

Design of 8-mer Peptides that Block *Clostridioides difficile* Toxin A in Intestinal Cells

Sudeep Sarma^{1†}, Carly M. Catella^{1†}, Ellyce T. San Pedro², Xingqing Xiao¹, Deniz Durmusoglu¹, Stefano Menegatti^{1,3}, Nathan Crook¹, Scott T. Magness², Carol K. Hall^{1*}

1. Department of Chemical Engineering, North Carolina State University, Raleigh NC 27695-7905, USA

2. Department of Medicine, University of North Carolina at Chapel Hill, NC 27514, United States

3. Biomanufacturing Training and Education Center (BTEC), North Carolina State University, Raleigh, NC 27695, United States

[†]These authors contributed equally to this manuscript.

*To whom correspondence should be addressed: hall@ncsu.edu.

Preprint Servers: The manuscript has been deposited as a preprint on bioRxiv.org

Keywords: Peptide binding design, *C. difficile* infection, toxin A glucosyltransferase domain, SPR, human epithelial cells

This PDF file includes:

Main Text

Figures 1 to 6

Tables 1 to 2

25 **Abstract**

26 Infections by *Clostridioides difficile*, a bacterium that targets the large intestine (colon), impact a large
27 number of people worldwide. Bacterial colonization is mediated by two exotoxins: toxins A and B. Short
28 peptides that can be delivered to the gut and inhibit the biocatalytic activity of these toxins represent a
29 promising therapeutic strategy to prevent and treat *C. diff.* infection. We describe an approach that
30 combines a *Peptide Binding Design* (PepBD) algorithm, molecular-level simulations, a rapid screening
31 assay to evaluate peptide:toxin binding, a primary human cell-based assay, and surface plasmon resonance
32 (SPR) measurements to develop peptide inhibitors that block Toxin A in colon epithelial cells. One
33 peptide, SA1, is found to block TcdA toxicity in primary-derived human colon (large intestinal) epithelial
34 cells. SA1 binds TcdA with a K_D of 56.1 ± 29.8 nM as measured by surface plasmon resonance (SPR).

35

36

37

38 Introduction

39 *Clostridioides difficile* (*C. diff.*) is a Gram-Positive, spore-forming bacterium that infects the intestinal
40 tract of humans and animals. In the last decade, *C. diff.* infection has been the leading cause of diarrhea
41 and inflammation of the colon in North America and in Europe¹. In many cases, *C. difficile* infection is
42 the consequence of a microbial imbalance caused by overtreatment with antibiotics such as penicillin,
43 carbapenem, and fluoroquinolone^{2,3}. These disrupt the gut microbiome, allowing the germination of *C.*
44 *diff.* spores and leading to the proliferation of bacteria and the subsequent release of virulent toxins. In
45 2017, more than 200K people were infected with *C. diff.* resulting in 12,800 deaths in the United States
46 alone^{4,5}. Most of the infections are associated with in-patient care, and more than 80% of the deaths occur
47 in people above 65 years in age⁶. The colonic epithelium is the primary site of infection as the epithelial
48 cells that line the gut wall are highly sensitive to the effects of *C. diff.* toxins and *C. diff.* preferentially
49 colonizes the colon^{7,8}.

50 The pathogenicity of *C. diff.* derives primarily from two major toxins: Toxin A (TcdA) and Toxin
51 B (TcdB)^{9,10}. *C. diff.* adheres to the gut wall using its surface layer proteins and produces two large Rho-
52 glucosylating toxins, TcdA and B, that share ~63% sequence homology^{11,12}. These toxins comprise four
53 domains: glucosyltransferase domain (GTD), autoprotease domain (APD), delivery domain and the
54 combined repetitive oligopeptides domain (CROP) (Figure 1a). The *C. diff.* toxins act via a four-step
55 intracellular mechanism (Figure 1b) : (1) The CROP domain, which is at the C-terminus of the toxins,
56 binds to carbohydrate molecules and proteins on the surface of colonic epithelial cells¹³⁻¹⁵ ; (2) the
57 delivery domain helps translocate the toxin into the cytosol of the target cells; (3) the APD cleaves the
58 GTD from the rest of the toxin; and (4) the GTD utilizes uridine diphosphate glucose (UDP-glucose) to
59 glucosylate Rho-family GTPases that are present in intestinal epithelial cells. The glucosylation of these
60 Rho-family GTPases disrupts transcription, cell cycle progression, apoptosis, and cytoskeleton regulation,
61 leading to cytopathic and cytotoxic effects¹⁶⁻¹⁹.

62 Multiple therapeutic approaches have been developed to treat *C. diff.* infection. The standard
63 practice is treatment with antibiotics (metronidazole and vancomycin), but in 20% of cases infection
64 reoccurs²⁰. Exposure to these antibiotics alters the microbial community in the gut, facilitating
65 colonization by *C. diff.*²¹. Merck introduced a monoclonal antibody, Bezlotoxumab, (marketed as
66 Zinplava) that targets *C. diff.* toxin B. While the rate of recurrent infection among patients receiving
67 Bezlotoxumab was substantially lower than for antibiotic-treated cohorts, the high cost of a single dose
68 (~ \$4K) and its intravenous infusion are burdensome^{22,23}. Another *C. diff.* treatment is Fecal Microbiota
69 Transplant (FMT), an investigational treatment not yet approved by the FDA²⁴. The methods of FMT

administration and optimal dosing strategies still vary from case to case. Additionally, FMT carries the risk of transmitting infectious diseases and antibiotic-resistant bacteria²⁵.

Short peptides are promising candidates for the prevention and treatment of *C. diff.* infection as they are cost effective and can be specific in action. Hence, the goal of this study is to identify peptide inhibitors that bind the catalytic domain of *C. diff.* Toxin A GTD by combining computational design, molecular-level simulations, and experimental refinement. To do this we employ PepBD, a computational *peptide binding design* (PepBD) algorithm developed in our group, which performs high-throughput screening of peptide binders to biomolecular targets²⁶⁻²⁹, e.g., proteins and RNA. The PepBD algorithm has been used successfully in the past to design 15-mer transfer RNA^{Lys3}-binding peptides³⁰, peptides that recognize cardiac troponin I³¹ and neuropeptide Y³², peptide ligands that bind to the Fc and Fab domains of immunoglobulin G^{33,34} and peptides that bind to the Receptor Binding Domain of the SARS-CoV-2 spike-protein³⁵. In an effort to rank and appraise the computationally suggested peptides, a microfluidic bead-based platform was used to rapidly identify peptides that exhibit the desired binding characteristics. This system uses fluorescence imaging and automated image analysis to measure the propensity for both on-target and off-target binding and has previously been applied to identify peptides that bind specifically to Cas9, VCAM-1, and IgG Fab fragments³⁶⁻³⁸. The efficacy of the peptides is tested via a trans-epithelial electrical resistance (TEER) assay on monolayers of the human gut epithelial culture model and via surface plasmon resonance measurements.

The starting point for the work described in this paper is our previously reported computationally identified 10-mer peptide, “NPA”, that binds to TcdA GTD³⁹. Our choice of protein target is TcdA GTD as it builds upon the research conducted in the Feig lab at Wayne State University and addresses the unmet need for efficacious therapeutics targeting TcdA⁴⁰. In their study, they employed phage display to identify short peptides that demonstrate binding affinity to TcdA GTD. This peptide neutralizes TcdA in differentiated small intestinal absorptive cells (SI) but has no effect on differentiated colon absorptive cells. While the mechanisms for this observation are unknown, a possible explanation is that proteases present on the brush border of SI cells cleave the 10-mer peptides into shorter, more-active forms that neutralize the toxins in the SI cells. Since the colon epithelial cells do not appreciably express proteases⁴¹, the 10-mer peptides are less likely to be cleaved in colon cells than in SI cells, and hence cannot neutralize the toxins in the colon cells.

In this work, we computationally design engineered variants of the NPA peptide shorter than 10 amino acids with the goal of identifying effective inhibitors of *C. diff.* TcdA. We begin by performing molecular dynamics simulations of fragments of NPA to see which of them can bind to the catalytic site of TcdA GTD. The simulation results predict that 8-mer peptide candidates are optimum. Hence, we apply the

103 PepBD algorithm to design 8-mer peptide sequences with 8-mer NPA as the “reference peptide”. Explicit
104 solvent atomistic MD simulations and binding free energy calculations are carried out to evaluate the
105 binding of the *in-silico*-suggested peptides to the TcdA GTD in solution. The peptides are rapidly screened
106 for TcdA binding and TcdA GTD binding through an in-house bead-based peptide display system to
107 eliminate weak peptide inhibitors. The efficacy of the peptides that make it through the bead-based peptide
108 display assay are tested using a trans-epithelial electrical resistance (TEER) assay on monolayers of the
109 human gut epithelial culture model. While conventional cellular toxicity assays for *C. diff* toxins use non-
110 physiologically relevant colon cancer cells or transformed kidney epithelial cells, here we use primary
111 human gut epithelial stem cells from the large intestine (descending colon) that are differentiated in the
112 main lineage (absorptive) of the gut lining. The experimental binding affinity of the top performing
113 peptide to TcdA is characterized using surface plasmon resonance (SPR).

114 Highlights of our results are as follows. Seven candidate peptide inhibitors (SA1-SA7) were identified
115 using our PepBD algorithm coupled with molecular level simulations. Based on the bead-based peptide
116 display screen for TcdA GTD binding, four peptides (SA1-SA4) were selected for further *in vitro*
117 assessment. SA1 was the only peptide that demonstrated neutralization properties of TcdA in the colon.
118 The dissociation constant, K_D , of SA1 to TcdA measured by SPR is 56.1 ± 29.8 nM. These findings
119 suggest that peptide SA1 might be an effective therapeutic drug to treat *C. diff.* infection.

