UNIVERSITY OF CALIFORNIA

Santa Barbara

Unveiling the Diversity of Contact-Dependent Inhibition Toxin Families:

A Focus on Ribonucleases and Proteases

A dissertation submitted in partial satisfaction of the

Requirements for the degree Doctor of Philosophy

In Molecular, Cellular, and Development Biology

by

Dinh Quan Nhan

Committee in charge:

Professor Christopher S. Hayes, Chair

Professor David Low

Professor Kathleen Foltz

December 2023

The dissertation of Dinh Quan Nhan is approved.

_____

David Low

_____

Kathleen Foltz

_____

Christopher S. Hayes, Committee Chair

November 2023

Unveiling the Diversity of Contact-Dependent Inhibition Toxin Families:

A Focus on Ribonucleases and Proteases

Acknowledgements

Getting a PhD is not an easy journey, as many are aware. I wouldn't be who I am today without the opportunities provided by Dr. *Christopher S. Hayes*, who allowed me to work as a lab technician and later become a PhD candidate in his group. I am profoundly grateful for everything he has done for me. His passion for science has had a positive and lasting impact on me over the past nine years. Without him constantly reigniting my passion for science, I wouldn't have been able to come this far. Thank you, my committee members, Dr. *David Low* and Dr. *Kathy Foltz*. Thank you for giving valuable feedback and support throughout the years.

I would like to express my gratitude to all the past and present members of the Hayes lab. Without your support and guidance, I would not have become an experienced senior member of the lab, ready to mentor the next generation. Three original members of LSB3211, formerly known as the "Troll cave," played pivotal roles in my scientific career. My beloved mentor, *Julia Willett*, has been both a great friend and an exceptional mentor. I appreciate her rigorous training, which has profoundly influenced my approach to research throughout my PhD career. *Greg Ekberg*, a dear friend for many years, has been a valuable resource within the lab. Whenever I faced research challenges, he was there to offer his expertise and support. I fondly remember our discussions about science and comic books and miss those moments dearly. Lastly, *Grant Gucinski*, a unique friend, has witnessed my growth from an undergraduate to a graduate student. I want to express my sincere gratitude for sharing valuable insights and tips, particularly in Northern blots.

The room of LSB3209 also housed some remarkable individuals that I could never forget. Firstly, I am grateful to have met *Sonya Donato*, who was my TA for MCDB131L.

She was the first person to offer me an opportunity to be part of the Hayes lab. *Kiho Song,* a great friend, provided crucial mental support. Whenever I encountered research obstacles, he would encourage me with words like, "Keep it up, Quan – you can do it!" *Tiffany Halvorsen,* I am grateful to meet you as the first member of the lab who inspired me to think out of the box and apply CDI system in different way. It is sad to say farewell to the late-night crew. *Steven Jensen¸* the closet lab mate, thank you for being a great buddy and I love the time that we could talk science, Pokémon, and most importantly Star Wars! The most important individual I consider my best friend in the lab is Dr. *Fernando Garza-Sánchez*, an encyclopedia of our lab. Without him, I wouldn't even have had the chance to be where I am today. He was the one who passed on valuable life lessons and experiences, with the most important research lesson being "never put all your eggs in one basket." This motivation led me to work on multiple projects, ultimately contributing to the chapters in this thesis.

I want to express my gratitude to *Jose Rojas* and *Ian Matthews*, two pre-med students I've had the privilege of working closely with. Thank you both for sharing your valuable knowledge in medicine, and it has been a great pleasure collaborating with you. I'd also like to extend my appreciation to some of the new recruits in our lab, including *Nicole Chan* and *Mike Costello*, for making the lab experience more enjoyable. A special thanks to Mike for introducing us to all the new and advanced tools in the lab.

My family and friends have been a great source of support throughout my graduate journey, and I would like to thank my mom, brother, aunts, and uncles for their unwavering encouragement. Finally, I thank my loving wife, *Thao Nguyen*. Thank you for providing me with unconditional love and support throughout my PhD. Thank you for being there for my lowest moments. Thank you for giving me a family that I have always wished for. I love you.

