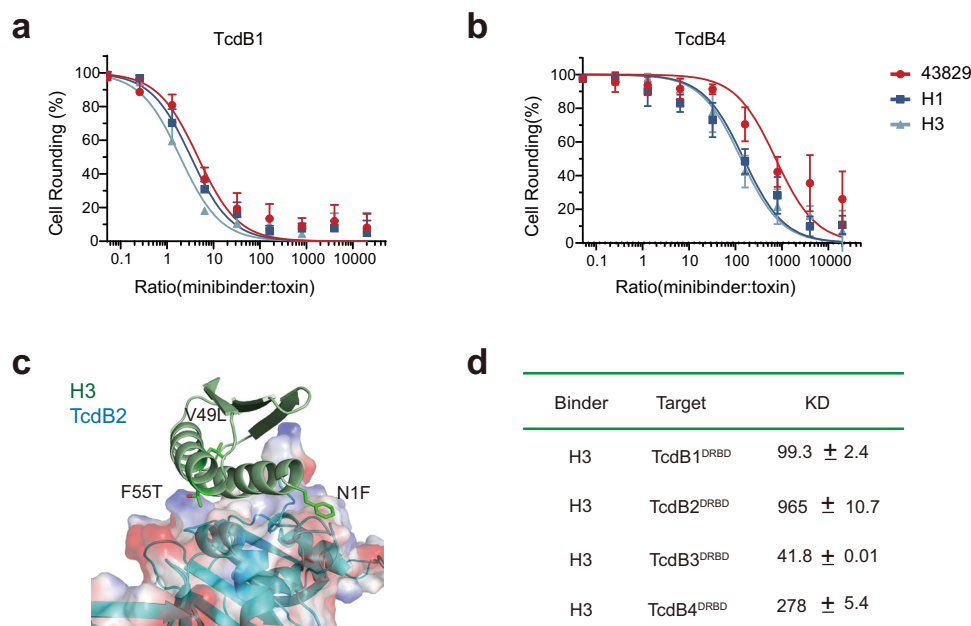


Fig. 4 | Evolved 43829 variant H3 with improved binding affinity against class II TcdB subtypes. a, b Neutralization of TcdB toxin variants by 43829 variants. The cytopathic effects induced by TcdB1 or TcdB4 at a concentration of 100 pM were quantified in HeLa *CSPG4*^{-/-} cells using the cell-rounding assay. Various concentrations of 43829, H1, and H3 were included, and the molar ratio of minibinder to TcdB was plotted on the X axis. Results from six experiments are presented as

mean ± SD. **c** Zoomed views of the H3 (green) and TcdB2 (cyan) complex structure model. The three mutated residues were depicted in sticks. **d** Binding affinity of the evolved minibinder H3 for TcdB toxin variants was assessed using BLI, with the resulting affinity values reported in picomolar (pM) units. Source data are provided as a Source Data file.

The heterodimeric binder offers optimal intestinal protection against TcdB in vivo

It has been shown that simultaneous blocking of both RBIs is key to offering optimal protection against TcdB in vivo^{18,25,36}. Therefore, we generated a heterodimeric binder called Fc_H3_5N2 through a bispecific combination of Fc_H3 and Fc_5N2 (Fig. 6a). When administrated the mice by intraperitoneal (i.p.) injection, Fc_H3_5N2 did not induce obvious variations of serum cytokines such as interleukin-2 (IL-2), interleukin-4 (IL-4), interleukin-6 (IL-6), interleukin-10 (IL-10), interferon-γ (IFN-γ), and tumor necrosis factor (TNF) (Supplementary Fig. 12a, b), suggesting Fc_H3_5N2 has minimal adverse immune effects in mice. We have assessed the stability of the minibinder in the gut lumen environment by incubating it with colon extracts. We found that no overt degradation of Fc_H3_5N2 was detected after incubating with the colon extracts for 8 hours (Supplementary Fig. 12c). Therefore, we conclude that Fc_H3_5N2 is stable in the colon lumen until it is excreted from the gastrointestinal tracts. To determine the intestinal protection of these binders against major clinical TcdB variants in the intestinal epithelium, we injected 2 μg of TcdB variants with or without 100 μg of binders into the ligated mouse colon. This method has the advantage of precisely controlling the toxin/binder dosages and time, allowing us to compare the relative differences in neutralizing efficacy among tested binders. Albeit the detailed manifestations varied, all four major clinical TcdBs (TcdB1-4) caused obvious damage to the mouse colonic epithelium, including epithelial disruption, edema, inflammation, and cryptic injury (Fig. 6). Specifically, TcdB2 induced severe mucosal and submucosal edema, TcdB3 generated a stronger inflammatory response, while TcdB4 caused more damage to the colon crypts (Supplementary Fig. 11). The addition of either Fc_H3 or Fc_5N2 partly alleviated intestinal damages. Impressively, compared to Fc_H3 or Fc_5N2, the same dose of Fc_H3_5N2 provided further enhanced protection against all four dominant TcdB variants to the colonic epithelium. These results suggest that the de novo-designed heterodimeric binder may serve as a general therapeutic avenue for CDI caused by diverse epidemic variants.

Discussion

The pathogenesis and disease progression of CDI mainly originate from the exotoxins produced by *C. difficile*, particularly TcdB. Neutralizing TcdB has proved to be a workable solution in treating CDI since intravenous infusion of anti-TcdB antibody bezlotoxumab lowered the CDI recurrence rate from 26–28% to 15–17% when co-administered with antibiotics⁵¹. However, the remaining CDI recurrence rate is still high (15–17%), likely because bezlotoxumab only blocks RBI-2 (for CSPG4) and fails to neutralize certain TcdB subtypes²⁵. In fact, there is a rising concern for the global spreading of *C. difficile* strains producing variant TcdB, which further causes challenges in CDI diagnosis and therapy. Therefore, developing anti-toxin proteins that can broadly recognize major clinical TcdB variants is in great demand.

Different TcdB subtypes exhibit remarkable primary sequence variations, with a maximum diversity of up to ~15%^{15,16}. Fortunately, recent advances in defining TcdB action mechanisms have clarified two key RBIs for TcdB-host recognition, which form the basis of our binder generation. Different strategies were adopted to generate neutralizing binders against these two RBIs. For RBI-1, which shows high diversity in varied toxin subtypes, we chose to de novo design mini-protein binders as the starting point. These mini-protein binders were later developed into a broad neutralizing binder H3, which binds to major clinical TcdB variants (TcdB1-4) with sub-nanomolar affinities. Cryo-EM structures of mini-protein binder 43829 bound to TcdB1 and TcdB4 not only provide validation for our design models, but also offer insights into the structural mechanism of such broad neutralization. Notably, the de novo design approach renders H3 high thermostability, making it possible to be further developed into temperature-insensitive medicines. For RBI-2, which recognizes CSPG4 with varied affinity in different toxin variants, we developed a binder 5N2 from its native receptor using directed evolution. 5N2 not only exhibits enhanced affinities to TcdB1 and TcdB2, but also binds to TcdB3 and TcdB4 with low nanomolar affinities, making it a universal neutralizer against TcdB RBI-2.

