

# Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to *Clostridium difficile* toxin A entry into cells

Liang Tao<sup>1,2,3,4,10\*</sup>, Songhai Tian<sup>1,2,3,10</sup>, Jie Zhang<sup>1,2,3,10</sup>, Zhuoming Liu<sup>3</sup>, Lindsey Robinson-McCarthy<sup>3</sup>, Shin-Ichiro Miyashita<sup>1,2,3</sup>, David T. Breault<sup>5,6,7</sup>, Ralf Gerhard<sup>8</sup>, Siam Oottamasathien<sup>9</sup>, Sean P. J. Whelan<sup>3</sup> and Min Dong<sup>1,2,3\*</sup>

***Clostridium difficile* toxin A (TcdA) is a major exotoxin contributing to disruption of the colonic epithelium during *C. difficile* infection. TcdA contains a carbohydrate-binding combined repetitive oligopeptides (CROPs) domain that mediates its attachment to cell surfaces, but recent data suggest the existence of CROPs-independent receptors. Here, we carried out genome-wide clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated protein 9 (Cas9)-mediated screens using a truncated TcdA lacking the CROPs, and identified sulfated glycosaminoglycans (sGAGs) and low-density lipoprotein receptor (LDLR) as host factors contributing to binding and entry of TcdA. TcdA recognizes the sulfation group in sGAGs. Blocking sulfation and glycosaminoglycan synthesis reduces TcdA binding and entry into cells. Binding of TcdA to the colonic epithelium can be reduced by surfen, a small molecule that masks sGAGs, by GM-1111, a sulfated heparan sulfate analogue, and by sulfated cyclodextrin, a sulfated small molecule. Cells lacking LDLR also show reduced sensitivity to TcdA, although binding between LDLR and TcdA are not detected, suggesting that LDLR may facilitate endocytosis of TcdA. Finally, GM-1111 reduces TcdA-induced fluid accumulation and tissue damage in the colon in a mouse model in which TcdA is injected into the caecum. These data demonstrate *in vivo* and pathological relevance of TcdA–sGAGs interactions, and reveal a potential therapeutic approach of protecting colonic tissues by blocking these interactions.**

The bacterium *Clostridium difficile* is a spore-forming opportunistic pathogen and one of the three ‘urgent threats’ classified by the Centers for Disease Control and Prevention of the United States. Disruption of gut flora by antibiotics allows *C. difficile* to colonize the colon, leading to diarrhoea and life-threatening pseudomembranous colitis<sup>1</sup>. The occurrence of *C. difficile* infection is exacerbated by the emergence of hypervirulent and antibiotic-resistant strains<sup>2–4</sup>. It is now the most common cause of antibiotic-associated diarrhoea and gastroenteritis-associated death in developed countries, accounting for around 500,000 cases and 29,000 deaths annually in the United States<sup>5</sup>.

Two homologous exotoxins, *C. difficile* toxin A and B (TcdA and TcdB), which target and disrupt the colonic epithelium, are the major virulent factors of *C. difficile*<sup>6–10</sup>. In addition, some hypervirulent strains also express a third toxin known as *C. difficile* transferase, which may suppress host eosinophilic responses<sup>11</sup>. TcdA (~308 kDa) and TcdB (~270 kDa) consist of four functional domains<sup>10,12</sup>: the N-terminal glucosyltransferase domain (GTD), a cysteine protease domain that mediates auto-cleavage and releases the GTD into the host cytosol, a central part containing both the transmembrane delivery domain and receptor-binding domain, and finally a C-terminal combined repetitive oligopeptides (CROPs) domain. The GTD glucosylates small GTPases of the Rho family, including Rho, Rac and CDC42, and inhibits

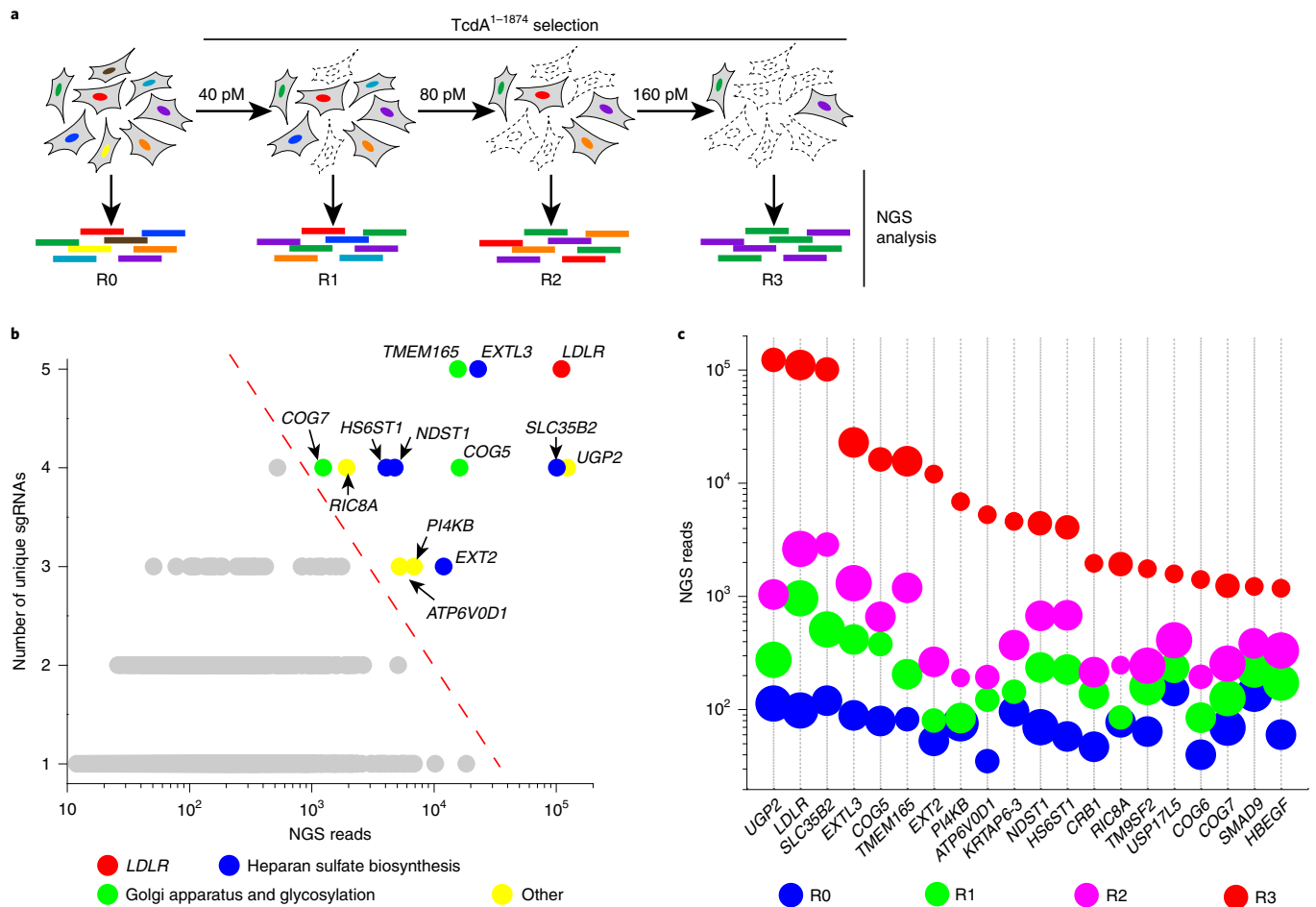
their function, resulting in cytopathic cell rounding and ultimately cell death.

The CROPs domains of TcdA and TcdB bear similarity with carbohydrate-binding proteins and may mediate toxin attachment to cell surfaces through various carbohydrate moieties. Particularly, CROPs from TcdA was shown to bind the trisaccharide Gal $\alpha$ 1,3Gal $\beta$ 1,4GlcNAc<sup>13</sup>. It has since been shown to also broadly recognize human I, Lewis X, and Lewis Y antigens, as well as glycosphingolipids, which all contain the Gal $\beta$ 1,4GlcNAc motif<sup>4,15</sup>.

Recent studies have shown that truncating the CROPs only modestly reduces the potency of TcdA and TcdB on cultured cells, suggesting the existence of CROPs-independent receptors<sup>16,17</sup>. Three candidate receptors have been reported for TcdB: chondroitin sulfate proteoglycan 4 (CSPG4), poliovirus receptor-like 3 (PVRL3) and the Wnt receptor frizzled proteins<sup>18–21</sup>. Two proteins have been previously suggested as potential receptors for full-length TcdA: sucrose-isomaltase and glycoprotein 96 (Gp96)<sup>22,23</sup>. However, sucrose-isomaltase is not expressed in the colon epithelium and Gp96 resides mainly in the endoplasmic reticulum.

Here we used a truncated TcdA lacking the majority of the CROPs domain and carried out genome-wide clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated protein 9 (Cas9)-mediated knockout (KO) screens, which identified sulfated glycosaminoglycans (sGAGs) and low-density lipoprotein

<sup>1</sup>Department of Urology, Boston Children’s Hospital, Boston, MA, USA. <sup>2</sup>Department of Surgery, Harvard Medical School, Boston, MA, USA. <sup>3</sup>Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Westlake University, Hangzhou, China. <sup>5</sup>Division of Endocrinology, Boston Children’s Hospital, Boston, MA, USA. <sup>6</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA. <sup>7</sup>Harvard Stem Cell Institute, Cambridge, MA, USA. <sup>8</sup>Institute of Toxicology, Hannover Medical School, Hannover, Germany. <sup>9</sup>Department of Surgery and Pediatric Urology, University of Utah/Primary Children’s Hospital, Salt Lake City, UT, USA. <sup>10</sup>These authors contributed equally: Liang Tao, Songhai Tian, Jie Zhang. \*e-mail: [taoliang@westlake.edu.cn](mailto:taoliang@westlake.edu.cn); [min.dong@childrens.harvard.edu](mailto:min.dong@childrens.harvard.edu)



**Fig. 1 | Genome-wide CRISPR-Cas9-mediated screen identifies host factors for *TcdA*.** **a**, Schematic of the screening process using *TcdA*<sup>1-1874</sup> on HeLa cells. Round zero (R0) represents cells at the beginning of the screen. Rounds 1, 2 and 3 (R1, R2 and R3) represent surviving cells after exposure to *TcdA*<sup>1-1874</sup> sequentially at the indicated toxin concentrations. **b**, Genes identified after R3 were ranked and plotted. The y axis shows the number of unique sgRNAs for each gene. The x axis represents the number of sgRNA reads for each gene. The top-ranking genes are colour-coded and grouped on the basis of their functions. The dashed red line indicates the top-ranked hits. **c**, The NGS reads from R0 to R3 for the top-20 ranked (ordered by NGS reads) genes in R3 were colour-coded and plotted. The diameter of the circle represents the number of unique sgRNAs detected for the gene. All top-20 ranked genes were progressively enriched from R0 to R3.

receptor (LDLR) as CROPs-independent host factors mediating binding and entry of *TcdA*.

