

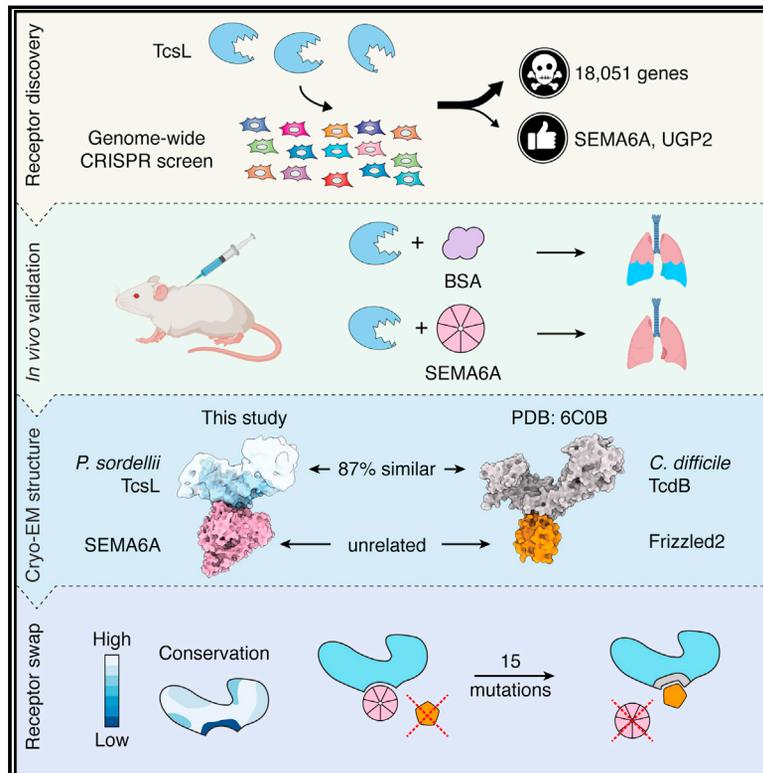


Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Recognition of Semaphorin Proteins by *P. sordellii* Lethal Toxin Reveals Principles of Receptor Specificity in Clostridial Toxins

Graphical Abstract



Authors

Hunsang Lee, Greg L. Beilhartz, Iga Kucharska, ..., Jean-Philippe Julien, Roman A. Melnyk, Mikko Taipale

Correspondence

jean-philippe.julien@sickkids.ca (J.-P.J.), roman.melnyk@sickkids.ca (R.A.M.), mikko.taipale@utoronto.ca (M.T.)

In Brief

A lethal bacterial toxin that plays a major role in toxic shock syndrome uses the host semaphorin proteins as receptors, and structural analysis shows a small binding interface that can be mutated to switch specificity between very different host receptors.

Highlights

- CRISPR screen identifies SEMA6A and SEMA6B as receptors for *P. sordellii* lethal toxin TcsL
- Soluble SEMA6A ectodomain protects mouse lungs from TcsL-induced edema
- 3.3 Å cryo-EM structure of TcsL bound to SEMA6A reveals atomic details of the interaction
- 15 mutations in the TcsL receptor-binding interface rewire receptor specificity



Article

Recognition of Semaphorin Proteins by *P. sordellii* Lethal Toxin Reveals Principles of Receptor Specificity in Clostridial Toxins

Hunsang Lee,^{1,9} Greg L. Beilhartz,^{2,9} Iga Kucharska,^{2,9} Swetha Raman,² Hong Cui,² Mandy Hiu Yi Lam,¹ Huazhu Liang,^{2,3} John L. Rubinstein,^{2,3,4} Daniel Schramek,^{5,6} Jean-Philippe Julien,^{2,3,7,8,10,*} Roman A. Melnyk,^{2,3,10,*} and Mikko Taipale^{1,6,8,10,11,*}

¹Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada

²Program in Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON M5G 0A4, Canada

³Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada

⁴Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada

⁵Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada

⁶Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

⁷Department of Immunology, University of Toronto, Toronto, ON M5S 1A8, Canada

⁸Molecular Architecture of Life Program, Canadian Institute for Advanced Research (CIFAR), Toronto, ON M5G 1M1, Canada

⁹These authors contributed equally

¹⁰Senior author

¹¹Lead Contact

*Correspondence: jean-philippe.julien@sickkids.ca (J.-P.J.), roman.melnyk@sickkids.ca (R.A.M.), mikko.taipale@utoronto.ca (M.T.)
<https://doi.org/10.1016/j.cell.2020.06.005>

SUMMARY

Pathogenic clostridial species secrete potent toxins that induce severe host tissue damage. *Paenibacillus sordellii* lethal toxin (TcsL) causes an almost invariably lethal toxic shock syndrome associated with gynecological infections. TcsL is 87% similar to *C. difficile* TcdB, which enters host cells via Frizzled receptors in colon epithelium. However, *P. sordellii* infections target vascular endothelium, suggesting that TcsL exploits another receptor. Here, using CRISPR/Cas9 screening, we establish semaphorins SEMA6A and SEMA6B as TcsL receptors. We demonstrate that recombinant SEMA6A can protect mice from TcsL-induced edema. A 3.3 Å cryo-EM structure shows that TcsL binds SEMA6A with the same region that in TcdB binds structurally unrelated Frizzled. Remarkably, 15 mutations in this evolutionarily divergent surface are sufficient to switch binding specificity of TcsL to that of TcdB. Our findings establish semaphorins as physiologically relevant receptors for TcsL and reveal the molecular basis for the difference in tissue targeting and disease pathogenesis between highly related toxins.

INTRODUCTION

Paenibacillus sordellii (also known as *Clostridium sordellii*) is an anaerobic gram-positive bacterium found in soil and in the gastrointestinal and vaginal tracts of animals and humans. *P. sordellii* is present in the rectal or vaginal tract of 3%–4% of women, but vaginal colonization rate after childbirth is as high as 29% (Aldape et al., 2016; Chong et al., 2016). Although the majority of carriers are asymptomatic, pathogenic *P. sordellii* infections arise rapidly and are highly lethal. The origin of pathogenic *P. sordellii* strains is unclear, but most infections occur in women after childbirth, medically induced abortion, or miscarriage, leading to a toxic shock syndrome with almost 100% mortality within days (Aldape et al., 2016; Chong et al., 2016; Fischer et al., 2005; Ho et al., 2009).

The primary cause of the high mortality associated with *P. sordellii* infections is the lethal toxin TcsL (Carter et al., 2011), which belongs to the large clostridial toxin (LCT) family (Orrell et al., 2017). LCTs enter the host cell by receptor-mediated endo-

cytosis into acidified endosomes followed by pH-dependent pore formation and translocation into the cytoplasm (Papatheodorou et al., 2010; Pfeifer et al., 2003; Zhang et al., 2014). After autoprocessing in the cytosol, the released cytotoxic glucosyltransferase enzymes potently modulate host cell function by inactivating small Rho-family GTPases by using uridine diphosphate (UDP)-glucose or UDP-*N*-acetylglucosamine as a co-substrate (Aktories and Barbieri, 2005; Aktories and Just, 1995; Reineke et al., 2007).

Although all LCTs are highly similar at the sequence level, they differ in their tissue specificity and in their effects on cell morphology, physiology, and viability. TcsL is most closely related to the *C. difficile* cytotoxin TcdB, sharing almost 90% sequence similarity. TcdB is the causal virulence factor behind gastrointestinal diseases associated with *C. difficile* infections. TcdB binds Frizzled family receptors FZD1, FZD2, and FZD7 expressed in the colonic epithelium (Chen et al., 2018; Tao et al., 2016), the primary site of *C. difficile* infection. In contrast, although present in the intestinal microbiota, *P. sordellii* does not infect or damage



the colonic epithelium, suggesting that TcsL binds a different cell surface receptor. This is supported by earlier competition experiments with recombinant TcdB and TcsL *in vitro*, showing differential sensitivity of cell lines to TcsL and TcdB (Chaves-Olarte et al., 1997). Although recent studies have begun to uncover receptors for other LCTs, the physiologically relevant TcsL receptors are still unknown. In addition, it is not understood at the molecular level how structurally and functionally similar LCTs can bind completely unrelated receptors.

Here, we employ genome-wide CRISPR/Cas9 screens to identify semaphorins SEMA6A and SEMA6B as the host cell receptors for TcsL. We show that recombinant SEMA6A ectodomain can protect lung endothelial cells from TcsL intoxication *in vitro* and mouse lungs from TcsL-induced edema *in vivo*. To further understand the molecular basis of LCT receptor specificity, we determine the cryogenic electron microscopy (cryo-EM) structure of TcsL bound to SEMA6A. Comparison of this structure to the TcsL-Frizzled2 structure (Chen et al., 2018) reveals that TcsL and TcdB bind their structurally unrelated receptors by using the same receptor-binding region. However, selective clustering of mutations explains how the two toxins have evolved distinct specificities. We formally demonstrate the role of the clustered mutations in the receptor-binding site by switching the receptor specificity of TcsL from SEMA6A to Frizzled2. These results have broad implications for our understanding of host receptor pathogenic toxins evolution at the molecular level.

RESULTS

Genome-wide CRISPR/Cas9 identifies SEMA6A as a host factor needed for TcsL toxicity

To identify potential cell-surface receptors for TcsL, we conducted a genome-wide CRISPR/Cas9 screen in Hap1 cells. We used the genome-wide TKOv3 guide RNA (gRNA) library that targets 18,053 genes with four different gRNAs per gene (Hart et al., 2017). Cells were infected with the library and treated with 0.1 nM or 1 nM TcsL. Sequencing of gRNAs from the surviving population revealed only two genes that were significantly enriched in both screens (Figures 1A and 1B and Table S1). One was UDP-glucose pyrophosphorylase 2 (UGP2), which was also a top hit in previous CRISPR/Cas9 screens with two other LCTs, TcdA and TcdB (Tao et al., 2016, 2019). UGP2 is needed for the synthesis of UDP-glucose, the sugar donor for the glucosylation activity of all LCTs (Aktorius and Just, 1995). The other hit was SEMA6A, encoding a transmembrane axon guidance molecule not previously linked to toxin function.

Notably, despite high sequence similarity between TcsL and TcdB, our screen did not identify Frizzled receptors, or CSPG4 or PVRL3, two other TcdB-associated receptors (LaFrance et al., 2015; Tao et al., 2016; Yuan et al., 2015). Neither did we identify known receptors for *C. perfringens* TpeL or *C. difficile* TcdA, other related LCTs (Schorch et al., 2014; Tao et al., 2019), although their receptors are expressed in Hap1 cells (Figure S1).

Validation of SEMA6A and SEMA6B as cellular receptors for TcsL

We first validated screen results by knocking out UGP2 and SEMA6A in Hap1 cells. UGP2^{KO} and SEMA6A^{KO} cells were high-

ly resistant to TcsL (Figure S1). Introduction of wild-type SEMA6A by lentiviral transduction to SEMA6A^{KO} cells rendered the cells more sensitive to the toxin (Figure S1). SEMA6A is a member of the semaphorin family, which in humans consists of twenty transmembrane and secreted proteins (Worzfeld and Offermanns, 2014). The four human SEMA6 class proteins are SEMA6A, SEMA6B, SEMA6C, and SEMA6D (Figure 1C). SEMA6A is most closely related to SEMA6B with 69% sequence similarity and 53% identity, followed by SEMA6D (60% similarity, 46% identity) and SEMA6C (56% similarity, 40% identity). We generated knockout Hap1 cells for each SEMA6 family member and assayed their sensitivity to TcsL. SEMA6C^{KO} and SEMA6D^{KO} cells were as sensitive to TcsL as control cells (50% growth inhibition [GI50]: 50 pM) (Figure 1D). In contrast, SEMA6B^{KO} cells were approximately 10-fold more resistant (GI50: 500 pM) than control cells, but not as resistant as SEMA6A^{KO} cells (50-fold; GI50: 2.5 nM). Moreover, SEMA6A-SEMA6B double knockout cells were more resistant to TcsL than either knockout alone, suggesting that SEMA6A and SEMA6B act redundantly (Figure 1D). Consistent with this, HeLa cells that do not express SEMA6A or SEMA6B were highly resistant to TcsL (Figure S1). Because SEMA6C and SEMA6D are expressed at low amounts in Hap1 and HeLa cells (Figure S1), we further assessed their role in TcsL intoxication by ectopically expressing each SEMA6 in SEMA6A^{KO} cells by lentiviral infection. In contrast to ectopic SEMA6A and SEMA6B expression, SEMA6C and SEMA6D did not render SEMA6A^{KO} cells more sensitive to the toxin (Figure 1E). Thus, SEMA6A and SEMA6B but not SEMA6C or SEMA6D regulate cellular sensitivity to TcsL. These results are consistent with the model that both SEMA6A and SEMA6B can act as TcsL receptors, and sensitivity to the toxin is determined by the total amount of SEMA6A and SEMA6B.

Recombinant SEMA6A ectodomain protects cells and mouse lungs from TcsL toxicity

We then expressed and purified the soluble recombinant extracellular domain (rECD) of SEMA6A and tested its effect on TcsL toxicity on Vero cells, a commonly used cell line for studying toxin function. Vero cells express SEMA6A but not SEMA6B (Figure S1). SEMA6A rECD inhibited TcsL toxicity in a dose-dependent manner (Figure 2A). Notably, SEMA6A rECD had a protective effect only when it was added to the cells before or simultaneously with TcsL (Figure S1). When cells were pre-treated for 1 h with TcsL, SEMA6A rECD had no effect on toxicity, suggesting that SEMA6A rECD must act before TcsL binds the cells. We also repeated the competition assay with soluble ectodomains of SEMA6B, SEMA6C, and SEMA6D. Consistent with our experiments in Hap1 cells (above), SEMA6B alleviated TcsL toxicity, whereas SEMA6C and SEMA6D had no effect (Figures 2B and 2C). These results strongly suggest that TcsL can directly bind SEMA6A and SEMA6B and this interaction is needed for TcsL entry into the cell.

A primary target of TcsL during *P. sordellii* infection is the vascular endothelium of the lung (Geny et al., 2007). Therefore, we used immortalized human lung microvascular cells (HULECs) as a physiologically relevant cell line to study the role of SEMA6A and SEMA6B in TcsL intoxication. Notably, HULECs express ~4-fold higher amounts of SEMA6A and ~11-fold higher

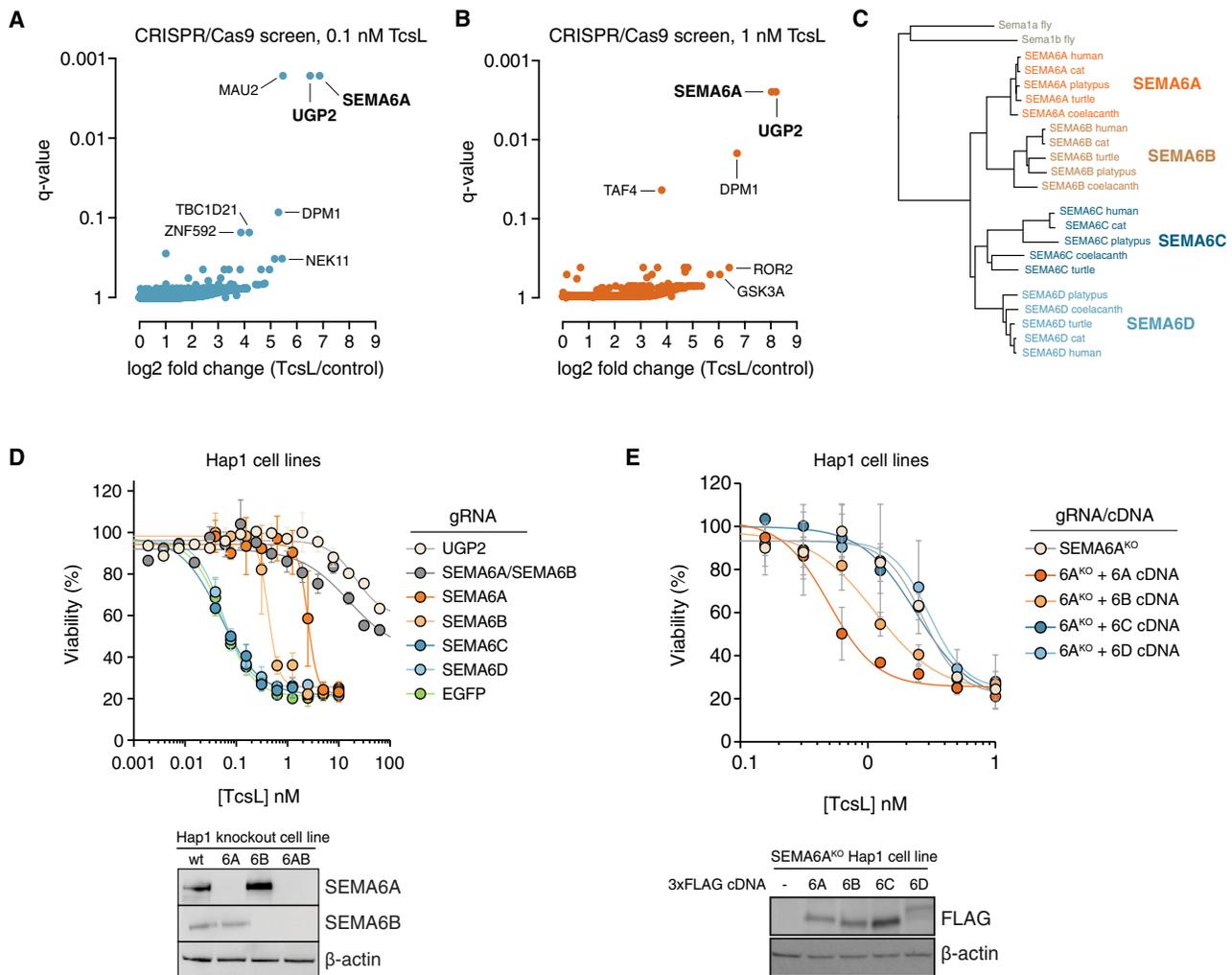


Figure 1. SEMA6A and SEMA6B are host cell receptors for *P. sordellii* lethal toxin TcsL

(A) Genome-wide CRISPR/Cas9 screen in Hap1 cells identifies factors regulating sensitivity to 0.1 nM TcsL. Hap1 cells were infected with a genome-wide TKOv3 gRNA library, treated with recombinant TcsL, and gRNAs from surviving cells were sequenced.