120 (Figure 1 should be placed here)

121

122 Results

123 Determining the Optimal Sequence Length and Initial Peptide Sequence for Designing Peptide 124 Inhibitors of TcdA GTD Catalytic Domain

125 We began by determining which fragment of NPA plays the most important role in binding to the
126 TcdA GTD catalytic domain, as this fragment can then serve as the reference peptide in our design
127 process. NPA was identified in our previous study³⁹, which reported computationally designed 10-mer
128 peptide sequences (Supplementary Table 1) that were experimentally tested using a functional cell culture
129 assay for their ability to neutralize TcdA in the small intestinal (jejunum) and large intestinal (colon) cells.
130 The reference peptide used to initialize the computational design in that study was RP: *EGWHAHTGGG*
131 (Figure 2a), discovered by Feig’s team at Wayne State University⁴¹ using phage display and verified
132 experimentally by them to inhibit the glucosyltransferase activity of TcdA. Using our PepBD algorithm<sup>26-
133 29</sup> and molecular dynamics simulations, we identified peptide NPA: *DYWFQRHGHR* (Figure 2c) that
134 binds to TcdA GTD. The critical residues on RP involved in binding to TcdA GTD are E1, W3, H4 and
135 H6 while the critical residues on NPA involved in binding to TcdA GTD are W3, R6 and H9. The key

interacting residues on TcdA GTD are within the reactive loop, viz. residues 509-526. A detailed analysis of the residue-residue interaction between RP:TcdA GTD and NPA:TcdA GTD can be found in our previous work³⁹. The amino acid sequences of residues 509-526 are provided in [Supplementary Note 1](#). Peptides RP and NPA showed toxin-neutralizing activity in jejunum cells but showed no effect in the colon cells. Since cells of the small intestine express proteases on the brush boarder of the cell to break down dietary proteins⁴¹, we speculated that when NPA was applied to jejunum cells, it was getting cleaved into smaller fragments with neutralizing activity. We performed LC-MS/MS on cell culture supernatants from colonic monolayers after peptide NPA and confirmed the presence of shorter peptides derived from full-length NPA. This is not surprising since expression of common proteases in the colon is nearly absent⁴². Thus, based on the assumption that peptides smaller than 10-mers might act as stronger toxin inhibitors, we modified our computational design approach.

Based on the experimental observations described above, we proceeded to test via molecular level simulations if peptides shorter than 10-mers can bind to the TcdA GTD. First, we performed atomistic molecular dynamics simulations of RP and NPA bound to the catalytic domain of TcdA GTD. Plots of the interaction energy (van der Waals + electrostatic + polar solvation energy terms) of each of the ten residues along the peptide sequence with the catalytic site of TcdA GTD were generated. The results indicate that residues on the N-terminus of the peptides have a higher contribution to the interaction energy with the catalytic site of TcdA GTD than the flanking residues on the C-terminus of the peptides ([Figure 2b and 2d](#)). Second, we simulated 6-, 7-, and 8-mer fragments of peptides RP and NPA bound to the catalytic domain of TcdA GTD. The simulations reveal that the 6-mer and 7-mer fragments of RP, *EGWHAH* and *EGWHAHT* are not stable (*i.e.*, $\Delta G_{binding} > 0$) near the binding site of TcdA GTD, whereas the 8-mer fragment *EGWHAHTG* binds to TcdA GTD with a binding free energy of $\Delta G_{binding} = -5.79 \frac{kcal}{mol}$). In contrast, both the 7-mer and 8-mer fragments of NPA, namely 7-mer *DYWFQRH* and 8-mer *DYWFQRHG* ($\Delta G_{binding}$ for 8-mer NPA = $-6.35 \frac{kcal}{mol}$) show good binding affinity for the TcdA GTD ([Figure 2e](#)). Accordingly, we resolved to design 8-mer peptide variants using *DYWFQRHG* as reference ligand for the PepBD algorithm.

(Figure 2 should be placed here)

***In-silico* Screening of TcdA GTD Binding Peptides and Evaluation of Binding Free Energies**

PepBD is a Monte Carlo-based *peptide binding design* algorithm that uses an iterative procedure to optimize the binding affinity and selectivity of peptides to a biomolecular target. The algorithm utilizes

as input the structure of the complex formed between an initial peptide sequence (reference peptide) and the target biomolecule and selects peptide variants by implementing sequence and conformation change moves on the peptide chain. The desired hydration properties of the designed peptides can be customized based on 6 residue types (hydrophobic, hydrophilic, positive, negative, other, and glycine). The classification of the 20 natural amino acids into the six residue types can be found in [Supplementary Table 2](#). A score function, Γ_{score} , which considers (i) the binding energy of the peptide to the receptor and (ii) the conformational stability of the peptide when bound to the receptor, is used to evaluate the acceptance of new peptide candidates. Details of the algorithm are provided in *Methods*.

We implemented the PepBD algorithm to identify an improved set of peptide inhibitors using the 8-mer NPA:TcdA GTD complex ([Figure 2f](#)) as the input structure. The $\Delta G_{binding}$ of 8-mer NPA bound to the TcdA GTD is -6.35 kcal/mol. Our goal is to utilize PepBD to generate new sequences that can bind to the TcdA GTD with higher binding affinity than 8-mer NPA. We investigated three cases with different sets of hydration properties for the peptide chain. For all three cases we ensure diversity in the amino acid composition of the peptide chain by allowing a balance among the various contributions to the binding energy, namely electrostatic, hydrophobic, hydrogen bonding, π - π , etc. which contributes to the peptide's affinity and selectivity. The three cases are as follows. Case One: $N_{hydrophobic} = 3$, $N_{hydrophilic} = 2$, $N_{positive} = 2$, $N_{negative} = 1$, $N_{other} = 0$ and $N_{glycine} = 0$, Case Two: $N_{hydrophobic} = 3$, $N_{hydrophilic} = 2$, $N_{positive} = 1$, $N_{negative} = 1$, $N_{other} = 0$ and $N_{glycine} = 1$ and Case Three: $N_{hydrophobic} = 1$, $N_{negative} = 3$, $N_{positive} = 2$, $N_{negative} = 1$, $N_{other} = 0$ and $N_{glycine} = 1$. For each case we perform the PepBD search with three different initial random seed numbers to randomize the initial peptide sequence. This enables our designs to proceed along different search pathways and sample peptides from a large pool of peptide sequences and conformations. During the *in-silico* evolution, new sequences and conformers are generated by mutating and exchanging amino acids on the peptide chain, which results in a fluctuation of the score. A lower Γ_{score} means stronger binding affinity of a peptide to the bound target. The root-mean-squared deviation, RMSD, of the new peptide conformers compared to the conformation of the initial peptide chain's conformation reflects the changes in the backbone scaffold of the peptide as the design process progresses. [Figure 3a](#) shows the Γ_{score} and the RMSD profile vs the number of sequence and conformation change moves performed with a distinct initial random seed for Case 1. [Figure 3b](#) shows the structure of one of the top performing peptides, SA1: *EFWWRRHN*, complexed with the TcdA GTD binding interface from Case 1 (random seed 1). Peptide SA1 has a $\Gamma_{score} = -44.59$ which was obtained at the 459th step of the sequence evolution. From Case 2 (random seed 3) the top performing peptide is SA3 *QEWGMRHW* ([Supplementary Note 2, Supplementary Figure 1A and 1C](#)), and from Case 3 (random seed 3) the top performing peptide is SA6 *EGWQHRHR* ([Supplementary Note 2, Supplementary Figure 1B and 1D](#)); more information on SA3 and SA6, and Cases

2 and 3 is provided in [Supplementary Note 2](#). A comprehensive list of the top peptide sequences obtained from PepBD, with their corresponding Case, random seed, Γ_{score} , $\Delta G_{binding}$ values and whether or not they were evaluated experimentally is provided in [Supplementary Table 3](#).

Once the *in-silico* evolution terminates, we perform explicit-solvent atomistic molecular dynamics (MD) simulations of the complexes formed between the lowest scoring peptides and TcdA GTD to predict their binding affinity. It is to be noted that Γ_{score} is not an accurate measure of the binding free energy of a peptide bound to TcdA GTD but is rather an important metric for us to use in selecting the peptides that we would like to move forward to perform explicit solvent MD simulations. Three independent simulations are carried out for each peptide:TcdA GTD complex for 100 ns to ensure that the system reaches an equilibrated state. The simulations are performed at 298K using the AMBER ff14SB forcefield and the AMBER18 package. We calculate the $\Delta G_{binding}$ of the peptide:receptor complex following the MD simulations using the MMGBSA protocol and variable dielectric constant method. Details of our atomistic MD simulation and $\Delta G_{binding}$ calculation procedure are provided in *Methods* and Supplementary material^{26-29,43}. [Table 1](#) reports the top peptides that we obtain from our computational procedure and their corresponding scores and $\Delta G_{binding}$ values. We performed a few 500 ns simulations on select peptide:protein complexes and concluded that there were no major conformational changes at long timescales ([see Supplementary Note 3, Supplementary Figure 2, Supplementary Tables 4 and 5](#)). SA1 is the most promising peptide with a $\Delta G_{binding}$ value of -15.94 kcal/mol. (Note: the lower the value of $\Delta G_{binding}$, the higher the binding affinity). The SA1:TcdA GTD complex obtained by performing a hierarchical clustering analysis on the last 5 ns of a 100 ns MD simulation is shown in [Figure 3c](#). Additionally, the plot of the residue-wise decomposition of the interaction energy between SA1 and the catalytic site of TcdA GTD is shown in [Figure 3d](#). The plot reveals that the critical SA1 residues involved in TcdA GTD binding are Trp3, Trp4, Arg5, Arg6, His7 and Asn8. Thus, tryptophan, arginine, histidine, and asparagine on SA1 are the four essential amino acids for TcdA GTD binding. A detailed discussion of key amino acid interactions of SA1 with TcdA GTD is provided in the [Supplementary Note 4 and Supplementary Figure 3](#). Amino acid sequence signatures in *C. diff* TcdA GTD binding peptides were derived from the residue composition of the top 1% of the lowest scoring peptides identified by PepBD for Cases 1 and 2 ([see Supplementary Note 5, Supplementary Figure 4, Supplementary Table 6](#)).