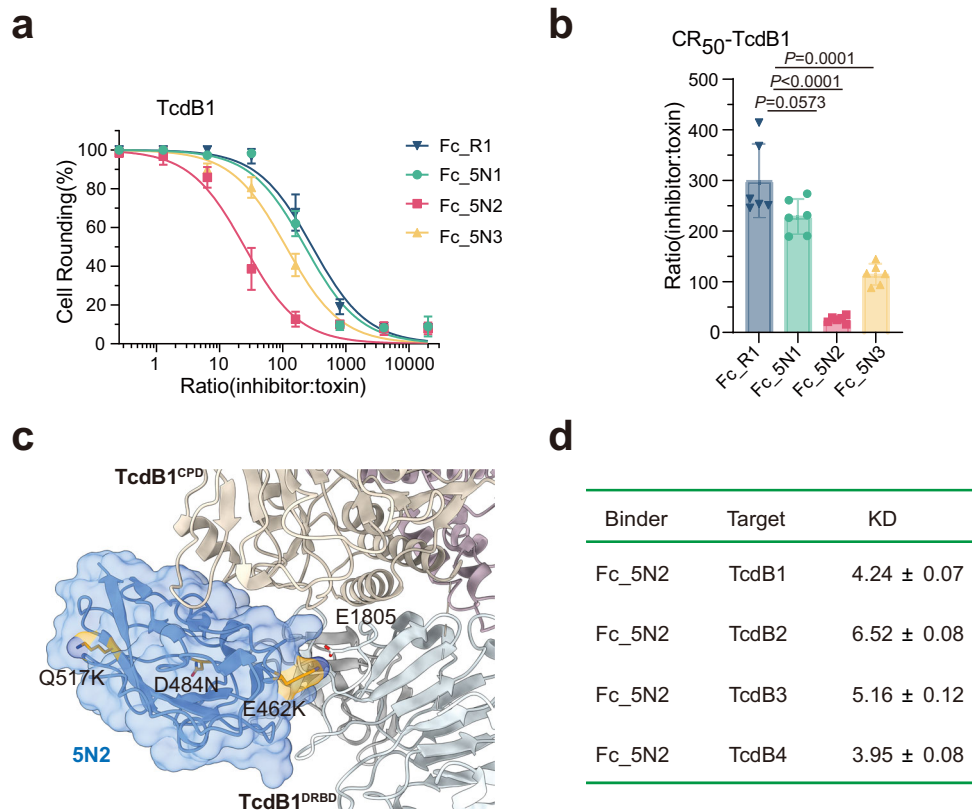


Fig. 5 | Characterization of evolved CSPG4-R1 variants. **a** The cytopathic effects of HeLa *FZD1/2/7*^{-/-} cells induced by TcdB1 (100 pM, ~4 hours) were quantified using the cell-rounding assay in the presence of various concentrations of Fc_5N1, Fc_5N2, Fc_5N3, or Fc_R1. The C-terminus of the Fc fragment of a human immunoglobulin G1 was fused with CSPG4-R1 variants (referred to as Fc-R1 variants). The concentrations of Fc-R1 variants relative to TcdB are shown on the X axis. Error bars ($n = 6$ biologically independent samples) indicate mean \pm SD. **b** The inhibitory effect of different Fc-R1 variants on the toxin was quantified. We illustrate the inhibitor-to-toxin ratio at the point when 50% of the cells exhibit a round phenotype (CR₅₀), as observed in **a**. The associated p-value serves as a statistical indicator of the

significant advantages observed in the evolved variants (Fc_5N1, Fc_5N2, Fc_5N3) compared to the original variant (Fc_R1) in terms of their effectiveness in neutralizing toxins. Error bars ($n = 6$ biologically independent samples) indicate mean \pm SD, two-tailed Student's *t* test. **c** Zoomed views of the 5N2 variant (depicted in blue) and the TcdB1 complex structure model. The three mutated residues are represented by yellow sticks. **d** The binding affinity between the evolved Fc_5N2 and various TcdB toxin variants was assessed by the BLI, with the resulting affinity values reported in nanomolar (nM) units. Source data are provided as a Source Data file.

Because blocking both RBIs has been proven the warrant of neutralizing TcdB in vivo, we next conjugated H3 and 5N2 through the Fc fragment and generated the heterodimer Fc_H3_5N2. Using the mouse colon-loop assays, we demonstrated that Fc_H3_5N2 can effectively protect the mice from all four major TcdB subtypes. Thus, we believe that Fc_H3_5N2 serves as a promising starting point to develop the next-generation TcdB neutralizer for CDI therapeutics.

De novo design of protein binders has emerged as a potent method for developing potential therapeutics to address various health challenges, including influenza virus^{40,41,52,53}, SARS-CoV-2 virus⁴⁶, botulinum neurotoxins⁴⁰, and tumors⁴¹. The computational design allows for the exploration of protein sequence and structure space on a much larger scale compared to conventional screening methods like yeast display or phage display. Consequently, this approach leads to the generation of binders that exhibit shape- and chemical-complementary interactions with specific structural epitopes of the target. These mini-protein binders demonstrate a strong binding affinity for the target, along with excellent thermostability and minimal immunogenicity. However, the design of pan-specific binders that can target a group of structurally similar but variant targets with unique interface residues presents a significant challenge. There is one study that has reported that the designer binders for influenza A H3 hemagglutinin incidentally also bind to influenza A H1 hemagglutinin⁴¹.

In this study, we employed a de novo design methodology accompanied by directed protein evolution and developed mini-

protein binders that potently and broadly neutralize all clinically important TcdB subtypes. We would like to emphasize that it is difficult to obtain a universal binder to the RBI-1 of TcdB using traditional methods, as RBI-1 is highly diverse in primary sequence and has a complex structure. Again, the de novo design methodology shows its unique strength in tackling such complicated situations. Our mini-protein binder design serves as another vivid example of generating immune molecules against notorious infectious agents.

Methods

Ethics statement

All the animal procedures reported herein were approved by the Institutional Animal Care and Use Committee at Westlake University (IACUC Protocol #22-018-2-TL). 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.