(B) Genome-wide CRISPR/Cas9 screen with 1 nM TcsL.

(C) Phylogenetic tree of SEMA6 family proteins.

(D) Hap1 cells were infected with Cas9 and gRNA targeting indicated genes and tested for sensitivity to TcsL. Data (n = 3) are represented as mean ± standard deviation. Shown at the bottom, expression of SEMA6A and SEMA6B in single and double knockout cell lines was assessed by western blotting.

(E) Hap1 SEMA6A^{KO} cells were infected with lentiviruses expressing 3xFLAG-tagged SEMA6 family proteins and tested for TcsL sensitivity. Data (n = 3) are represented as mean ± standard deviation. Shown at the bottom, expression of SEMA6 proteins in infected cell lines was validated with western blotting.

See also Figure S1 and Table S1.

amounts of SEMA6B than Hap1 cells (Figure 3A). We assayed the sensitivity of HULECs to four related clostridial toxins: TcdA and TcdB from *C. difficile*, TpeL from *C. perfringens*, and TcsL (Figure 3B). TpeL showed low toxicity, and TcdA and TcdB were several orders of magnitude less toxic to HULECs than what is reported for other cell types (Chaves-Olarte et al., 1997; Gupta et al., 2017). Remarkably, the GI₅₀ of TcsL was just ~50 femtomolar, indicating that ~200 toxin molecules per cell are lethal to HULECs. Furthermore, recombinant ectodomain of mouse Sema6a fused to Fc fragment (rSema6a-Fc) but not mouse Sema6c ectodomain-Fc (rSema6c-Fc) could block TcsL-induced rounding of HULECs in a dose-dependent manner

(Figures 3C and 3D). These results further suggest that SEMA6A and SEMA6B are the physiologically relevant TcsL receptors in endothelial cells.

We then addressed the role of semaphorins *in vivo* in a mouse model of TcsL intoxication. We first examined the expression of Sema6a and Sema6b in mouse lung tissue by immunohistochemistry. Both proteins were highly expressed in lung endothelium and pneumocytes (Figure 3E). Mice were injected intraperitoneally with a lethal dose of TcsL together with rSema6a-Fc, rSema6c-Fc, or bovine serum albumin (BSA) (n = 3 for each group). Mice co-injected with BSA or Sema6c-Fc rapidly developed symptoms of TcsL intoxication, including decreased

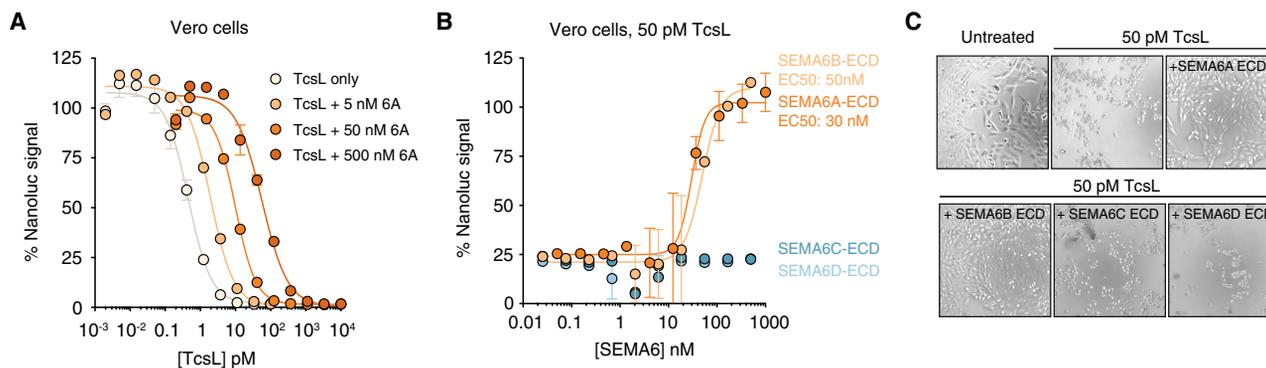


Figure 2. SEMA6A and SEMA6B ectodomains protect cells from TcsL intoxication

(A) Vero cells stably expressing Nanoluciferase viability reporter were treated with increasing amounts of TcsL in the presence of recombinant SEMA6A ectodomain. Data ($n = 2$) are represented as mean \pm standard deviation.

(B) Vero cells stably expressing Nanoluciferase viability reporter were treated with 50 pM TcsL with increasing amounts of recombinant SEMA6 family ectodomains for 24 h. Data ($n = 2$) are represented as mean \pm standard deviation.

(C) Microscopy images of Vero cells treated with TcsL (50 pM) and SEMA6 family ectodomains (1 μ M) for 24 h.

See also Figure S1.

mobility and signs of ataxia and dehydration. Within 4 h, all symptomatic mice had a buildup of fluid in the lungs (Figure 3F). Histopathologically, symptomatic mice showed edema surrounding pulmonary vessels (Figure 3F). In contrast, rSema6a-Fc protected the mice from TcsL-induced symptoms. Mice co-injected with rSema6a-Fc had no pleural effusion and did not show signs of edema after 4 h (Figures 3F and 3G). Altogether, these data support that SEMA6A and SEMA6B are the physiologically relevant receptors for TcsL and that this interaction can be inhibited *in vivo*.

Cryo-EM structure of the TcsL_{1285–1804}-SEMA6A complex

Having established SEMA6A and SEMA6B as the cellular receptors of TcsL, we next structurally characterized the interaction between TcsL and SEMA6A by using cryo-EM. We used a shortened TcsL fragment spanning amino acids (aa) 1285–1804 (TcsL_{1285–1804}), analogous to the TcdB fragment previously used to determine the TcdB-Frizzled2 complex structure (Chen et al., 2018). In a biolayer interferometry assay with immobilized ligand, which measures the avidity of interactions, the TcsL_{1285–1804} fragment bound recombinant SEMA6A ectodomain with nanomolar apparent affinity (Figures 4A and S3). Glutaraldehyde cross-linking was used to prevent dissociation of the TcsL_{1285–1804}-SEMA6A complex during cryo-EM grid preparation, as described previously (Dong et al., 2019; Kastner et al., 2008; Yan et al., 2015). Initial analysis of the dataset revealed that approximately 50% of the SEMA6A dimers were bound to TcsL by using this sample preparation approach. Consequently, 2D and 3D classification resulted in two homogeneous datasets corresponding to the TcsL-SEMA6A complex (179,188 particle images) (Figures S2 and S3) and unliganded SEMA6A (281,207 particle images) (Figures S2 and S3), which were used to produce final density maps to overall resolutions of 3.3 and 3.1 Å, respectively (Table S2). Local resolution of the TcsL-SEMA6A map ranged from 2.8 Å at the binding interface to > 30 Å at the C terminus of TcsL (Figure S4). Indeed, 3D vari-

ability analysis (Punjani et al., 2017) revealed continuous motion of the C and N termini of TcsL in contrast to a largely rigid SEMA6A dimer and TcsL-SEMA6A interface (Video S1). A molecular model was built for the well-resolved central domain of TcsL (residues 1400–1637), corresponding to approximately 40% of the TcsL construct (Figure 4A).

The structure of TcsL is highly similar to previously determined structures of other clostridial toxins, particularly TcdB (root-mean-square deviation [RMSD] = 1.3 Å) (Chen et al., 2018, 2019; Chumbler et al., 2016; Simeon et al., 2019) (Figure S5). Similarly, unliganded and TcsL-bound SEMA6A structures are in remarkable agreement (Figure S5), indicating that TcsL binding is not accompanied by significant conformational changes in the receptor and that glutaraldehyde treatment had a negligible effect on the SEMA6A structure.

SEMA6A interacts with TcsL at two predominant sites (Figure 4B), burying a total surface area of 1,236 Å². The major contact is through a short α -helix (residues 101–110) between blades 1 and 2, in a characteristic “extrusion” of semaphorin family beta-propellers (Janssen et al., 2010; Love et al., 2003; Nogi et al., 2010). The other site is in the flexible loop between beta sheets 3B and 3C (Asp189, Phe190, Leu191, and Ile193) (Figures 4B and S6). Notably, the TcsL-binding interface is where SEMA6A’s cognate ligand Plexin A2 also binds (Figure 4C) (Janssen et al., 2010; Nogi et al., 2010). Indeed, over half of the TcsL-binding interface in SEMA6A is shared with Plexin A2 (Figures 4C and S6). Another notable similarity between Plexin A2 and TcsL is their semaphorin specificity: both bind SEMA6A and SEMA6B but not SEMA6C or SEMA6D (Worzfeld and Offermanns, 2014). In the SEMA6A-TcsL structure, two SEMA6A residues (Arg108 and Ile193) that engage in contact with TcsL are identical in SEMA6B but not conserved in SEMA6C and SEMA6D (Figure S6), likely explaining the specificity of TcsL. Ile193 forms critical contacts also with Plexin A2 (Figure S6), suggesting that the semaphorin specificity of TcsL and Plexin A2 are based on similar molecular principles.

On TcsL, the SEMA6A interaction interface is located on a convex edge of the delivery domain (Figure 5A). The highly

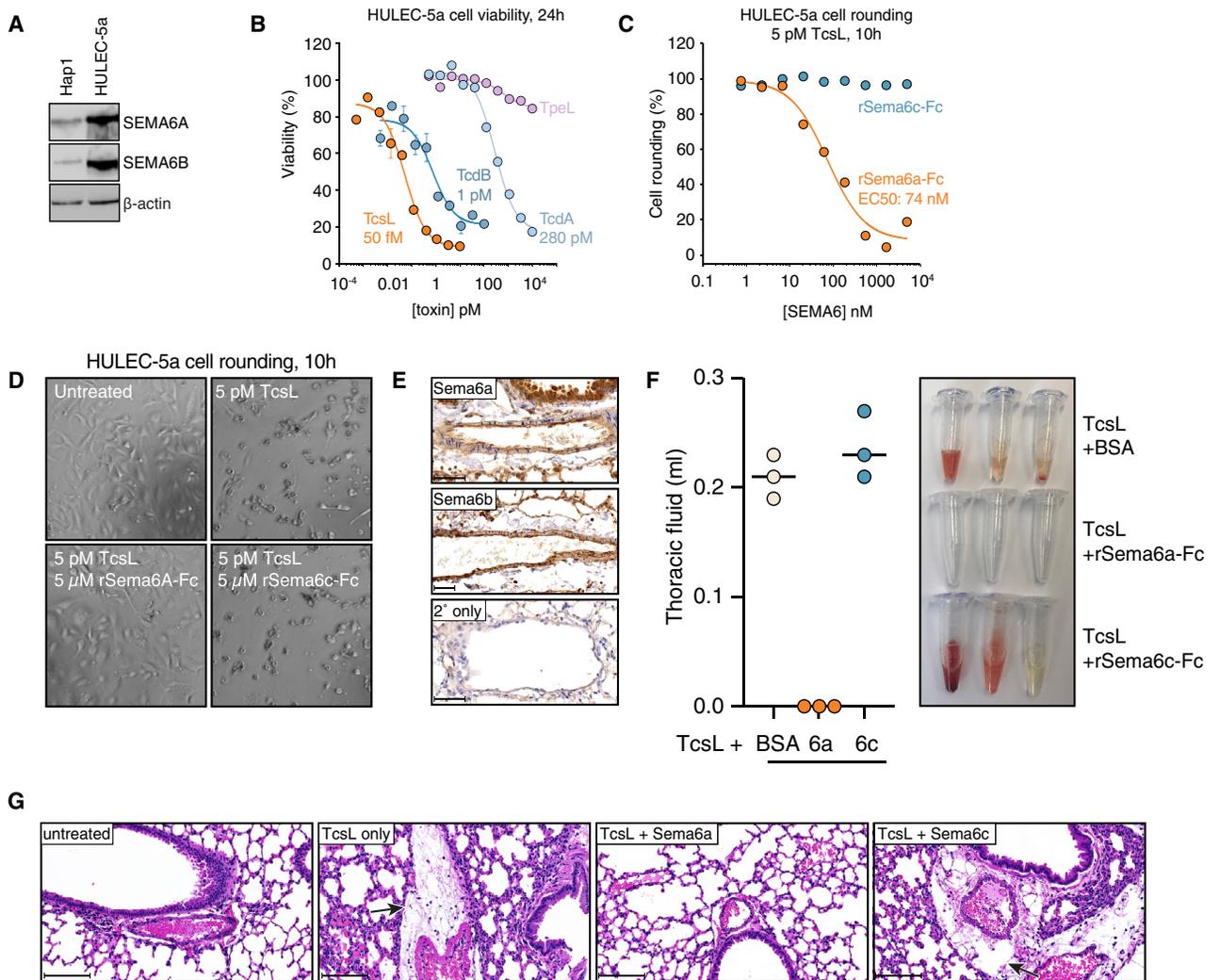


Figure 3. SEMA6A ectodomain protects lung endothelial cells and mouse lungs from TcsL-induced toxicity

- (A) SEMA6A and SEMA6B protein expression in human lung endothelial cells.
- (B) HULECs are extremely sensitive to TcsL. HULEC-5a cells were treated with increased amounts of indicated Clostridial toxins.
- (C) Recombinant mouse Sema6a ectodomain fused to Fc domain protects HULEC-5a cells from TcsL-induced cell rounding. Cells were treated with 5 pM TcsL and increasing amounts of Sema6a and Sema6c ectodomains.
- (D) Microscopy images of HULEC-5a cells treated with TcsL and recombinant Sema6a and Sema6c ectodomains.
- (E) Immunohistochemistry images of Sema6a and Sema6b expression in mouse lung tissue sections.
- (F) Mice were intraperitoneally injected with 15 ng TcsL and 1,000-fold molar excess of Sema6a ectodomain, Sema6c ectodomain or BSA. Thoracic fluid was collected and measured from symptomatic mice 4 h after injection.
- (G) Lung tissue sections of mice treated with indicated conditions. Arrows indicate lung edema induced by TcsL.

hydrophobic interface is mainly formed by residues located in an antiparallel beta sheet (residues 1466–1511), flanked by an α -helix (1433–1438) on one side and a loop (1596–1601) on the other. Strikingly, this location on *C. difficile* cytotoxin TcdB is also the one used to bind Frizzled receptors, which are structurally unrelated to semaphorins (Chen et al., 2018) (Figure 5A). Nevertheless, both toxins bind their receptors in a similar manner. The hydrophobic nature of this pocket is conserved in TcsL but the amino acid identities are not (Figure 5A). Moreover, the interaction of TcdB with FZD2 is bridged by a palmitoleic acid (PAM) moiety, which is buried in a hydrophobic pocket of TcdB (Fig-

ure 5B) (Chen et al., 2018). In TcsL, the pocket buries the hydrophobic side chain of SEMA6A Met109 in an analogous manner to the palmitoleic acid moiety in the TcdB-FZD2 structure (Figure 5B and Table S3). In fact, although none of the six residues that in TcdB contact PAM through hydrophobic interactions are conserved in TcsL, all six corresponding residues in TcsL form similar hydrophobic contacts with SEMA6A Met109 (Figure 5B and Table S3).