VITA OF DINH QUAN NHAN
December 2023

EDUCATION

MA/Ph.D. student in Molecular, Cellular, and Developmental Biology   2017-Present
University of California, Santa Barbara, CA
GPA: 3.96/4.00
Bachelor of science, Microbiology   2014-2016
University of California, Santa Barbara, CA
GPA: 3.36/4.00
Associate in Arts, Biology, Chemistry, and Biomedical Sciences   2011-2014
Santa Barbara City College, CA
GPA: 3.86/4.00

RESEARCH EXPERIENCE

Graduate Student Researcher- University of California, Santa Barbara   2017-Present
Department of Molecular, Cellular, and Developmental Biology
Advisor: Christopher S. Hayes (christopher.hayes@lifesci.ucsb.edu)

- Used biochemical approaches to characterize contact-dependent growth inhibition (CDI) system with emphasis on ribonuclease and protease.
- Identified the cleavage of tRNAs targeted by CDI toxin.
- Used genetic approaches to identify the inner-membrane protein partner required for CDI toxin translocation.
- Studied the cytoplasmic-entry domain on CDI toxin.

Lab Assistant I- University of California, Santa Barbara.   2016-2017
Department of Molecular, Cellular, and Developmental Biology
Advisor: Christopher S. Hayes (christopher.hayes@lifesci.ucsb.edu)

- Screened RNase activities of different toxins via 32P-γ[ATP] northern blot.
- Assisted on DNA cloning for different projects.

Undergraduate Researcher- University of California, Santa Barbara.   2014-2016
Department of Molecular, Cellular, and Developmental Biology
Advisor: Christopher S. Hayes (christopher.hayes@lifesci.ucsb.edu)

- Studied on inner membrane proteins and Colicin related toxins in _E. coli_.

Undergraduate Researcher- University of California, Santa Barbara.   2014-2015
Department of Ecology, Evolution, and Marine Biology
Advisor: Ryoko Oono (ryoko.oono@lifesci.ucsb.edu)

- Studied relationship of endophyte infection frequency and fog gradient

Maxwell Intern, Santa Barbara City College   2013-2014
Advisor: *Jennifer L. Maupin* (jlmaupin@pipeline.sbcc.edu)

- Assisted on DNA barcoding project in spiders.
- Created a spider collection for local species at Robert J. Profant teaching museum.

TEACHING EXPERIENCE

Teaching Assistant-University of California, Santa Barbara

Introduction of Biology- MCDB1A- Fall   2023

- Taught sophomore-level lecture section and graded assignments.
- Covered major and current general biology.

Microbiology-MCDB131-Fall   2021-2022

- Taught senior-level lecture section and graded assignments.
- Covered major and current microbiology and molecular biology.

Molecular Cell Biology for Brain Sciences-MCDB 6-Summer   2022

- Taught students from Psychological and Brain Sciences majors
- Covered introductory levels of molecular and cell biology.

Microbiology Laboratory- MCDB131L-Fall   2017-2022

- Taught senior-level laboratory and graded assignments.
- Organized and wrote laboratory manual for student microbiology course.

Bacterial Pathogenesis Laboratory- MCDB132L-Winter   2018-2019

- Taught senior-level laboratory and graded assignments.
- Covered bacterial genetics materials including phage transduction, transposon mutagenesis, bacterial mating, and transformation.

Medical Microbiology- MCDB138-Spring   2020-2021

- Taught senior-level lecture section and graded assignments.
- Covered medical aspects of common bacterial pathogens.

PUBLICATION

1) Jensen SJ, Ruhe ZC, Williams AF, **Nhan DQ,** Garza-Sánchez F, Low DA, Hayes CS. Paradoxical Activation of a Type VI Secretion System Phospholipase Effector by Its Cognate Immunity Protein. J Bacteriol. 2023 May 22:e0011323. doi: 10.1128/jb.00113-23. Epub ahead of print. PMID: 37212679.