## Results

**CRISPR screens identify host factors for *TcdA*.** To identify the CROPs-independent receptors involved in *TcdA* actions, we used a truncated *TcdA* (*TcdA*<sup>1-1874</sup>) lacking the majority of the CROPs domain (Supplementary Fig. 1a), which has previously been shown to retain high levels of toxicity to multiple cell lines<sup>17</sup>. We first validated the toxicity of *TcdA*<sup>1-1874</sup> on various human cell lines using the standard cytopathic cell-rounding assay, which measures the percentages of rounded cells after incubation with a series of concentrations of toxins for 24 h (Supplementary Fig. 1b,c). The toxin concentration that induces 50% of cells to become round is defined as CR<sub>50</sub>, and is used to compare the sensitivity of different cell lines to *TcdA*<sup>1-1874</sup>. HeLa cells are one of the most sensitive human cell lines to *TcdA*<sup>1-1874</sup>, and were selected to carry out genome-wide CRISPR-Cas9 mediated KO screens.

HeLa cells stably expressing Cas9 were transduced with a lentiviral single guide RNA (sgRNA) library (GeCKO v.2) targeting 19,052 human genes<sup>24</sup>. The cells were subjected to three rounds of selection with *TcdA*<sup>1-1874</sup> (40, 80 and 160 pM, Fig. 1a). The genes targeted

by sgRNAs in surviving cells were identified via next-generation sequencing (NGS). We ranked the target genes on the basis of the number of unique sgRNAs (y axis) and the total NGS reads (x axis) (Fig. 1b). All top-ranked genes were enriched over the three rounds, suggesting that mutations in these genes offered survival advantages in the presence of *TcdA*<sup>1-1874</sup> (Fig. 1c).

The top-ranked gene encodes LDLR, a well-known receptor for low-density lipoproteins. Many other top-ranked genes encode key players in heparan sulfate biosynthesis and sulfation pathways<sup>25</sup>, including the glycosyltransferases exostosin-2 (EXT2) and exostosin-like 3 (EXTL3), the sulfotransferases heparan sulfate 6-*O*-sulfotransferase 1 (HS6ST1), *N*-deacetylase and *N*-sulfotransferase 1 (NDST1), and solute carrier family 35 member B2 (SLC35B2), which transports the activated form of sulfate into Golgi. Several other enzymes involved in glycosaminoglycan (GAG) synthesis were also identified (Supplementary Fig. 2a). Heparan sulfate is usually attached to core proteins as heparan sulfate proteoglycans (HSPGs). Both HSPGs and LDLR are widely expressed on the surface of various cells, and are therefore promising receptor candidates for *TcdA*.

Among the top-50 ranked genes, three (*UGP2*, *PI4KB* and *ATP6V0D1*) were also found in the top list of genes in our previous

genome-wide CRISPR–Cas9 screen using TcdB<sup>1–1830</sup> (Supplementary Fig. 2b). *UGP2* encodes uridine diphosphate-glucose pyrophosphorylase, which synthesizes uridine diphosphate-glucose, a co-factor required for TcdA and TcdB to glucosylate small GTPases<sup>26</sup>. *ATP6V0D1* is a component of vacuolar-type H<sup>+</sup>-ATPase for acidification of endosomes, which is an essential condition to trigger translocation of TcdA and TcdB<sup>27,28</sup>. *PI4KB* is a key player in phospholipid metabolism and signalling, and its role in toxin action remains to be established.

Other notable top hits include *COG5*, *COG7*, *TMEM165* and *RIC8A*. *COG5* and *COG7* are members of the conserved oligomeric Golgi (COG) complex<sup>29</sup>. In fact, all eight COG members were identified in the final round of screening (Supplementary Fig. 2c). *TMEM165* is a multi-pass transmembrane protein localized to the Golgi. Although the exact function of the COG complex and *TMEM165* remains to be fully established, mutations in COG complex and *TMEM165* both result in congenital disorders of glycosylation<sup>29,30</sup>, and affect multiple glycosylation pathways including biosynthesis of heparan sulfate<sup>31–33</sup>. *RIC8A* is a guanine nucleotide exchange factor and its role in TcdA action remains to be validated.

We also performed a parallel genome-wide CRISPR–Cas9-mediated KO screen using full-length TcdA on HeLa cells (Supplementary Fig. 2d). However, this screen only yielded *UGP2* as the top hit. Two other hits, *SGMS1* and *ZNF283*, were barely over our threshold. *SGMS1* regulates lipid raft formation and may affect the endocytosis process. *ZNF283* is a cytosolic protein, and its role in TcdA action remains to be validated. Lack of potential receptor candidates in the top hits suggests that full-length TcdA may utilize multiple receptors and entry pathways.

**sGAGs contribute to cellular entry of TcdA<sup>1–1832</sup>.** TcdA<sup>1–1874</sup> still contains a short fragment of the CROPs domain. Therefore, we further generated a truncated TcdA (TcdA<sup>1–1832</sup>) that deletes the entire CROPs to exclude any potential contribution from the residual CROPs domain (Supplementary Fig. 1a). TcdA<sup>1–1874</sup> and TcdA<sup>1–1832</sup> showed similar potency on HeLa cells in the cytopathic cell-rounding assays (Supplementary Fig. 1b).

Using TcdA<sup>1–1832</sup>, we first validated the role of *EXT2* and *EXTL3*, as they are specifically required for the elongation of the heparan sulfate chain, but not other types of GAGs. We generated *EXT2* and *EXTL3* KO HeLa cell lines using the CRISPR–Cas9 system. Both cell lines showed a reduction of cell surface heparan sulfate levels compared with wild-type cells, as measured by flow cytometry analysis using a heparan sulfate antibody (Supplementary Fig. 3a). Both *EXT2* and *EXTL3* KO cells showed a modest four-to-fivefold reduction in sensitivity to TcdA<sup>1–1832</sup> compared with wild-type cells, whereas their sensitivities towards TcdB<sup>1–1830</sup> remained the same as wild-type cells (Fig. 2a).

Several top-ranked genes identified in our screen, including *SLC35B2*, *NDST*, *HS6ST*, *HS2ST* and *HS3ST*, are involved in sulfation of GAGs<sup>25</sup> (Supplementary Fig. 2a). To examine the role of sulfation, we generated three single clones of *SLC35B2* KO HeLa cells using the CRISPR–Cas9 approach. Reduction of heparan sulfate in these cells was confirmed by flow cytometry analysis (Supplementary Fig. 3b). These cell lines all showed around tenfold reduction in sensitivity towards TcdA<sup>1–1832</sup> compared with wild-type cells, whereas their sensitivities towards TcdB<sup>1–1830</sup> were not changed (Fig. 2b). The reduced sensitivity of *SLC35B2* KO cells to TcdA<sup>1–1832</sup> was further confirmed by immunoblotting for RAC1 glucosylation (Supplementary Fig. 4a). Finally, *SLC35B2* KO cells also showed approximately threefold reduction in sensitivity to full-length TcdA (Fig. 2c).

**Characterizing the specificity of TcdA–sGAGs interactions.** We next carried out competition assays to further validate the role of sGAGs. First, we used surfen (bis-2-methyl-4-amino-quinolyl-6-carbamide), which is a small molecule that binds to and neutralizes

negative charges on all sGAGs<sup>34</sup>. Pre-incubation of cells with surfen protected HeLa cells from TcdA<sup>1–1832</sup> in a concentration-dependent manner, whereas it offered no protection from TcdB<sup>1–1830</sup> (Fig. 2d and Supplementary Fig. 5a). Similar results were observed with Huh7 cells (Supplementary Fig. 5b).