Cells and antibodies

HeLa cells (H1, CRL-1958) were originally obtained from ATCC, and Expi293F cells (Expi293F™, A14527) were purchased from Thermo-Fisher Scientific (U.S.). HeLa *CSPG4*^{-/-} and *FZD1/2/7*^{-/-} cells were previously generated laboratory stocks. HEK293F cells were cultivated at

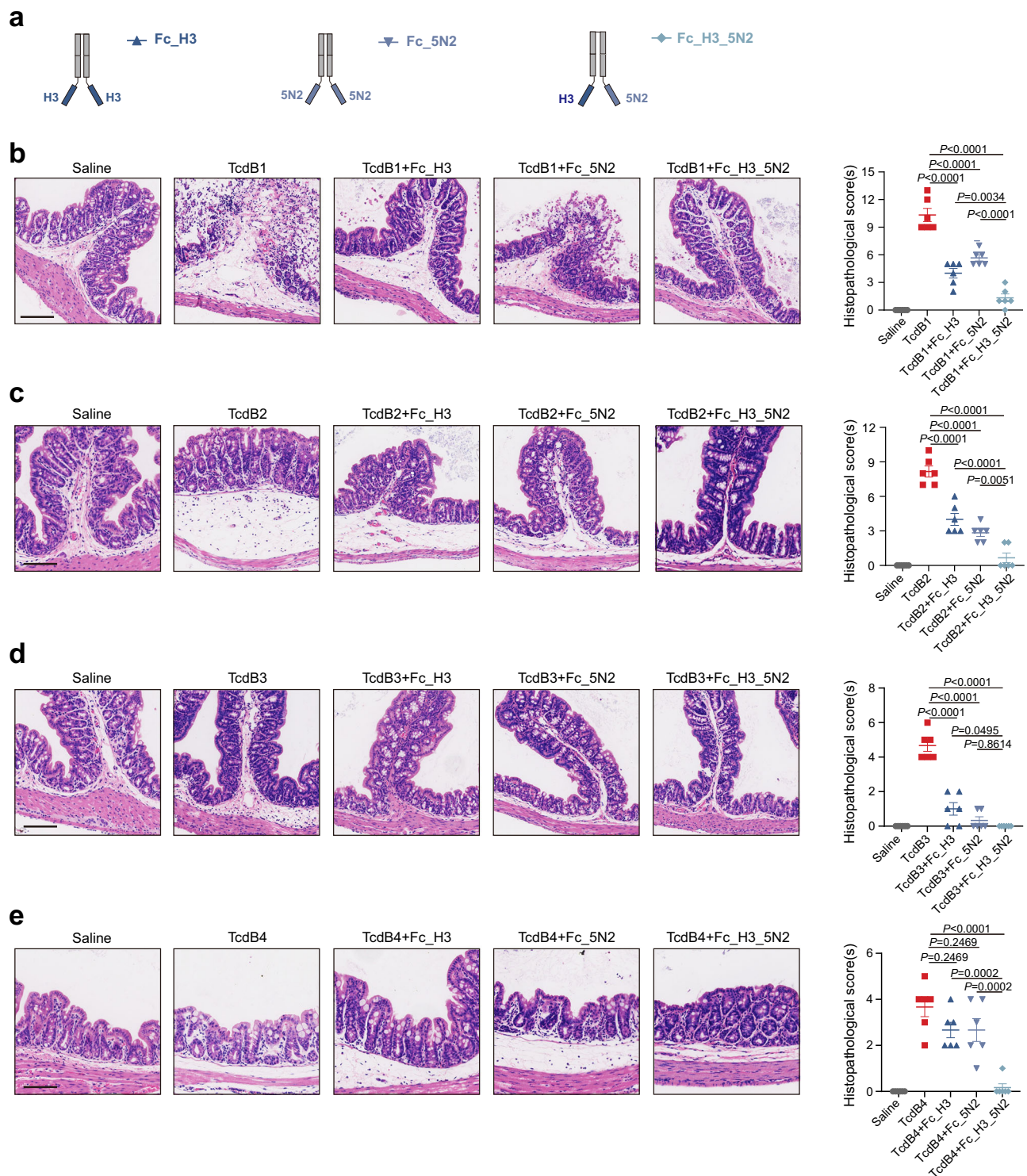


Fig. 6 | In vivo toxin neutralization. **a** Fc_H3 and Fc_5N2 refers to H3 and 5N2 fused to the IgG1-Fc fragment, respectively. Fc_H3_5N2 denotes the fusion of H3 and 5N2 to each arm of the bispecific IgG1-Fc. We used saline as the control. Mouse colonic tissues harvested after colon-loop ligation assays were evaluated for the histopathology induced by TcdB1 (**b**), TcdB2 (**c**), TcdB3 (**d**), and TcdB4 (**e**), both with and without TcdB binders, through H&E staining. Representative images are depicted,

with the scale bar indicating a length of 100 μ m. The overall histological scores were determined by assessing edema, inflammatory cell infiltration, epithelial disruption, and crypt damage. Six mice were analyzed, and error bars indicate the mean \pm SEM. One-way ANOVA using the Tukey test for multiple comparisons. Source data are provided as a Source Data file.

37 $^{\circ}$ C in a humidified environment with 95% air and 5% CO₂ in SMM 293-TIS media (Sino Biological Inc., Beijing, China) supplemented with 1% penicillin-streptomycin. Other cells were cultivated at 37 $^{\circ}$ C in a humidified environment with 95% air and 5% CO₂ in DMEM media supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-

streptomycin. HeLa cells were confirmed to be legitimate using STR profiling after mycoplasma contamination tests came back negative (Shanghai Biowing Biotechnology Co. LTD, Shanghai, China).

The goat monoclonal anti-human IgG-Fc antibody was purchased from a commercial vendor (Sino Biological, SSA001, 1:10,000). The

c-myc FITC antibody was purchased from a commercial vendor (Miltenyi Biotec, 130-116-485, 1:250).

Mice

C57BL/6J mice (female, 6–8 weeks) were acquired from Jiangsu Jicuiyao Kang Laboratory Animal Care Co., Ltd. (Jiangsu, China). Mice were kept under the care of full-time staff members in specific-pathogen-free micro-isolator cages with free access to drinking water and food. All mice were housed at 20–24 °C with 40–60% humidity and were subjected to a 12-hour cycle of light/dark (7 a.m. to 7 p.m.).

Computational design of pan-specific mini-protein binders

The de novo design protocol consists of structure preparation, global search, focused search, and model selection^{41,46}.

In the structure preparation step. The crystal structure (PDB: 6COB) of the FZD protein bound to TcdB1^{DRBD} was refined using the Rosetta FastRelax protocol with coordinate constraints. Subsequently, the TcdB1^{DRBD} subunit (residues 1310–1600) was extracted from the refined structure for further analysis²⁰. The complex cryo-EM structure (PDB: 7VIN) of TFPI bound to TcdB4 was refined using the RosettaCM protocol^{18,54}. To expedite the in silico calculations, the RBI-1 part of TcdB4^{DRBD} (residues 1300–1600) was extracted from the refined structure using Coot⁵⁵. Furthermore, the interface residues were mutated to match those of TcdB2, and the side chains were relaxed using the Rosetta mutation and relax protocol^{56,57}.