We experimentally validated the TcsL-SEMA6A interaction surface by first generating and purifying a variant of TcsL (TcsL^{4mut}) that carried four mutations in core interacting residues

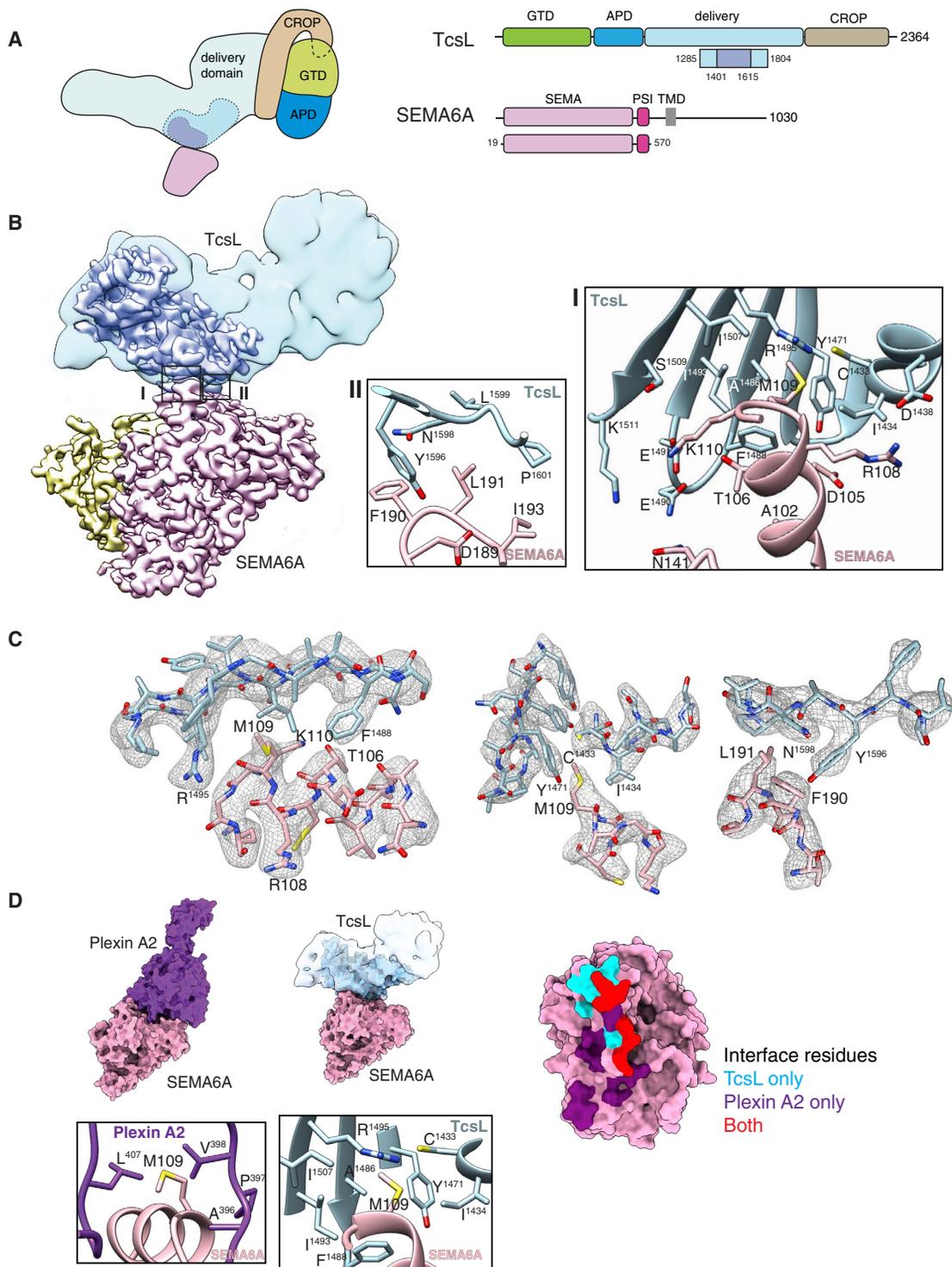


Figure 4. Cryo-EM structure of the TcsL₁₂₈₅₋₁₈₀₄-SEMA6A complex

(A) On the left, domain structure of TcsL and SEMA6A. Constructs used in cryo-EM analysis are indicated below. Darker blue color indicates TcsL region resolved to medium and high resolution ($< 7 \text{ \AA}$) in cryo-EM. Shown on the right is a schematic indicating the location of the TcsL fragment used for cryo-EM, based on full-length TcdB structure (PDB: 6OQ5). Abbreviations are as follows: GTD, glucosyltransferase domain; APD, autoprocessing domain; delivery, delivery domain; CROP, combined repetitive oligopeptides; SEMA, semaphorin domain, PSI, plexin-semaphorin-integrin domain; TMD, transmembrane domain.

(legend continued on next page)

(Cys1433Asp, Ile1434Lys, Ala1486Ser, and Tyr1596Arg) (shown in Figure 5A). TcsL Cys1433, Ile1434, and Ala1486 form key hydrophobic interactions in the pocket accommodating SEMA6A Met109, whereas TcsL Tyr1596 interacts with several residues in the flexible loop of SEMA6A through hydrophobic interactions and through a hydrogen bond with SEMA6A Phe190 backbone amide (Figure 5A and Table S3). TcsL^{4mut} was highly soluble and retained its inositol hexaphosphate (IP6)-dependent auto-processing activity *in vitro*, suggesting that it is properly folded (Figure 5C). However, TcsL^{4mut} was 2,500-fold less toxic in Vero cells than the wild-type TcsL, indicating that the four interface residues are physiologically relevant (Figure 5C). We then validated the relevance of SEMA Met109 in TcsL toxicity. We introduced 3xFLAG-tagged wild-type SEMA6A or Met109Asp mutant by lentiviral infection into SEMA6A/6B double knockout Hap1 cells and assessed their effect on TcsL toxicity. In contrast to wild-type SEMA6A, the Met109Asp mutant did not make SEMA6A/6B double knockout cells more sensitive to TcsL, although both mutants were expressed at similar levels (Figure 5D).

Rewiring the receptor specificity of TcsL

Our results establish that TcsL and TcdB have evolved to bind different host receptors through the same interacting region. Consistent with this, the surface is highly diverged between the two toxins: of the 25 SEMA6A-interacting residues in TcsL, only five are identical and nine are similar compared with those in TcdB. Conversely, of the 19 residues in TcdB that interact with FZD2 or PAM, only three are identical and eight are similar in TcsL. In stark contrast, 76% of non-interacting surface residues in TcdB are identical in TcsL ($p < 0.0001$, Fisher's exact test), indicating that the divergence is specific to the receptor-binding surface. This surface might represent a more general receptor-binding site in clostridial toxins, given that the divergence extends to all members of the family (Figure S6). In particular, the five-stranded beta sheet that forms the base of the hydrophobic pocket for SEMA6A Met109 in TcsL and for the palmitoleic acid moiety in the FZD2-TcdB complex is highly variable between LCTs (Figure 6A and Figure S6).

We directly tested the role of the evolutionarily divergent surface in receptor specificity with recombinant proteins. As expected, TcsL_{1285–1804} interacted with SEMA6A, and TcdB_{1285–1804} formed a stable complex with FZD7 in size exclusion chromatography (Figure 6B). However, TcsL did not interact with FZD7, consistent with the extensive divergence of the receptor-binding interface (Figure 6B). We then generated a TcsL frizzled-binding domain (FBD) variant TcsL(FBD)_{1285–1804} that had a TcdB-like interface, by changing 15 TcsL residues in the

receptor-binding interface to those in TcdB (Figure 6C). Remarkably, these mutations were sufficient to switch TcsL binding specificity: the TcsL_{1285–1804} (FBD) hybrid protein robustly interacted with FZD7 but no longer bound SEMA6A (Figure 6D). Thus, selective clustering of mutations on a single surface of the delivery domain of TcdB and TcsL explains the divergent receptor specificity of the two toxins.

DISCUSSION

Identifying the cellular receptors of bacterial toxins that are the primary determinants of human diseases is of major importance both for elucidating the fundamental strategies used by pathogens to cause disease and for informing the development of therapeutics to block their uptake. Our discovery and structural characterization of semaphorins as *P. sordellii* lethal toxin receptors provides a potential therapeutic avenue for these devastating infections and reveals how evolution has sculpted the receptor-binding interfaces of large clostridial toxins to recognize distinct host receptors.

SEMA6 family proteins coordinate axon repulsion and attraction during development (Suto et al., 2007; Tawarayama et al., 2010; Xu et al., 2000), and as such their role in toxin entry in endothelial cells might seem unexpected. However, semaphorins regulate several other cellular processes outside the central nervous system (Neufeld et al., 2012; Worzfeld and Offermanns, 2014). For example, SEMA6A promotes angiogenesis through vascular endothelial growth factor (VEGF) signaling in vascular endothelial cells, and SEMA6B positively regulates endothelial cell proliferation (Kigel et al., 2011; Segarra et al., 2012). It seems that *P. sordellii* and TcsL have evolved to exploit the non-neuronal semaphorin function to their own advantage during infection. In canonical semaphorin-plexin signaling, semaphorins act as ligands for plexins either through cell-cell contacts or *in cis*. However, signaling can also occur through the reverse route where semaphorins are the receptors that induce downstream signaling events (Haklai-Topper et al., 2010; Jongbloets and Pasterkamp, 2014; Perez-Branguli et al., 2016). Our cryo-EM structure of the TcsL-SEMA6A complex shows that the binding site of TcsL on SEMA6A significantly overlaps with the native Plexin A2 binding site (Janssen et al., 2010; Nogi et al., 2010). It is highly unlikely that both ligands could simultaneously bind SEMA6A because of steric clashes. However, disruption of the plexin-semaphorin axis is unlikely to contribute to acute toxicity of TcsL. For example, Hap1 and Vero cells that are highly sensitive to the toxin do not express the Plexin A2 or A4 (Figure S1), and we did not identify known downstream regulators of reverse SEMA6A signaling, such as Abl or Mena (Perez-Branguli et al.,

(B) Composite cryo-EM map of the TcsL-SEMA6A complex. SEMA6A monomers are colored pink and yellow, TcsL domain resolved to medium and high resolution ($< 7 \text{ \AA}$) is colored dark blue. Low-pass filtered density (10 \AA) of the TcsL protein is shown in light blue. Insets I and II show contact residues between SEMA6A (pink) and TcsL (blue).

(C) Atomic model built into the SEMA6A-TcsL map. Cryo-EM density of SEMA6A-TcsL is shown as gray mesh with the model built shown as sticks (pink for SEMA6A and blue for TcsL).

(D) Comparison of Plexin A2-SEMA6A (PDB: 3OKY) (Janssen et al., 2010) and TcsL-SEMA6A binding interactions. M109 of SEMA6A interacts with hydrophobic residues of Plexin A2, including L407 and V398 (left). TcsL buries M109 of SEMA6A in a binding pocket containing several hydrophobic residues (middle). An overlay of TcsL and Plexin A2 binding surfaces (shown in blue and purple, respectively) reveals a subset of SEMA6A residues participating in binding to both protein ligands (red).

See also Figures S2–S5.

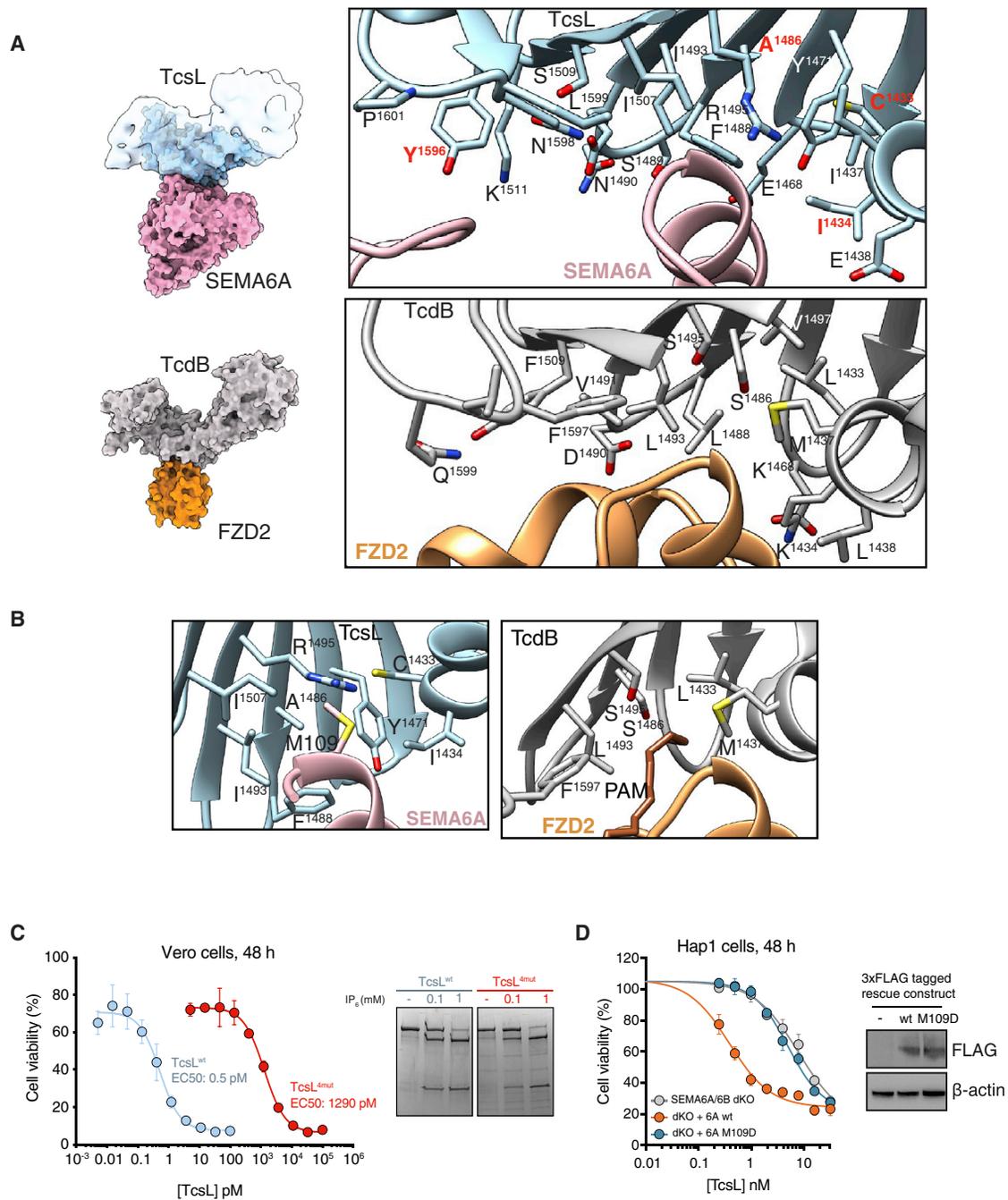


Figure 5. Clostridial toxins use the same region to bind their cognate receptors

(A) Comparison of the TcsL-SEMA6A and TcdB-FZD2 (PDB:6C0B) (Chen et al., 2018) binding interface. Residues mutated in TcsL^{4mut} are indicated in red. (B) TcsL buries M109 of SEMA6A in a hydrophobic binding pocket (left), whereas TcdB utilizes a similar hydrophobic binding pocket to interact with the palmitoleic acid moiety of FZD2 (right). (C) Experimental validation of the TcsL-SEMA6A interaction interface. The cytotoxicity of wild-type TcsL and TcsL^{4mut} variant with four mutated interaction interface residues (C1433D-I1434K-A1486S-Y1596R; shown in A) was assessed in Vero cells. (D) Validation of SEMA6A M109 as a critical interacting residue with TcsL. SEMA6A/SEMA6B double knockout cells were infected with 3xFLAG-tagged wild-type SEMA6A or M109D mutant and assayed for sensitivity to TcsL. Protein expression levels were confirmed by western blotting (right). See also Figures S2–S5.