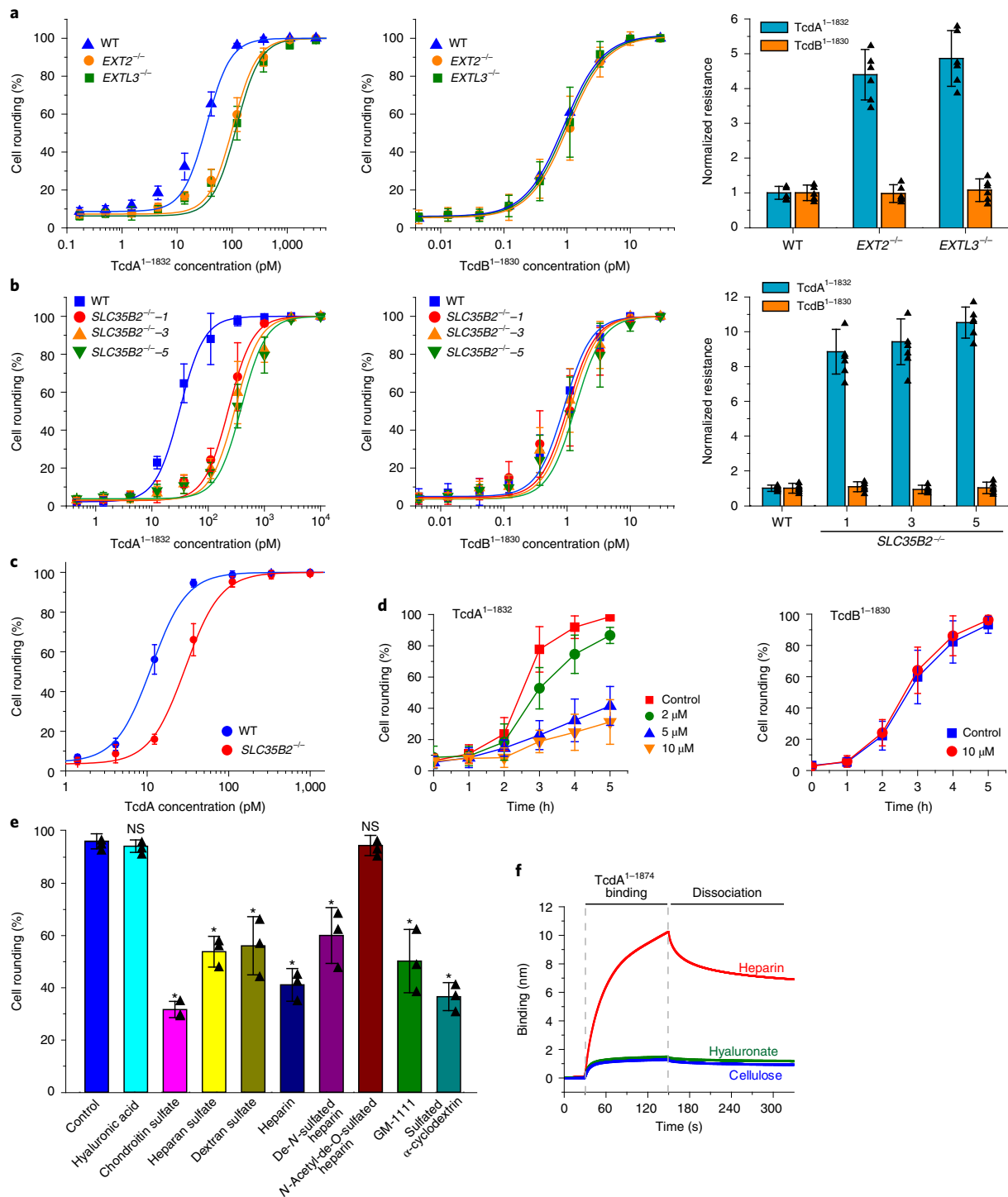
To understand the selectivity of TcdA–GAG interactions, we carried out competition assays using a panel of GAGs including heparan sulfate, heparin, de-*N*-sulfated heparin, *N*-acetyl-de-*O*-sulfated heparin, chondroitin sulfate and dextran sulfate. Heparin is a highly sulfated variant of heparan sulfate and it is widely utilized as an anticoagulant. In addition, we also tested synthetic sulfated molecules GM-1111 and sulfated cyclodextrin. GM-1111 contains the same carbohydrate moieties and sulfation groups as heparan sulfate, but with distinct glycosidic bonds. It has been developed as a heparan sulfate mimic with reduced anticoagulation activities<sup>35</sup>. Sulfated cyclodextrin is a small molecule that is distinct from GAGs. Non-sulfated GAG hyaluronic acid and polysaccharide cellulose were also examined. These molecules are shown in Supplementary Fig. 6.

Pre-incubation of TcdA<sup>1–1832</sup> with heparan sulfate, heparin, chondroitin sulfate, dextran sulfate, GM-1111 and sulfated cyclodextrin all reduced the level of cell rounding, whereas hyaluronic acid showed no effect (Fig. 2e). These results suggest that TcdA may not recognize heparan sulfate specifically, but rather interacts mainly with the sulfation group. Furthermore, the finding that de-*N*-sulfated heparin protected cells from TcdA<sup>1–1832</sup>, whereas *N*-acetyl-de-*O*-sulfated heparin did not offer any protection (Fig. 2e), suggests that TcdA preferentially recognizes *O*-sulfation.

To further characterize direct TcdA–sGAG interactions, we used bio-layer interferometry (BLI) assay by immobilizing biotinylated heparin onto the probe. Binding of TcdA to the immobilized heparin would result in a shift in the light interference pattern that can be monitored in real time. Biotinylated hyaluronate and cellulose were analysed in parallel as controls. Both full-length TcdA and TcdA<sup>1–1874</sup> showed robust binding to biotin–heparin, but not to biotin–hyaluronate and biotin–cellulose (Fig. 2f and Supplementary Fig. 7a). TcdA–heparin interactions appear to be influenced by the ionic strength of the buffer: higher salt concentrations reduce heparin–TcdA interactions (Supplementary Fig. 7b). At 150 mM salt concentration, the apparent dissociation constants ( $K_D$ ) for TcdA–heparin and TcdA<sup>1–1874</sup>–heparin are at similar levels (85.5 nM for TcdA<sup>1–1874</sup> versus 23.2 nM for full-length TcdA; Supplementary Fig. 7c–e).

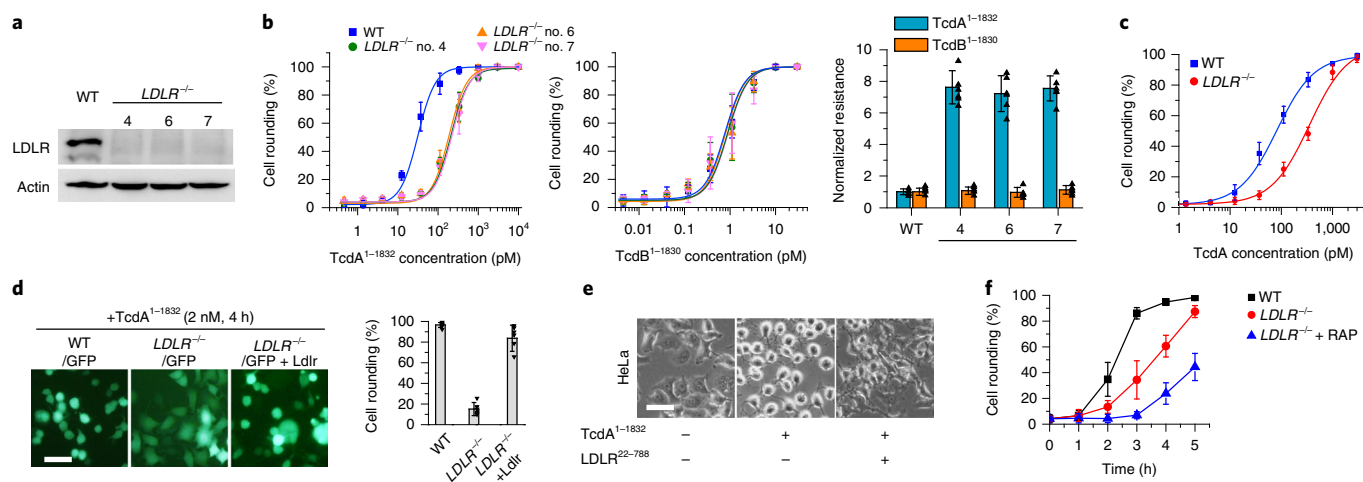
**LDLR contributes to cellular entry of TcdA<sup>1–1832</sup>.** To validate the role of LDLR, we generated *LDLR* KO HeLa cells using the CRISPR–Cas9 system. Three single-KO clones were established and the loss of LDLR expression was confirmed in the clones by immunoblot analysis (Fig. 3a). All three KO lines showed reduced sensitivity by about sevenfold to TcdA<sup>1–1832</sup>, whereas their sensitivity to TcdB<sup>1–1830</sup> remained the same as that of wild-type cells (Fig. 3b). The reduced sensitivity of *LDLR* KO cells to TcdA<sup>1–1832</sup> was also confirmed by immunoblot against RAC1 glucosylation (Supplementary Fig. 4b). *LDLR*<sup>−/−</sup> cells also showed around threefold reduction in sensitivity to full-length TcdA, thus validating the role of LDLR in cytotoxic activity of full-length TcdA (Fig. 3c). The sensitivity to TcdA was restored when *LDLR* KO cells were transfected with mouse *Ldlr* (Fig. 3d), which is not targeted by the sgRNA. Furthermore, Huh7 *LDLR*<sup>−/−</sup> cells, which were previously generated and validated<sup>36</sup>, also showed reduced sensitivity to TcdA<sup>1–1832</sup> compared with wild-type Huh7 cells (Supplementary Fig. 8).

We further carried out a competition assay using the soluble extracellular domain of LDLR (residues 22–788, LDLR<sup>22–788</sup>). Co-incubation of LDLR<sup>22–788</sup> with TcdA<sup>1–1832</sup> (200:1) reduced the percentage of rounded cells (Fig. 3e). LDLR belongs to a large family of proteins including VLDLR, LRP1, LRP1b, LRP2 (also known as megalin), LRP5, LRP6 and LRP8 (also known as ApoER2), which



**Fig. 2 | sGAGs contribute to cellular entry of TcdA<sup>1-1832</sup>.** **a**, The sensitivities of *EXT2*<sup>-/-</sup> and *EXTL3*<sup>-/-</sup> HeLa cells to TcdA<sup>1-1832</sup> (left) and TcdB<sup>1-1830</sup> (middle) were quantified using the cytopathic cell-rounding assay. The percentage of rounded cells was quantified, plotted and fitted. The toxin concentration resulting in 50% cell rounding is defined as CR<sub>50</sub> and is used for comparisons by normalizing to the level of wild-type (WT) HeLa cells as normalized resistance (right; y axis, each data point is also shown as triangle in the bar graph). **b**, The sensitivities of three *SLC35B2*<sup>-/-</sup> HeLa cell lines to TcdA<sup>1-1832</sup> and TcdB<sup>1-1830</sup> were quantified using the cytopathic cell-rounding assay and normalized to the level of wild-type HeLa cells. Each data point is also shown as triangle in the bar graph (right). **c**, The sensitivities of wild-type and *SLC35B2*<sup>-/-</sup> (clone no. 5) HeLa cells to full-length TcdA were evaluated using the cytopathic cell-rounding assay. The percentage of rounded cells was quantified, plotted and fitted. **d**, Pre-incubation of surfen in the medium reduced the potency of TcdA<sup>1-1874</sup> but not TcdB<sup>1-1830</sup> on HeLa cells in a concentration-dependent manner, as measured by the cytopathic cell-rounding assay over time. **e**, Competition assay on HeLa cells by pre-incubating TcdA<sup>1-1874</sup> (2 nM) with the indicated GAGs, polysaccharides and synthetic sulfated molecules (all at 1 mg ml<sup>-1</sup>). The degree of protection from TcdA was evaluated by the cytopathic cell-rounding assay 4 h later (\**P* < 0.005; *P* > 0.05 is considered as non-significant (NS); two-sided Student's *t*-test, *n* = 3). Each data point is also shown as a triangle in the bar graph. **f**, BLI assays showing that TcdA<sup>1-1874</sup> (1 μM) strongly bound to biotin-heparin but not to biotin-hyaluronate or biotin-cellulose. Experiments were repeated three times. In **a-d**, *n* = 6; data are mean ± s.d. The experiments were repeated three times independently with similar results.