(Figure 3 should be placed here)

(Table 1 should be placed here)

Bead-Based Pre-Screening for TcdA Binding

The candidate peptide inhibitors suggested by PepBD were pre-screened *in vitro* for potent and selective TcdA GTD binding to eliminate weak inhibitors prior to the cell-based assays. Robust inhibition of TcdA glucosyltransferase activity relies on the ability of the peptides to outcompete the TcdA GTD's substrate, UDP-Glucose. Given TcdA's relatively low Michaelis constant ($K_M \sim 4.5 \mu\text{M}$) compared to the cellular concentration of its substrate UDP-Glucose ($92 \mu\text{M}$), it is especially important that the peptides exhibit high binding strength and selectivity to TcdA GTD (see supplementary material for detailed analysis)^{11, 44-47}.

Using a microfluidic screening system developed by our team in prior work^{36,38}, we implemented a dual-fluorescence assay to evaluate the TcdA GTD inhibitory activity of the peptide candidates SA1-SA7, 8-mer NPA and 8-mer RP (Figure 4a). Each peptide sequence was immobilized on a translucent ChemMatrix Aminomethyl bead that was then contacted with red fluorescently labeled TcdA (TcdA-AF594) and with a green-fluorescent analog of UDP-Glucose (UDP-Glucose-Fluorescein). Beads displaying peptides that bind to TcdA exhibit a red fluorescence signal; beads exhibit green fluorescence when UDP-Glucose-Fluorescein binds to the catalytic site of the TcdA GTD. Thus, peptides that can selectively bind to the TcdA GTD and displace the UDP-Glucose-Fluorescein from the catalytic site will exhibit red, but not green, fluorescence. Beads are analyzed by a high-throughput assay (350 beads per hr) that correlates the fluorescence intensity of the beads to the binding strength and selectivity of the peptides displayed. Images of each bead are analyzed via a custom algorithm that ensures consistent and objective bead characterization within a screened ensemble.

Beads displaying TcdA-binding peptides accumulate TcdA-AF594 on their surface, leading to a characteristic red halo fluorescence. This is because TcdA is quite large (MW=308 kDa, 2710 residues) and hence, poorly diffuses into the narrow pores of the beads. The intensity of the halo correlates with the amount of TcdA-AF594 bound³⁸ (Figure 4b). Beads displaying selective TcdA GTD-binding peptides displace UDP-Glucose-Fluorescein from the GTD, and thus peptide:TcdA GTD binding, can be observed as a loss of green fluorescence⁴⁴ from the TcdA binding region (Figure 4c).

The performance of the peptides (SA1-SA7, 8-mer NPA and 8-mer RP) was inferred based on the intensity of the red halo fluorescence, and on the loss of green fluorescence of the beads, respectively (Figure 4d). SA1 was the most promising TcdA binding peptide with the highest mean red halo fluorescence. Despite having high green fluorescence, there was a significant ($p = 0.0011$) reduction in the green fluorescence from the center of the beads to the peptide:TcdA binding interface compared to the no-TcdA-AF594 no-UDP-Glucose-Fluorescein control (RP(-T -U)) (Figure 4c), meaning that SA1 bound selectively to TcdA GTD. SA1, SA2, SA3 and SA4 had higher mean red halo fluorescence, and

thus higher TcdA binding, than 8-mer NPA. SA1, SA3 and SA4 had higher mean red halo fluorescence, and thus higher TcdA binding, than 8-mer RP. Notably, all of the peptides that were screened showed binding to TcdA, however the peptides from Cases 1 (SA1, SA2, SA4) and 2 (SA3, SA5) performed significantly better than those from Case 3 (SA6, SA7). All peptides except SA7 showed a significant ($p < 0.005$) decrease in green fluorescence from the center to the halo (TcdA binding) region, indicating relatively universal exclusion of UDP-Glucose from the peptide:TcdA interface (Figure 4d). Despite differences in background green fluorescence between peptides, the uniform decrease in green fluorescence at the TcdA halo provided little information to differentiate between peptides. The selective binding of most PepBD designed peptides to the TcdA GTD demonstrates the robustness of the peptide design algorithm to reliably identify peptide binders to particular pockets. Peptides SA1-SA4, which had higher red fluorescence than NPA and thus promising TcdA binding, were selected for further evaluation.

(Figure 4 should be placed here)

Functional testing of SA1 on human colonic epithelium.

Peptides SA1-SA4 were screened on a human colon epithelial culture system to preliminarily evaluate neutralizing capabilities. In this assay system, colonic epithelial stem cells are applied to transwell inserts and cultured to confluence on the permeable membrane of the insert (Figure 5A,B). Once the cell barrier is achieved after about 4 days of ISC (intestinal stem cell) expansion, the media is changed to promote differentiation into the primary absorptive lineage of the colon and the cell type most exposed to *C. difficile* toxins⁴⁸⁻⁵⁰. Tight junction proteins are upregulated thereby increasing the barrier function of the cellular monolayer and decreasing the ion flux, which is measured as Trans-Epithelial Electrical Resistance (TEER)⁵¹. Toxicity is measured over time as a drop in TEER indicating TcdA-dependent changes in cytoskeleton, which cause leaky tight junctions. For the quick screen, peptides were pre-incubated with differentiated colon epithelial monolayers for 2 hours, then TcdA was added at a concentration of 30 pM, which is the concentration observed in stool of *C. diff* patients⁵¹. SA1 demonstrated some neutralizing capabilities in the primary screen as observed by preservation of TEER compared to conditions with TcdA without SA1. Because SA1 showed promising neutralizing activity, it was more extensively tested for efficacy using the same TcdA toxicity assay. As expected, colon monolayers responded with a near complete loss of TEER at 12 hours post-TcdA exposure, however, pre-exposure of monolayers with SA1 produced ~79% protection from TcdA toxicity (Figure 5C). These data demonstrate the ability of SA1 to protect barrier function in the presence of clinically relevant concentrations of TcdA.

(Figure 5 should be placed here)

300

301 **Measuring SA1 Kinetic Parameters by Surface Plasmon Resonance**

302 Surface plasmon resonance was used to measure key kinetic parameters for SA1:TcdA binding.
 303 SA1 was covalently bound to a mixed bio-resistant thiol self-assembled monolayer (SAM) on the gold
 304 sensor surface. Characterization of the surface thickness and composition can be found in [Supplementary](#)
 305 [Note 6, Supplementary Figure 5, 6 and 7](#). Various concentrations of the analyte, TcdA, were injected over
 306 the surface and the net angular response in degrees was recorded. The amount of analyte, TcdA, bound to
 307 the surface was calculated using a device-specific conversion factor previously described⁵². The net
 308 equilibrium response was fit to a Langmuir isotherm ([Equation 1](#)) as shown in [Figure 6a](#), resulting in an
 309 equilibrium dissociation constant, K_D , of 56.1 ± 29.8 nM and a maximum binding capacity, Q_{max} , of 12.0
 310 ± 2.2 nmol/m².

$$311 \quad Q = \frac{Q_{max}[TcdA]}{K_D + [TcdA]} \quad 1$$

312 SA1:TcdA binding is a moderate-to-high affinity interaction based on its mid-nanomolar K_D . A
 313 fuller picture of the binding emerges by considering the kinetics in addition to the equilibrium state. Rapid
 314 recognition, which is characterized by a high adsorption rate constant (k_a), is important in the context of
 315 competitive inhibitors. Additionally, complex stability, as reflected by a low desorption rate constant (k_d),
 316 is desired. To succeed as a potent competitive TcdA inhibitor, SA1 needs to outcompete UDP-Glucose for
 317 the active site and slowly dissociate from it. Dynamic response measurements over TcdA injections
 318 allowed for the calculation of k_a and K_D using Equation 2, which are summarized in [Table 2](#), and shown
 319 in [Figure 6b](#).

$$320 \quad Q_t = \frac{Q_{max}[TcdA]}{K_D + [TcdA]} \left(1 - \frac{1}{e^{([TcdA] + K_D)k_a t}} \right) \quad 2$$

321 The average k_a was found to be $7.0 \pm 2.5 \times 10^{-5}$ nmol⁻¹ s⁻¹, which is high for a peptide inhibitor, given its
 322 relative backbone flexibility. Review of the literature indicates that peptide inhibitor adsorption rate
 323 constants typically fall between 10^{-7} and 10^{-6} nmol⁻¹ s⁻¹ [53,54](#). The average k_d was found to be $4.0 \pm 1.4 \times 10^{-3}$
 324 s⁻¹, which is consistent with that of other peptide ligands⁵³⁻⁵⁷. Compared to antibody:antigen binding
 325 interactions, which are known to be high affinity (typical K_D are between 10^{-8} and 10^{-11} M), SA1:TcdA
 326 ($K_D \sim 10^{-8}$ M) exhibits remarkably high binding affinity for non-antibody:antigen biorecognition⁵⁸.