2) Bartelli NL, Passanisi VJ, Michalska K, Song K, **Nhan DQ**, Zhou H, Cuthbert BJ, Stols LM, Eschenfeldt WH, Wilson NG, Basra JS, Cortes R, Noorsher Z, Gabraiel Y, Poonen-Honig I, Seacord EC, Goulding CW, Low DA, Joachimiak A, Dahlquist FW, Hayes CS. Author Correction: Proteolytic processing induces a conformational switch required for antibacterial toxin delivery. Nat Commun. 2022 Oct 21;13(1):6251. doi: 10.1038/s41467-022-34174-z. Erratum for: Nat Commun. 2022 Aug 29;13(1):5078. PMID: 36271082; PMCID: PMC9587248.

3) Michalska K, **Nhan DQ**, Willett JLE, Stols LM, Eschenfeldt WH, Jones AM, Nguyen JY, Koskiniemi S, Low DA, Goulding CW, Joachimiak A, Hayes CS. Functional plasticity of antibacterial EndoU toxins. Mol Microbiol. 2018 Aug;109(4):509-527. doi: 10.1111/mmi.14007. Epub 2018 Aug 12. PMID: 29923643; PMCID: PMC6173971

HONORS AND AWARDS

Graduate
2022   2022-23 Doctoral Student Travel Grant
2019   22nd BMSE/MCDB Annual Retreat-Best poster award
2019   Molecular Genetics of Bacteria and Phages Meeting-Outstanding poster award
Undergraduate
2016   Undergraduate Research and Creative Activities (URCA) Grant
2016   Departmental honors in Microbiology from UCSB
2015   Robert J. Profant scholarship from SBCC
2014   Natural Journal Scholarship from the Santa Barbara Museum of Natural History
2013   Outstanding student in Cell Biology class at SBCC

ORAL PRESENTATION

Nhan DQ and Hayes CS. Unveiling a novel protease toxin from *Citrobacter rodentium* that targets Gyrase and Topoisomerase IV.
Bacterial Toxin Delivery Meeting. Sigtuna, Sweden. Augst 2023.

POSTER PRESENTATIONS

Graduate

Nhan DQ and Hayes CS. A novel contact-dependent growth inhibition (CDI) toxin from *Citrobacter rodentium* inactivates DNA gyrase and topoisomerase IV.
Molecular Genetics of Bacteria and Phage Meeting. Madison, WI. August 2022.

Nhan DQ, Michalska K, and Hayes CS. A peptide antibiotic transporter SbmA is required for O32:H37 CDI RNase translocation.
Molecular Genetics of Bacteria and Phage Meeting. Madison, WI. August 2019.

Undergraduate

Nhan DQ, Willett JLE, and Hayes CS. The role of inner membrane proteins in contact-dependent Inhibition system in *E. coli*.
UCSB Undergraduate Research Colloquium. Santa Barbara, CA. April 2016

MENTORING

Graduate Student Mentor

Ian Matthews (Fall 2021)
A mutagenesis approach for identifying receptor recognition and inhibitory mechanisms of contact dependent growth inhibition systems.
Christine Joyce (Spring 2020)
An investigation on a putative permissive factor for a ribonuclease from *Yersinia kristensenii*.

Undergraduate Student Mentor

Cecilia Cheng (Fall 2022)
An investigation on fimbrial delivery system from *Citrobacter rodentium*.

Abstract

Unveiling the Diversity of Contact-Dependent Inhibition Toxin Families:

A Focus on Ribonucleases and Proteases

by

Dinh Quan Nhan

Bacteria have evolved numerous strategies to establish their communities, enabling them to compete and survive in diverse environments. Cell density and complex conditions often trigger bacteria to regulate specific gene expressions, providing them with a competitive advantage. One prevalent mechanism observed among Gram-negative bacteria is the Contact-dependent Growth inhibition (CDI) system, which requires close contact with neighboring cells to deliver toxic effectors. The CDI system utilizes the two-partner secretion system to facilitate toxin delivery. The cycle starts with the biogenesis of a beta-barrel protein, CdiB on the CDI+ cell outer membrane. Its periplasmic domain pushes a large filamentous CdiA protein through the lumen of CdiB and allows it to secrete until it reaches a secretion arrest signal. In this conformation, the receptor binding domain (RBD) is located at the distal end of the filament, waiting for the target cell outer membrane receptor recognition. Furthermore, a cognate immunity protein, CdiI is encoded within the CDI+ cell that safeguards cells from self- and kin-intoxications.