**Fig. 3 | LDLR contributes to cellular entry of TcdA<sup>1-1832</sup>.** **a**, The absence of LDLR expression in three *LDLR*<sup>-/-</sup> HeLa cell lines was confirmed by immunoblot analysis. Actin served as a loading control. The experiments were repeated three times independently with similar results. **b**, The sensitivities of three *LDLR*<sup>-/-</sup> HeLa cell lines to TcdA<sup>1-1832</sup> and TcdB<sup>1-1830</sup> were quantified using the cytopathic cell-rounding assay and normalized to the levels of wild-type HeLa cells (*n* = 6). Each data point is also shown as a triangle in the bar graph. **c**, The sensitivities of wild-type and *LDLR*<sup>-/-</sup> HeLa cells to full-length TcdA were evaluated using cytopathic cell-rounding assays (*n* = 6). The percentage of rounded cells was quantified, plotted and fitted. **d**, Ectopic expression of a mouse *Ldlr* in *LDLR*<sup>-/-</sup> (no. 4) cells restored the sensitivity of these cells to TcdA<sup>1-1832</sup> and resulted in cell rounding under the assay conditions (2 nM, 4 h). Green fluorescent protein (GFP) was co-transfected to mark transfected cells. Left, representative images showing fluorescence of transfected cells; right, percentage of rounded cells (*n* = 6). Each data point is also shown as triangle in the bar graph. **e**, Pre-incubation of the ectodomain of LDLR (residues 22–788, 400 nM) with TcdA<sup>1-1832</sup> (2 nM, 4 h) protected HeLa cells from the toxin and prevented cell rounding. The experiments were repeated three times independently with similar results. **f**, Pre-incubation of RAP (4 μM) in culture medium further protected *LDLR*<sup>-/-</sup> (no. 4) cells from TcdA<sup>1-1832</sup> (10 nM) as measured by the cell-rounding assay over time (*n* = 6). Scale bars in **d** and **e**, 50 μm. Data are mean ± s.d. Experiments were repeated two times independently with similar results.

share similar domains with LDLR and often act as redundant receptors for many LDLR ligands. Receptor associated protein (RAP) binds tightly to most LDLR family members and its binding inhibits binding of LDL and many other ligands<sup>37–39</sup>. Adding RAP to the medium further reduced the sensitivity of *LDLR* KO cells to TcdA<sup>1-1832</sup> (Fig. 3f), suggesting that other LDLR family members also contribute to entry of TcdA<sup>1-1832</sup> into cells.

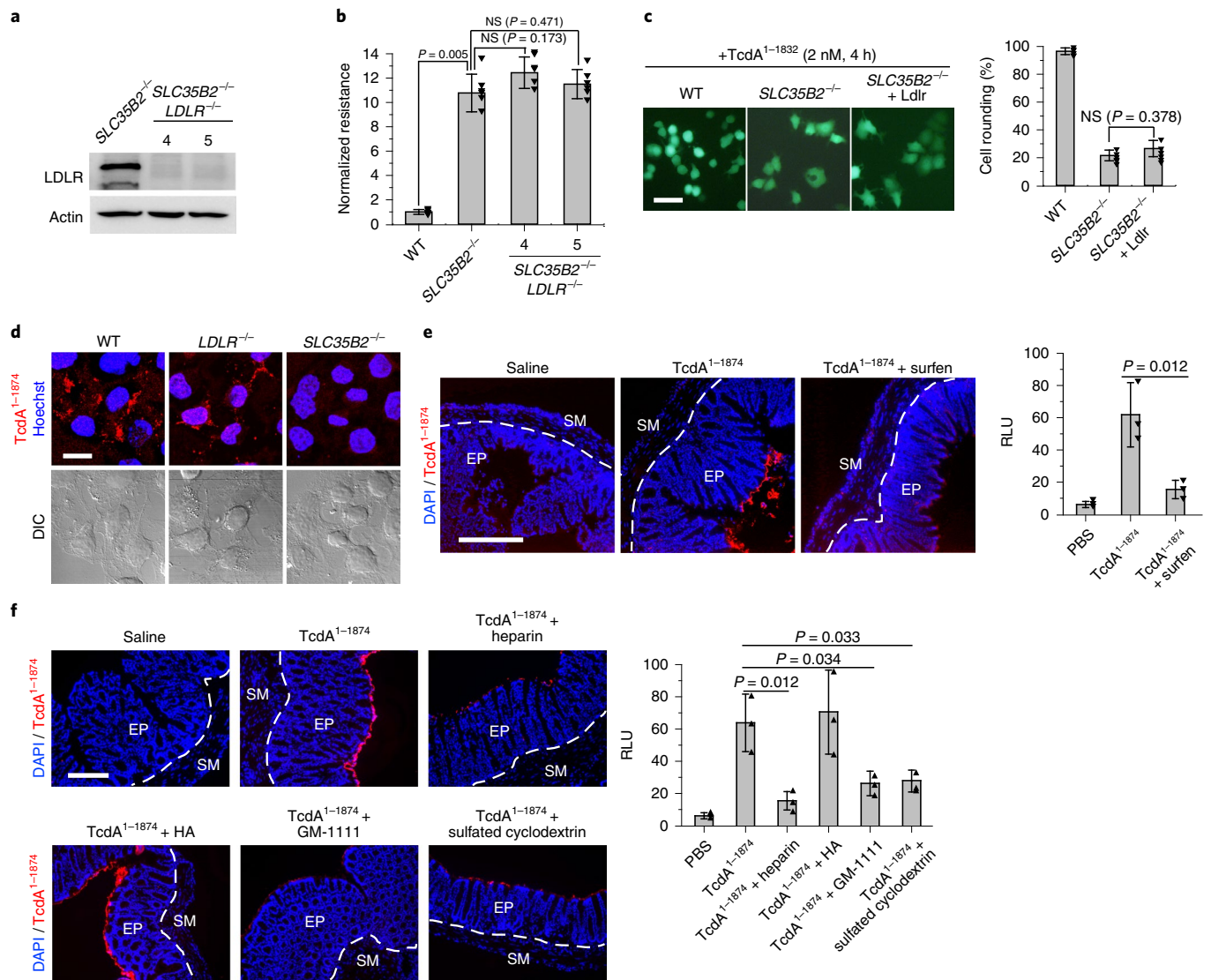
To examine binding of TcdA<sup>1-1874</sup> to LDLR *in vitro*, we used purified Fc-tagged extracellular domain of LDLR (LDLR<sup>22-788</sup>-Fc) produced in HEK293 cells. This LDLR<sup>22-788</sup>-Fc mediated strong binding of RAP, but we did not detect direct binding of TcdA<sup>1-1874</sup> to LDLR<sup>22-788</sup>-Fc in either BLI assays or an alternative dot blot assay (Supplementary Fig. 9). These results suggest that either that TcdA<sup>1-1874</sup> binding to LDLR is weak or that their interactions may require additional cellular factors.

**sGAGs are major cellular attachment factors for TcdA<sup>1-1874</sup>.** To further understand the role of LDLR and sGAGs, we generated *LDLR*<sup>-/-</sup> *SLC35B2*<sup>-/-</sup> double-KO cell lines by knocking out LDLR from HeLa *SLC35B2*<sup>-/-</sup> cells using the CRISPR–Cas9 approach. Two single-cell clones were established, and lack of LDLR expression was confirmed by immunoblot (Fig. 4a). However, these two double-KO cell lines did not further increase their resistance to TcdA<sup>1-1832</sup> compared with LDLR and *SLC35B2* single-KO cells (Fig. 4b and Supplementary Fig. 4c). Moreover, overexpression of exogenous mouse *Ldlr* by transient transfection did not increase the sensitivity of *SLC35B2*<sup>-/-</sup> cells to TcdA<sup>1-1832</sup> (Fig. 4c). These data suggest that LDLR and sGAGs are not redundant receptors, and that they could act cooperatively. We therefore examined binding of TcdA<sup>1-1874</sup> to wild-type versus *LDLR*<sup>-/-</sup> and *SLC35B2*<sup>-/-</sup> HeLa cells, using TcdA<sup>1-1874</sup> directly labelled with a fluorescent dye. As shown in Fig. 4d, *LDLR*<sup>-/-</sup> cells showed similar overall TcdA<sup>1-1874</sup> binding as wild-type cells. By contrast, binding of TcdA<sup>1-1874</sup> to *SLC35B2*<sup>-/-</sup> cells was diminished. These results suggest that sGAGs are the major attachment

factor mediating binding of TcdA<sup>1-1874</sup> on cell surfaces under our assay conditions.

**sGAGs are attachment factors for TcdA<sup>1-1874</sup> in the colonic epithelium.** The colonic epithelium is the pathologically relevant target of TcdA. sGAGs are abundant both in the intestinal mucosa and on the basolateral side of the epithelium<sup>40–42</sup>. To examine the contribution of sGAGs to TcdA binding to the colonic epithelium, we used a colon loop ligation assay<sup>20</sup>. In brief, fluorescence-labelled TcdA<sup>1-1874</sup> was injected into a ligated colon segment and incubated for 30 min. Colon tissues were then dissected and fixed. TcdA<sup>1-1874</sup> showed strong binding to the apical side of the colonic epithelium and binding appears to extend into the lumen (Fig. 4e). Co-injecting surfen reduced binding of TcdA<sup>1-1874</sup> (Fig. 4e). Similarly, heparin, GM-1111 and sulfated cyclodextrin all reduced binding of TcdA<sup>1-1874</sup>, whereas hyaluronic acid showed no effect (Fig. 4f). These results suggest that sGAGs are major attachment factors in the colon epithelium for TcdA<sup>1-1874</sup>.