During the global search stage. We implemented the PatchDock and RIFDock protocols^{48,58}. In this stage, billions of individual amino acids were independently docked against the RBI-1 region of TcdB1, and the contact residues were highlighted by magenta rectangles in Supplementary Fig. 1. A specific energy cutoff value (−1.5 Rosetta energy unit) was applied to filter out the docking results that met the criteria. These selected results were stored, and the corresponding inverse rotamers were generated. In the PatchDock protocol, we initially mutated approximately 4000 mini-protein scaffolds into poly-Valine. This was done to facilitate the positioning of the docked scaffolds on the target protein. The moderate size of valine's sidechain enables a designable distance between the docked scaffold and the target protein. Using the default parameters of PatchDock, for each scaffold we generated 3000 initial docking conformations. Subsequently, in RIFDock, we integrated the side-chain interface information provided by RIFGen and leveraged the positional data obtained from PatchDock. This integration aimed to optimize the interface of the RIFDock output, maximizing the potential for interactions with the target.

For local search stage. After conducting the RIFDock algorithm, we obtained a total of 103,936 outputs. We utilized a machine learning (ML) predictor model to select 10,000 outputs for FastDesign. Subsequently, we selected 500 designs from the FastDesign results based on computational interface metrics, such as ddG and contact molecular surface. We extracted interface structural motifs with varying lengths, ranging from 5 to 11 residues. Motifs that surpassed a specific energy threshold (ddG < −2) were identified as privileged motifs. These privileged motifs were then grafted by the mini-protein scaffolds in the library (~9000), resulting in a total of ~6,000,000 designs with different scaffolds and different docked positions. These grafted scaffolds were further designed by using the FastDesign protocol to design the interaction between binders and TcdB1.

For Model selection. After FastDesign, designed models were filtered based on the computational metrics, and 69,752 designs were selected. These criteria involved parameters such as a solvent-accessible

surface area (SASA) greater than 1400, absence of vbuns (buried unsatisfied polar atoms), and a CMS score (contact molecular surface) exceeding 250. Following this, we superimposed TcdB2^{DRBD} onto the position of TcdB1^{DRBD} of each designed complex model to make the TcdB2^{DRBD}-binder complex and performed Rosetta relax. Subsequently, the models were filtered based on the computational metrics again, with modified parameters with a SASA score >1200 and a CMS score greater than 250. We identified a total of 3434 mini-binders for further experimental tests as the pan-specific minibinder designs. Similarly, we selected 15,289 designs from the 69,752 designs as the TcdB1-targeting design.

Experimental materials and methods

DNA constructs. Genes encoding the mini-protein binder library were synthesized by Agilent Technologies, Inc. For *E. coli* expression, DNA fragments encoding mini-protein binder were PCR amplified and cloned into a pET28a vector with an N-terminal His-tag. For HEK293F expression, DNA fragments encoding mini-protein binder were PCR amplified and cloned into a pC3.4 vector with an N-terminal HI-signal peptides (MELGLSWIFLLAILKGVQC) and IgG1-Fc or IgG1-Fc^{Hole} tag, with a His-tag fused to their C termini⁵⁹. DNA fragments encoding TcdB1^{1285–1661}, TcdB1^{1–1967}, TcdB2^{1285–1804}, TcdB3^{1286–1805}, and TcdB4^{1285–1834} were PCR amplified and cloned into a pET28a vector with a N-terminal AVI tag and C-terminal His-tag. DNA fragments encoding CSPG4-R1 and related mutants were PCR amplified and cloned into a pCAG or pC3.4 vector, with an N-terminal HI-signal peptide (MELGLSWIFLLAILKGVQC) and Fc or Fc^{Knob} tag, and with His-tag or Flag-tag fused to their C termini.

Recombinant proteins. Recombinant full-length TcdB proteins were expressed in *Bacillus subtilis* SL401 as described previously^{18,20}. Briefly, genes encoding different subtypes of TcdB were cloned into the pHT01 vector, with a 6xHis tag introduced at their C-terminus. The TcdB proteins were then expressed in the *Bacillus subtilis* strain SL401. The bacteria were cultured at 37 °C until the OD₆₀₀ reached 0.6, and expression was induced by adding 1 mM isopropyl-b-D-thiogalactopyranoside at 25 °C for 16 hours. The purification of the His-tagged TcdB was performed using Ni-affinity chromatography and size-exclusion chromatography. All the minibinder proteins were expressed in *E. coli* BL21(DE3) and purified by Ni²⁺-NTA affinity column and size-exclusion chromatography (Superdex 75 Increase 10/300 GL, GE health). Recombinant Fc_CSPG4-R1 (Fc_R1) and Fc_minibinders were expressed in HEK293F cells in secreted form and purified as His- or FLAG-tagged proteins.

DNA library preparation. To ensure unbiased amplification of short DNA fragments during PCR reactions, all protein sequences were extended to 65 amino acids by adding a (GGGS)_n linker at the C-terminus of the designs. The protein sequences were then reverse-translated and optimized using DNAtools2.0, employing the codon frequency table specific to *Saccharomyces cerevisiae*⁶⁰. Oligonucleotide pools encoding the de novo designs and the SSM libraries were procured from Agilent Technologies (USA). Homemade combinatorial libraries were generated using SwiftLib, which optimized the degenerate codons to encompass the desired range of mutational amino acids⁶¹. Subsequently, oligomers were ordered to obtain DNA containing combinatorial mutations via PCR. Error-prone libraries were prepared in-house using GeneMorph II (Agilent, USA). Specifically, 1 ng of DNA was utilized as a template in a 25 µl reaction, resulting in an average of 1–2 mutations per 100 residues after 40 cycles.

All libraries were amplified using KOD polymerase (Takara) and a PCR machine (Bio-Rad, C1000 Touch Thermal Cycler). Specifically, the libraries were amplified in four to five 50 µl reactions, with 22 cycles performed to prevent overamplification. The resulting PCR product

was then subjected to electrophoresis on a DNA agarose gel. Subsequently, the band corresponding to the expected size was excised, and DNA fragments were extracted using Tiangen kits (Tiangen, China). The final PCR product was purified using a Tiangen clean-up kit (Tiangen). For yeast electroporation, 3–4 μg of linearized modified pETcon vector (pETcon3) and 3–4 μg of the inserted fragments were co-transformed into the EBY100 yeast strain following a previously established protocol⁴⁰.

DNA libraries for deep sequencing were prepared using the same PCR protocol, with the initial step starting from yeast plasmid prepared from 5×10^7 to 1×10^8 cells using a Tiangen kit (Tiangen). Illumina adapters and 6-bp pool-specific barcodes were incorporated during the second PCR step. Gel extraction was performed to obtain the final DNA product for sequencing. All distinct sorting pools were subjected to Illumina NextSeq sequencing by Novogene (China).