The CdiA filament features an intricate design with domains designated for different functions. Strikingly, the filament is well conserved across species up until the VENN motif at the C-terminus. Sequence alignment revealed that post-VENN sequences encode the effector proteins targeting different components of the bacterial cells. This thesis reports several aspects of CDI, including the two unique RNase effectors isolated from *Escherichia coli* STEC_O31 and O32:H37, a distinctive protein translocation to import CdiA-CT, and a novel protease toxin found in *Citrobacter rodentium* DBS100.

Chapter I provides a general introduction to contract-dependent growth inhibition and other inhibitory systems, offering an overview of our current understanding of key features of the CDI. Chapter II presents a detailed characterization of an EndoU toxin from STEC_O31. In Chapter II, we introduce a different class of RNase toxin that utilizes an elongation factor (EF-Tu) as a co-factor for activity from *Escherichia coli* O32:H37. Chapter IV focuses on a translocation mechanism of CdiA-CT$^{O32:H37}$ via SbmA in *E. coli*. Finally, in Chapter V, we expand our search for new toxin effectors, leading to the discovery of a cysteine protease isolated from *Citrobacter rodentium* DBS100.

.

Table of Contents

Chapter I: Introduction

The prevalence of bacteria spanning the entirety of Earth is a well-acknowledged fact, yet the true significance they hold in our daily lives remains somewhat elusive to many. Modern advancements in molecular biology and biochemistry have granted us a new perspective to look at these microorganisms. The dynamics of microbial interactions can be broadly classified into two main categories, interactions between microorganisms and their host, and interaction between different types of microorganisms themselves.

Among these interactions, those between microorganisms and their host exhibit a spectrum of effects, both beneficial and detrimental. Certain microbes establish symbiotic relationships by residing within or on their host organisms. For instance, *Rhizobium* species resides in the root nodules of legumes to facilitate nitrogen fixation (Zahran, 1999), *Pseudonocardia* species provides antibiotics to protect its host, Attine ants from environmental pathogens (Goldstein & Klassen, 2020) , and *Vibrio fischeri* colonizes the light organ of the squid *Euprymna scolopes* to disguise its silhouette from predators on the ocean floor with bioluminescence (Nyholm & McFall-Ngai, 2021). However, some critters exhibit pathogenic tendencies to the host. For example, *Escherichia coli* equipped with many virulent factors that can cause intestinal and extraintestinal diseases (Kaper et al., 2004).

In recent times, a growing body of evidence has caused many researchers to propose that comprehending the interplay among microbes might hold the key to balancing the beneficial and pathogenic bacteria within our own gut microbiomes. This insight could shed light on immune response and inflammatory conditions and lead to novel biotherapeutic interventions (Ogunrinola et al., 2020; Pickard et al., 2017). Central to this understanding is

the recognition that bacteria are always competing for limited nutrients and resources for survival. Microbial communities have evolved many competitive mechanisms for harvesting scarce resources, including rapid growth to outpace competitors, combative strategies to eliminate neighboring rivals, and alternative metabolic pathways to outcompete adversaries (Ghoul & Mitri, 2016).

This exploration delves into the realm of combative strategies adopted by bacteria, which can be further categorized into two subtypes: those operating in a contact-less environment, exemplified by bacteriocins, and those reliant on direct contact scenarios, typified by various secretion systems.