327 **(Figure 6 should be placed here)**

328 **(Table 2 should be placed here)**

329 **Discussion**

330 The goal of this work was to identify lead peptide candidates that bind to the catalytic site of TcdA
 331 glucosyltransferase domain and hence can inhibit the glucosyltransferase activity of toxin TcdA. In our

332 previous work, we reported 10-mer peptides that were able to neutralize TcdA in differentiated small
333 intestinal absorptive cells (SI) but showed no effect on differentiated colon absorptive cells. A likely
334 explanation for this observation is that proteases present on the brush border of SI cells cleaved the 10-
335 mer peptides into shorter, more-active forms that neutralize the toxins in the SI cells. To probe this
336 possibility, we performed mass spectrophotometry on the media containing the 10-mer reference peptide
337 (EGWHAHTGGG⁴⁰) that had been incubated with SI cells. After 20 hours of exposure to the SI cells the
338 reference peptide showed 4 predominant and smaller peptides and the full-length 10-mer reference peptide
339 was the least abundant species suggesting the 10-mer was being cleaved by SI proteases (Supplementary
340 Figure 8). We would not expect this effect in colon cells as they do not generally express proteases⁵⁹. We
341 performed atomistic molecular dynamics simulations to test if shorter peptides might be more effective
342 inhibitors of *C. diff.* TcdA than the 10-mers. The simulations predicted that 8-mer peptides are at an
343 optimum peptide length. We applied our PepBD algorithm and combined it with molecular dynamics
344 simulations and binding free energy calculations to find seven 8-mer peptides (SA1-SA7). These peptides
345 were rapidly screened using an in-house bead-based microfluidic screening technique to check for
346 selective TcdA- GTD binding, and the weak peptide inhibitors (SA5-SA7) were eliminated. The efficacies
347 of peptides SA1-SA4 were tested using a trans-epithelial electrical resistance (TEER) assay on
348 monolayers of the human gut epithelial stem cells from the large intestine (descending colon). Peptide
349 SA1 blocked TcdA toxicity in colon epithelial cells. The binding affinity of this peptide to TcdA was
350 characterized using surface plasmon resonance (SPR) and the dissociation constant, K_D , was found to be
351 56.1 ± 29.8 nM.

352 One of our future goals is to develop peptide inhibitors that bind at the catalytic site of TcdB GTD,
353 as TcdB is an even more challenging and clinically relevant target than TcdA. We computationally
354 evaluated the binding affinity of SA1 for the TcdB GTD (method summarized in the Supplementary Note
355 7). Our computational predictions suggest that peptide SA1 exhibits low binding affinity ($\Delta G_{binding} =$
356 -1.56 kcal/mol) for the TcdB GTD catalytic site. Hence, SA1 is selective towards Tcd A.

357 Anti-toxin drugs (such as SA1 developed here) are therapeutically relevant because they eliminate
358 the causative agent of disease (e.g., the toxin). Anti-toxin drugs are also appealing because they are highly
359 specific (limiting off-target effects on the host or microbiota) and impose low, if any, fitness cost on the
360 toxin-producing pathogen (thereby reducing the pressure for resistance to develop). In this way, anti-
361 toxin drugs are complementary to (and potentially synergistic with) standard-of-care antibiotics. By
362 binding to the toxin's catalytic site, SA1 competitively inhibits the key disease-causing biochemical
363 reaction employed by *C. diff.* Further, because the toxin active site is the most highly conserved region
364 of the toxin, we hope that SA1 will be active on a variety of *C. diff* strains and maintain robust activity as

new strains arise. SA1 represents an improvement in anti-TcdA GTD peptide therapeutics that is similarly potent, but mechanistically distinct from recently published small molecule glucosyltransferase inhibitors⁶⁰⁻⁶¹. Looking forward, the ease with which peptides can be manufactured at scale promises to improve the equitability of access to anti-toxin therapies, and also allows them to be manufactured at the site of disease via engineered gut microbes. Taken together, this work illustrates a structure-guided, rational approach to designing anti-toxin peptides that is readily generalizable to other toxins secreted by *C. diff* and other antibiotic-resistant pathogens.

Methods

1. Materials

N, N-dimethylformamide (DMF), dichloromethane (DCM), Alexa Fluor 594 NHS Ester, N-hydroxysuccinimide (NHS), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC), sodium chloride, HPLC-grade acetonitrile, HPLC-grade formic acid, LCMS-grade acetonitrile, LCMS-grade formic acid, Pierce™ Dye Removal Columns, glacial acetic acid, and hydrochloric acid were purchased from ThermoFisher Scientific (Waltham, MA). Triisopropylsilane-silane (TIS), sodium hydroxide (NaOH), phosphate-buffered saline (PBS) pH 7.4, 1,2-ethanedithiol (EDT), Tween 20, thioanisole, phenol, 3 kDa MWCO Amicon Ultra centrifugal filters, and Glucose-UDP-Fluorescein conjugate were purchased from Millipore-Sigma (Burlington, MA). 30% Hydrogen peroxide, Kaiser test kits, and Magnesium chloride were purchased from Sigma Aldrich (St. Louis, MO). 96% sulfuric acid was purchased from Macron Fine Chemicals (Randor, PA). Petroleum ether, and ethyl ether were purchased from EMD Millipore Corporation (Darmstadt, Germany). Tris HCl was purchased from IBI Scientific (Peosta, IA). Piperidine, Diisopropylethylamine (DIPEA), trifluoroacetic acid (TFA), 1-methyl-2-pyrrolidinone (NMP), 2-(7-aza-1Hbenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), Fmoc-L-Arg(Pbf)-Wang resin, Fmoc-L-Trp(Boc)-Wang resin, Fmoc-L-Thr(tBu)-Wang resin, Fmoc-L-Pro-Wang resin, Fmoc-L-Asn(Trt)-Wang resin, Rink-Amide resin, and all Fmoc protected amino acids were purchased from ChemImpex, Inc. (Wood Dale, IL). ChemMatrix Aminomethyl resin (0.7 mmol/g functional density, 100-200 mesh) was purchased from PCAS Biomatrix, Inc. (Saint-Jean-sur-Richelieu, Quebec, Canada). Toxin A from *Clostridioides difficile* was purchased from List Biological Labs, Inc. (Campbell, CA). Bioresistant alkanetiols hydroxyl terminated ($\text{HSC}_{11}(\text{EG})_3\text{OH}$, 2-{2-[2-(1-mercaptoundec-11-yloxy)-ethoxy]-ethoxy}-ethanol) and carboxyl terminated ($\text{HSC}_{11}(\text{EG})_6\text{OCH}_2\text{COOH}$, (2-(2-(2-(2-(2-(11-mercaptoundecyloxy)-ethoxy)-ethoxy)-ethoxy)-ethoxy)-ethoxy)-ethoxy-acetic acid)) were obtained from Prochimia Surfaces (Poland). Gold sensor slides were obtained from BioNavis Ltd. (Tampere, Finland). Ethanol (200 proof) was obtained from Decon

Labs, Inc (King of Prussia, PA). Milli-Q water (MQ water, resistivity > 18 MΩ cm) was obtained by using a Millipore water purification system (Billerica, MA). Nitrogen gas and liquid nitrogen were obtained from Airgas National Welders (Raleigh, NC).

2. Computational Peptide Design

The PepBD algorithm uses an iterative procedure that optimizes peptide sequences to bind with higher affinity and specificity to a biomolecular target than a known reference ligand. The design process is summarized below.

(a) Generate input peptide:TcdA GTD structure: The input structure for the PepBD algorithm was an 8-mer fragment of an *in-silico* peptide, NPA, identified earlier and experimentally-verified to neutralize TcdA in jejunum cells complexed with the TcdA GTD.

(b) Compute initial score of random peptide:TcdA structure: A random peptide sequence is generated and draped on the backbone scaffold of the initial peptide (NPA 8-mer) bound to the TcdA GTD and its Γ_{score} is calculated.

(c) Iteration of peptide sequence-change and conformation-change moves: The design algorithm performs 10,000 evolution steps and generates variants of the original peptide that bind to the target protein by two kinds of moves: sequence change (mutation) and conformation change.

(d) Evaluate score Γ_{score} of new peptide sequence/conformer: The score of the newly generated peptide sequence or conformer in complex with the TcdA GTD is evaluated. The score function, Γ_{score} , that we use to evaluate newly generated peptide candidates is given by:

$$\Gamma_{score} = \Delta E_{binding} + \lambda(E_{peptide-VDW}^{bound} + E_{peptide-ELE}^{bound} + E_{peptide-EGB}^{bound}) \quad (3)$$

The first term of Eq. (3), $\Delta E_{binding}$, accounts for the difference in the energy of the complex and the energies of the peptide and target biomolecule prior to binding. The second term is the peptide stability term and accounts for the energy of the free peptide in the bound-state configuration. λ is a weighting factor for the peptide stability term with a value of 0.01. Lower scores mean better binders. The force field parameters are taken from the Amber 14SB force field.

(e) Monte Carlo Metropolis Algorithm: The Monte Carlo Metropolis algorithm is used to accept or reject new trial peptides.

More details regarding the PepBD algorithm and Γ_{score} can be found in our previous work²⁶⁻²⁹.

3. Atomistic Molecular Dynamics Simulation

Explicit-solvent atomistic MD simulations are carried out in the canonical (NVT) ensemble using the AMBER 18 package to investigate the dynamics of the binding process between the peptide sequences and the TcdA GTD. The starting configurations of the peptide:TcdA GTD complexes in each MD

simulation are the output from the searches in the PepBD algorithm. We carry out three independent simulations for each peptide:TcdA GTD complex for 100 ns to ensure that the system reaches an equilibrated state. Each peptide-receptor complex is solvated in a periodically-truncated octahedral box containing a 12 Å buffer of TIP3P water (~ 36,000 water molecules) surrounding the complex in each direction. Counterions such as Na⁺ or Cl⁻ were added to neutralize the peptide:protein complex prior to running the MD simulations. No additional salt ions were added. The implicit-solvent molecular mechanics/generalized Born surface area (MM/GBSA) approach with the variable internal dielectric constant model is used to post-analyze the last 5 ns simulation trajectories of the peptide:TcdA GTD complexes to calculate the binding free energies. Details of the computational procedures and post-analysis of the atomistic MD simulations can be found in our previous work²⁶⁻³¹. MD simulation parameters are described in [Supplementary Table 7](#).