Yeast surface display. *S. cerevisiae* EBY100 strain cultures were grown in C-Trp-Ura media and induced in SGCAA media following the previous protocol⁴⁰. The cells were washed with PBSF (PBS with 1% BSA), then incubated with the biotinylated target protein (e.g., TcdB1^{1285–1661}, TcdB2^{1285–1804}, TcdB4^{1285–1834}) and the c-myc FITC antibody (Miltenyi Biotec). After a wash, the cells were incubated with streptavidin phycoerythrin (ThermoFisher), washed again, and sorted in PBS buffer.

For the original de novo design minibinder library, two approaches were used for sorting. In the first approach, the library was screened for 6 rounds with 1 μM TcdB1, 1 μM TcdB2, 200 nM TcdB1, 100 nM TcdB2, 10 nM TcdB1, and 1 nM TcdB2. In the second approach, the library was screened for five rounds with TcdB1 concentrations of 1 μM , 200 nM, 50 nM, 5 nM, and 1 nM. Each sorting pool was sequenced using Illumina NovaSeq 6000 to identify good de novo design minibinders. The SSM library was screened twice with TcdB2 and TcdB4. There were five rounds of the TcdB2 screening, with concentrations of 500 nM, 50 nM, 5 nM, 1 nM, and 200 pM. There were three rounds of the TcdB4 screening, with concentrations of 20 nM, 4 nM, and 1 nM. The combinatorial libraries were sorted to for binding to TcdB1 and TcdB4, with concentrations of 20 nM TcdB4, 4 nM TcdB4, 1 nM TcdB4, 1 nM TcdB1, 400 pM TcdB4, and 100 pM TcdB4. The sorting pool of the SSM library and the combinatorial libraries was sequenced using the Illumina NovaSeq 6000 to identify beneficial mutations.

For the CSPG4-R1 error-prone libraries, they were sorted using biotinylated TcdB1^{1–1967}. The sorting process began with a concentration of 40 nM TcdB1. Then DNA was extracted to make the second error-prone library, with a gradually decreasing concentration of TcdB1^{1–1967} at 20 nM, 15 nM, and 5 nM. To quickly test the function of these mutations, the final sorting pools from the libraries were plated on C-Trp-Ura plates, and the sequences of clones were determined using Sanger sequencing.

BLI assays. The binding affinities between mini-binders and the recombinant TcdB1^{1285–1661}, TcdB2^{1285–1804}, TcdB3^{1286–1805}, and TcdB4^{1285–1834} were assessed using the Octet RED96 system (ForteBio) through a BLI assay. The truncated TcdB proteins were genetically modified with AVI-tags and His-tag, followed by biotinylating using BirA, and subsequently purified as His-tagged proteins. All proteins were appropriately diluted in a balancing buffer containing 150 mM NaCl, 20 mM Tris-HCl at pH 8.0, and 0.05% Tween-20. To elaborate, Biotin-labeled TcdB proteins at a concentration of 10 ng/ μL were immobilized onto Streptavidin biosensors (ForteBio), then equilibrated with the buffer. Subsequently, the biosensors were exposed (association) to the mini-binders at varying concentrations for association, followed by washing with the balancing buffer for dissociation.

The binding affinity between Toxin and CSPG4-R1 and CSPG4-R1 variants, such as Fc_R1, Fc_R1^{E462K}, and Fc_5N2 were measured by the Octet RED96 system (ForteBio) with AHC biosensors (anti-human

IgG-Fc capture biosensors, ForteBio). Before use, biosensors were soaked in the assay buffer (20 mM HEPES pH 7.5, 400 mM NaCl, 10 mM CaCl_2 , 0.05% Tween-20) for at least 10 min. Briefly, Fc fused protein (e.g. Fc_R1, Fc_5N2) was immobilized onto biosensors and balanced with the assay buffer. The biosensors were then exposed to different concentrations of full-length TcdB variants, followed by the dissociation in the same assay buffer. All the binding affinities were calculated using the Data Analysis software provided by Forte Bio.

Circular dichroism far ultraviolet. CD measurements were performed using a Chirascan V100 spectropolarimeter equipped with a temperature-controlled multi-cell holder. Wavelength scans were conducted from 280 to 195 nm at 25 °C and 95 °C, followed by a scan at 25 °C after a rapid refolding process that lasted approximately 5 minutes. The temperature melts were monitored by measuring the dichroism signal at 222 nm, with increments of 2 °C per minute after a 30-second equilibration. The protein, at a concentration of 0.15 mg/ml, was prepared in PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , 1.8 mM KH_2PO_4 , pH 7.4) and placed in a cuvette with a path length of 1 mm. Notably, design 43829 exhibited more than half of the mean residue ellipticity values, indicating that its melting temperature (T_m) exceeded 95 °C.

Cryo-EM specimen preparation and data acquisition. 4 μL of the purified TcdB1-minibinder complex at a concentration of 20 mg/mL or the TcdB4-minibinder complex at a concentration of 6 mg/mL were applied onto glow-discharged holey carbon grids (Quantifoil Au R2.1/3.1, 300 mesh). The grid was blotted with filter paper for 3.5 s in a chamber set with 100% humidity at 4 °C to remove the excess sample and then plunge-frozen in liquid ethane cooled by liquid nitrogen with the VitroBot Mark IV system (ThermoFisher Scientific). The grids were transferred to a 300 kV Titan Krios equipped with Gatan K3 Summit detector and a GIF Quantum energy filter (slit width 20 eV). Micrographs were recorded in a defocus range from -2.0 to $-1.4 \mu\text{m}$. Each stack of 32 frames was exposed for 2.56 s, with an exposing time of 0.08 s per frame. The total dose was $\sim 50 \text{ e}^-/\text{\AA}^2$ for each stack. EPU was used for the fully automated data collection⁶².

Cryo-EM data processing. All the 32 frames in each stack were aligned and summed using MotionCor2, a whole-image motion correction program⁶³. The micrographs of TcdB1-minibinder and TcdB4-minibinder were binned to pixel sizes of 1.087 Å and 1.0773 Å, respectively. Dose-weighted micrographs were utilized for subsequent steps in cryoSPARC (v4.2.1+230427)⁶⁴. After CTF estimation by Patch-CTF in cryoSPARC, initial particles were selected from all micrographs using a template picker created from the map of 7vIn with a pixel size of 1.087 Å. To expedite the initial calculation steps, particles were extracted using a box size of 320 pixels and subsequently cropped to 80 pixels, generating 2D averages. After excluding poor-quality classes, the selected particles were recentered using aligned shifts, extracted using a box size of 320 pixels, and cropped to 160 pixels, followed by the generation of 2D averages. These selected particles were then used for ab-initio reconstruction. After excluding incomplete classes, particles were extracted using a box size of 320 pixels.