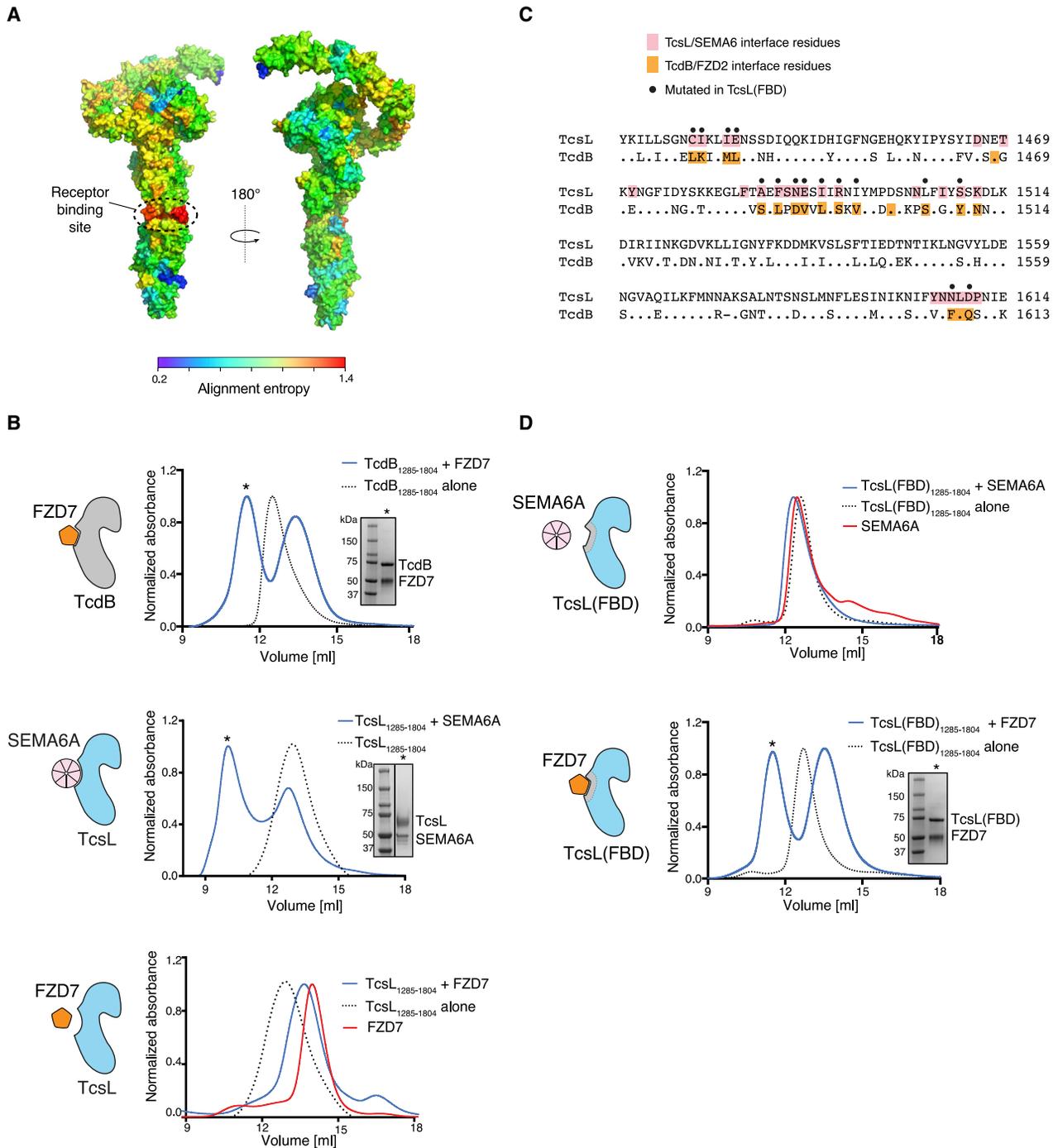


Figure 6. TcsL and TcdB bind different host receptors through the same interface region

(A) Sequence alignment entropy in large clostridial toxin family shown as a rainbow spectrum on the TcdB full-length cryo-EM structure (PDB ID: 6OQ5) (Chen et al., 2019). Entropy was calculated as a 10-aa moving window. The receptor-binding surface is indicated.

(B) SEC profiles and SDS-PAGE analysis of FZD7-TcdB₁₂₈₅₋₁₈₀₄ (top), SEMA6A-TcsL₁₂₈₅₋₁₈₀₄ (middle), and FZD7-TcsL₁₂₈₅₋₁₈₀₄ (bottom). SEC fractions used for SDS-PAGE analysis are highlighted with an asterisk.

(C) Sequence alignment between TcsL and TcdB. TcsL residues interacting with SEMA6A are highlighted in pink and TcdB residues contacting FZD2 are highlighted in orange. Black dots denote the 15 mutations introduced in TcsL (FBD)₁₂₈₅₋₁₈₀₄ variant that resulted in shifting the TcsL binding specificity from SEMA6A to FZD7.

(D) SEC profiles and SDS-PAGE analysis of FZD7-TcsL (FBD)₁₂₈₅₋₁₈₀₄ (TcsL variant with a TcdB-like binding interface) (top) and SEMA6A-TcsL (FBD)₁₂₈₅₋₁₈₀₄ (bottom). SEC fractions used for SDS-PAGE analysis are highlighted with an asterisk.

See also Figure S6.

2016), in our CRISPR/Cas9 screens. Future work can address whether disruption of this signaling axis signaling modulates TcsL toxicity *in vivo*.

SEMA6A and SEMA6B add to the list of proteins that have been identified as receptors or host factors for large clostridial toxins, alongside Frizzled, CSPG4, and Nectin3 for TcdB; LRP1 for TpeL; and sulfated glycosaminoglycans and LDLR for TcdA (LaFrance et al., 2015; Schorch et al., 2014; Tao et al., 2016, 2019; Yuan et al., 2015). Given the high degree of conservation between all LCTs, we cannot exclude the possibility that TcsL also binds additional receptors on host cells. However, we have not identified additional factors in CRISPR/Cas9 screens in cells that do not express SEMA6A or SEMA6B. Moreover, if such receptors exist, they must be distinct from those of other LCTs.

TcsL shares almost 90% similarity with *C. difficile* cytotoxin TcdB, which binds Frizzled family receptors in the colonic epithelium (Chen et al., 2018; Tao et al., 2016). Previous work had suggested that TcsL and TcdB use different receptors (Chaves-Olarte et al., 1997), which is consistent with their distinct target tissues during infection. Given the sequence conservation, one might have expected that their receptors are also related to each other. This was not the case: semaphorins and Frizzled receptors are structurally and evolutionarily unrelated. However, the cryo-EM structure of the TcsL-SEMA6A complex revealed that TcsL and TcdB use exactly the same region to interact with their cognate receptors. Comparison of the structures of TcsL and TcdB bound to their cognate receptors provides an elegant molecular explanation for the distinct receptor specificity: selective clustering of adaptive mutations in the receptor-binding interface. The high sequence diversity in this region between all clostridial toxins indicates that this interface is likely a more general receptor-binding site for these toxins. Indeed, the receptor-binding site of TpeL, a more distantly related clostridial toxin, has been mapped to a region that includes the hypervariable site (Schorch et al., 2014).

The striking diversity in the receptor-binding specificity of clostridial toxins is reminiscent of receptor-binding proteins of several viruses. For example, closely related coronaviruses can bind evolutionarily unrelated host receptors through highly variable loops in the spike protein (Wong et al., 2017; Wu et al., 2020). Thus, in a similar manner to rapidly evolving viral proteins, a sculptable receptor-binding interface in large clostridial toxins provides an evolutionarily flexible platform that might allow pathogenic clostridial species to adapt to and attack novel host organisms and tissues.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- **EXPERIMENTAL MODELS AND SUBJECT DETAILS**
 - Cell lines

● METHOD DETAILS

- Genome-wide CRISPR screen
- Next-generation sequencing library preparation
- CRISPR screen validation
- Site-directed mutagenesis
- Vero cell experiments
- CPD Cleavage assay
- Hulec-5a cells
- Recombinant proteins
- Biolayer interferometry (BLI)
- TcsL-SEMA6A complex formation and glutaraldehyde crosslinking
- Cryo-EM data collection and image processing
- Model building

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2020.06.005>.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Amy Tong, Mr. Christopher Mogg, Mr. Syed Haider, and Ms. Melissa Geng for their help with this project. We are grateful to Dr. Samir Benlekbir for help with cryo-EM data collection and for advice regarding specimen preparation. M.T. and H.L. were supported by the University of Toronto's Medicine by Design initiative (Mbd New Ideas Award to M.T. and Mbd Postdoctoral fellowship to H.L.). Medicine by Design receives funding from the Canada First Research Excellence Fund (CFREF). D.S. is supported by a Krembil Award. R.A.M. is supported by a CIHR Project Grant (RN292258 – 366017) and an NSERC Discovery grant (RGPIN-2017-06817). This research was supported by the CIFAR Azrieli Global Scholar program (to J.-P.J. and M.T.), the Ontario Early Researcher Awards program (to J.-P.J. and M.T.), and the Canada Research Chairs program (J.L.R., J-P.J., and M.T.). Cryo-EM data were collected at the Toronto High Resolution High Throughput cryo-EM facility, supported by the Canada Foundation for Innovation and Ontario Research Fund.

AUTHOR CONTRIBUTIONS

Conceptualization, methodology: Hunsang Lee, Greg Beilhartz, Iga Kucharska, Daniel Schramek, Jean-Philippe Julien, Roman Melnyk, Mikko Taipale. Investigation: Hunsang Lee, Greg Beilhartz, Iga Kucharska, Huazhu Liang, Mandy Lam, Daniel Schramek, Mikko Taipale. Writing: Hunsang Lee, Greg Beilhartz, Iga Kucharska, Daniel Schramek, Jean-Philippe Julien, Roman Melnyk, Mikko Taipale. Resources: Swetha Raman, Hong Cui, John Rubinstein. Supervision: Jean-Philippe Julien, Roman Melnyk, Mikko Taipale. Funding acquisition: Jean-Philippe Julien, Roman Melnyk, Mikko Taipale.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 6, 2020

Revised: May 6, 2020

Accepted: June 1, 2020

Published: June 25, 2020

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010).

- PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.
- Aktories, K., and Barbieri, J.T. (2005). Bacterial cytotoxins: targeting eukaryotic switches. *Nat. Rev. Microbiol.* **3**, 397–410.
- Aktories, K., and Just, I. (1995). Monoglucosylation of low-molecular-mass GTP-binding Rho proteins by clostridial cytotoxins. *Trends Cell Biol.* **5**, 441–443.
- Aldape, M.J., Bayer, C.R., Bryant, A.E., and Stevens, D.L. (2016). A novel murine model of *Clostridium sordellii* myonecrosis: Insights into the pathogenesis of disease. *Anaerobe* **38**, 103–110.
- Carter, G.P., Awad, M.M., Hao, Y., Thelen, T., Bergin, I.L., Howarth, P.M., Seemann, T., Rood, J.I., Aronoff, D.M., and Lyras, D. (2011). TcsL is an essential virulence factor in *Clostridium sordellii* ATCC 9714. *Infect. Immun.* **79**, 1025–1032.
- Chaves-Olarte, E., Weidmann, M., Eichel-Streiber, C., and Thelestam, M. (1997). Toxins A and B from *Clostridium difficile* differ with respect to enzymatic potencies, cellular substrate specificities, and surface binding to cultured cells. *J. Clin. Invest.* **100**, 1734–1741.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21.
- Chen, P., Tao, L., Wang, T., Zhang, J., He, A., Lam, K.H., Liu, Z., He, X., Perry, K., Dong, M., and Jin, R. (2018). Structural basis for recognition of frizzled proteins by *Clostridium difficile* toxin B. *Science* **360**, 664–669.
- Chen, P., Lam, K.H., Liu, Z., Mindlin, F.A., Chen, B., Gutierrez, C.B., Huang, L., Zhang, Y., Hamza, T., Feng, H., et al. (2019). Structure of the full-length *Clostridium difficile* toxin B. *Nat. Struct. Mol. Biol.* **26**, 712–719.
- Chong, E., Winikoff, B., Charles, D., Agnew, K., Prentice, J.L., Limbago, B.M., Platais, I., Louie, K., Jones, H.E., and Shannon, C.; NCT01283828 Study Team (2016). Vaginal and Rectal *Clostridium sordellii* and *Clostridium perfringens* Presence Among Women in the United States. *Obstet. Gynecol.* **127**, 360–368.
- Chumbler, N.M., Rutherford, S.A., Zhang, Z., Farrow, M.A., Lisher, J.P., Farquhar, E., Giedroc, D.P., Spiller, B.W., Melnyk, R.A., and Lacy, D.B. (2016). Crystal structure of *Clostridium difficile* toxin A. *Nat. Microbiol.* **1**, 1–6.
- Dong, D., Zheng, L., Lin, J., Zhang, B., Zhu, Y., Li, N., Xie, S., Wang, Y., Gao, N., and Huang, Z. (2019). Structural basis of assembly of the human T cell receptor-CD3 complex. *Nature* **573**, 546–552.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
- Fischer, M., Bhatnagar, J., Guarner, J., Reagan, S., Hacker, J.K., Van Meter, S.H., Poukens, V., Whiteman, D.B., Iton, A., Cheung, M., et al. (2005). Fatal toxic shock syndrome associated with *Clostridium sordellii* after medical abortion. *N. Engl. J. Med.* **353**, 2352–2360.
- Geny, B., Khun, H., Fitting, C., Zaranonelli, L., Mazuet, C., Cayet, N., Szatanik, M., Prevost, M.-C., Cavaillon, J.-M., Huerre, M., and Popoff, M.R. (2007). *Clostridium sordellii* lethal toxin kills mice by inducing a major increase in lung vascular permeability. *Am. J. Pathol.* **170**, 1003–1017.
- Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25.
- Gupta, P., Zhang, Z., Sugiman-Marangos, S.N., Tam, J., Raman, S., Julien, J.-P., Kroh, H.K., Lacy, D.B., Murgolo, N., Bekkari, K., et al. (2017). Functional defects in *Clostridium difficile* TcdB toxin uptake identify CSPG4 receptor-binding determinants. *J. Biol. Chem.* **292**, 17290–17301.
- Haklai-Topper, L., Mlechkovich, G., Savariego, D., Gokhman, I., and Yaron, A. (2010). Cis interaction between Semaphorin6A and Plexin-A4 modulates the repulsive response to Sema6A. *EMBO J.* **29**, 2635–2645.
- Hart, T., Tong, A.H.Y., Chan, K., Van Leeuwen, J., Seetharaman, A., Aregger, M., Chandrashekar, M., Hustedt, N., Seth, S., Noonan, A., et al. (2017). Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda)* **7**, 2719–2727.
- Ho, C.S., Bhatnagar, J., Cohen, A.L., Hacker, J.K., Zane, S.B., Reagan, S., Fischer, M., Shieh, W.-J., Guarner, J., Ahmad, S., et al. (2009). Undiagnosed cases of fatal *Clostridium*-associated toxic shock in Californian women of childbearing age. *Am. J. Obstet. Gynecol.* **201**, 459.e1–459.e7.
- Janssen, B.J.C., Robinson, R.A., Pérez-Brangulí, F., Bell, C.H., Mitchell, K.J., Siebold, C., and Jones, E.Y. (2010). Structural basis of semaphorin-plexin signalling. *Nature* **467**, 1118–1122.
- Jongbloets, B.C., and Pasterkamp, R.J. (2014). Semaphorin signalling during development. *Development* **141**, 3292–3297.
- Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., et al. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858.
- Kigel, B., Rabinowicz, N., Varshavsky, A., Kessler, O., and Neufeld, G. (2011). Plexin-A4 promotes tumor progression and tumor angiogenesis by enhancement of VEGF and bFGF signaling. *Blood* **118**, 4285–4296.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
- LaFrance, M.E., Farrow, M.A., Chandrasekaran, R., Sheng, J., Rubin, D.H., and Lacy, D.B. (2015). Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity. *Proc. Natl. Acad. Sci. USA* **112**, 7073–7078.
- Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology* **15**, 554.
- Love, C.A., Harlos, K., Mavaddat, N., Davis, S.J., Stuart, D.I., Jones, E.Y., and Esnouf, R.M. (2003). The ligand-binding face of the semaphorins revealed by the high-resolution crystal structure of SEMA4D. *Nat. Struct. Biol.* **10**, 843–848.
- Lyumkis, D. (2019). Challenges and opportunities in cryo-EM single-particle analysis. *J. Biol. Chem.* **294**, 5181–5197.
- Marr, C.R., Benlekbir, S., and Rubinstein, J.L. (2014). Fabrication of carbon films with ~ 500nm holes for cryo-EM with a direct detector device. *J. Struct. Biol.* **185**, 42–47.
- Neufeld, G., Sabag, A.D., Rabinowicz, N., and Kessler, O. (2012). Semaphorins in angiogenesis and tumor progression. *Cold Spring Harb. Perspect. Med.* **2**, a006718–a006718.
- Nogi, T., Yasui, N., Mihara, E., Matsunaga, Y., Noda, M., Yamashita, N., Toyofuku, T., Uchiyama, S., Goshima, Y., Kumanogoh, A., and Takagi, J. (2010). Structural basis for semaphorin signalling through the plexin receptor. *Nature* **467**, 1123–1127.
- Orrell, K.E., Zhang, Z., Sugiman-Marangos, S.N., and Melnyk, R.A. (2017). *Clostridium difficile* toxins A and B: Receptors, pores, and translocation into cells. *Crit. Rev. Biochem. Mol. Biol.* **52**, 461–473.
- Papatheodorou, P., Zamboglou, C., Genisyuerk, S., Guttenberg, G., and Aktories, K. (2010). Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. *PLoS ONE* **5**, e10673.
- Perez-Brangulí, F., Zagar, Y., Shanley, D.K., Graef, I.A., Chédotal, A., and Mitchell, K.J. (2016). Reverse Signaling by Semaphorin-6A Regulates Cellular Aggregation and Neuronal Morphology. *PLoS ONE* **11**, e0158686.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- Pfeifer, G., Schirmer, J., Leemhuis, J., Busch, C., Meyer, D.K., Aktories, K., and Barth, H. (2003). 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.

- Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296.
- Raman, S., Beiltschmidt, M., To, M., Lin, K., Lui, F., Jmeian, Y., Ng, M., Fernandez, M., Fu, Y., Mascal, K., et al. (2019). Structure-guided design fine-tunes pharmacokinetics, tolerability, and antitumor profile of multispecific frizzled antibodies. *Proc. Natl. Acad. Sci. USA* **116**, 6812–6817.
- Reineke, J., Tenzer, S., Rupnik, M., Koschinski, A., Hasselmayer, O., Schratzenholz, A., Schild, H., and von Eichel-Streiber, C. (2007). Autocatalytic cleavage of *Clostridium difficile* toxin B. *Nature* **446**, 415–419.
- Rubinstein, J.L., and Brubaker, M.A. (2015). Alignment of cryo-EM movies of individual particles by optimization of image translations. *J. Struct. Biol.* **192**, 188–195.
- Schorch, B., Song, S., van Diemen, F.R., Bock, H.H., May, P., Herz, J., Brummelkamp, T.R., Papatheodorou, P., and Aktories, K. (2014). LRP1 is a receptor for *Clostridium perfringens* TpeL toxin indicating a two-receptor model of clostridial glycosylating toxins. *Proc. Natl. Acad. Sci. USA* **111**, 6431–6436.
- Segarra, M., Ohnuki, H., Maric, D., Salvucci, O., Hou, X., Kumar, A., Li, X., and Tosato, G. (2012). Semaphorin 6A regulates angiogenesis by modulating VEGF signaling. *Blood* **120**, 4104–4115.
- Simeon, R., Jiang, M., Chamoun-Emanuelli, A.M., Yu, H., Zhang, Y., Meng, R., Peng, Z., Jakana, J., Zhang, J., Feng, H., and Chen, Z. (2019). Selection and characterization of ultrahigh potency designed ankyrin repeat protein inhibitors of *C. difficile* toxin B. *PLoS Biol.* **17**, e3000311.
- Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501.
- Suto, F., Tsuboi, M., Kamiya, H., Mizuno, H., Kiyama, Y., Komai, S., Shimizu, M., Sanbo, M., Yagi, T., Hiromi, Y., et al. (2007). Interactions between plexin-A2, plexin-A4, and semaphorin 6A control lamina-restricted projection of hippocampal mossy fibers. *Neuron* **53**, 535–547.
- Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796.
- Tao, L., Zhang, J., Meraner, P., Tovaglieri, A., Wu, X., Gerhard, R., Zhang, X., Stallcup, W.B., Miao, J., He, X., et al. (2016). Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. *Nature* **538**, 350–355.
- Tao, L., Tian, S., Zhang, J., Liu, Z., Robinson-McCarthy, L., Miyashita, S.-I., Breault, D.T., Gerhard, R., Oottamasathien, S., Whelan, S.P.J., and Dong, M. (2019). Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to *Clostridium difficile* toxin A entry into cells. *Nat. Microbiol.* **4**, 1760–1769.
- Tawarayama, H., Yoshida, Y., Suto, F., Mitchell, K.J., and Fujisawa, H. (2010). Roles of semaphorin-6B and plexin-A2 in lamina-restricted projection of hippocampal mossy fibers. *J. Neurosci.* **30**, 7049–7060.
- Terwilliger, T.C., Ludtke, S.J., Read, R.J., Adams, P.D., and Afonine, P.V. (2019). Improvement of cryo-EM maps by density modification (Biochemistry).
- Tivol, W.F., Briegel, A., and Jensen, G.J. (2008). An improved cryogen for plunge freezing. *Microsc. Microanal.* **14**, 375–379.
- Visser, J.J., Cheng, Y., Perry, S.C., Chastain, A.B., Parsa, B., Masri, S.S., Ray, T.A., Kay, J.N., and Wojtowicz, W.M. (2015). An extracellular biochemical screen reveals that FLRTs and Unc5s mediate neuronal subtype recognition in the retina. *eLife* **4**, e08149.
- Wong, A.H.M., Tomlinson, A.C.A., Zhou, D., Satkunarajah, M., Chen, K., Sharon, C., Desforges, M., Talbot, P.J., and Rini, J.M. (2017). Receptor-binding loops in alphacoronavirus adaptation and evolution. *Nat. Commun.* **8**, 1735.
- Worzfeld, T., and Offermanns, S. (2014). Semaphorins and plexins as therapeutic targets. *Nat. Rev. Drug Discov.* **13**, 603–621.
- Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*.
- Xu, X.-M., Fisher, D.A., Zhou, L., White, F.A., Ng, S., Snider, W.D., and Luo, Y. (2000). The transmembrane protein semaphorin 6A repels embryonic sympathetic axons. *J. Neurosci.* **20**, 2638–2648.
- Yan, C., Hang, J., Wan, R., Huang, M., Wong, C.C.L., and Shi, Y. (2015). Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**, 1182–1191.
- Yuan, P., Zhang, H., Cai, C., Zhu, S., Zhou, Y., Yang, X., He, R., Li, C., Guo, S., Li, S., et al. (2015). Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B. *Cell Res.* **25**, 157–168.
- Zhang, Z., Park, M., Tam, J., Auger, A., Beilartz, G.L., Lacy, D.B., and Melnyk, R.A. (2014). Translocation domain mutations affecting cellular toxicity identify the *Clostridium difficile* toxin B pore. *Proc. Natl. Acad. Sci. USA* **111**, 3721–3726.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--|-------------------------------------|-----------------------------------|
| Antibodies | | |
| anti-SEMA6A, polyclonal | R&D Systems | Cat#AF1615; RRID: AB_2185995 |
| anti-SEMA6B, polyclonal | R&D Systems | Cat#AF2094; RRID: AB_2254388 |
| anti-FLAG, HRP conjugate | Sigma-Aldrich | Cat#A8592, RRID: AB_439702 |
| Goat Anti-Rabbit IgG Antibody (H+L), Biotinylated | Vector Laboratories | Cat#BA-1000; RRID: AB_2313606 |
| Peroxidase IgG Fraction Monoclonal Mouse Anti-Goat IgG, light chain specific | Jackson ImmunoResearch Laboratories | Cat#205-032-176; RRID: AB_2339056 |
| anti-Hsp90, monoclonal | Santa Cruz | Cat#sc-13119; RRID: AB_675659 |
| Anti-mouse IgG, HRP conjugate | Cell Signaling Technology | Cat#7076S; RRID: AB_330924 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| Human recombinant SEMA6A-ECD-Fc | R&D Systems | Cat#1146-S6-025 |
| Human recombinant SEMA6B-ECD-Fc | R&D Systems | Cat#2094-S6-050 |
| Human recombinant SEMA6C-ECD-Fc | R&D Systems | Cat#2219-S6-050 |
| Human recombinant SEMA6D-ECD-Fc | R&D Systems | Cat#2095-S6-050 |
| Tcsl | This paper | N/A |
| Tcsl ^{4mut} | This paper | N/A |
| TcdB | This paper | N/A |
| TpeL | Borden Lacy lab | N/A |
| TcdA | This paper | N/A |
| PrestoBlue Cell Viability Reagent | Thermo Fisher Scientific | Cat#A13262 |
| DMEM | Thermo Fisher Scientific | Cat#11995073 |
| IMDM | Thermo Fisher Scientific | Cat#12440053 |
| GIBCO™ FreeStyle™ 293 Expression Medium | Thermo Fisher Scientific | Cat#12338026 |
| Lipofectamine® 2000 Transfection Reagent | Thermo Fisher Scientific | Cat#11668019 |
| FectoPRO™ DNA Transfection Reagent | VWR | Cat#10118-444 |
| Puromycin | Wisent | Cat#450-162-XL |
| Hexadimethrine bromide | Sigma-Aldrich | Cat#107689-10G |
| Blasticidin S HCl | Thermo Fisher Scientific | Cat#A1113903 |
| FBS, Qualified | Thermo Fisher Scientific | Cat#12483-020 |
| TurboFectin Transfection Reagent | Origene Technologies | Cat#TF81001 |
| Inositol hexakisphosphate | Sigma-Aldrich | Cat#P0109 |
| Dithiothreitol | BioBasic Inc. | Cat#B0058 |
| Penicillin-Streptomycin | Thermo Fisher Scientific | Cat#15140122 |
| Glutaraldehyde | Sigma-Aldrich | Cat#G7526-10ML |
| Critical Commercial Assays | | |
| Nano-Glo Luciferase Assay System | Promega | Cat#N1130 |
| Qubit dsDNA BR Assay Kit | Thermo Fisher Scientific | Cat#Q32850 |
| Bacillus megaterium Protein Production System | MoBiTec | Cat#BMEG02 |
| Presto™ Mini Plasmid Kit | FroggaBio | Cat # PDH300 |
| Geneaid™ Midi Plasmid Kit | FroggaBio | Cat#PI025 |

(Continued on next page)

Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|-----------------------------------|-----------------------|
| GenepHlow™ Gel/PCR Kit | FroggaBio | Cat#DFH300 |
| QIAamp DNA Blood Maxi Kit | QIAGEN | Cat#51194 |
| DpnI | NEB | Cat#R0176L |
| CellTiter-Glo Luminescent Cell Viability Assay | Promega | Cat#G7572 |
| alamarBlue™ Cell Viability Reagent | Thermo Fisher Scientific | Cat#DAL1100 |
| NEBNext Ultra II Q5 Master Mix | NEB | Cat#M0544L |
| Gateway™ BP Clonase™ Enzyme Mix | Thermo Fisher Scientific | Cat#11789013 |
| Gateway™ LR Clonase™ Enzyme mix | Thermo Fisher Scientific | Cat#11791019 |
| Ni-NTA biosensors | ForteBio | Cat#18-5102 |
| Deposited Data | | |
| CRISPR screen data | This paper | Table S1 |
| Tcsl-SEMA6A map | This paper | EMDB: EMD-21898 |
| SEMA6A map | This paper | EMBD: EMD-21899 |
| Tcsl-SEMA6A structure | This paper | PDB: 6WTS |
| Experimental Models: Cell Lines | | |
| Vero Cells | ATCC | CCL-81 |
| Hulec-5a cells | ATCC | CRL-3244 |
| Hap1 cells | Horizon Discovery | C631 |
| HeLa Kyoto cells | Iain Cheeseman lab | N/A |
| FreeStyle™ 293-F Cells | Thermo Fisher Scientific | Cat# R79007 |
| Oligonucleotides | | |
| Oligonucleotides and synthetic DNA used in the study | This paper | Table S4 |
| Recombinant DNA | | |
| pHIS1522-TcsL | This paper | N/A |
| pHIS1522-TcsL(C1433D/Y1596R) | This paper | N/A |
| pHIS1522-TcsL(C1433D/Y1596R/I1434K/A1486S) | This paper | N/A |
| pET-SUMO-StrepTag-TEV-TcsL(1285-1804)-TEV-HIS | This paper | N/A |
| pET-SUMO-StrepTag-TEV-TcsL(1285-1804)FBD-TEV-HIS | This paper | N/A |
| pET-SUMO-StrepTag-TEV-TcdB(1285-1804)-TEV-HIS | This paper | N/A |
| pcDNA3.1-SEMA6A (M109D)-3xFLAG | This paper | N/A |
| TKOv3 gRNA-Cas9 library (Toronto KnockOut (TKO) CRISPR Library - Version 3) | Hart et al., 2017 | Addgene plasmid 90294 |
| pHLsec-human-SEMA6A-Strep | This paper | N/A |
| pDONR221 SEMA6A | This paper | N/A |
| pDONR221 SEMA6B | This paper | N/A |
| pDONR221 SEMA6C | This paper | N/A |
| pDONR221 SEMA6D | This paper | N/A |
| pcDNA3.1 SEMA6A-3xFLAG-V5 | This paper | N/A |
| pcDNA3.1 SEMA6B-3xFLAG-V5 | This paper | N/A |
| pcDNA3.1 SEMA6C-3xFLAG-V5 | This paper | N/A |

(Continued on next page)

Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--------------------------------------|--|---|
| pcDNA3.1 SEMA6D-3xFLAG-V5 | This paper | N/A |
| pLenti6.2 SEMA6A-3xFLAG-V5 | This paper | N/A |
| pLenti6.2 SEMA6B-3xFLAG-V5 | This paper | N/A |
| pLenti6.2 SEMA6C-3xFLAG-V5 | This paper | N/A |
| pLenti6.2 SEMA6D-3xFLAG-V5 | This paper | N/A |
| pX459 SEMA6A gRNA-Cas9 | This paper | N/A |
| pX459 SEMA6B gRNA-Cas9 | This paper | N/A |
| pX459 SEMA6C gRNA-Cas9 | This paper | N/A |
| pX459 SEMA6D gRNA-Cas9 | This paper | N/A |
| pX459 UGP2 gRNA-Cas9 | This paper | N/A |
| pLentiCRISPRv2 SEMA6A gRNA-Cas9 | This paper | N/A |
| pLentiCRISPRv2 SEMA6B gRNA-Cas9 | This paper | N/A |
| pLentiCRISPRv2 SEMA6C gRNA-Cas9 | This paper | N/A |
| pLentiCRISPRv2 SEMA6D gRNA-Cas9 | This paper | N/A |
| pLentiCRISPRv2 UGP2 gRNA-Cas9 | This paper | N/A |
| pCMV-VSV-G | Stewart et al., 2003 | Addgene plasmid 8454 |
| pPAX2 | Didier Trono lab | Addgene plasmid 12260 |
| Sema6a.a-Fc-His | Visser et al., 2015 | Addgene plasmid 72163 |
| Sema6c-Fc-His | Visser et al., 2015 | Addgene plasmid 72167 |
| pHLsec-FZD7(42-179)-mVenus | Raman et al., 2019 | N/A |
| Software and Algorithms | | |
| Octet Data Analysis Software 9.0.0.6 | ForteBio | https://www.fortebio.com/products/octet-systems-software |
| GraphPad Prism | GraphPad Software | https://www.graphpad.com/ |
| EPU | ThermoFisher Scientific | https://www.fei.com/software/e-pu-automated-single-particles-software-for-life-sciences/ |
| SBGrid | SBGrid Consortium | https://sbgrid.org/ |
| Phyre2 | Kelley et al., 2015 | http://www.sbg.bio.ic.ac.uk/~phyre2/ |
| cryoSPARCv2 | Punjani et al., 2017 | https://cryosparc.com/ |
| Phenix | Adams et al., 2010 | https://www.phenix-online.org/ |
| UCSF Chimera | Pettersen et al., 2004 | https://www.cgl.ucsf.edu/chimera/ |
| UCSF ChimeraX | Goddard et al., 2018 | https://www.cgl.ucsf.edu/chimerax/ |
| Coot | Emsley et al., 2010 | https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/ |
| PyMOL | The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. | https://pymol.org/2/#products |
| PDBePISA | Krissinel and Henrick, 2007 | https://www.ebi.ac.uk/pdbe/pisa/ |
| Other | | |
| Homemade holey gold grids | Marr et al., 2014 | N/A |

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Mikko Taipale (mikko.taipale@utoronto.ca).