**Blocking sGAG–TcdA interactions reduces TcdA toxicity in the colon.** We next examined the contribution of sGAGs-mediated binding in the context of full-length TcdA *in vivo*. Injecting fluorescence-labelled full-length TcdA into the ligated colon loop for 30 min resulted in robust binding to the apical side of the colonic epithelium (Fig. 5a). Co-injecting recombinantly produced CROPs fragment reduced binding of TcdA, consistent with the finding that the CROPs region mediates TcdA binding to cells<sup>43</sup>. Co-injecting surfen with TcdA reduced binding of TcdA, confirming that sGAGs contribute to binding of full-length TcdA to the colonic epithelium (Fig. 5a). Similarly, co-injection with GM-1111 or sulfated cyclodextrin also reduced TcdA binding to the colonic epithelium (Fig. 5b). Interestingly, combining CROPs and surfen together largely abolished binding of TcdA to the colonic epithelium (Fig. 5a). Thus, both CROPs-mediated and sGAGs-mediated binding contribute to TcdA binding to the colonic epithelium.

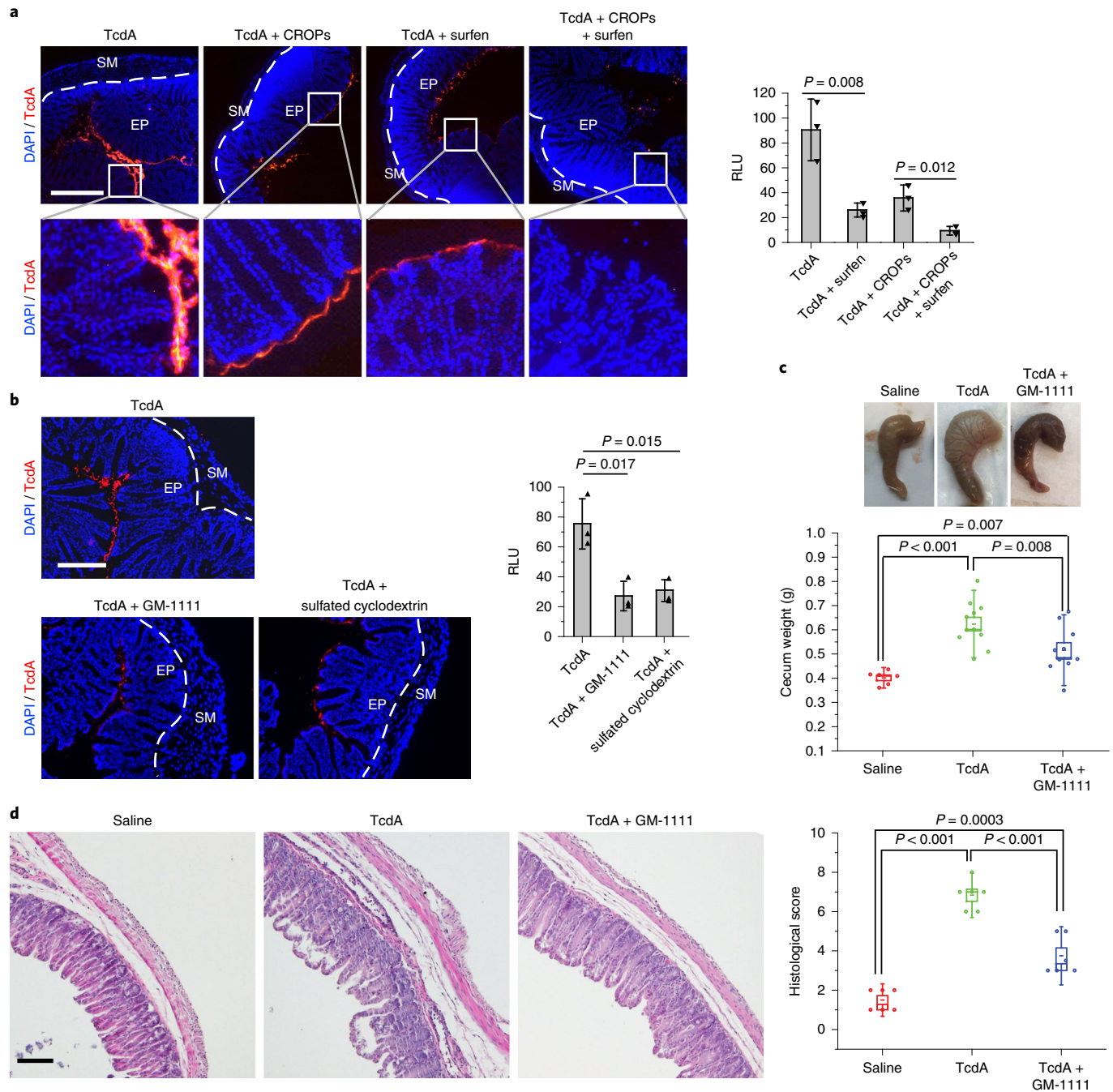


**Fig. 4 | sGAGs are major attachment factors for TcdA.** **a**, The absence of LDLR expression in two *SLC35B2*<sup>-/-</sup>*LDLR*<sup>-/-</sup> HeLa cell lines was confirmed by immunoblot analysis. The experiments were repeated three times independently with similar results. **b**, The sensitivities of two *SLC35B2*<sup>-/-</sup>*LDLR*<sup>-/-</sup> HeLa cell lines and their parental cell line *SLC35B2*<sup>-/-</sup> (no. 5) to TcdA<sup>1-1832</sup> were quantified using the cytopathic cell-rounding assay, and normalized to the level of wild-type HeLa cells. **c**, Ectopic expression of a mouse *Ldlr* did not restore TcdA<sup>1-1832</sup> (2 nM, 4 h) entry into *SLC35B2*<sup>-/-</sup> cells under our assay conditions. Left, representative images; transfected cells are marked by GFP expression. Right, quantification of cell rounding. **d**, Immunofluorescence analysis showing Alexa Fluor 555-labelled TcdA<sup>1-1874</sup> (5 nM) robustly bound to wild-type and *LDLR*<sup>-/-</sup> (no. 4) HeLa cells, but not to *SLC35B2*<sup>-/-</sup> (no. 5) cells. Cell nuclei were labelled with Hoechst dye. DIC, differential interference contrast image. The experiments were repeated three times. **e**, Co-injection of surfen (50 μM) with Alexa Fluor 555-labelled TcdA<sup>1-1874</sup> (5 nM, red) into ligated colon prevented TcdA<sup>1-1874</sup> binding to the colonic epithelium. Cell nuclei were labelled with DAPI dye (blue). EP, epithelial cells; SM, smooth muscles. **f**, Co-injection of heparin, GM-1111 or sulfated cyclodextrin, but not hyaluronic acid (HA) (all at 1 mg ml<sup>-1</sup>) with Alexa Fluor 555-labelled TcdA<sup>1-1874</sup> (5 nM) into the ligated colon reduced TcdA<sup>1-1874</sup> binding to the colonic epithelium. Cell nuclei were labelled with DAPI dye (blue). In **b** and **c**, *n* = 6; NS, *P* > 0.05. Mann-Whitney test (two-sided). In **e** and **f**, *n* = 3; binding of TcdA was quantified using ImageJ; two-sided Student's *t*-test. Each data point is also shown as a triangle in the bar graph. Data are mean ± s.d. Scale bars represent 50 μm in **c**, 20 μm in **d**, and 200 μm in **e** and **f**.

To further examine the relevance of sGAG–TcdA interactions for TcdA-induced pathogenesis *in vivo*, we utilized a mouse caecum-injection model that was previously established to assess pathogenesis of TcdA and TcdB<sup>14</sup>. In brief, TcdA or TcdA premixed with inhibitors was injected into the caecum. Mice were allowed to recover for 6 h before euthanization. The caecum and the ascending colon were collected and weighed to assess the degree of fluid accumulation. The caecum tissue was also fixed and subjected to hematoxylin and eosin staining and histological score analysis based on four criteria (disruption of the epithelium, haemorrhagic congestion, mucosal oedema and inflammatory cell infiltration) on

a scale of 0–3 (normal, mild, moderate or severe). Injection of TcdA induced fluid accumulation in the colon tissues, severe mucosal oedema, mild-to-moderate disruption of the epithelium, haemorrhagic congestion and inflammatory cell infiltration (Fig. 5c,d).

Finding a suitable inhibitor for use in the caecum-injection model was challenging, as heparan sulfate and many sGAG mimics induced haemorrhage in the intestine and colon after incubation for 6 h. This is likely to be caused by their anticoagulation activity. Surfen alone at the concentration required to reduce TcdA binding also induced damage to colonic tissues after incubation for 6 h. After surveying many different sGAG mimics, we found that GM-1111, which was



**Fig. 5 | Blocking sGAG–TcdA interactions reduces TcdA toxicity in the colon. a**, Co-injecting either surfen (50  $\mu$ M) or TcdA CROPs (150 nM) with Alexa Fluor 555-labelled full-length TcdA (5 nM) partially reduces TcdA binding to the colonic epithelium. Co-injecting both surfen and TcdA CROPs with TcdA largely abolished toxin binding. Representative images (left), and quantification of binding (right;  $n = 3$ ). **b**, Co-injection of GM-1111 or sulfated cyclodextrin with TcdA reduced TcdA binding to the colonic epithelium ( $n = 3$ ). **c**, TcdA (4  $\mu$ g), TcdA premixed with GM-1111 (0.5 mg ml<sup>-1</sup>) or saline was injected into the caecum of mice. After 6 h, the caecum tissue was excised. Representative caecum tissues are shown, and the weight of each caecum was measured and plotted. Boxes represent mean  $\pm$  s.e.m.; bars represent s.d.; two-sided Student's *t*-test;  $n = 6$  (saline), 11 (TcdA) or 10 (TcdA + GM-1111). **d**, Caecum tissues from **c** were sectioned and subjected to haematoxylin and eosin staining. Representative images are shown and the histological scores were assessed on the basis of disruption of the epithelia, haemorrhagic congestion, mucosal oedema and inflammatory cell infiltration. Boxes represent mean  $\pm$  s.e.m.; bars represent s.d.; Student's *t*-test. In **a** and **b**, binding of TcdA was quantified using ImageJ; two-sided Student's *t*-test with multiple comparisons. Data are mean  $\pm$  s.d. Each data point is also shown as a triangle in the bar graph. Scale bars represent 200  $\mu$ m in **a** and **b**, and 100  $\mu$ m in **d**.

specifically developed to reduce anticoagulation activity, can be used at the dose that reduces TcdA binding without itself inducing visible tissue damage. Co-injecting GM-1111 with TcdA significantly reduced fluid accumulation in the colon (caecum weight; Fig. 5c) and overall tissue damage as evidenced by histological scoring (Fig. 5d).