For the TcdB1-minibinder dataset 1, a total of 1,277,125 particles were extracted from 2824 micrographs using an 80-pixel box size. After two rounds of 2D classification, ab-initio reconstruction, and non-uniform refinement, 171,823 particles were extracted with a pixel size of 1.087 Å. Subsequently, a 3.26 Å map was generated, clearly showing the CPD domain and GTD domain of the toxin. The DRBD domain of the toxin was used as the target for two rounds of 3D classification, resulting in 71,399 high-quality particles that were used to generate 2D templates for further study. For the TcdB1-minibinder dataset 2, a total of 6,238,406 particles were extracted from 6120 micrographs using an 80-pixel box size following template pick. After two rounds of 2D classification, 1,532,509 particles were

selected using a 160-pixel box size. Subsequently, a further 2D classification was performed, resulting in the extraction of 1,094,456 particles for ab-initio modeling. After hetero-refinement, 3D classification, and non-uniform refinement, a total of 294,171 particles were extracted with a pixel size of 1.087 Å. In conjunction with the 71,399 high-quality particles, these particles were used to generate a 2.93 Å resolution map, revealing clear density of the main-chain of the mini-protein binder. Following this, a local 3D classification was performed around the mini-protein binder, resulting in 204,441 particles that generated a 3.13 Å resolution map. This map exhibited clear density around the two helices of the 43829 binder, rough density around the beta-sheet of 43829, and poor density around the loop region. To better depict the 43829 binder, a local mask was added, and after local refinement, the minibinder exhibited clear main-chain density at the loop region and clear side-chain density at the interface (Supplementary Fig. 5).

For the TcdB4-minibinder dataset 1, a total of 902,003 particles were extracted from 1,321 micrographs using an 80-pixel box size. After two rounds of 2D classification, ab-initio reconstruction, and non-uniform refinement, 231,253 particles were extracted with a pixel size of 1.0773 Å. Subsequently, a 3.22 Å map was generated, clearly depicting the CPD domain and GTD domain of the toxin, that were used to generate 2D templates for further study. For dataset 2 of the TcdB4-minibinder, a total of 6,238,093 particles were extracted from 6,821 micrographs using an 80-pixel box size after template selection. Following two rounds of 2D classification, 2,461,206 particles were chosen with a 160-pixel box size. Subsequently, an additional round of 2D classification was conducted, resulting in the extraction of 1,994,138 particles for ab-initio modeling. After hetero-refinement, 3D classification, and non-uniform refinement, a total of 540,192 particles were extracted with a pixel size of 1.0773 Å. These particles, in combination with the 231,253 high-quality particles from dataset 1, were employed to generate a 2.74 Å resolution map, which displayed distinct density of the main-chain of the mini-protein binder. Following this, a local 3D classification was performed specifically around the mini-protein binder, resulting in 228,588 particles that produced a 3.04 Å resolution map. This map exhibited clear density surrounding the two helices of the 43829 binder. To enhance the resolution of the 43829 binder, a local mask was applied. Subsequent local refinement yielded a minibinder with evident main-chain density at the loop region and discernible side-chain density at the interface (Fig. 3 and Supplementary Fig. 8).

Model building and refinement

For the complex structure of TcdB1-minibinder. The initial structure model of the mini-protein binder was imported from the design model. This was combined with the structure of the TcdB1 DRBD domain (residues 1300–1800, PDB code: 6COB) and the structure of TcdB1 (PDB code: 7ML7), which includes the GTD domain, CPD domain, and parts of the CROPs domain. These models were fitted into the cryo-EM map and combined using UCSF Chimera and Coot, followed by iterative rounds of real-space refinement in PHENIX, cartesian and torsion adjustments using RosettaCM, and further manual adjustments in COOT^{54,55,65}.

For building the model of TcdB4-minibinder. The initial structure model for TcdB4 was imported from the TcdB4-TFPI complex (PDB code: 7VIN), and the initial minibinder structure model was generated by Rosetta in the design phase. The model was fitted into the map by the dock_in_map process in PHENIX⁶⁵. Like the TcdB1-minibinder complex, iterative rounds of adjustments were made through real-space refinement, RosettaCM, and manual operations in COOT^{54,55,65}.

The complex structures of TcdB1-minibinder and TcdB4-minibinder were validated by examining the Molprobit scores and

analyzing the statistics of the Ramachandran plots⁶⁶. Please refer to Supplementary Table 2 for detailed statistics on the 3D reconstructions and model refinement. All structure figures were prepared by using PyMOL (<https://www.pymol.org>) and UCSF ChimeraX⁶⁷.

Cytopathic cell-rounding assay

The cytopathic effect (cell-rounding) of TcdB was analyzed using the gold-standard cell-rounding assay. The cells were incubated with a mixture of toxin and inhibitor in different proportions in vitro, following the addition of a complete medium. The cells were photographed four to six hours after they were incubated. 10×–20× objectives on an Olympus IX73 microscope equipment were used to get the phase-contrast pictures. A zone of 300 × 300 μm was selected randomly, which contains 20–100 cells. Round-shaped and normal-shaped cells were counted manually, and the percentage of round-shaped cells was calculated.

Mouse cytokine assay

Mice were injected intraperitoneally with a certain amount of Fc_H3_5N2 or PBS. Whole blood was collected from the heart 2 hours or 48 hours later. The blood samples were centrifuged at 4 °C for 15 minutes, and the supernatant, which is the serum, was collected. Cytokines in the serum samples were measured using the BD cytometric bead array and Mouse Th1/Th2/Th17 Cytokine Kit (BD Biosciences, CA, USA). According to the manufacturer's instructions, we tested the cytokines, including mouse IL-2, IL-4, IL-6, IL-10, IFN-γ, and TNF. 50 μL of the mixed captured beads, 50 μL of the serum sample or standard dilutions, and 50 μL of phycoerythrin (PE) detection reagent were added consecutively to each assay tube and incubated for 2 hours at room temperature in the dark. The samples were measured on the Beckman Flow Cytometer CytoFLEX LX-5L1 and analyzed using FCAP Array Software (BD Biosciences).

Evaluation of the stability of the binder

After the mice were euthanized, the colon segments were excised and homogenized. The homogenate was then centrifuged, and the supernatant was collected. The supernatant was incubated with the binder Fc_H3_5N2 at 37 °C for varying durations. The changes in the abundance of Fc_H3_5N2 within the mixture were then detected using a Western blot assay.

Colon-loop ligation assay

All procedures were conducted following the guidelines approved by the Institutional Animal Care and Use Committee at Westlake University (IACUC Protocol #22-018-2-TL). Female mice aged 6–8 weeks were given an intraperitoneal injection of 1% pentobarbital sodium to induce anesthesia. To find the ascending colon and use silk ligatures to close a ~2 cm loop, a midline laparotomy was done. Using an insulin syringe, 2 μg of toxin with or without 100 μg inhibitor in 100 μL PBS, or 100 μL PBS alone were injected into the sealed colon segment, and the skin incision was then sutured. Mice were housed in a 37 °C thermostatic plate to recuperate. The mice were sedated and allowed to sleep. After six hours, they were humanely euthanized using specialized CO₂ euthanasia equipment specifically designed for small animals, and the ligated colon parts were removed. Following fixing, sectioning, staining, and paraffin embedding, the colon segments were analyzed histopathologically.