Materials Availability

Plasmids generated in this study are available from Addgene or from the authors with a Material Transfer Agreement.

Data and Code Availability

The cryo-EM maps of the TcsL1285-1804/SEMA6A complex and unliganded SEMA6A have been deposited to the EMDB under accession numbers EMD-21898 and EMD-21899, respectively, and the TcsL1285-1804/SEMA6A model to the PDB under accession number 6WTS.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Animal husbandry, ethical handling of mice and all animal work were carried out according to guidelines approved by Canadian Council on Animal Care and under protocols approved by the Centre for Phenogenomics Animal Care Committee (23-0350H). Female C67/Bl6J mice, 8 weeks of age, were intraperitoneally injected with TcsL together with mouse Sema6a-Fc, mouse Sema6c-Fc, or BSA. Treatment-naive, littermate female mice were group-housed and randomly assigned to experimental groups. In brief, 15 ng of TcsL was mixed with a 1000-fold molar excess of each protein at 4°C on a rotator for one h prior to injection. Four h after intoxication, animals were euthanized and fluid present in the thoracic cavity was collected for analysis. For immunohistochemistry, lung tissue was fixed 4% paraformaldehyde for 48 h, processed through an ethanol series to xylene and then paraffin using a Leica ASP300 automatic tissue processor and embedded in paraffin wax using Leica Histocore Arcadia H. Samples were sectioned at 4.5 µm with a Leica RM2255 semi-automatic microtome. Slides for IHC were dewaxed, rehydrated and sections were treated with 3% H₂O₂ in PBS to kill endogenous peroxidase activity. Antigen retrieval was performed using a 10 mM Tris/1mM EDTA/0.05% Tween 20 (pH 9) buffer solution in a microwave for 15 min. Primary antibody was applied at 4°C overnight using SEMA6A (R&D, AF1615, goat, 1:100 dilution) and SEMA6B (R&D, AF2094, goat, 1:100 dilution). Rabbit anti-goat secondary (Vector Labs BA-1000) was applied for 30 min followed by ABC reagent (Vector Labs PK-6100) for 25 min and developed with DAB (Vector Labs SK-4100) for 4 min or less. Tissue was counterstained with Harris Hematoxylin (HHHS-128, Sigma) for 8 min and mounted with Shur Mount (Electron Microscopy Sciences 17991-01). Stained slides were digitized at 40x using a Nanozoomer 2.0 HT (Hamamatsu Photonics) and images of stained samples were analyzed using NPD.view2 software (Hamamatsu U12388-01).

Cell lines

All cell lines were grown at 37°C in 5% CO₂. Vero-NlucP cells stably express a destabilized nanoluciferase (Promega) were generated by infecting Vero cells (ATCC CCL-81; female cell line) with a lentivirus expressing nanoluciferase with a C-terminal PEST sequence. Stable clones were selected by puromycin and limiting dilution. These cells are referred to as Vero-NlucP. Cells were grown in DMEM with 10% FBS and 1% penicillin/streptomycin. Hulec-5a human lung microvascular endothelial cells were purchased from ATCC (CRL-3244; male) and cultured in Vascular Cell Basal Medium (ATCC PCS-100-030) with the Endothelial Cell Growth Kit-VEGF (ATCC PCS-100-041). Hap1 cells were purchased from Horizon Discovery (C631; male) and cultured in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% FBS and 1% penicillin/streptomycin. HeLa Kyoto cells (female) used in this study were a gift from the Cheeseman laboratory (Whitehead Institute, Cambridge, MA) and cultured in DMEM with 10% FBS and 1% penicillin/streptomycin. HEK293F cells used in this study were purchased from Thermo Fisher Scientific (R79007; female) and cultured in GIBCO FreeStyle 293 Expression Medium. Cell lines obtained from commercial sources were not further authenticated. HeLa Kyoto cells were authenticated with STR profiling (GenePrint 24 System, Promega) at The Centre for Applied Genomics, The Hospital for Sick Children, Toronto.

METHOD DETAILS

Genome-wide CRISPR screen

50 million Hap1 cells were infected with TKOv3 library at 200-fold library coverage at multiplicity of infection (MOI) of 0.3. The infected cells were selected with 2 µg/mL puromycin for 3 days (T0) and maintained at a minimum of 100-fold gRNA library coverage at any time point. Cells were passaged until the fourth day (T4) to allow sufficient time for protein turnover. 7 million cells were seeded into two 10 cm plates per condition and allowed to recover overnight after trypsinization. At T5, TcsL was added at a final concentration of 0.1 nM or 1 nM, and the cells were further incubated for 48 h. At T7, cells were washed with 1xPBS and allowed to repopulate in normal growth media (IMDM/10% FBS). The untreated population of cells (negative controls) were passaged every three days in parallel. The medium of toxin treated cells were replenished every 3 days. Once surviving colonies were visible, cells were trypsinized and re-seeded to facilitate repopulation. Once the toxin treated cells reached 100% confluence, the cells were collected with the corresponding untreated population. Genomic DNA was extracted from the frozen pellets using QIAamp DNA Blood Maxi Kit (QIAGEN) following the manufacturer's protocol.

Next-generation sequencing library preparation

The gRNA sequences were PCR amplified from the extracted genomic DNA. Each amplified sample was then barcoded and processed on Illumina Next-seq high-output mode at a read depth of at least 5 million reads per sample. MAGeCK software (Li et al., 2014) was used to generate rankings for positively enriched genes.

CRISPR screen validation

The gRNAs targeting SEMA6A, SEMA6B, SEMA6C, SEMA6D and UGP2 were chosen from TKOv3 library and were cloned into pX459 (Addgene plasmid #62988; a gift from Feng Zhang) and lentiCRISPRv2 (Addgene plasmid #52961; a gift from Feng Zhang). Plasmids were transfected into HAP1 cells seeded on a 6-well plate with Turbofectin (OriGene) following the manufacturer's protocol. One day post-transfection, the medium was changed to medium containing 2 $\mu\text{g}/\text{mL}$ puromycin and was further selected for 3 days. Cells were washed and moved to a 10 cm plate with fresh growth medium with no antibiotics.

Wild-type and knock-out cells were seeded on a 96-well plate a day before toxin application at < 40% confluency. Toxins were serially diluted in 1xPBS with 10% glycerol before applying to cells. Cells were incubated with toxins for 24 to 48 h. Cell viability was measured either using AlamarBlue dye (Invitrogen) or CellTiter-Glo reagent (Promega) following the manufacturer's protocol.

For ectopic expression of SEMA6A, SEMA6B, SEMA6C, and SEMA6D, full-length genes were cloned into a pLenti6.2 plasmid with a C-terminal FLAG and V5 tags. Lentiviruses were packaged by transfecting the lentiviral plasmid with VSV-G and psPAX packaging plasmids into HEK293T cells. Virus-containing media was collected three days post-transfection. The media containing packaged virus was added to HAP1-SEMA6A knock-out cells in a 1:10 ratio and selected on 10 $\mu\text{g}/\text{mL}$ blasticidin for 7 to 10 days. The expression of the target construct was confirmed by SDS-PAGE and western blotting with an anti-FLAG-HRP antibody (Sigma).

Site-directed mutagenesis

The mutation in SEMA6A was generated with oligonucleotides containing the mutation of interest and a standard site-directed mutagenesis protocol, followed by digestion of the template plasmid with DpnI (NEB). Mutations in TcsL(1285-1804) and TcsL were made by synthesizing the mutant fragment and cloning by NebBuilder (NEB).

Vero cell experiments

Vero-NlucP cells were plated at 4000 cells/well in a white-walled clear bottom 96-well plate (Corning) and incubated overnight to attach. For toxicity experiments, TcsL was added at the indicated concentrations and incubated for 24 or 48 h (indicated in the figures) at 37°C, 5% CO₂. Nanoluciferase levels were measured using the Nano-Glo Luciferase Assay System (Promega) as per the manufacturer's instructions. For SEMA6 competition experiments, recombinant SEMA6 proteins were added immediately before TcsL at the indicated concentrations and the assay proceeded as above. Cell viability experiments with Vero cells (without Nluc reporter) were conducted exactly as above, except 10 μl PrestoBlue reagent was added to all wells, and fluorescence was read after 4 h on a plate reader (Molecular Devices M5e).

CPD Cleavage assay

A total of 3 μg TcsL or TcsL^{4mut} was incubated with the indicated amounts of inositol hexakisphosphate in a buffer containing 20 mM Tris pH 8, 300 mM NaCl, 5 mM DTT for 3 h at 37°C. Samples were analyzed via SDS-PAGE.

Hulec-5a cells

Hulec-5a human lung microvascular endothelial cells were plated in white-walled clear-bottom 96-well plates at 4,000 cells/well and left to attach overnight. TcsL and/or SEMA6 proteins were added as indicated. For cell viability experiments, cells were incubated for 48 h at 37°C, 5% CO₂ and cell viability was measured using CellTiter-Glo (Promega) as per the manufacturer's instructions. For cell rounding experiments, media was replaced by complete media containing 1 μM CellTracker Orange CMRA (Molecular Probes). After 60 min, excess dye was removed by media exchange with complete media. Cells were then incubated with 5 pM TcsL and various concentrations of SEMA6 proteins as indicated. Cells were incubated for 10 h before imaging. CellTracker-labeled cells were evaluated on a Cellomics ArrayScan VTI HCS reader (Thermo Scientific) using the target acquisition mode, a 10x objective, and a sample rate of 150 objects per well. After recording all image data, the cell rounding and shrinking effects of TcsL intoxication were calculated using the cell rounding index (CRI), a combined measure of the length-to-width ratio (LWR) and area parameters. Dose response curves were plotted and fit to a sigmoidal function (variable slope) to determine EC₅₀ using Prism software (GraphPad Software).

Recombinant proteins

Recombinant human SEMA6A, SEMA6B, SEMA6C, and SEMA6D ectodomains fused to Fc domain were purchased from R&D Systems. Plasmid pHis1522 encoding His-tagged TcsL was synthesized and codon optimized for *Bacillus megaterium* (Genscript). To express and isolate recombinant TcsL, transformed *B. megaterium* was inoculated into LB containing tetracycline and grown to an A₆₀₀ of 0.7, followed by overnight xylose induction at 37°C. Bacterial pellets were collected, resuspended with 20 mM Tris pH 8, 0.5 M NaCl, and passed twice through an EmulsiFlex C3 microfluidizer (Avestin) at 15,000 psi, then clarified by centrifuging at 18,000 x g for 20 min. TcsL was purified by nickel affinity chromatography followed by anion exchange chromatography using HisTrap FF and HiTrap Q columns (GE Healthcare), respectively. Fractions containing TcsL were verified by SDS-PAGE, then pooled and diafiltered with a 100,000 MWCO ultrafiltration device (Corning) into 20 mM Tris pH 7.5, 150 mM NaCl. Finally, glycerol was added to 15% v/v, the protein concentration was estimated by A₂₈₀ (using extinction coefficient 300205), and stored at -80°C.

TcsL fragments (1285-1804) and all mutants thereof were synthesized (IDT) and cloned into pET Champion vector with an N-terminal 6xHIS-SUMO-Strep-TEV sequence and a C-terminal TEV-6xHIS. Positive clones were verified by sequencing. NiCo21 (DE3) competent *E. coli* (NEB) were transformed and inoculated in LB media with kanamycin and grown to an A₆₀₀ of 0.6. Protein

expression was induced by the addition of 1 mM IPTG for 4 h at 23°C. Cells were pelleted and protein was purified similarly to TcsL, except after nickel affinity chromatography, they were passed through a size exclusion column and eluted into 20 mM Tris pH 8, 150 mM NaCl. Mouse Sema6A-Fc and Sema6c-Fc expression constructs were a gift from Woj Wojtowicz (Addgene plasmids 72163 and 72167, respectively) and human SEMA6A was gene-synthesized (GeneArt). All SEMA6 constructs were transiently expressed in suspension HEK293F cells. Proteins were purified by either Protein A affinity or by using ion exchange chromatography (MonoQ 10/100 GL column, GE Healthcare) with a 0 mM–1 M NaCl gradient. These steps were followed by gel filtration chromatography (Superdex 200 Increase, GE Healthcare) in 20 mM Tris pH 9.0 and 150 mM NaCl buffer.

Recombinant FZD7 was expressed and purified as previously described (Raman et al., 2019). Briefly, a human FZD7 (residues 42–179)-mVenus construct was transiently expressed in suspension HEK293F cells and purified using Ni-NTA affinity chromatography. The protein was eluted with an increasing gradient of imidazole with a maximum concentration of 500 mM, in a buffer containing 20 mM Tris pH 8.0, 500 mM NaCl and 5% (v/v) glycerol. This step was followed by gel filtration chromatography (Superdex 200 Increase, GE Healthcare) in 20 mM Tris pH 8.0, 150 mM NaCl buffer. To test the formation of toxin-receptor complexes, TcsL1285–1804, TcsL (FBD)1285–1804 and TcdB1285–1804 (positive control) were mixed with 5-fold molar excess of SEMA6A or FZD7 and incubated at room temperature for 30 min. This was followed by gel filtration chromatography in 20 mM Tris pH 9.0, 150 mM NaCl buffer.

Biolayer interferometry (BLI)

To determine the binding kinetics for recombinant His-tagged TcsL1285–1804 and human SEMA6A, BLI experiments were performed on an Octet Red96 instrument (FortéBio) at 25°C. All proteins were diluted in kinetics buffer (1xPBS pH 7.4, 0.01% (w/v) BSA, 0.002% (v/v) Tween-20). After 10–30 µg/mL of His-tagged TcsL1285–1804 was immobilized on Ni-NTA sensors, the baseline was established for 60 s. Subsequently, the loaded biosensors were dipped into wells containing serial dilutions of human SEMA6A, to determine the rate of association. Sensors were then dipped back into kinetics buffer to establish the dissociation rate. The curves were fitted to a 1:1 binding model and the apparent dissociation constant (K_D) was evaluated using FortéBio's Data Analysis software 9.0. Reported values represent the average of four independent experiments with standard error of the mean (SEM).

TcsL-SEMA6A complex formation and glutaraldehyde crosslinking

TcsL_{1285–1804} was combined with excess of SEMA6A and purified by gel filtration chromatography (Superdex 200 Increase, GE Healthcare) in 20 mM HEPES pH 7.0 and 50 mM NaCl. Fractions containing the complex were concentrated to 0.12 mg/mL and proteins were crosslinked by addition of 0.05% (v/v) glutaraldehyde (Sigma Aldrich) and incubated at room temperature for 60 min. The reaction was stopped by addition of Tris-HCl pH 7.0 to a final concentration of 50 mM. Subsequently, the complex was concentrated to 0.9 mg/mL, spun down for 30 min at 14,500 x g and directly used for cryo-EM grid preparation.

Cryo-EM data collection and image processing

Homemade holey gold grids (Marr et al., 2014) were glow discharged in air for 15 s before use. TcsL-SEMA6A (3 µl, 0.9 mg/mL) was applied to grids, blotted for 12 s, and frozen in a mixture of liquid ethane and propane (Tivol et al., 2008) using a modified FEI Vitrobot (maintained at 4°C and 100% humidity). Data collection was performed on a Thermo Fisher Scientific Titan Krios G3 operated at 300 kV with a Falcon 4 camera automated with the EPU software. A nominal magnification of 75,000 (calibrated pixel size of 1.03 Å) and defocus range between 1.6 and 2.2 µm were used for data collection. Exposures were fractionated as movies of 30 frames with an exposure rate of 5 electrons/pixel/second and total exposure of 45.2 electrons/Å².

A total of 4,581 raw movies were obtained for the TcsL-SEMA6A sample. Image processing was carried out in cryoSPARC v2 (Punjani et al., 2017). Motion correction was performed with Patch Motion algorithm and CTF parameters were estimated from the average of aligned movie frames with Patch CTF. 3,896,638 particle images were selected by template matching and individual particle images were corrected for beam-induced motion using local motion algorithm (Rubinstein and Brubaker, 2015) within cryoSPARC v2. *Ab initio* structure determination and classification revealed that ~50% particle images corresponded to a SEMA6A dimer and the remaining particles to a SEMA6 dimer-TcsL complex. The overwhelming majority of particles for the SEMA6A-TcsL complex had one TcsL molecule bound to the SEMA6A dimer, with no evidence of two TcsL molecules bound to the dimer that could be clearly identified for this cross-linked sample.