## Discussion

The presence of numerous negatively charged sulfate groups in sGAGs provides an ideal multivalent landing pad for proteins and macromolecules through electrostatic interactions. These sulfate groups are known to interact with a large array of endogenous



ligands, such as fibroblast growth factors, vascular endothelial growth factor, transforming growth factor  $\beta$ , chemokines and cytokines<sup>45</sup>. Unsurprisingly, these proteoglycans are also exploited by a long list of viral, bacterial and parasitic pathogens as attachment factors<sup>46</sup>. As TcdA is capable of binding to isolated sGAGs, it should be able to bind both to proteoglycans containing sGAGs as well as to free sGAGs on the cell surface and in the extracellular matrix. The exact binding sites for sGAGs in TcdA remain to be determined and it is possible that multiple positively charged surface regions of TcdA are involved.

LDLR belongs to a family of structurally related receptors, many of which act as redundant receptors for various ligands and viruses<sup>47</sup>. Interestingly, the LDLR family member LRP1 was previously established as the receptor for TpeL toxin<sup>39</sup>, which belongs to the same toxin family as TcdA but naturally lacks the CROPs domain. It is likely that LDLR family members other than LDLR can also contribute to TcdA<sup>1-1832</sup> entry, as RAP further reduces the sensitivity of LDLR KO cells.

LDLR family receptors rapidly and constitutively recycle between cell membranes and endosomes. This provides an ideal mechanism by which to mediate endocytosis into cells. Indeed, LDLR has been exploited as a receptor for many viruses, such as vesicular stomatitis virus (VSV), hepatitis C virus and the minor group common cold virus<sup>36,38,48</sup>. Although it remains unknown whether TcdA is capable of recognizing LDLR family members directly on cell surfaces, the major contribution of LDLR members is likely to occur through facilitation of endocytosis of TcdA bound to sGAGs. Similar synergistic actions between proteoglycans and LDLR family members are common for endogenous ligands. For instance, HSPG contributes to the capture of PCSK9 on cell surfaces and subsequently presents PCSK9 to LDLR<sup>49</sup>. Furthermore, many viruses that utilize HSPG as an initial attachment factor recruit additional protein receptors to mediate their endocytosis<sup>50</sup>. For instance, respiratory syncytial virus uses HSPG as an attachment factor and ICAM1 and VLDLR as additional protein receptors<sup>50</sup>. Such a 'two-step' model allows the pathogens and toxins to both maximize their chance of landing on the cell surface and take advantage of rapid endocytosis and recycling of LDLR family members.

A combination of surfen and the CROPs domain protein largely abolished binding of full-length TcdA to the colonic epithelium, demonstrating that TcdA attaches to the colonic epithelium via at least two independent binding interfaces: interactions with sGAGs in a CROPs-independent manner and interactions with carbohydrate moieties via the CROPs. These results are consistent with the previous finding that TcdA<sup>1-1874</sup> and TcdA<sup>1875-2710</sup> do not compete with each other, whereas both can reduce binding of full-length TcdA to cells<sup>17</sup>. These results further support a previously proposed 'two-receptor' model for TcdA<sup>10,16,39</sup>. Finally, GM-1111 alone reduced the toxicity of TcdA in the caecum-injection model *in vivo*, demonstrating the therapeutic potential of protecting colonic tissues from TcdA by targeting TcdA-sGAGs interactions.

## Methods

**Materials.** HeLa (H1, CRL-1958), HT-29 (HTB-38), CHO-C6 and 293T (CRL-3216) cells were originally obtained from ATCC. They tested negative for mycoplasma contamination, but have not been authenticated. Huh7 and Huh7 LDLR<sup>-/-</sup> cells were provided by Y. Matsuura (Osaka University)<sup>36</sup>. The following mouse monoclonal antibodies were purchased from the indicated vendors: RAC1 (23A8, Abcam), non-glucosylated RAC1 (clone 102, BD Biosciences),  $\beta$ -actin (AC-15, Sigma) and heparan sulfate (F58-10E4, mouse IgM, Amsbio). Rabbit monoclonal IgG against LDLR (EP1553Y) was purchased from Abcam. Chicken polyclonal IgY (753A) against TcdA was purchased from List Biological Labs. Statistical analysis was performed using OriginPro 8 (v.8.0724, OriginLab) software.

**Protein purification.** Recombinant TcdA (from *C. difficile* strain VPI 10463), TcdA<sup>1-1874</sup>, TcdA<sup>1-1832</sup> and CROPs (TcdA<sup>1875-2710</sup>) were cloned into modified pWH1520 vector, and TcdB<sup>1-1830</sup> was cloned into pHIS1522 vector, expressed in

*Bacillus megaterium* and purified as His<sub>6</sub>-tagged proteins. The expression plasmid pQTEV-LRPAP1 (31327) encoding RAP was obtained from Addgene and RAP was purified as a His<sub>6</sub>-tagged protein. Genes encoding the ectodomain of human LDLR (residues 22-788) and IgG1 Fc were fused and cloned into pHLsec vector (provided by A. Jonathan (Harvard Medical School)). For the expression of Fc-tagged LDLR<sup>22-788</sup>, HEK293T cells were transfected with Lipofectamine 3000 (Invitrogen). Transfected cells were grown for 5 h, and the culture medium was then replaced with serum-free medium for 4 d. LDLR<sup>22-788</sup>-Fc in the culture medium was collected and purified.

**Genome-wide CRISPR-Cas9 screening with TcdA<sup>1-1874</sup>.** The HeLa CRISPR genome-wide knockout library was generated as previously described<sup>20</sup>. In brief, the GeCKO v2 library is composed of two sublibraries. Each sublibrary contains three unique sgRNA per gene and was independently prepared and screened. HeLa-Cas9 cells were transduced with sgRNA lentiviral library at a multiplicity of infection of 0.2. For each CRISPR sublibrary,  $7.9 \times 10^7$  cells were plated onto three 15-cm cell culture dishes to ensure sufficient sgRNA coverage, with each sgRNA being represented around 1,200 times. These cells were exposed to TcdA<sup>1-1874</sup> for 48 h. Cells were then washed three times to remove loosely attached cells. The remaining cells were cultured with toxin-free medium to ~70% confluence and subjected to the next round of screening with higher concentrations of toxins. Three rounds of screenings were performed with TcdA<sup>1-1874</sup> (40, 80 and 160 pM). Remaining cells from each round were collected and their genomic DNA was extracted using the Blood and Cell Culture DNA mini kit (Qiagen). DNA fragments containing the sgRNA sequences were amplified by PCR using primers lentiGP-1\_F (AATGGACTATCATATGCTTACCCTGAACTTGAAGTATTTCCG) and lentiGP-3\_R (ATGAATACTGCCATTGTCTCAAGATCTAGTACGC). NGS (Illumina MiSeq) was performed by Genewiz.

**Generating HeLa KO cell lines.** To generate *EXT2*<sup>-/-</sup>, *EXTL3*<sup>-/-</sup> and *LDLR*<sup>-/-</sup> cells, the following sgRNA sequences were cloned into LentiGuide-Puro vectors (Addgene) to target the indicated genes: 5'-CGATTACCCACAGGTGCTAC-3' (*EXT2*), 5'-GAGGTGAGCATCGTCATCAA-3' (*EXTL3*) and 5'-CCAGCTGGACCCACACGA-3' (*LDLR*). HeLa-Cas9 cells were transduced with lentiviruses that express the sgRNAs. Mixed populations of infected cells were selected with puromycin (2.5  $\mu$ g ml<sup>-1</sup>). For *LDLR* knockouts, three single colonies were isolated (*LDLR*<sup>-/-</sup> no. 4, *LDLR*<sup>-/-</sup> no. 6 and *LDLR*<sup>-/-</sup> no. 7). To generate *SLC35B2*<sup>-/-</sup> cells, a sgRNA targeting exon 1 of *SLC35B2* (5'-GCTTTATGTTACCTGGCTAC-3') was cloned into lentiCRISPR v.2-Blast (Addgene plasmid 83480). Lentivirus was generated by transfecting 293T cells with lentiCRISPR v.2-Blast-*SLC35B2*sgRNA, pCD/NL-BH<sup>+</sup>DDD and pCAGGS-VSV-G. HeLa-Cas9 cells were transduced with the lentivirus and selected with 10  $\mu$ g ml<sup>-1</sup> blasticidin. Three single colonies were isolated and validated (*SLC35B2*<sup>-/-</sup> no. 1, *SLC35B2*<sup>-/-</sup> no. 3 and *SLC35B2*<sup>-/-</sup> no. 5). *SLC35B2*<sup>-/-</sup>*LDLR*<sup>-/-</sup> double-KO cells were generated from *SLC35B2*<sup>-/-</sup> no. 5 by transducing lentiviruses that express *LDLR* sgRNA (5'-CCAGCTGGACCCACACGA-3'). Stable *SLC35B2*<sup>-/-</sup>*LDLR*<sup>-/-</sup> cells were selected with puromycin (2.5  $\mu$ g ml<sup>-1</sup>) and hygromycin B (200  $\mu$ g ml<sup>-1</sup>). The deficiency of LDLR in *LDLR*<sup>-/-</sup> and *SLC35B2*<sup>-/-</sup>*LDLR*<sup>-/-</sup> cells was validated by immunoblot.

**FACS analysis.** In brief, cells were collected with 1 mM EDTA in PBS and subsequently re-suspended in PBS with 1% BSA. Cells were incubated with either the 10E4 monoclonal antibody against heparan sulfate (1:400), or mouse IgM (1:200; ab18401, Abcam) for 1 h on ice. Cells were washed twice with PBS and incubated with goat anti-mouse IgG/IgM Alexa488 (1:1,000; A10680, Molecular Probes) for 1 h on ice, washed twice, and followed by single-cell sorting using a FACS MoFlo Astrios EQ cell sorter (Beckman Coulter). Data were analysed using FlowJo software (FlowJo).

**Cytopathic cell-rounding assay.** The cytopathic effect of TcdA and TcdB was analysed using the standard cell-rounding assay. In brief, cells were exposed to TcdA, TcdA<sup>1-1874</sup>, TcdA<sup>1-1832</sup> or TcdB<sup>1-1830</sup> for 24 h, and phase-contrast images of cells were recorded (Olympus IX51,  $\times 10$ - $\times 20$  objectives). A zone of 300  $\times$  300  $\mu$ m was selected randomly, containing 50-150 cells. The numbers of normal and round-shaped cells were counted manually. The percentage of round-shaped cells was analysed using the Origin software.