4. Bead-Based Screening Assay

4.1 Solid Phase Peptide Synthesis

SA1-SA7, NPA (8-mer), and RP (8-mer) were synthesized on ChemMatrix Aminomethyl resin following a GSG linker on a Biotage Syro I peptide synthesizer (Biotage, Uppsala, Sweden) following the Fmoc/tBu protecting strategy. The GSG linker aids in displaying the peptide on the surface of the resin. The resin was swelled in DMF for 30 min. Amino acids were coupled by incubating the resin with 3 equivalents (relative to functional density of resin) protected amino acid, 3 eq. HATU, and 6 eq. DIPEA in dry DMF for 15 minutes at 45 °C. The coupling of each amino acid coupling was monitored by Kaiser test. Fmoc removal of the resin and after each amino acid conjugation was performed using 5 ml 20% v/v piperidine in DMF at room temperature for 3 min and then again for 10 min. Protected resin was stored dried under nitrogen at 4 °C. The completed peptide was deprotected immediately before use through acidolysis by incubating the resin in the deprotection cocktail, Reagent K, 82.5% v/v TFA, 5% v/v phenol, 5% v/v water, 5% v/v thioanisole, 2.5% v/v EDT for 3 hr at room temperature and under mixing. The deprotected resin was rinsed with DCM, DMF, then DCM and dried under nitrogen.

4.2 Screening

Peptide sequences were screened for binding to TcdA on an in-house microfluidic bead imaging system originally designed for sorting solid phase peptide libraries^{36,38}. Exclusion of a fluorescent UDP-Glucose co-factor analog (UDP-Glucose-Fluorescein) from the UDP-Glucose binding pocket of TcdA was used as a proxy for visualizing peptide binding in the desired position.

Tris-buffer was removed from the TcdA solution with 3 kDa MWCO Amicon Ultra centrifugal filters, through 5 rounds of 10-fold concentration and dilution following the manufacturer's recommended protocol. TcdA was labeled with Alexa Fluor 594 through NHS chemistry. 1 µl 10 mg/ml NHS-Alexa Fluor 594 was added to 100 µl 1 mg/ml TcdA. After 1 h incubation, unbound dye was removed with

Pierce Dye Removal Columns per the manufacturer's instructions. UDP-Glucose-Fluorescein was dissolved at 1 mg/ml in PBS, 10 mM Magnesium Chloride (Binding Buffer, BB).

Resin was incubated with 0.2 μ M fluorescently labeled TcdA (TcdA-AF594) overnight. Immediately prior to screening, UDP-Glucose-Fluorescein was spiked into the resin solution at a final concentration of 0.5 μ M for 15 min. The resin was gently washed 4 times with BB + 0.2% Tween 20 (Screening Buffer, SB). Resin beads were imaged in the red and green channels on the microfluidic system previously developed^{36,38}. Beads were visualized on an Olympus IX81 Motorized Trinocular Inverted Fluorescence Phase Contrast Microscope fitted with FITC and RFA8 Chroma filter cubes and were imaged with a Hamamatsu C13440 camera. Each resin solution was diluted in excess SB prior to loading on the microfluidic device. Resin beads were flown through the imaging chamber one at a time and imaged in the red (TcdA-AF594) and green (UDP-Glucose-Fluorescein) channels. Approximately 30 individual beads from each peptide sequence were imaged.

Image processing and analysis was performed on the red and green channel images for each bead. Because the red channel displayed prominent halo fluorescence, the 90th percentile of pixels in the bead were used for the analysis. This reduced bias from variance in the center of the bead. Additionally, the concentration of TcdA used for incubation was previously tuned to allow for a range of intensities in the halo, which allowed for differentiation between moderate and strong binders. The mean green fluorescence of the 90th percentile red area and the mean 50th percentile green fluorescence were calculated to determine UDP-Glucose exclusion from TcdA. The average and standard deviation of the 90th percentile of the red channel and the green channel were computed.

5. Free Peptide Synthesis and Purification for Functional Testing of SA1 on Human Gut Epithelium

SA1-SA4 were synthesized on Wang resin pre-loaded with the first amino acid on an Initiator+ Alstra (Biotage, Uppsala, Sweden) following the Fmoc/tBu protecting strategy. The pre-loaded Wang resin was end capped with 1 ml 5M acetic anhydride in 2.5 ml 2M DIPEA for 30 minutes at room temperature. Amino acids were coupled by incubating the resin with 5 equivalents (relative to functional density of resin) protected amino acid, 5 eq. HATU, and 10 eq. DIPEA in dry DMF for 5 minutes at 75 °C. The coupling of each amino acid coupling was monitored by Kaiser test. Fmoc removal after each amino acid conjugation was performed using 5 ml 20% v/v piperidine in DMF at room temperature for 3 min and then again for 10 min. The completed peptide was cleaved and side chain protecting groups were removed through acidolysis by incubating the resin in the deprotection cocktail, Reagent K, 82.5% v/v TFA, 5% v/v phenol, 5% v/v water, 5% v/v thioanisole, 2.5% v/v EDT for 3 h at room temperature and under mixing. The deprotected peptide dissolved in the deprotection cocktail was precipitated in ice cold 50% v/v ethyl ether and 50% v/v petroleum ether (ether solution). The solution was cooled at -80 °C for

30 min, pelleted, and washed three times with ice cold ether solution. The crude peptide pellet was dried under nitrogen, redissolved in 50% v/v acetonitrile and 50% v/v water, and dried on the Biotage V-10 Touch (Biotage, Uppsala, Sweden) prior to purification and analysis.

SA1-SA4 were purified via flash chromatography on an Isolera Prime (Biotage, Uppsala, Sweden) with a Biotage Sfär Bio C18 column. Crude peptide was dissolved in 10% v/v acetonitrile, 90% v/v water, and 0.1% v/v formic acid and applied to a column samplet. Reverse phase chromatography was performed with a gradient from 5% to 70% acetonitrile in water. 0.1% formic acid was used as a modifier. Fractions were collected above a threshold 220 nm and 280 nm absorbance. Fractions were analyzed by LC-MS to identify and quantify the desired peptide. Purified peptide fractions were lyophilized. For final polishing, the peptides were dissolved in 50% v/v water and 50% v/v acetonitrile, filter sterilized, aliquoted, and lyophilized.

6. Cell-based Assay

6.1 Culture of primary colonic stem cells. Donor Selection. Human transplant-grade donor intestines were obtained from HonorBridge (Durham, NC) and exempted from human subject's research by the UNC Office of Human Research Ethics. Donor acceptance criteria were as follows: age 65 years or younger, brain-dead only, negative for human immunodeficiency virus, hepatitis, syphilis, tuberculosis, or COVID-19, as well as no prior history of severe abdominal injury, bowel surgery, cancer, or chemotherapy. Colonic tissue from a 34-year-old Hispanic male was used for all studies. Colonic ISCs from this donor were isolated from primary tissues. Surgical specimens of human colon were obtained from donors at HonorBridge (Durham, NC). Crypts from the colon were removed from the specimen by incubation in a chelating buffer for 75 min at 20°C followed by vigorous shaking in a 50 mL conical tube. The chelating buffer was composed of EDTA (2 mM), dithiothreitol (DTT, 0.5 mM, freshly added), Na₂HPO₄ (5.6 mM), KH₂PO₄ (8.0 mM), NaCl (96.2 mM), KCl (1.6 mM), sucrose (43.4 mM), D-sorbitol (54.9 mM), pH 7.4. Liberated crypts were expanded as a monolayer on a neutralized collagen hydrogels. Crypts were placed on the top of 1 mg/mL collagen hydrogels (1 ml into each well of 6-well plate) at a density of 10,000 crypts/well, overlaid with 3 mL of Expansion Media (EM)⁶² containing 10 mmol/L Y-27632 (S1049; SelleckChem, Houston, TX), and incubated at 37°C in 5% CO₂. EM was used to expand the epithelial cell numbers as monolayers; media was changed the day after seeding and every 48 hours afterwards. When the cell coverage was greater than 80% (typically 4 to 6 days), the epithelium was dissociated to fragments⁶³ to seed onto either 6-well tissue-culture plates coated with collagen hydrogels for continued expansion, or onto 12-well Transwell inserts (3460; Corning, Corning, NY) coated with 1% Matrigel for experiments.

⁶². To generate differentiated enterocytes for the toxicity assays, 2 wells of a 6-well colonic ISC expansion plate, where cells were ~90% confluent, were dissociated ⁶² as described above and plated on 12-well transwell inserts coated with 1% Matrigel. Colonic ISCs were expanded in colon expansion media (EM, see **reference 62** for all media formulations) until confluent (~4-days). After colonic ISCs were confluent, the media was changed to differentiation media (DM) and transepithelial electrical resistance (TEER) was measured every 24-hours using the EVOM2 (World Precision Instruments, FL). At 3-4 days of differentiation when TEER was >1000 ohms/cm², the toxicity assays were initiated. All cells were incubated at 37°C in a humidified environment containing 5% CO₂.

6.2 TcdA and SA1 exposure to differentiated human colonic epithelial cells.

Peptides were diluted at a concentration of 1.0 mM in 500 µl of DM and added to the apical reservoir. The peptides were allowed to preincubate with the colonic monolayers for 2 hours prior to the addition of TcdA (SML1154; Sigma-Aldrich, St. Louis, MO). TcdA at 30 pMol final concentration was added directly to the apical reservoir media containing the peptide and incubated at 37°C in a humidified environment containing 5% CO₂. TEER was measured before and after the peptides were applied. TEER was then measured at the timepoints indicated.

7. Surface Plasmon Resonance (SPR)

7.1 Synthesis and purification of modified SA1 for SPR experiments

Modified peptide SA1-K-amide was synthesized to facilitate grafting to gold sensors through the C-terminal lysine residue. Amide functionalization of the C-terminus prevented electrostatic repulsion between the C-terminus of the peptide and carboxylic acid groups on the SAM. SA1-K-amide was synthesized on Rink Amide resin on an Initiator+ Alstra (Biotage, Uppsala, Sweden) following the Fmoc/tBu protecting strategy as described above. A modified deprotection cocktail, 91.5% w/w TFA, 2.5% w/w water, 2.5% w/w TIS, 2.5% w/w DTT, and 1% w/w indole, was used. SA1-K-amide was collected, purified, and analyzed through the protocol described above.