H&E staining, and histopathological analysis

Colon specimens were embedded in paraffin after being fixed in formalin for 12–24 hours, dried with an alcohol gradient, and cleaned with xylene. Sections of paraffin blocks with a thickness of 5 μm were cut. Hematoxylin and eosin were used to stain the sections. Two pathologists graded the H&E staining sections blindly on a scale of 0–4

(moderate to severe) based on edema, inflammatory cell infiltration, epithelial enterocyte damage, and crypt damage. The average scores were plotted on the charts.

Statistics and reproducibility

For biochemical experiments, data are presented as mean \pm standard deviation (SD), and for pathological experiments, as mean \pm standard error of the mean (SEM). The sample size (n) and statistical hypothesis testing method are described in the legends for the associated figures. Statistical analyses of data were performed with GraphPad Prism v10 and FCAP Array Software (BD Bioscience). Experiments in Supplementary Fig. 12c have been repeated at least twice 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 cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession number [EMD-39072](#) (TcdB1 in complex with mini-protein binder), [EMD-39073](#) (TcdB4 in complex with mini-protein binder) with local refinement map provided as additional map, and the associated models have been deposited in the RSCB Protein Data Bank under accession number [8Y9B](#) (TcdB1 in complex with mini-protein binder), [8Y9C](#) (TcdB4 in complex with mini-protein binder). The raw data of DNA high-throughput sequencing is submitted to NCBI SRA in the BioProject [PRJNA1154555](#), and the processed data are provided in FigShare [<https://doi.org/10.6084/m9.figshare.26562841>]. Source data are provided with this paper.

Code availability

The data that support this study and Rosetta scripts are available as source data provided in Figshare, under the accession code [<https://doi.org/10.6084/m9.figshare.26562841>]. Rosetta Modeling Suit 2019.47.61047 (<https://www.rosettacommons.org/>) is available to academic and non-commercial users for free. The source code for RIF docking is available at <https://github.com/rifdock/rifdock>. Binding data was collected in Octet RED96 and processed using Octet Analysis software. Flow cytometry data was collected in Beckman Flow Cytometer CytoFLEX LX-5L and BD FACSMelody™ Cell Sorter and processed by Flowjo v9. The figures were generated by Adobe Illustrator. The cryo-EM data are processed by CryoSPARC 4.2.1, Coot-0.9.5 and Phenix-1.19.2.

References

- Schnitzlein M. K., Young V. B. Capturing the environment of the *Clostridioides difficile* infection cycle. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 508–520 (2022).
- Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. *Clostridium difficile* infection. *Nat. Rev. Dis. Prim.* **2**, 16020 (2016).
- Guh, A. Y. et al. Trends in U.S. burden of *Clostridioides difficile* infection and outcomes. *N. Engl. J. Med.* **382**, 1320–1330 (2020).
- Lessa, F. C. et al. Burden of *Clostridium difficile* infection in the United States. *N. Engl. J. Med.* **372**, 825–834 (2015).
- Buddle, J. E. & Fagan, R. P. Pathogenicity and virulence of *Clostridioides difficile*. *Virulence* **14**, 2150452 (2023).
- Kuehne, S. A. et al. The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* **467**, 711–713 (2010).
- Lyras, D. et al. Toxin B is essential for virulence of *Clostridium difficile*. *Nature* **458**, 1176–1179 (2009).
- Carter, G. P. et al. Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during *Clostridium difficile* infections. *MBio* **6**, e00551 (2015).
- Kuehne, S. A. et al. Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. *J. Infect. Dis.* **209**, 83–86 (2014).
- Kordus, S. L., Thomas, A. K. & Lacy, D. B. *Clostridioides difficile* toxins: mechanisms of action and antitoxin therapeutics. *Nat. Rev. Microbiol.* **20**, 285–298 (2022).
- Aktories, K., Schwan, C. & Jank, T. *Clostridium difficile* toxin biology. *Annu. Rev. Microbiol.* **71**, 281–307 (2017).
- 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).
- Knight, D. R. et al. Major genetic discontinuity and novel toxigenic species in *Clostridioides difficile* taxonomy. *Elife* **10**, e64325 (2021).
- Li, Z. et al. Ribotype classification of *Clostridioides difficile* isolates is not predictive of the amino acid sequence diversity of the toxin virulence factors TcdA and TcdB. *Front. Microbiol.* **11**, 1310 (2020).
- Shen, E. et al. Subtyping analysis reveals new variants and accelerated evolution of *Clostridioides difficile* toxin B. *Commun. Biol.* **3**, 347 (2020).
- Mansfield, M. J. et al. Phylogenomics of 8,839 *Clostridioides difficile* genomes reveals recombination-driven evolution and diversification of toxin A and B. *PLoS Pathog.* **16**, e1009181 (2020).
- Aktories, K. Another surprise in receptor binding of *C. difficile* toxins. *Innovation* **3**, 100261 (2022).
- Luo, J. et al. TFPI is a colonic crypt receptor for TcdB from hyper-virulent clade 2 *C. difficile*. *Cell* **185**, 980–994 e915 (2022).
- Tao, L. et al. Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. *Nature* **538**, 350–355 (2016).
- Chen, P. et al. Structural basis for recognition of frizzled proteins by *Clostridium difficile* toxin B. *Science* **360**, 664–669 (2018).
- 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).
- Manion, J. et al. *C. difficile* intoxicates neurons and pericytes to drive neurogenic inflammation. *Nature* **622**, 611–618 (2023).
- 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).
- Pan, Z. et al. Functional analyses of epidemic *Clostridioides difficile* toxin B variants reveal their divergence in utilizing receptors and inducing pathology. *PLoS Pathog.* **17**, e1009197 (2021).
- 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).
- Kroh, H. K. et al. A neutralizing antibody that blocks delivery of the enzymatic cargo of toxin TcdB into host cells. *J. Biol. Chem.* **293**, 941–952 (2018).
- Yang, Z. et al. A novel multivalent, single-domain antibody targeting TcdA and TcdB prevents fulminant *Clostridium difficile* infection in mice. *J. Infect. Dis.* **210**, 964–972 (2014).
- Koulaouzidis, A., Tatham, R., Moschos, J. & Tan, C. W. Successful treatment of *Clostridium difficile* colitis with intravenous immunoglobulin. *J. Gastrointest. Liver Dis.* **17**, 353–355 (2008).
- Babcock, G. J. et al. Human monoclonal antibodies directed against toxins A and B prevent *Clostridium difficile*-induced mortality in hamsters. *Infect. Immun.* **74**, 6339–6347 (2006).
- Aboudola, S. et al. *Clostridium difficile* vaccine and serum immunoglobulin G antibody response to toxin A. *Infect. Immun.* **71**, 1608–1610 (2003).
- Kyne, L., Warny, M., Qamar, A. & Kelly, C. P. Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. *Lancet* **357**, 189–193 (2001).