Since the TcsL-SEMA6A particles exhibited partially preferred orientation on vitrified cryo-EM grids, we collected an additional dataset of 2,843 images under 40° tilt (Lyumkis, 2019; Tan et al., 2017). Data collection and image processing parameters were identical with untilted dataset. 1,465,298 particle images were selected by template matching and merged with the previously obtained dataset of 3,896,638 particle images. Multiple rounds of heterogeneous refinement followed by non-uniform refinement resulted in a 3.3 Å resolution map of the TcsL-SEMA6A complex (179,188 particle images) and a 3.1 Å resolution map of unliganded SEMA6A (281,207 particle images). To improve the quality of the final TcsL-SEMA6A map, density modification was used in Phenix (Adams et al., 2010; Terwilliger et al., 2019).

Model building

An initial model for TcsL was created using the Phyre2 server (Kelley et al., 2015) with the TcdB crystal structure (PDB: 6C0B) as a reference. The atomic coordinates of SEMA6A dimer (PDB: 3OKW) were manually fitted into the density map using UCSF Chimera

(Pettersen et al., 2004) to generate a starting model, followed by manual rebuilding using Coot (Emsley et al., 2010). All models were refined using the phenix.real_space_refine (Adams et al., 2010) with secondary structure and geometry restraints. The final models were evaluated by MolProbity (Chen et al., 2010). Statistics of the map reconstruction and model refinement are presented in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

TcsL intoxication experiments were conducted in triplicate unless otherwise noted. All figures show arithmetic means of replicate experiments and error bars denote standard deviations. Fisher's exact test was used to assess the difference between receptor-binding domain conservation and the non-receptor binding surface residues.

Supplemental Figures

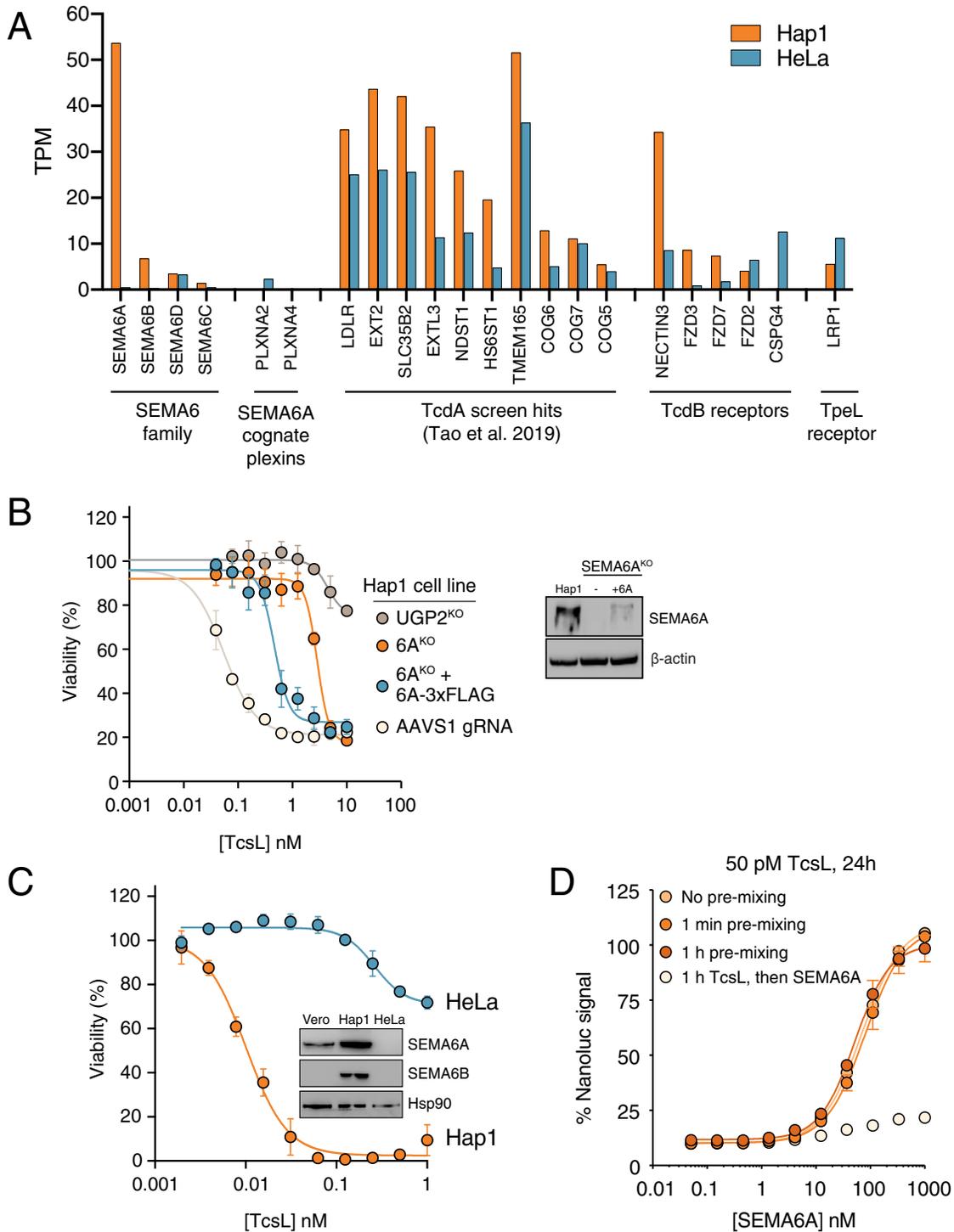


Figure S1. Validation of SEMA6A and SEMA6B as host factors required for TcsL intoxication, related to Figure 1 and Figure 2
 (A) Expression of SEMA6 family genes, the cognate SEMA6A/6B ligands Plexin A2 and Plexin A4, and known clostridial toxin receptors and host cell factors in Hap1 and HeLa cells based on Human Protein Atlas (proteinatlas.org).

(legend continued on next page)

(B) Left, Sensitivity of SEMA6A and UGP2 knockout cells to TcsL. SEMA6A and UGP2 knockout cells were generated with CRISPR/Cas9. SEMA6A-3xFLAG was ectopically expressed in SEMA6A knockout cells by lentiviral infection. Data (n = 3) are represented as mean \pm standard deviation. Right, SEMA6A expression in wild-type Hap1 cells, SEMA6A^{KO} cells, and in SEMA6A^{KO} cells ectopically expressing SEMA6A-3xFLAG.

(C) Hap1 and HeLa cells were treated with increasing concentrations of TcsL and cell viability measured 24 h later. Data (n = 3) are represented as mean \pm standard deviation. Inset, expression of SEMA6A in Vero, Hap1, and HeLa cells was assessed by western blot.

(D) SEMA6A ectodomain protects Vero cells from TcsL toxicity only when added simultaneously with the toxin. SEMA6A and TcsL were added to Vero cells simultaneously or after 1-min or 1 h pre-incubation. Alternatively, TcsL was added for 1 h prior to treatment with SEMA6A. Data (n = 2) are represented as mean \pm standard deviation

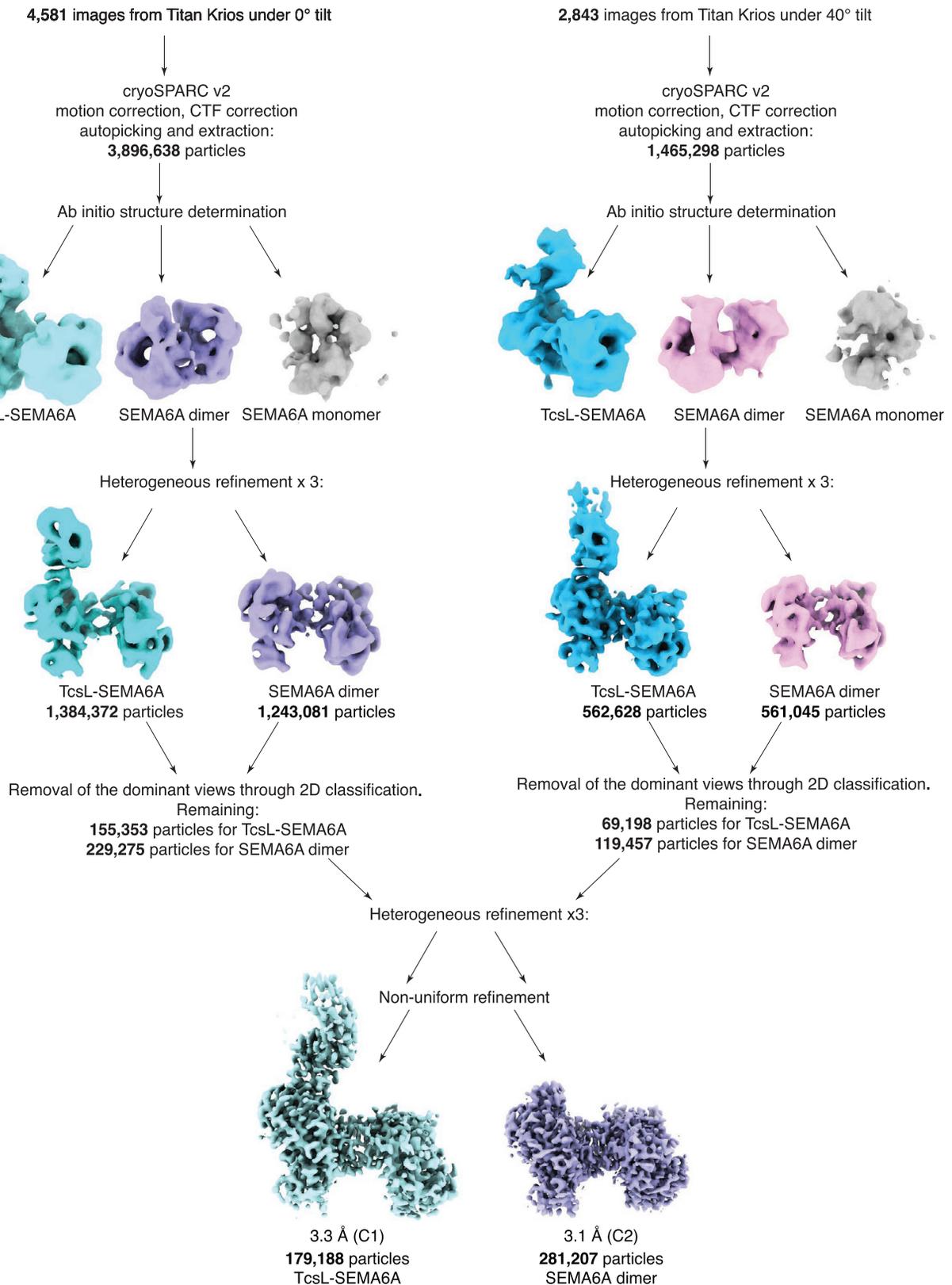


Figure S2. Cryo-EM data processing workflow in cryoSPARC v2, related to Figure 4 and Figure 5

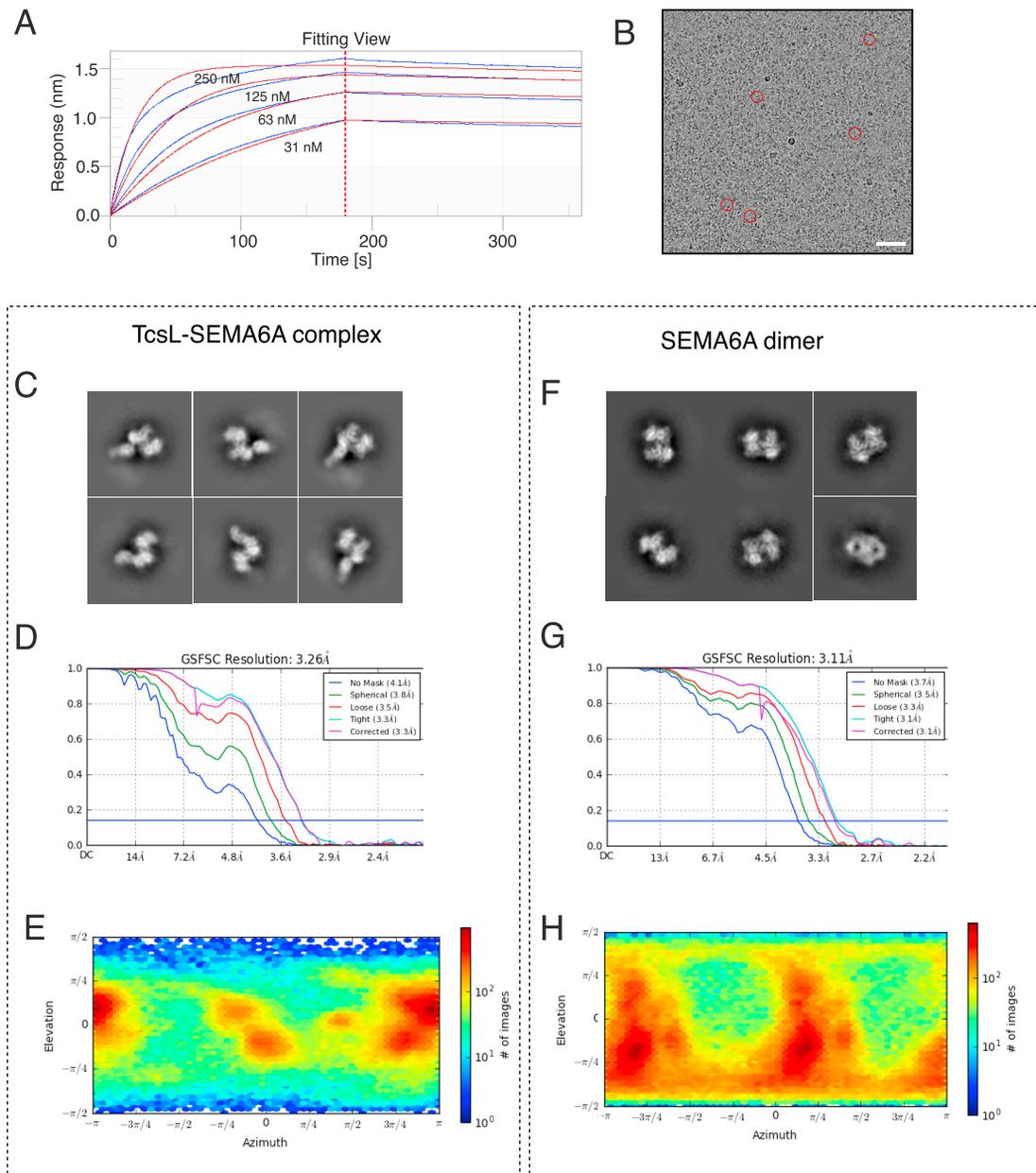


Figure S3. Biochemical and cryo-EM analysis of the SEMA6A/TcsL complex, related to Figure 4 and Figure 5

(A) Representative binding curves for the TcsL₁₂₈₅₋₁₈₀₄/SEMA6A interaction. In this analysis, His-tagged TcsL₁₂₈₅₋₁₈₀₄ was immobilized on the Ni-NTA biosensor. The average apparent binding affinity from four independent experiments is 2.4 ± 0.8 nM. The data (blue) were fitted using a 1:1 binding model (red).

(B) An example of cryo-EM micrograph. Scale bar is 50 nm.

(C) Selected 2D class averages of the TcsL/SEMA6A complex.

(D) Gold standard Fourier shell correlation (GSFSC) curve of the final 3D non-uniform refinement of the TcsL/SEMA6A complex in cryoSPARC v2.

(E) Viewing direction distribution of the TcsL/SEMA6A data.

(F) Selected 2D class averages of SEMA6A dimer.

(G) GSFSC curve of the final 3D non-uniform refinement of the SEMA6A dimer in cryoSPARC v2.

(H) Viewing direction distribution of the SEMA6A dimer data.