7.2 Cleaning gold sensors

Gold sensor slides (12 x 20 x 0.5 mm³), 50 nm gold adhered on glass sensors with 2 nm chromium, were cleaned prior to use by soaking in Piranha solution (98% H₂SO₄ and 30% H₂O₂ at 3:1 v/v) for 15-20 min, followed by profuse rinsing in MQ water, rinsing in 200 proof ethanol, and drying under a stream of nitrogen. The gold sensor slides were briefly immersed in 200 proof ethanol prior to modification. (*Warning: Piranha solution reacts violently with organic materials and should be handled with extreme caution.*)

7.3 Self-assembled monolayer (SAM) on gold surfaces followed by peptide grafting

A mixed thiol SAM was produced by dissolving $\text{HSC}_{11}(\text{EG})_3\text{OH}$ and $\text{HSC}_{11}(\text{EG})_6\text{OCH}_2\text{COOH}$ at a 3:1 molar ratio, 1 mM total concentration, in 200 proof ethanol. The gold sensors were immersed in the thiol solution under nitrogen and protected from light for 24 h. The surfaces were vigorously rinsed with 200 proof ethanol to disrupt multilayers, and then dried under nitrogen for subsequent characterization, peptide grafting, or SPR experiments. Modified peptide SA1-K-amide was grafted to carboxylic acid groups on the SAM through NHS/EDC chemistry. 200 mM EDC and 50 mM NHS dissolved in MQ water was applied to the surfaces of the gold sensor for 1 h. The activated gold sensors were rinsed in MQ water and dried under nitrogen. 1 mg/ml peptide SA1-K-amide was dissolved in MQ water and conjugated to the activated gold sensors for 1 h. The peptide grafted gold sensors were rinsed in MQ water and dried under nitrogen. Remaining activated carboxylic acid sites were blocked by incubating 3 M ethanolamine in MQ water on the surface of the gold sensors for 30 min. The gold sensors were rinsed in MQ water and dried under nitrogen for subsequent characterization and SPR experiments.

7.4 Surface density of SAMs and grafted peptide

The surface thickness of the mixed thiol SAM and SAM with grafted peptide SA1 was measured using an M-2000 DI Spectroscopic Ellipsometer (J.A. Woollam, Lincoln, NE) at three angles of incidence $\Phi = 55^\circ, 65^\circ, 75^\circ$. From the bottom up, the samples were modeled as 3 mm Cauchy substrate (glass), 2 nm chromium, and 50 nm gold per the manufacturer's specifications. The initial value of the Cauchy substrate (polymer layer) on the gold surface was estimated to be 25 nm and the surface thickness was measured after the model was generated. Absorbance interference was mitigated by condensing the wavelength range from 600-1000 nm.

7.5 Time-of-flight secondary ion mass spectrometry (ToF-SIMS) analysis of SPR Sensors

Time-of-flight secondary ion mass spectrometry (ToF-SIMS) characterization of gold sensor slides with self-assembled monolayers (SAMs) and SA1 covalently grafted to SAM. Experiments were performed on an TOF SIMS V instrument (ION TOF, Inc., Chestnut Ridge, NY). Positive and negative spectra were recorded for the mixed-thiol SAM and SAM-SA1. The positive secondary ion mass spectra were calibrated using H^+ , C^+ , C_2H_3^+ , C_3H_5^+ , and C_4H_7^+ . The negative secondary ion mass spectra were calibrated using C^- , O^- , OH^- , and C_n^- , respectively.

7.6 Surface Plasmon Resonance (SPR) experiments

A KSV SPR 200 instrument (BioNavis Instruments, Helsinki, Finland) was used to detect changes in the refractive index at the sensor interface. Sensors were equilibrated with 30 $\mu\text{l}/\text{min}$ 50 mM Tris, 50 mM NaCl, 10 mM MgCl_2 (Running Buffer, RB) for more than 5 minutes. After a stable baseline was established, 250 μl TcdA in RB at various concentrations was injected at 30 $\mu\text{l}/\text{min}$ followed by RB. New

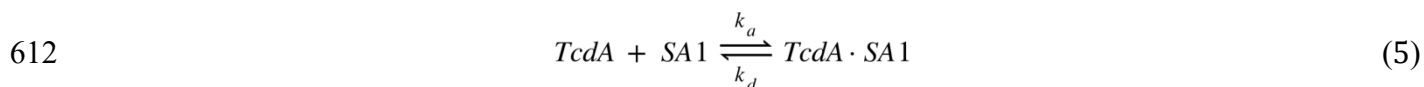
sensors were used for each measurement due to the difficulty disrupting the peptide:TcdA binding interaction.

7.7 Equilibrium and kinetic parameters

The equilibrium dissociation constant, K_D , was found by measuring the equilibrium net change in SPR signal (degrees) after each injection of TcdA at concentrations from approximately $0.1K_D$ to $10K_D$. The net change in degrees was converted to mass of protein adsorbed per unit area through a previously measured conversion factor, 1 degree response = 7.31 mg/m^2 . The data was fit to a Langmuir isotherm using the adsorbed mass per unit area, $Q \text{ (nmol/m}^2\text{)}$, and the solution TcdA concentration, $[TcdA] \text{ (nM)}$,

$$Q = \frac{Q_{max}[TcdA]}{K_D + [TcdA]} \quad (4)$$

where K_D is the equilibrium dissociation constant (nM) and Q_{max} is the maximum binding capacity (nmol/m^2). A Langmuir model is appropriate when there is a 1:1 interaction between the ligand and analyte, there are no mass transfer limitations, and binding events are independent. Assuming reversible binding,



where k_a is the second order association (adsorption) constant and k_d is the first order dissociation (desorption) constant. Assuming no mass transfer limitations, the concentration of TcdA in the bulk and at the surface is equal. The SPR response and amount adsorbed, Q , are proportional to the concentration of TcdA bound to the peptide, $[TcdA \cdot SA1]$. The maximum SPR response and maximum binding capacity, Q_{max} , are proportional to the maximum bound ligand $[TcdA]_{tot}$. Substituting these into the rate equation yields,

$$\frac{dQ}{dt} = k_a[TcdA](Q_{max} - Q_t) - k_d Q_t \quad (6)$$

Integrating Equation 6 and substituting for k_d ,

$$Q_t = \frac{Q_{max}[TcdA]}{K_D + [TcdA]} \left(1 - \frac{1}{e^{([TcdA] + K_D)k_a t}} \right) \quad (7)$$

The first term in Equation 7 determines the equilibrium level, and the second term determines the time to reach equilibrium. The dissociation constant can be determined from K_D and k_a .

624

8. Statistics and reproducibility

Details of the experiments are provided in the main text and in “Methods”. For the bead-based assay, statistical analyses were performed using a two-tailed T-test in conjunction with the RP (-T) control.

Significance was determined at $p < 0.05$. The number of replicates for each sample was as follows: SA1 (23), SA2 (20), SA3 (25), SA4 (19), SA5 (23), SA6 (21), SA7 (15), NPA (36), RP (14), RP (-T) (8), and RP (-T -U) (7). Approximately 40 beads were imaged for each sample, with images containing out-of-focus beads or bead aggregates being excluded. For the SPR assay, the data collected were generally single measurements, except for two conditions (10 nM, $n = 2$; 25 nM, $n = 3$). Each SPR sensor was used to run two samples across two separate channels, leading to a total of 14 measurements. However, one measurement was excluded due to the loss of a stable signal. The measurements were carried out over two sessions, and one condition (25 nM TcdA) was repeated in both sessions as a control.

Acknowledgements

This research was supported by funds from the National Science Foundation (CBET-1934284), from the National Institutes of Health (R01DK109559), from the North Carolina Biotechnology Center (2019-FLG-3841), UNC Center for Gastrointestinal Biology and Disease, P30DK034987 and from the Novo Nordisk Foundation (AIM-Bio), NNF19SA0035474. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by National Science Foundation (ACI-1548562). S. Sarma, X. Xiao and C.K. Hall thank the San Diego Supercomputer Center (SDSC) for computing time.

Data Availability

Data is available in the manuscript and/or supporting information. Amber topology file and final PDB file of each MD simulation run of SA1-SA7, raw data files to reproduce plots are available at: <https://github.com/CarolHall-NCSU-CBE/8-mer-peptides-TcdA>.

Code Availability

The PepBD code and the matlab code for image processing in the bead-assay is available at: <https://github.com/CarolHall-NCSU-CBE/8-mer-peptides-TcdA>.

Competing Interest Statement

S.S, C.M.C, X.X, S.M, N.C, S.T.M and C.K.H have a competing interest. S.S, C.M.C, X.X, S.M, N.C, S.T.M and C.K.H have filed a patent application for this work. E.S.P and D.D have no competing interests.