A) Bacteriocins: colicins

Bacteriocins, potent antimicrobial peptides, emerge from the realms of both Gram-negative and Gram-positive bacteria (Cascales et al., 2007; Jack et al., 1995). Their discovery dates back nearly a century, as ribosomal synthesized, antimicrobial peptides. Operating as pore-forming toxins, these peptides exert their effect on the cell membranes of target bacteria. The primary provenance of these bacteriocins is within the lactic acid bacteria (LAB), such as *Lactobacillus, Enterococcus, Streptococcus, and Bifidobacterium.* LAB exhibit remarkable antimicrobial properties to foodborne pathogens (Darbandi et al., 2022).

The illustrious family of bacteriocins, known as colicins, takes center stage in this narrative. Emanating from select strains of *E. coli*, colicins carry lethal intent directed exclusively at closely related *E. coli* kin. The first colicin molecule was first described by Gratia in 1925 from *E. coli* V with its potency against *E. coli* φ (Gratia, 1925). Subsequently, akin proteins resembling colicins were unearthed from diverse members of the enteric bacterial community, such as cloacins from *Enterobacter cloacae,* pyocins from

Pseudomonas pyogenes, and marcesins from *Serratia marcescens* (Cascales et al., 2007).

Within the scope of this review, our spotlight centers on the captivating realm of colicins.

Their biology and behavior serve as a foundational framework for comprehending bacterial

inhibition within an environment devoid of direct contact.

Colicin regulation and release

  Colicins, weighing in at 40-80 kDa, originate from colicinogenic *E. coli* strains,

where they are encoded by the colicinogenic plasmid known as pColD. (Hardy et al., 1973;

Herschman & Helinski, 1967). Most described colicins exert their activities by degrading

nucleic acids and forming pores in the membranes of the target cells. A few other unique

colicins are found to have alternative substrates, such as colicin M, which primarily targets

the peptidoglycan (Cascales et al., 2007). The genetic layout of most known colicins resides

within an operon contains two to three genes that encode for the structural toxin product, an

immunity protein that neutralizes the toxin, and in some specific colicins, a lysis protein is

also encoded to lyse the cells for the release of colicins (Riley, 1993a, 1993b).

  The purpose of colicin lies in their ability to selectively eliminate neighboring

bacteria of the same species, thereby gaining a growth advantage to their producer.

Remarkably, colicins are designed to avoid harming their own originators. This strategic

objective requires strict gene regulation of the operon; a task accomplished through the

coordination of two distinct promoters. The first of these promoters is an SOS box (P_{SOS}) that

is typically bound to the LexA repressor during normal condition (Cascales et al., 2007),

However, under circumstances of DNA damage, triggered by events such as RecA-mediated

auto-proteolysis, the LexA repressor undergoes alteration (Little & Mount, 1982; Maslowska

et al., 2019). This alteration prompts the expression of SOS response genes and colicin in this

instance. The second promoter, constitutively active, resides upstream to the immunity gene. This strategic layout guarantees a steady supply of immunity protein, a crucial antidote for neutralizing active colicin (Cascales et al., 2007).

Colicin classification

To uncover the receptors recognized by colicins, researchers were able to isolate colicin resistant mutants (Nagel de Zwaig & Luria, 1967; Nomura & Witten, 1967). This method yielded a variety of mutants, ultimately categorized to either acquire mutations within the *tol* or *tonB* gene, which can be further used to group colicins into two distinct groups, groups A and B(Davies & Reeves, 1975; Davis et al., 2004).

Group A colicins, including colicins A, E1-E9, K, L, N, S4, U, and Y, have harnessed the Tol system for their translocation mechanisms. In contrast, group B colicins, comprising colicins B, D, H, Ia, Ib, M, 5, and 10, operate via the TonB system for their entry into target cells (El Ghachi et al., 2006). Strikingly, another distinct characteristic emerges between these two groups: group A colicins are found to be encoded by smaller plasmids and released into medium, and group B colicins have larger plasmids without secreting to the medium. This division not only sheds light on the diversity of colicin strategies but also offers insights into their distinctive genetic and physiological underpinnings.