32. Salcedo, J. et al. Intravenous immunoglobulin therapy for severe *Clostridium difficile* colitis. *Gut* **41**, 366–370 (1997).
33. Bartlett, J. G. Bezlotoxumab - a new agent for *Clostridium difficile* Infection. *N. Engl. J. Med.* **376**, 381–382 (2017).
34. Orth, P. et al. Mechanism of action and epitopes of *Clostridium difficile* toxin B-neutralizing antibody bezlotoxumab revealed by X-ray crystallography. *J. Biol. Chem.* **289**, 18008–18021 (2014).
35. Murase, T. et al. Structural basis for antibody recognition in the receptor-binding domains of toxins A and B from *Clostridium difficile*. *J. Biol. Chem.* **289**, 2331–2343 (2014).
36. Simeon, R. et al. Selection and characterization of ultrahigh potency designed ankyrin repeat protein inhibitors of *C. difficile* toxin B. *PLoS Biol.* **17**, e3000311 (2019).
37. Peng Z. et al. Designed ankyrin repeat protein (DARPin) neutralizers of TcdB from *Clostridium difficile* ribotype 027. *mSphere* **4**, e00596-19 (2019).
38. Li, S. et al. Critical roles of *Clostridium difficile* toxin B enzymatic activities in pathogenesis. *Infect. Immun.* **83**, 502–513 (2015).
39. Kordus, S. L. et al. Nanobodies against *C. difficile* TcdA and TcdB reveal unexpected neutralizing epitopes and provide a toolkit for toxin quantitation in vivo. *PLoS Pathog.* **19**, e1011496 (2023).
40. Chevalier, A. et al. Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).
41. Cao, L. et al. Design of protein-binding proteins from the target structure alone. *Nature* **605**, 551–560 (2022).
42. Listov D., Goverde C. A., Correia B. E., Fleishman S. J. Opportunities and challenges in design and optimization of protein function. *Nat. Rev. Mol. Cell Biol.* **25**, 639–653 (2024).
43. Marchand, A., Van Hall-Beauvais, A. K. & Correia, B. E. Computational design of novel protein-protein interactions - an overview on methodological approaches and applications. *Curr. Opin. Struct. Biol.* **74**, 102370 (2022).
44. Pan, X. & Kortemme, T. Recent advances in de novo protein design: principles, methods, and applications. *J. Biol. Chem.* **296**, 100558 (2021).
45. Baker, E. G., Bartlett, G. J., Porter Goff, K. L. & Woolfson, D. N. Miniprotein design: past, present, and prospects. *Acc. Chem. Res.* **50**, 2085–2092 (2017).
46. Cao, L. et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
47. Henkel D. et al. Receptor binding domains of TcdB from *Clostridioides difficile* for chondroitin sulfate proteoglycan-4 and frizzled proteins are functionally independent and additive. *Toxins (Basel)* **12**, 736 (2020).
48. Dou, J. et al. De novo design of a fluorescence-activating beta-barrel. *Nature* **561**, 485–491 (2018).
49. Silva, D. A., Correia, B. E. & Procko, E. Motif-driven design of protein-protein interfaces. *Methods Mol. Biol.* **1414**, 285–304 (2016).
50. Berger, S. et al. Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *Elife* **5**, e20352 (2016).
51. Wilcox, M. H. et al. Bezlotoxumab for prevention of recurrent *Clostridium difficile* infection. *N. Engl. J. Med.* **376**, 305–317 (2017).
52. Strauch, E. M. et al. Computational design of trimeric influenza-neutralizing proteins targeting the hemagglutinin receptor binding site. *Nat. Biotechnol.* **35**, 667–671 (2017).
53. Fleishman, S. J. et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011).
54. Song, Y. et al. High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013).
55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
56. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E. & Baker, D. Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Sci.* **23**, 47–55 (2014).
57. Nivon, L. G., Moretti, R. & Baker, D. A Pareto-optimal refinement method for protein design scaffolds. *PLoS One* **8**, e59004 (2013).
58. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005).
59. Spreter Von Kreudenstein, T., Lario, P. I. & Dixit, S. B. Protein engineering and the use of molecular modeling and simulation: the case of heterodimeric Fc engineering. *Methods* **65**, 77–94 (2014).
60. Hoover, D. M. Lubkowski J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30**, e43 (2002).
61. Jacobs, T. M., Yumerefendi, H., Kuhlman, B. & Leaver-Fay, A. SwiftLib: rapid degenerate-codon-library optimization through dynamic programming. *Nucleic Acids Res.* **43**, e34 (2015).
62. Thompson, R. F., Iadanza, M. G., Hesketh, E. L., Rawson, S. & Ranson, N. A. Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy. *Nat. Protoc.* **14**, 100–118 (2019).
63. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
64. 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).
65. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* **74**, 531–544 (2018).
66. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
67. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Acknowledgements

We would like to thank the flow cytometry facility, microscopy facility, cryo-EM facility, the Westlake University HPC Center for computation assistance, and the Protein Characterization and Crystallography Facility of Westlake University for help in sample analysis. This work was funded by the Ministry of Science and Technology of the People's Republic of China project 2020YFA0909200 (P.L.) and project 2023YFC2308403 (L.T.), "Pioneer" and "Leading Goose" R&D Program of Zhejiang (Grant no. 2024SSYS0036), Zhejiang Provincial Natural Science Foundation of China Grant No. LR23C050001, National Natural Science Foundation of China project 22137005 (P.L.), project 32430002 (L.T.), and project 32301219 (X.L.), Westlake University-Muyuan Joint Research Institute program WU2024MY001 (L.T.), Westlake Center for Genome Editing project 21200000A992210 (L.T.), and the Research Center for Industries of the Future (RCIF) at Westlake University.

Author contributions

P.L. and L.T. conceived and supervised the project; X.L. and Y.Z. contributed equally to this work; X.L. designed the minibinder, performed directed evolution, solved the cryo-EM structure with the assistance of K.S.. Y.Z. performed the cell-rounding assay and the colon-loop assay with the help of Q.Y. and J.L.. P.L. and X.L. wrote the original draft and all authors participated in manuscript revision.

Competing interests

X.L., Y.Z., L.T., and P.L. are listed as inventors on a patent application submitted by Westlake University to protect the intellectual property rights of the TcdB toxin binders developed in this study. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52582-1>.

Correspondence and requests for materials should be addressed to Liang Tao or Peilong Lu.

Peer review information *Nature Communications* thanks Amedee des Georges, Timothy Jenkins, 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024