661

662 **Author contributions**

663 S.S, C.M.C, X.X, D.D, S.M, N.C, S.T.M and C.K.H designed research, S.S, C.M.C and E.S.P performed
664 research, S.S, C.M.C, E.S.P, S.M, N.C, S.T.M and C.K.H analyzed data, S.S, C.M.C, E.S.P, S.M, N.C,
665 S.T.M and C.K.H wrote the paper

666

667

- 669 1. E. G. McDonald, T.C. Lee, Clostridium difficile Infection. *N. Engl. J. Med.* **373**, 286-288 (2015).
- 670 2. J. Czepiel, *et al.*, Clostridium difficile infection: review. *Eur. J. of Clin. Microbiol. Infect. Dis.* **38**, 1211-1221 (2019).
- 671 3. Centers for Disease Control and Prevention, Nearly half a million Americans suffered from Clostridium difficile
672 infections in a single year. Page last reviewed: March 22, 2017 (archived document); Available from:
673 <https://www.cdc.gov/media/releases/2015/p0225-clostridium-difficile.html> (accessed 2 November 2022).
- 674 4. Balsells E, Shi T, Leese C, et al. Global burden of Clostridium difficile infections: A systematic review and meta-
675 analysis. *J. Glob. Health.* **9**, 010407 (2019).
- 676 5. Centers for Disease Control and Prevention, Clostridioides difficile Infection. Page last reviewed: November 13, 2019;
677 Available from: https://www.cdc.gov/hai/organisms/cdiff/cdiff_infect.html (accessed 2 November 2022).
- 678 6. F. C. Lessa, *et al.*, Burden of Clostridium difficile Infection in the United States. *N. Engl. J. Med.* **372**, 825-834 (2015).
- 679 7. L. M. Napolitano, C. E. Edminston, Clostridium difficile disease: Diagnosis, pathogenesis, and treatment update.
680 *Surgery.* **162**, 325-348 (2017).
- 681 8. M. Rupnik, M. H. Wilcox, D. N. Gerding, Clostridium difficile infection: new developments in epidemiology and
682 pathogenesis. *Nat. Rev. Microbiol.* **7**, 526-536 (2009).
- 683 9. K. Aktories, C. Schwan, T. Jank, Clostridium difficile Toxin Biology. *Annu. Rev. Microbiol.* **71**, 281-307 (2017).
- 684 10. M. C. Abt, P. T. McKenney, E. G. Pamer, Clostridium difficile colitis: Pathogenesis and host defence. *Nat. Rev.*
685 *Microbiol.* **14**, 609-620 (2016).
- 686 11. W. J. Bradshaw, A. K. Roberts, C. C. Shone, K. R. Acharya, The structure of the S-layer of Clostridium difficile. *J.*
687 *Cell. Commun. Signal.* **12**, 319-331 (2018).
- 688 12. R. P. Fagan, *et al.*, Structural insights into the molecular organization of the S-layer from Clostridium difficile. *Mol.*
689 *Microbiol.* **71**, 1308-1322 (2009).
- 690 13. J. G. S. Ho, A. Greco, M. Rupnik, K. K. S. Ng, Crystal structure of receptor-binding C-terminal repeats from
691 Clostridium difficile toxin A. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18373-18378 (2005).
- 692 14. T. Murase, *et al.*, Structural basis for antibody recognition in the receptorbinding domains of toxins a and B from
693 clostridium difficile. *J. of Biol. Chem.* **289**, 2331-2343 (2014).
- 694 15. L. Tao, *et al.*, Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to Clostridium difficile
695 toxin A entry into cells. *Nat. Microbiol.* **4**, 1760-1769 (2019).
- 696 16. M. O. P. Ziegler, T. Jank, K. Aktories, G. E. Schulz, Conformational Changes and Reaction of Clostridial
697 Glycosylating Toxins. *J. Mol. Biol.* **377**, 1346-1356 (2008).
- 698 17. P. Chen, *et al.*, Structure of the full-length Clostridium difficile toxin B. *Nat. Struct. Mol. Biol.* **26**, 712-719 (2019).
- 699 18. D. Lyras, *et al.*, Toxin B is essential for virulence of Clostridium difficile. *Nature.* **458**, 1176-1179 (2009).
- 700 19. S. A. Kuehne, *et al.*, The role of toxin A and toxin B in Clostridium difficile infection. *Nature.* **467**, 711-713 (2010).
- 701 20. C. Slimings, T. V. Riley, Antibiotics and hospital-acquired Clostridium difficile infection: update of systematic review
702 and meta-analysis. *J. of Antimicrob. Chemother.* **69**, 881-891 (2014).
- 703 21. B. B. Lewis, *et al.*, Loss of Microbiota-Mediated Colonization Resistance to Clostridium difficile Infection with Oral
704 Vancomycin Compared with Metronidazole. *J. Infect. Dis.* **212**, 1656-1665 (2015).
- 705 22. Y. Lee, *et al.*, Bezlotoxumab (Zinplava) for Clostridium Difficile Infection: The First Monoclonal Antibody Approved
706 to Prevent the Recurrence of a Bacterial Infection. *P. T.* **42**, 735-738 (2017).

707 23. K. Kinoshita, Preclinical and clinical properties of bezlotoxumab (ZINPLAVA® 25 mg/ml concentrate for solution
708 for infusion), novel therapeutic agent for clostridium difficile infection. *Folia Pharmacol. Jpn.* **152**, 39-50 (2018).

709 24. B. J. Kelly, P. Tebas, Clinical Practice and Infrastructure Review of Fecal Microbiota Transplantation for Clostridium
710 difficile Infection. *Chest.* **153**, 266-277 (2018).

711 25. Z. DeFillip, *et al.*, Drug-Resistant *E. coli* Bacteremia Transmitted by Fecal Microbiota Transplant. *N. Engl. J. Med.*
712 **381**, 2043-2050 (2019)

713 26. X. Xiao, C. K. Hall, P. F. Agris, The design of a peptide sequence to inhibit HIV replication: A search algorithm
714 combining Monte Carlo and self-consistent mean field techniques. *J. Biomol. Struct. and Dyn.* **32**, 1523-1536 (2014).

715 27. X. Xiao, M. E. Hung, J. N. Leonard, C. K. Hall, Adding energy minimization strategy to peptide-design algorithm
716 enables better search for RNA-binding peptides: Redesigned λ N peptide binds boxB RNA. *J. Comp. Chem.* **37**, 2423-
717 2435 (2016).

718 28. X. Xiao, P. F. Agris, C. K. Hall. Introducing folding stability into the score function for computational design of RNA-
719 binding peptides boosts the probability of success. *Proteins: Struct. Funct. Genet.* **84**, 700-711 (2016).

720 29. X. Xiao, Y. Wang, J. N. Leonard, C. K. Hall. Extended Concerted Rotation Technique Enhances the Sampling
721 Efficiency of the Computational Peptide-Design Algorithm. *J. Chem. Theory Comput.* **13**, 5709-5720 (2017).

722 30. J. L. Spears, X. Xiao, C. K. Hall, P. F. Agris. Amino acid signature enables proteins to recognize modified tRNA.
723 *Biochemistry.* **53**, 1125-1133 (2014).

724 31. X. Xiao, *et al.*, Advancing Peptide-Based Biorecognition Elements for Biosensors Using in-Silico Evolution. *ACS*
725 *Sensors.* **3**, 1024-1031 (2018).

726 32. X. Xiao, *et al.*, In Silico Discovery and Validation of Neuropeptide-Y-Binding Peptides for Sensors. *J. Phys. Chem.*
727 *B.* **124**, 61-68 (2020).

728 33. X. Xiao, *et al.*, Novel peptide ligands for antibody purification provide superior clearance of host cell protein
729 impurities. *J. Chromatogr. A.* **1625**, 461237 (2020).

730 34. X. Xiao, *et al.* De novo discovery of peptide-based affinity ligands for the fab fragment of human immunoglobulin G.
731 *J. Chromatogr. A.* **1669**, 462941 (2022).

732 35. S. Sarma, S. M. Herrera, X. Xiao, G. A. Hudalla, C. K. Hall, Computational Design and Experimental Validation of
733 ACE2-Derived Peptides as SARS-CoV-2 Receptor Binding Domain Inhibitors. *J. Phys. Chem. B.* **126**, 8129-8139
734 (2022)

735 36. K. Day, *et al.*, Discovery and Evaluation of Peptide Ligands for Selective Adsorption and Release of Cas9 Nuclease
736 on Solid Substrates. *Bioconjugate Chem.* **30**, 3057-3068 (2019).

737 37. R. Prodromou, *et al.*, Engineering Next Generation Cyclized Peptide Ligands for Light-Controlled Capture and Release
738 of Therapeutic Proteins. *Adv. Funct. Mater.* **31**, 2101410 (2021).

739 38. S. Saberi-Bosari, *et al.*, Affordable Microfluidic Bead-Sorting Platform for Automated Selection of Porous Particles
740 Functionalized with Bioactive Compounds. *Scientific Reports.* **9**, 7210 (2019).

741 39. X. Xiao, *et al.* In Silico Identification and Experimental Validation of Peptide-Based Inhibitors Targeting Clostridium
742 difficile Toxin A. *ACS Chem. Biol.* **17**, 118-128 (2022).

743 40. S. J. Abdeen, R. J. Swett, A. L. Feig, Peptide inhibitors targeting clostridium difficile toxins A and B. *ACS Chem. Biol.*
744 **5**, 1097-1103 (2010).

745 41. N. C. Chumbler, *et al.*, Crystal structure of Clostridium difficile toxin A. *Nat. Microbiol.* **15002**, 1-6 (2016).