Group A colicins utilizing the Tol system were best studied using the DNase colicin E9. When colicin binds to BtuB, an outer membrane protein, the unstructured N-terminal domain of colicin E9 recognizes a secondary outer membrane protein (OmpF) to mediate periplasmic translocation and recruit TolB, a membrane-associated periplasmic protein. Different group A colicins require distinct subsets of the TolABQR proteins for translocation. In contrast, group B colicins require single outer membrane protein for

translocation, TonB. TonB is found to interact with ExbB and ExbD for transducing energy of the proton motive force and facilitate active transport.

Colicin cytotoxic activity

While colicin M, however, utilizes an unique mechanism to degrade peptidoglycan precursor molecules and inhibit cell wall synthesis (El Ghachi et al., 2006), most colicins pass through the membranes of the target cells using either the Tol or Ton systems. Intoxicated cells are then killed through two basic mechanisms: 1) formation of pores in the inner membrane to disrupt the proton motive force or 2) hydrolyzation of the nucleic acids, respectively. The design of the pore-forming colicins is orientation-dependent. It remains inert and does not require an immunity protein for survival when it inserts into the membrane from the intracellular compartment. Most of the pore-forming domains (P-domains) arranged their ~200 amino acids in ten α-helices with the two-core hydrophobic α-helices along with eight amphipathic helices. Although the complete model of pore-forming remains unsolved, it is clear that the P-domains must find a way to rearrange the amphipathic helices and insert the two-core hydrophobic helices into the inner membrane of the target cells to form a lipid-soluble voltage-gated channel (Stroukova & Lakey, 2015).

Nucleases, the other significant category of colicins, exhibit their nuclease activities by cleaving various nucleic acid substrates, such as DNA (colicins E2, E7, E8, E9), 16s rRNA (colicins E3, E4, and E6), and tRNA (colicins D and E5). DNase colicins employ a transition metal cofactor, including Zn^{2+}, Mg^{2+}, and Ca^{2+}. to make random cuts in genomic DNA, ultimately causing double-stranded breaks leading to cell death. In the case of colicins E7 and E9 DNases, studies show that they bind to the minor groove of double-stranded DNA, causing distortion of the phosphodiester backbone towards the coordinated metal ion

(Cascales et al., 2007). Additionally, two classes of RNAs are found to be the common substrates for colicins, 16s rRNA and tRNAs. For instance, colicin E3 targets 16s rRNA near the A site of the ribosome (Ogle et al., 2003), while E5 and D cleave the anticodon loops of specific tRNA molecules. Colicin E5 primarily targets tRNA^{His}, tRNA^{Asn}, tRNA^{Tyr}, and tRNA^{Asp} whereas colicin D cleaves only four arginine isoacceptors (Masaki et al., 1997; Masaki & Ogawa, 2002).

B) Bacterial Secretion Systems

In contrast to the contactless property of colicins and other bacteriocins, Gram-negative bacteria employ various secretion systems for close contact elimination or commonly known as contact-dependent inhibitions (CDIs). Many types of secretion systems have been characterized, including those found in Gram-negative bacteria, types I-VI secretion systems (T1SS-T6SS) and the unique secretion type VII secretion system (T7SS) described in Gram-positive bacteria. T1SS, T3SS, and T4SS are known as one-step transporter systems, requiring single step pathway to secrete the effector molecules across both the inner and outer-membranes. In this thesis, I mainly focus on the type V secretion system.

Contact-dependent growth inhibition (CDI) system

Contact-dependent growth inhibition (CDI) system was first described in *E. coli* strain EC93, an intestinal isolate from a rat colony. It was found that EC93 strain had growth advantage over the laboratory strain *E. coli* K12 strain after several hours. It was also confirmed that the toxin effector was not a diffusible element like colicins, but rather it required direct physical contact with the target cells (Aoki et al., 2005). Hence, this bacteriocin system was coined as contact-dependent growth inhibition. Subsequently, the gene locus responsible for this inhibition event was identified as *cdiBAI* (Aoki et al., 2005).