**Competition assays with GAGs or ecto-domain of LDLR.** TcdA<sup>1-1832</sup> (2 nM) was pre-mixed with or without 1 mg ml<sup>-1</sup> heparan sulfate (Sigma, H7640), chondroitin sulfate (Sigma, C9819), dextran sulfate (Sigma, D4911), hyaluronic acid (Sigma, 53747), heparin (Fisher Bioreagents, BP252450), de-N-sulfated heparin (Carbosynth, YD58544), N-acetyl-de-O-sulfated heparin (Carbosynth, YD58545), sulfated cyclodextrin (Sigma-Aldrich, 494542-5G), GM-1111 (Glycomira) or 400 nM LDLR<sup>22-788</sup> in fresh DMEM medium and incubated at 37 °C for 20 min. The mixture was then added to the cells. Cells were further incubated at 37 °C and the percentage of rounded cells over time was recorded and analysed.

**Competition assays with RAP or surfen.** The cells were pre-incubated with RAP or surfen in the medium at indicated concentrations at 37 °C for 20 min.



The medium was then supplemented with 2 nM TcdA<sup>1-1832</sup> and cells were incubated further at 37 °C and the percentage of rounded cells over time was recorded and analysed.

**Dot blot assay.** The indicated amounts of RAP, TcdA<sup>1-1832</sup>, and TcdB<sup>1-1830</sup> were spotted onto a nitrocellulose membrane and allowed to dry completely in air. The membrane was then blocked with 5% skimmed milk for 1 h at room temperature followed by overnight incubation with LDLR<sup>22-788</sup>-Fc at 4 °C. The bound LDLR<sup>22-788</sup>-Fc was detected with a monoclonal antibody against human Fc fragment. The experiments were repeated in triplicate.

**Surface binding of TcdA<sup>1-1874</sup> on HeLa cells.** TcdA and TcdA<sup>1-1874</sup> were labelled using an Alexa 555 antibody labelling kit (A20187, ThermoFisher Scientific) following the manufacturer's instruction. Wild-type, *SLC35B2*<sup>-/-</sup> or *LDLR*<sup>-/-</sup> HeLa H1-Cas9 cells were incubated with 5 nM Alexa 555-labelled TcdA<sup>1-1874</sup> in PBS for 30 min on ice. Cells were washed three times with ice-cold PBS and fixed with 4% paraformaldehyde. Cell nuclei were labelled with Hoechst dye. Confocal images were captured with the Ultraview Vox Spinning Disk Confocal System.

**BLI assay.** The binding affinities between TcdA<sup>1-1874</sup> and heparin were measured by BLI assay using the Blitz system (ForteBio). In brief, biotinylated heparin (20 µg ml<sup>-1</sup>, B9806, Sigma-Aldrich), biotin-cellulose (Creative PEGWorks, CE501) or biotin-hyaluronate-biotin (Sigma, B1557) was immobilized onto capture biosensors (Dip and Read Streptavidin, ForteBio) and balanced with indicated buffers. The biosensors were then exposed to TcdA<sup>1-1874</sup>, followed by washing. Binding affinities ( $K_D$ ) were calculated using the Blitz system software (ForteBio). The experiments were repeated in triplicate.

**Colon loop ligation assay.** All animal studies were conducted in accordance with ethical regulations under protocols approved by the Institute Animal Care and Use Committee (IACUC) at Boston Children's Hospital (no. 3028). Statistical consideration was not used to determine the sample size of mice. Animals were distributed to each experimental group randomly. Experiments and data analysis were carried out without blinding. Colons from adult CDI mice (6–8 weeks, both male and female, from Envigo) were dissected out and incubated in PBS on ice. A ~2 cm loop in the ascending colon was sealed with silk ligatures. One hundred microlitres of Alexa555-labelled TcdA<sup>1-1874</sup> or TcdA (5 nM each) in PBS was injected through an intravenous catheter into the sealed colon segment with or without TcdA<sup>1875-2710</sup> (150 nM) and/or surfen (5 µM), GM-1111 (1 mg ml<sup>-1</sup>), sulfated and cyclodextrin (1 mg ml<sup>-1</sup>). The colon segments were incubated on ice for 30 min, then cut open, washed with PBS, fixed with paraformaldehyde and subjected to cryosectioning into sections 10 µm thick. Confocal images were captured with the Ultraview Vox Spinning Disk Confocal System. Toxin binding was quantified using ImageJ software. The binding signal intensity was averaged based on the length of the epithelium. Three images were analysed and the *P* value was calculated by Student's *t*-test.

**Caecum-injection assay.** Mice (CD1, 6–8 weeks of age, male and female, Envigo) were anaesthetised with 3% isoflurane after overnight fasting. A midline laparotomy was performed. TcdA (4 µg in 100 µl saline), TcdA premixed with GM-1111 (4 µg TcdA + 0.5 mg ml<sup>-1</sup> GM-1111), or saline was injected into the caecum through the ileocaecal junction. The gut was then returned to the abdomen. The incision was closed with stitches and mice were allowed to recover. After 6 h, mice were euthanized and the caecum plus the ascending colon (~1.5 cm) was excised and weighed. The caecum tissue was fixed with 10% phosphate buffer formalin and embedded in paraffin. Tissue sections were subjected to haematoxylin and eosin staining for histological score analysis based on four criteria (disruption of the epithelium, haemorrhagic congestion, mucosal oedema and inflammatory cell infiltration) on a scale of 0–3 (normal, mild, moderate or severe).

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

## Data availability

The data that support the findings of this study are available from the corresponding authors upon request.

Received: 31 January 2018; Accepted: 19 April 2019;

Published online: 03 June 2019

## References

- Theriot, C. M. & Young, V. B. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. *Annu. Rev. Microbiol.* **69**, 445–461 (2015).
- Collins, J. et al. Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. *Nature* **553**, 291–294 (2018).
- McDonald, L. C. et al. An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N. Engl. J. Med.* **353**, 2433–2441 (2005).
- Hunt, J. J. & Ballard, J. D. Variations in virulence and molecular biology among emerging strains of *Clostridium difficile*. *Microbiol. Mol. Biol. Rev.* **77**, 567–581 (2013).
- Lessa, F. C. et al. Burden of *Clostridium difficile* infection in the United States. *N. Engl. J. Med.* **372**, 825–834 (2015).
- 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. The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* **467**, 711–713 (2010).
- 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).
- Aktories, K., Schwan, C. & Jank, T. *Clostridium difficile* toxin biology. *Annu. Rev. Microbiol.* **71**, 281–307 (2017).
- Cowardin, C. A. et al. The binary toxin CDT enhances *Clostridium difficile* virulence by suppressing protective colonic eosinophilia. *Nat. Microbiol.* **1**, 16108 (2016).
- Chumbler, N. M. et al. Crystal structure of *Clostridium difficile* toxin A. *Nat. Microbiol.* **1**, 15002 (2016).
- Krivan, H. C., Clark, G. F., Smith, D. F. & Wilkins, T. D. Cell surface binding site for *Clostridium difficile* enterotoxin: evidence for a glycoconjugate containing the sequence Gal $\alpha$ 1-3Gal $\beta$ 1-4GlcNAc. *Infect. Immun.* **53**, 573–581 (1986).
- Tucker, K. D. & Wilkins, T. D. Toxin A of *Clostridium difficile* binds to the human carbohydrate antigens I, X, and Y. *Infect. Immun.* **59**, 73–78 (1991).
- Teneberg, S. et al. Molecular mimicry in the recognition of glycosphingolipids by Gal $\alpha$ 3 Gal $\beta$ 4 GlcNAc $\beta$ -binding *Clostridium difficile* toxin A, human natural anti- $\alpha$ -galactosyl IgG and the monoclonal antibody Gal-13: characterization of a binding-active human glycosphingolipid, non-identical with the animal receptor. *Glycobiology* **6**, 599–609 (1996).
- Genisyurek, S. et al. Structural determinants for membrane insertion, pore formation and translocation of *Clostridium difficile* toxin B. *Mol. Microbiol.* **79**, 1643–1654 (2011).
- Olling, A. et al. The repetitive oligopeptide sequences modulate cytopathic potency but are not crucial for cellular uptake of *Clostridium difficile* toxin A. *PLoS ONE* **6**, e17623 (2011).
- Yuan, P. et al. Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B. *Cell Res.* **25**, 157–168 (2015).
- LaFrance, M. E. et al. Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity. *Proc. Natl Acad. Sci. USA* **112**, 7073–7078 (2015).
- 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).
- Pothoulakis, C. et al. Rabbit sucrase-isomaltase contains a functional intestinal receptor for *Clostridium difficile* toxin A. *J. Clin. Invest.* **98**, 641–649 (1996).
- Na, X., Kim, H., Moyer, M. P., Pothoulakis, C. & LaMont, J. T. gp96 is a human colonocyte plasma membrane binding protein for *Clostridium difficile* toxin A. *Infect. Immun.* **76**, 2862–2871 (2008).
- Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
- Kreuger, J. & Kjellen, L. Heparan sulfate biosynthesis: regulation and variability. *J. Histochem. Cytochem.* **60**, 898–907 (2012).
- Chaves-Olarte, E. et al. UDP-glucose deficiency in a mutant cell line protects against glycosyltransferase toxins from *Clostridium difficile* and *Clostridium sordellii*. *J. Biol. Chem.* **271**, 6925–6932 (1996).
- Barth, H. et al. Low pH-induced formation of ion channels by *Clostridium difficile* toxin B in target cells. *J. Biol. Chem.* **276**, 10670–10676 (2001).
- Qa'Dan, M., Spyres, L. M. & Ballard, J. D. pH-induced conformational changes in *Clostridium difficile* toxin B. *Infect. Immun.* **68**, 2470–2474 (2000).
- Smith, R. D. & Lupashin, V. V. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydr. Res.* **343**, 2024–2031 (2008).
- Foulquier, F. et al. TMEM165 deficiency causes a congenital disorder of glycosylation. *Am. J. Hum. Genet.* **91**, 15–26 (2012).
- Jae, L. T. et al. Deciphering the glycosylome of dystroglycanopathies using haploid screens for Lassa virus entry. *Science* **340**, 479–483 (2013).
- Tanaka, A. et al. Genome-wide screening uncovers the significance of *N*-sulfation of heparan sulfate as a host cell factor for chikungunya virus infection. *J. Virol.* **91**, e00432-17 (2017).
- Tian, S. et al. Genome-wide CRISPR screens for Shiga toxins and ricin reveal Golgi proteins critical for glycosylation. *PLoS Biol.* **16**, e2006951 (2018).
- Schuksz, M. et al. Surfen, a small molecule antagonist of heparan sulfate. *Proc. Natl Acad. Sci. USA* **105**, 13075–13080 (2008).