42. J. Burclaff, *et al.*, A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. *Cell Mol. Gastroenterol. Hepatol.* **13**, 1544-1589 (2022).
43. X. Xiao, B. Zhao, P. F. Agris, C. K. Hall, Simulation study of the ability of a computationally-designed peptide to recognize target tRNA^{Lys3} and other decoy tRNAs. *Protein Sci.* **25**, 2243-2255 (2016).
44. J. W. Loughney, *et al.*, Development of a non-radiolabeled glucosyltransferase activity assay for *C. difficile* toxin A and B using ultra performance liquid chromatography. *J. Chromatogr. A.* **1498**, 169-175 (2017).
45. W. P. Ciesla, D. A. Bobak. Clostridium difficile toxins A and B are cation-dependent UDP-glucose hydrolases with differing catalytic activities. *J. Biol. Chem.* **273**, 16021-16026 (1998).
46. M. R. Laughlin, W. A. Petit, J. M. Dizon, R. G. Shulman, E. J. Barrett. NMR measurements of in vivo myocardial glycogen metabolism. *J. Biol. Chem.* **263**, 2285-2291 (1988).
47. S. Bhattacharyya, A. Kerzmann, A. L. Feig, Fluorescent analogs of UDP-glucose and their use in characterizing substrate binding by toxin A from Clostridium difficile. *Eur. J. Biochem.* **269**, 3425-3432 (2002).
48. D. B. Gunasekara, *et al.*, A Monolayer of Primary Colonic Epithelium Generated on a Scaffold with a Gradient of Stiffness for Drug Transport Studies. *Anal. Chem.* **90**, 13331-13340 (2018).
49. Y. Wang, *et al.*, Self-renewing Monolayer of Primary Colonic or Rectal Epithelial Cells. *Cell. Mol. Gastroenterol. Hepatol.* **4**, 165-182 (2017).
50. I. Gomez-Martinez, *et al.*, A Planar Culture Model of Human Absorptive Enterocytes Reveals Metformin Increases Fatty Acid Oxidation and Export. *Cell. Mol. Gastroenterol. Hepatol.* **14**, 409-434 (2022).
51. L. Song, *et al.*, Development and Validation of Digital Enzyme-Linked Immunosorbent Assays for Ultrasensitive Detection and Quantification of Clostridium difficile Toxins in Stool. *J. Clin. Microbiol.* **53**, 3204-3212 (2015).
52. N. Islam, F. Shen, P. V. Gurgel, O. J. Rojas, R. G. Carbonell, Dynamic and equilibrium performance of sensors based on short peptide ligands for affinity adsorption of human IgG using surface plasmon resonance. *Biosens. Bioelectron.* **58**, 380-387 (2014).
53. Y. Wang, *et al.*, PD-1-targeted discovery of peptide inhibitors by virtual screening, molecular dynamics simulation, and surface plasmon resonance. *Molecules.* **24**, 3784 (2019).
54. P. V. Ershov, *et al.*, Kinetic and thermodynamic analysis of dimerization inhibitors binding to HIV protease monomers by surface plasmon resonance. *Biochem. (Mosc.) Suppl. B: Biomed. Chem.* **6**, 94-97 (2012).
55. M. C. Weiger, *et al.*, Quantification of the binding affinity of a specific hydroxyapatite binding peptide. *Biomaterials.* **31**, 2955-2963 (2010).
56. E. Weiss, *et al.*, Determination of Kinetic Constants for the Interaction between a Monoclonal Antibody and Peptides Using Surface Plasmon Resonance. *Biochemistry.* **31**, 6298-6304 (1992).
57. M. M. Morelock, R. H. Ingraham, R. Betageri, S. Jakes, Determination of Receptor-Ligand Kinetic and Equilibrium Binding Constants using Surface Plasmon Resonance: Application to the Ick SH2 Domain and Phosphotyrosyl Peptides. *Journal of Medicinal Chemistry.* **38**, 1309-1318 (1995).
58. S. Lin, A. Shih-Yuan Lee, C. C. Lin, C. K. Lee, Determination of Binding Constant and Stoichiometry for Antibody-Antigen Interaction with Surface Plasmon Resonance. *Curr. Proteom.* **3**, 271-282 (2007).
59. J. Burclaff, *et al.* A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. *CMGH.* **13**, 1554-1589 (2022).
60. J. Tam, G. L. Beilhartz, A. Auger, P. Gupta, A. G. Therien, R. A. Melnyk, Small Molecule Inhibitors of Clostridium difficile Toxin B-Induced Cellular Damage. *Chem. Biol.* **22**, 175-185 (2015).

786 61. A. S. Paparella, B. L. Aboulache, R. K. Harijan, K. S. Potts, P. C. Tyler, V. L. Schramm, Inhibition of *Clostridium*
787 *difficile* TcdA and TcdB Toxins with Transition State Analogues. *Nat. Commun.* **12**, 6285, 2021.

788 62. M. T. Ok, J. Liu, R. J. Bliton, C. M. Hinesley, E. E. T. San Pedro, K. A. Breau, I. Gomez-Martinez, J. Burclaff, S. T.
789 Magness, A Leaky Colon Model Reveals Uncoupled Apical/Basal Cytotoxicity in Early *Clostridioides difficile* Toxin
790 Exposure. *Am. J. Physiol. Gastrointest.* **324**, 262-280 (2023).

791 63. K. A. Breau, M. T. Ok, I. Gomez-Martinez, J. Burclaff, N. P. Kohn, S. T. Magness, Efficient transgenesis and
792 homology-directed gene targeting in monolayers of primary human small intestinal and colonic epithelial stem cells.
793 *Stem Cell Rep.* **17**, 1493-1506, 2022.

794
795
796

Figures and Tables

Figure 1. *Clostridoides difficile* Toxin A pathogenicity. (a) The crystal structure of Toxin A with the glucosyltransferase domain (GTD, red), autoprotease domain (APD, blue) and delivery domain (orange) (PDB ID: 4R04) is illustrated. (b) Schematic of the TcdA induced toxicity in human epithelial cells. (c) The catalytic site of TcdA (shown in blue) in the GTD (red) plays an important role in inducing *C. diff.* infection.

Figure 2. Determining initial peptide sequence to design peptides that will block *C. diff.* Toxin A in colon cells. (a) Peptide RP (EGWHAHTGGG) at the catalytic site of the Toxin A Glucosyltransferase Domain and (b) the residue wise decomposition of the interaction energy plot of peptide RP. (c) Peptide NPA (DYWFQRHGHR) at the catalytic site of the Toxin A Glucosyltransferase Domain and (d) the residue wise decomposition of the interaction energy plot of peptide NPA. The TcdA GTD is shown in red, and the peptide is colored in blue. (e) Table showing predictions from molecular dynamics simulation of whether or not the 6-mer, 7-mer, and 8-mer fragments on RP and NPA bind to TcdA GTD (f) Starting structure for PepBD: Reference peptide NPA (8-mer) bound to the TcdA GTD at the catalytic site with a $\Delta G_{binding}$ of -6.35 kcal/mol.

Figure 3. Peptide sequence SA1 was obtained from PepBD and molecular dynamics simulations. (a) The score/RMSD vs the number of sequence and conformation steps for Case 1 with a distinct initial random seed results in (b) peptide SA1: EFWWRRHN (c) Snapshot of peptide SA1 bound to TcdA GTD obtained from molecular dynamics simulation. The conformation of the peptide at the binding interface is shown. (d) Plot showing the residue-wise decomposition of the interaction energy (van der Waals + Electrostatic + Polar solvation energy contribution) in the SA1:TcdA GTD complex.

Figure 4. Bead based screening for selective TcdA GTD binding for each peptide sequence using TcdA-AF594 (in red) and UDP-Glucose-Fluorescein (in green). (a) Schematic for bead capture and imaging on custom microfluidic device and microscope. (b) Mean red and green intensity (0-255 pixel value) of 90th percentile red area (TcdA halo) of beads displaying each peptide (number of beads (n) = 23, 20, 25, 19, 23, 21, 15, 36, 14, 8, 7 respectively). RP (-T) and RP (-T -U) are no TcdA-AF594 and no TcdA-AF594/UDP-Glucose-Fluorescein controls, respectively. (c) Percent change in green fluorescence from 90th percentile red halo region to 50th percentile green; more negative values indicate exclusion of UDP-Glucose from peptide:TcdA interface, and thus selective peptide binding to the TcdA GTD (* p<0.005). (d) Representative red and green composite images for each peptide.

Figure 5. Peptide SA1 has functional neutralizing effects on TcdA in a human colonic epithelial culture model. (a) Summary of experimental design and schematic of human ISC expansion in defined media (EM) of human colonic epithelial monolayers on transwells (*a*). Forced differentiation with defined media (DM) (*b*). (c) TEER measurements on differentiated Colon (Descending) monolayers treated with DM (media with vehicle), TcdA, and TcdA with SA1. Values are expressed as a % of TEER at t=0 when TcdA was added to cultures. One way ANOVA $p < 0.005$, multiple comparisons tests 12hr-20hr $p < 0.05$.

Figure 6. Surface plasmon resonance adsorption isotherms and dynamic responses for SA1:TcdA binding. (a) Maximum TcdA adsorption and TcdA concentration fit to a Langmuir isotherm (Equation 1) yielding K_D and Q_{max} . (b) Experimental dynamic response (black solid lines), individual theoretical fit (black dashed lines), and average theoretical fit (solid grey lines, 95% confidence interval in light grey) for TcdA concentrations above, below, and approximately at K_D .

856 **Table 1.** The initial peptide sequence (8-mer NPA) and the list of peptide sequences identified by PepBD
857 screening with their corresponding ΔF_{score} and $\Delta G_{binding}$ values with S.E (n =1250 snapshots).

Peptide	Case	Sequence	$r_{score} \left(\frac{kcal}{mol} \right)$	$\Delta G_{binding} \left(\frac{kcal}{mol} \right)$
8-mer NPA		DYWFQRHG		-6.35 ± 0.51
SA1	Case 1	EFWWRRHN	-44.59	-15.94 ± 0.40
SA2	Case 1	QDWMRRHW	-50.20	-13.19 ± 0.39
SA3	Case 2	QEWMMRRHW	-43.21	-11.76 ± 0.42
SA4	Case 1	MFWEHRHR	-46.74	-11.01 ± 0.49
SA5	Case 2	EFWMGRHH	-42.83	-6.16 ± 0.45
SA6	Case 3	EGWQHRHR	-44.37	-12.54 ± 0.49
SA7	Case 3	HEWGRRHN	-44.90	-9.56 ± 0.47

858

859 **Table 2.** Association and dissociation constants for various TcdA concentration ranges.

Concentration Range	<i>Average k_a at flow rate $30 \frac{\mu l}{min} \left[\frac{n}{M s} \right]$</i>	<i>Average k_d at flow rate $30 \frac{\mu l}{min} \left[\frac{1}{s} \right]$</i>
[TcdA] < K_D	8.1×10^{-5}	4.6×10^{-3}
[TcdA] $\sim K_D$	7.1×10^{-5}	4.0×10^{-3}
[TcdA] > K_D	5.4×10^{-5}	3.0×10^{-3}
All [TcdA]	$7.0 \pm 2.5 \times 10^{-5}$	$4.0 \pm 1.4 \times 10^{-3}$

860

861

862

863