6

The inhibition process starts with synthesis of the first gene in the locus, *cdiB*. It encodes as a CdiB, β-barrel protein that recognizes the two-partner secretion (TPS) of CdiA, filamentous protein and allows it to thread through its lumen. This filament delivers a polymorphic toxin, CdiA-CT separated by a conserved VE(D)NN motif. The CdiA-CT is equipped with a cytoplasmic-entry domain crucial for recognizing inner-membrane proteins (Aoki et al., 2008; Ruhe et al., 2018). Most importantly, an immunity factor encoded by *cdiI*, is always encoded at the end of the genetic locus to protect the CDI+ cells from self- and kin-intoxications.

CDI uptakes in the target cells

Similar to other bacteriocins, CdiA possesses the ability to specifically recognize outer-membrane proteins in *Escherichia coli*. For instance, group A colicins, target proteins like OmpA and OmpF, while certain group B colicins recognize proteins such as FepA or BtuB. These bacteriocins subsequently translocate to the periplasmic space and make use of inner-membrane proteins for their function. The CdiA protein binds to specific outer-membrane receptors on neighboring target cells. For example, CdiA[EC93] binds to the BamA protein (Aoki et al., 2008; Ruhe, Wallace, et al., 2013), and CdiA[STEC3] binds to the Tsx protein (Ruhe et al., 2018).

Once binding occurs, CdiA delivers a polymorphic toxin derived from its C-terminal domain (CdiA-CT). This CdiA-CT features a cytoplasmic-entry domain essential for recognizing specific inner-membrane proteins. These inner membrane proteins facilitate the translocation of various CdiA-CT variants: CdiA-CT[MHI813] via MetI, CdiA-CT[TTO1] via GltJ/K, CdiA-CT[NC101]/CdiA-CT[EC3006] via PtsG, CdiA-CT[Dd3937] via RbsC, and CdiA-CT[EC869-o11] via YciB (Willett et al., 2015). Additionally, the *Burkholderia* CDI system, BcpA-CT, relies on GltJ/K for its recognition (Myers-Morales et al., 2021).

Many CdiA-CTs require co-factors for their toxicity

Certain CdiA-CTs from specific strains have the capacity to utilize regular cytoplasmic proteins as their enabling factors, either to enhance or trigger the activities of the toxins. Take, for example, the CdiA-CT[UPEC536] domain in *Escherichia coli* strain 536 (UPEC536), which remains latent until it comes into contact with the biosynthetic enzyme O-acetylserine sulfhydrylase A, CysK. In typical *E. coli* cells, the O-acetylserine sulfhydrylase isoenzymes CysK and CysM facilitate the final step of cysteine synthesis from serine. The interaction between CysK and CdiA-CT[UPEC536] closely resembles the one found in CysK/CysE (serine O-acetyltransferase) and activates the latent tRNase activity of CdiA-CT. Remarkably, target cells lacking CysK display complete resistance to the toxin (Diner et al., 2012).

Moreover, various common translation factor proteins, including EF-Tu (elongation factor thermal unstable) and EF-Ts (elongation factor thermos stable), have been observed to form associations with several CdiA-CTs. For instance, CdiA-CT[NC101] binds to domain 2 of EF-Tu to cleave the single-stranded 3'-end of tRNAs containing guanine discriminator nucleotides(Michalska et al., 2017). In the case of CdiA-CT[EC869], it strongly interacts with EF-Tu, and with the support of EF-Ts, it stabilizes the GTP·EF-Tu·tRNA ternary complexes, which in turn facilitate the cleavage of $tRNA^{Gln}$ and $tRNA^{Asn}$ (Jones et al., 2017). Lastly, CdiA-CT[Kp342] necessitates both EF-Tu and EF-Ts to cleave the acceptor stem of deacylated $tRNA_{GAU}^{Ile}$ (Gucinski et al., 2019).

Various CdiA-CTs exhibit diverse characteristics, contributing to the extensive scope of CDI toxins explored in this thesis. Throughout my research, I have identified distinct toxin systems originating from different bacterial species, enriching our understanding of CDI toxins. Chapter II provides an in-depth analysis of an EndoU toxin found in STEC_O31.