35. Zhang, J. et al. Novel sulfated polysaccharides disrupt cathelicidins, inhibit RAGE and reduce cutaneous inflammation in a mouse model of rosacea. *PLoS ONE* **6**, e16658 (2011).
36. Yamamoto, S. et al. Lipoprotein receptors redundantly participate in entry of hepatitis C virus. *PLoS Pathog.* **12**, e1005610 (2016).
37. Fisher, C., Beglova, N. & Blacklow, S. C. Structure of an LDLRRAP complex reveals a general mode for ligand recognition by lipoprotein receptors. *Mol. Cell* **22**, 277–283 (2006).
38. Finkelshtein, D., Werman, A., Novick, D., Barak, S. & Rubinstein, M. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. *Proc. Natl Acad. Sci. USA* **110**, 7306–7311 (2013).
39. Schorch, B. et al. 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 (2014).
40. Griffin, C. C. et al. Isolation and characterization of heparan sulfate from crude porcine intestinal mucosal peptidoglycan heparin. *Carbohydr. Res.* **276**, 183–197 (1995).
41. Bode, L. et al. Heparan sulfate and syndecan-1 are essential in maintaining murine and human intestinal epithelial barrier function. *J. Clin. Investig.* **118**, 229–238 (2008).
42. Yamamoto, S. et al. Heparan sulfate on intestinal epithelial cells plays a critical role in intestinal crypt homeostasis via Wnt/ $\beta$ -catenin signaling. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G241–G249 (2013).
43. Sauerborn, M., Leukel, P. & von Eichel-Streiber, C. The C-terminal ligand-binding domain of *Clostridium difficile* toxin A (TcdA) abrogates TcdA-specific binding to cells and prevents mouse lethality. *FEMS Microbiol. Lett.* **155**, 45–54 (1997).
44. Zhang, Y. et al. The role of purified *Clostridium difficile* glucosylating toxins in disease pathogenesis utilizing a murine cecum injection model. *Anaerobe* **48**, 249–256 (2017).
45. Lindahl, U., Couchman, J., Kimata, K. & Esko, J. D. *Essentials of Glycobiology* 3rd edn (eds Varki, A. et al.) Ch. 17 (Cold Spring Harbor Laboratory Press, 2015).
46. Kamhi, E., Joo, E. J., Dordick, J. S. & Linhardt, R. J. Glycosaminoglycans in infectious disease. *Biol. Rev. Camb. Phil. Soc.* **88**, 928–943 (2013).
47. Jeon, H. & Blacklow, S. C. Structure and physiologic function of the low-density lipoprotein receptor. *Annu. Rev. Biochem.* **74**, 535–562 (2005).
48. Agnello, V., Abel, G., Elfahal, M., Knight, G. B. & Zhang, Q. X. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc. Natl Acad. Sci. USA* **96**, 12766–12771 (1999).
49. Gustafsen, C. et al. Heparan sulfate proteoglycans present PCSK9 to the LDL receptor. *Nat. Commun.* **8**, 503 (2017).

50. Bomsel, M. & Alfsen, A. Entry of viruses through the epithelial barrier: pathogenic trickery. *Nat. Rev. Mol. Cell Biol.* **4**, 57–68 (2003).

## Acknowledgements

We thank Y. Matsuura (Osaka University) and A. Jonathan (Harvard Medical School) for providing cDNA and cell lines, H. Tatge (Hannover Medical School) for toxin purification, J. Savage (Glycomira) for providing GM-1111 and C. Araneo (Harvard Medical School) for assisting flow cytometry analysis. This study was partially supported by National Institute of Health (NIH) grants (R01NS080833, R01AI132387, R01AI139087, and R21NS106159 to M.D.). R.G. acknowledges support by the Federal State of Lower Saxony, Niedersächsisches Vorab (VWZN2889, VWZN3215 and VWZN3266). M.D. and D.T.B. acknowledge support by the NIH-funded Harvard Digestive Disease Center (P30DK034854) and Boston Children's Hospital Intellectual and Developmental Disabilities Research Center (P30HD18655). L.T. acknowledges support by the National Natural Science Foundation of China (Grant no. 31800128). M.D. and S.P.J.W. hold the Investigator in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund.

## Author contributions

L.T. and M.D. initiated and designed the project. L.T. and S.T. carried out the CRISPR–Cas9 screen. L.T., S.T. and J.Z. carried out colon loop ligation assays. S.T. and J.Z. carried out caecum-injection assays. Z.L., L.R.-M. and S.P.J.W. generated heparan sulfate-deficient cells, analysed cell surface heparan sulfate levels and provided related reagents. S.M. purified LDLR–Fc. R.G. provided TcdA and performed the experiment on CHO cells. D.T.B. and S.O. provided key reagents and advice. L.T. and M.D. wrote the manuscript with input from all co-authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41564-019-0464-z>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to L.T. or M.D.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

*Our web collection on [statistics for biologists](#) may be useful.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

ImageJ, Origin Lab, Excel, FlowJo software (FlowJo Inc, Ashland, OR). (all commercially available).

Data analysis

ImageJ, Origin Lab, Excel, FlowJo software (FlowJo Inc, Ashland, OR). (all commercially available).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available on request from the corresponding author (MD).



## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

### Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<i>Sample size was determined based on authors' past experience and what is commonly accepted sample size number in the literature. Each sample size was selected so that a reasonable scientist would conclude that the size is sufficient to draw a statistical conclusion. At least two biological replicates were carried out.</i>
Data exclusions	No data exclusions.
Replication	All experiments were replicated at least twice. All attempts at replication were successful.
Randomization	Samples were allocated into experimental groups randomly.
Blinding	Blinding was not performed, as virtually all the data are quantitative and not easily subject to operator bias.

### Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).</i>
Research sample	<i>State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i>
Sampling strategy	<i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i>
Data collection	<i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

### Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>
-------------------	---

**Research sample** *Describe the research sample (e.g. a group of tagged *Passer domesticus*, all *Stenocereus thurberi* within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*

**Sampling strategy** *Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*

**Data collection** *Describe the data collection procedure, including who recorded the data and how.*

**Timing and spatial scale** *Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*

**Data exclusions** *If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*

**Reproducibility** *Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*

**Randomization** *Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*

**Blinding** *Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work?  Yes  No

## Field work, collection and transport

**Field conditions** *Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*

**Location** *State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*

**Access and import/export** *Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*

**Disturbance** *Describe any disturbance caused by the study and how it was minimized.*

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involvement
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

### Methods

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials  All unique materials are readily available from the authors or from standard commercial sources (noted in the method section)

## Antibodies

Antibodies used  The following mouse monoclonal antibodies were purchased from the indicated vendors: RAC1 (23A8, Abcam), non-glycosylated RAC1 (Clone 102, BD Biosciences),  $\beta$ -actin (AC-15, Sigma), and HS (F58-10E4, mouse IgM, Amsbio). Rabbit monoclonal IgG against LDLR (EP1553Y) was purchased from Abcam. Chicken polyclonal IgY (#753A) against TcdA was purchased from List Biological Labs.

Validation

Each primary antibody was validated using WB, IHC, IFC, or flow cytometry according to Manufacturer's instructions.

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

HeLa (H1, #CRL-1958), HT-29 (#HTB-38), CHO-C6, and 293T (#CRL-3216) cells were originally obtained from ATCC.

Authentication

All cell lines are from laboratory-derived frozen ampoules produced from ATCC vials.

Mycoplasma contamination

They tested negative for mycoplasma contamination.

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

## Palaeontology

Specimen provenance

*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*

Specimen deposition

*Indicate where the specimens have been deposited to permit free access by other researchers.*

Dating methods

*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*
 Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Fadul CD-1 mice (6-8 weeks, both male and female, from Envigo, NJ).

Wild animals

No wild animals were used.

Field-collected samples

No field-collected samples.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

Recruitment

*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

## ChIP-seq

Data deposition

 Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

 Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

*May remain private before publication.**For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*

Files in database submission

*Provide a list of all files available in the database submission.*Genome browser session  
(e.g. [UCSC](#))*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*



## Methodology

Replicates	<i>Describe the experimental replicates, specifying number, type and replicate agreement.</i>
Sequencing depth	<i>Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.</i>
Antibodies	<i>Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.</i>
Peak calling parameters	<i>Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.</i>
Data quality	<i>Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.</i>
Software	<i>Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.</i>

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation	Cells were harvested and stained with primary antibody and single cells were sorted using a FACS .
Instrument	FACS MoFlo Astrrios EQ(cell sorter-Beckman Coulter)
Software	FlowJo software (FlowJo Inc, Ashland, OR)
Cell population abundance	Single cell clones were cultured and confirmed by heparan sulfate staining and analyzed using FlowJo software
Gating strategy	Forward and side scatter gating. the major density of events is captured by this gate. The events with very low FSC and SSC, as well as those with low FSC and high SSC are eliminated. These events represent debris, cell fragments and pyknotic cells.
	<input type="checkbox"/> Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type	<i>Indicate task or resting state; event-related or block design.</i>
Design specifications	<i>Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.</i>
Behavioral performance measures	<i>State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).</i>

### Acquisition

Imaging type(s)	<i>Specify: functional, structural, diffusion, perfusion.</i>
Field strength	<i>Specify in Tesla</i>
Sequence & imaging parameters	<i>Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.</i>
Area of acquisition	<i>State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.</i>

Diffusion MRI  Used  Not used

## Preprocessing

Preprocessing software

*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization

*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template

*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal

*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring

*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

## Statistical modeling & inference

Model type and settings

*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested

*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis:  Whole brain  ROI-based  Both

Statistic type for inference  
(See [Eklund et al. 2016](#))

*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction

*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

## Models & analysis

n/a | Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis

*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis

*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*