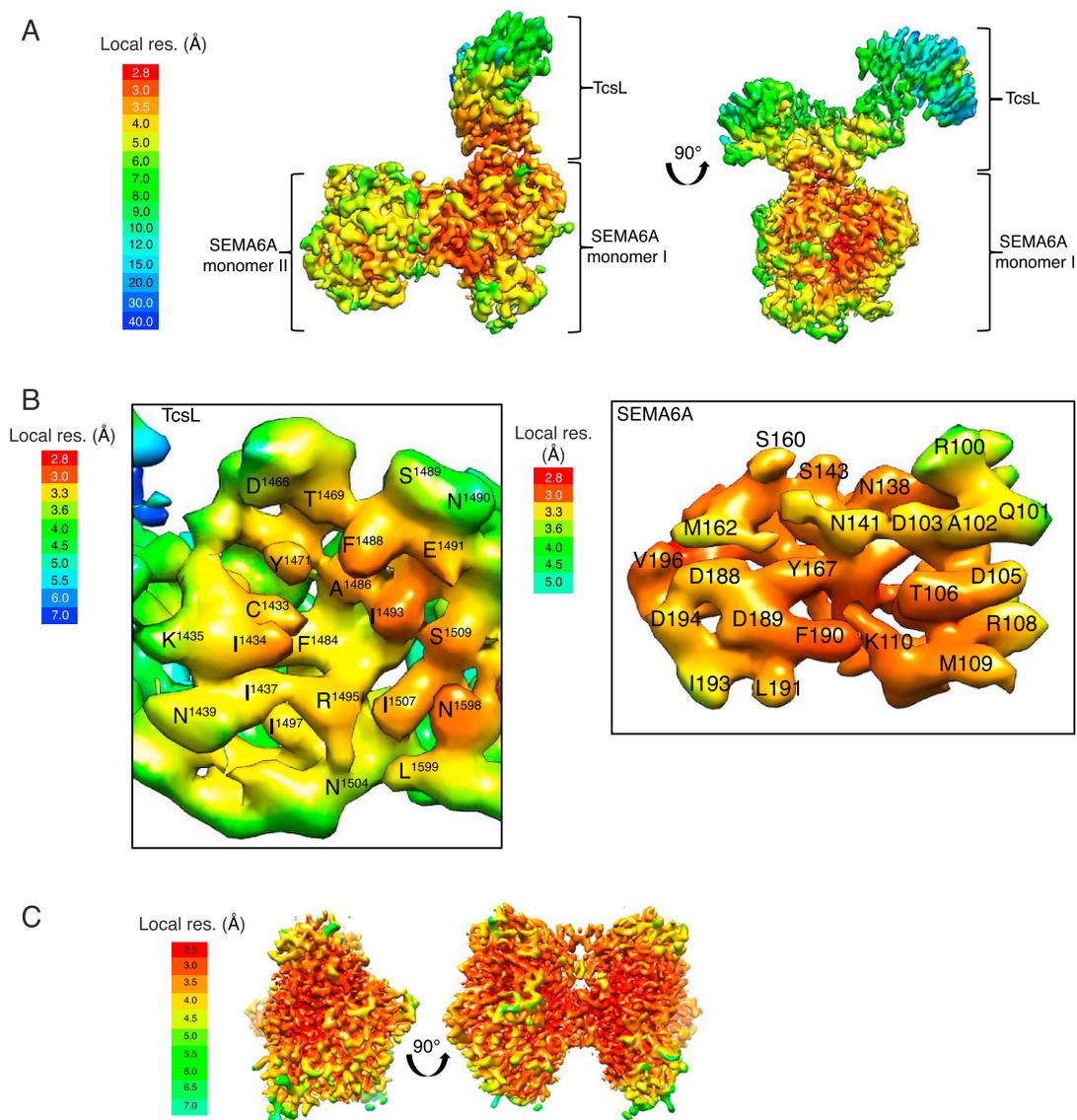


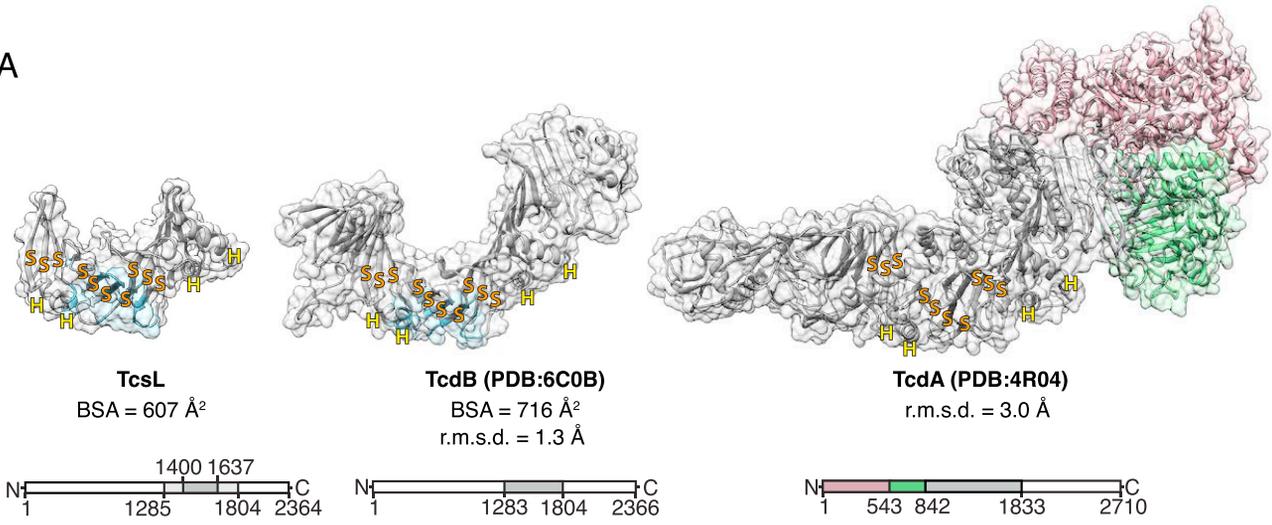
Figure S4. Local resolution (Å) plotted on the surface of cryo-EM map of the SEMA6A/TcsL complex and SEMA6A dimer, related to [Figure 4](#) and [Figure 5](#)

(A) Local resolution of the SEMA6A/TcsL ranges from 2.8 Å at the core of the SEMA6A to > 30 Å at the flexible terminus of TcsL.

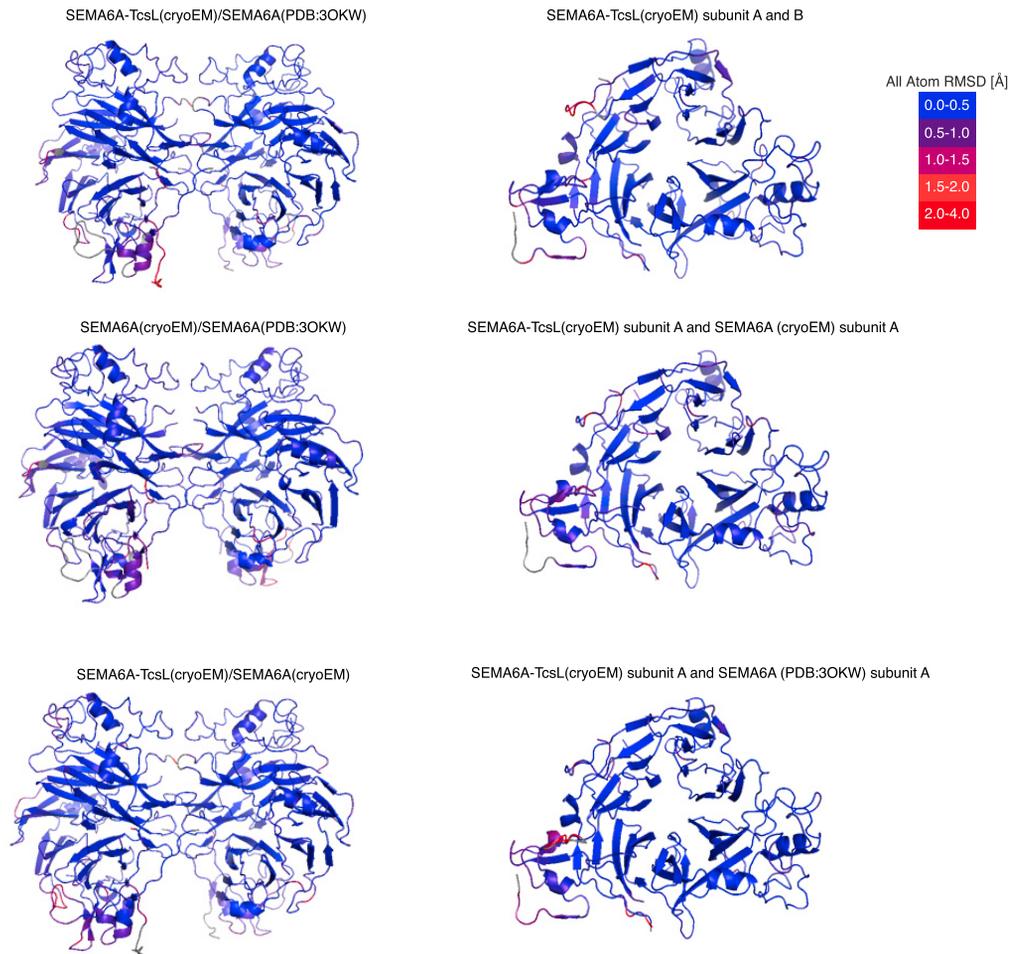
(B) Local resolution plotted on the surface of TcsL (left) and SEMA6A (right) binding interfaces.

(C) Local resolution plotted on the surface of the SEMA6A cryo-EM map.

A



B



(legend on next page)

Figure S5. Comparison of published structures of clostridial toxins and SEMA6A, related to Figure 4 and Figure 5

(A) Left, Cryo-EM structure of TcsL toxin fragment (residues in construct: 1283-1804, residues resolved and modeled: 1400-1637); middle, X-ray structure of TcdB toxin fragment (residues in the model 1283-1804; PDB ID: 6C0B; [Chen et al., 2018](#)); right, X-ray structure of TcdA toxin fragment (residues in the model 1-1832; PDB ID: 4R04; [Chumbler et al., 2016](#)). Binding interfaces of TcsL and TcdB to their respective receptors are highlighted in light blue. "H" and "S" letters denote the positions of selected α helices and β strands conserved in all three toxin structures. The glucosyltransferase domain is highlighted in pink, the autoprotease domain is highlighted in green and the delivery domains are highlighted in gray.

(B) Comparison of SEMA6A in different structures. All atom RMSD values were calculated using Pymol and are plotted on the surface of SEMA6A dimers (left) and monomers (right).

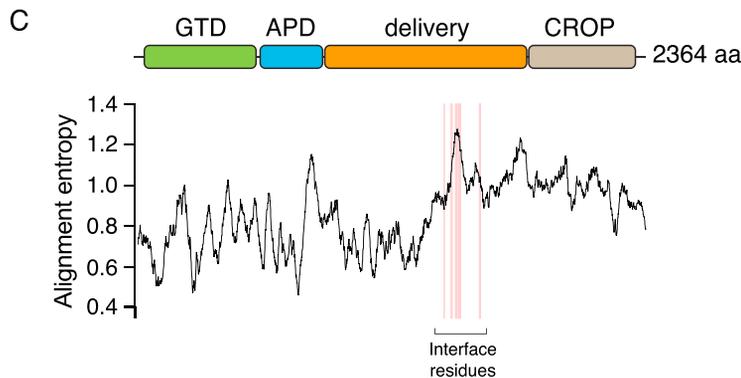
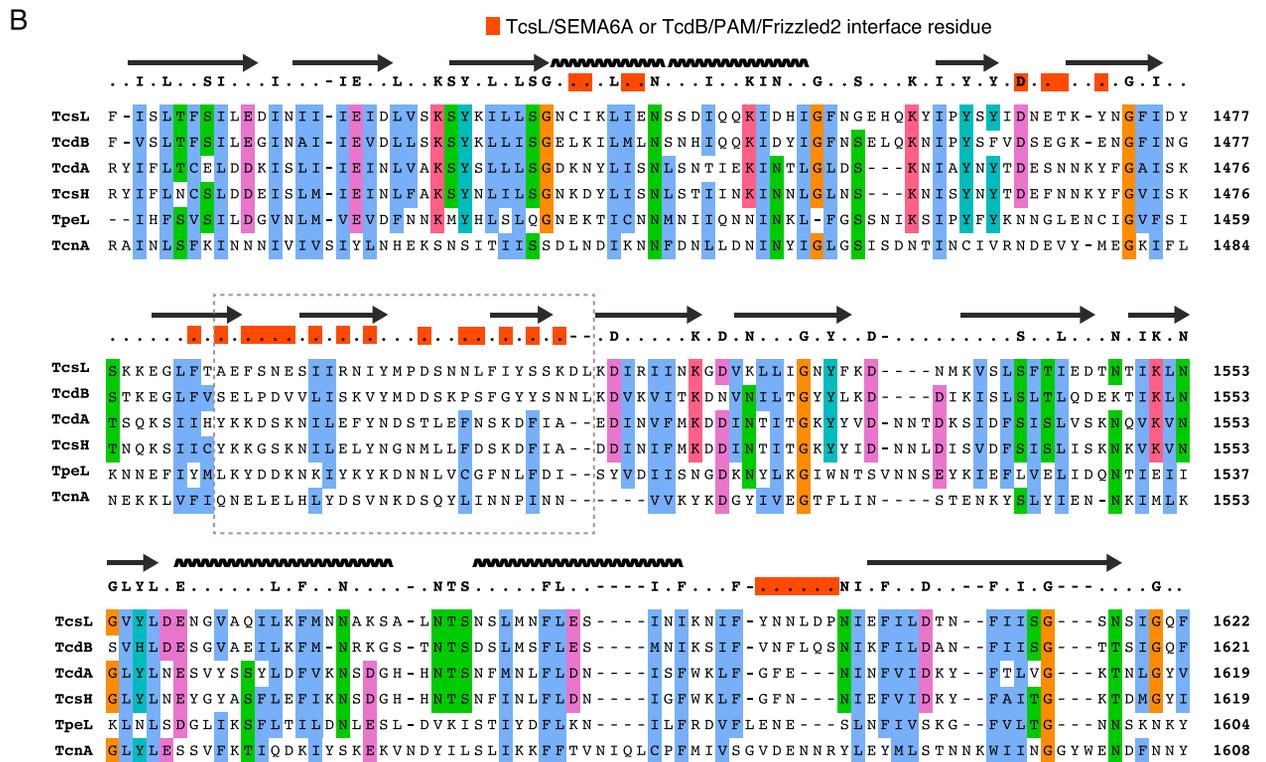
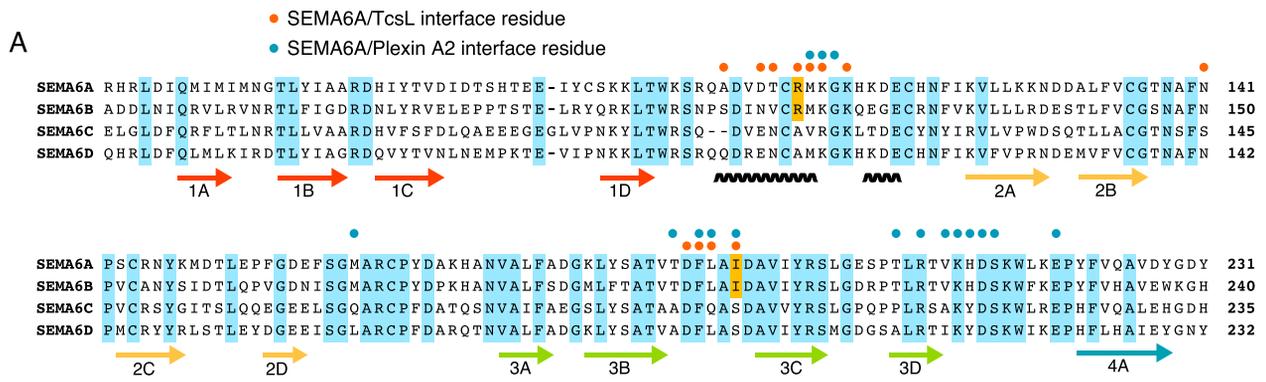


Figure S6. Conservation of SEMA6A and TcsL interface residues in semaphoring and large clostridial toxin families, related to Figure 6

A, Sequence alignment of SEMA6 family proteins. Residues conserved in all four SEMA6 family proteins are denoted in light blue. SEMA6A residues forming contacts with TcsL are indicated as red circles, and those interacting with Plexin A2 are indicated as blue circles. The two TcsL-interacting residues that differ between SEMA6A/SEMA6B and SEMA6C/SEMA6D are colored orange. Secondary structure elements with numbered beta-propeller blades are shown below the alignment.

(B) The receptor-binding surface of TcsL and TcdB is highly divergent between all clostridial toxins. Partial sequence alignment of six known large clostridial toxins. Secondary structure elements and the consensus sequence (at least 4/6 identical residues) are shown above the alignment. Amino acids are colored based on their biophysical properties (ClustalX coloring) if at least 4/6 residues are similar in each column. Red boxes indicate interface residues in TcsL/SEMA6A or TcdB/Fzd2 complexes. The gray box highlights the evolutionarily divergent beta sheet in the receptor-binding interface.

(C) Alignment entropy was calculated as a 20-aa moving window along the alignment of six known clostridial toxins. Red bars indicate the location of TcsL/SEMA6A interface residues.