Further expanding the repertoire, Chapter III introduces a novel class of RNase toxin that relies on the participation of an elongation factor (EF-Tu) as a co-factor for its activity, sourced from *Escherichia coli* O32:H37. Dedicated to elucidating translocation mechanisms, Chapter IV delves into the CdiA-CT$^{O32:H37}$ translocation process via SbmA in *E. coli*. Lastly, Chapter V broadens the exploration by uncovering a cysteine protease, isolated from *Citrobacter rodentium* DBS100, as a new toxin effector in our pursuit of expanding the understanding of CDI toxins

Chapter II: Functional plasticity of antibacterial EndoU toxins

<u>Note:</u> This research was originally published in *Molecular Microbiology.* I collaborated on this project with Julia Willett. Christopher Hayes and Karolina Michalska contributed to the writing of the manuscript which is presented here in thesis format.

Michalska, K., **Nhan D.Q.,** Willett, J., Stols, L. M., Eschenfeldt, W. H., Jones, A. M., Nguyen, J. Y., Koskiniemi, S., Low, D. A., Goulding, C. W., Joachimiak, A., & Hayes, C. S. (2018). Functional plasticity of antibacterial EndoU toxins. *Molecular microbiology*, *109*(4), 509–527. https:/A/doi.org/10.1111/mmi.14007

Supplementary data can also be found at: https:/A/doi.org/10.1111/mmi.14007

A) Introduction

Bacteria are social organisms that engage in cooperative and antagonistic relationships with their neighbors. Many of these interactions are mediated by small signaling molecules, like quorum sensing auto-inducers, which coordinate group activities like biofilm formation and virulence gene expression (Lopez *et al.*, 2010, Ng & Bassler, 2009). Diffusible inhibitory factors, such as antibiotics and bacteriocins, are exploited as weapons to eliminate rivals (Cascales *et al.*, 2007, Ghequire & De Mot, 2014). Research over the past decade has revealed that bacteria also antagonize their neighbors through direct delivery of protein toxins (Willett *et al.*, 2015b, Souza *et al.*, 2015, Whitney *et al.*, 2017, Vassallo *et al.*, 2017). This phenomenon was first discovered as "contact-dependent growth inhibition" (CDI) between different strains of *Escherichia coli* (Aoki *et al.*, 2005). CDI is mediated by a sub-family of type V secretion systems (T5SS) that are widely distributed throughout proteobacteria, fusobacteria and negativicutes (Aoki *et al.*, 2010, Zhang *et al.*, 2012, Willett *et al.*, 2015b, Jones *et al.*, 2017b). T5SS/CDI[+] bacteria use CdiB transporters to export and present filamentous CdiA effector proteins on the cell surface. CdiA effectors carry a variety of polymorphic C-terminal toxin domains (CdiA-CT), which are transferred directly into neighboring bacteria upon binding specific receptors (Aoki *et al.*, 2008, Beck *et al.*, 2016, Ruhe *et al.*, 2017). T5SS/CDI loci also encode immunity proteins that neutralize CdiA-CT toxin activity and protect the cell against self-inhibition. After the discovery of CDI, type VI secretion systems (T6SS) were also found to mediate proximity-dependent inter-bacterial competition (Hood *et al.*, 2010, MacIntyre *et al.*, 2010). The T6SS is a bacteriophage-like contractile apparatus that injects protein effectors directly into nearby target bacteria (Basler *et al.*, 2012). In contrast to CdiA effectors, which carry a single toxin domain, the T6SS apparatus supports simultaneous delivery of multiple toxins with distinct

activities (Russell *et al.*, 2014). Other secretion systems were first predicted to mediate inter-bacterial competition based on genetic linkage to toxin/immunity gene pairs. Zhang *et al.* predicted that *Neisseria* species export MafB toxins through type II secretion systems (T2SS), and that Gram-positive bacteria use type VII secretion systems (T7SS) to deploy LXG/WXG100/ESAT-6 effectors (Zhang *et al.*, 2012). Those predictions were subsequently supported by experimental work showing that both systems produce toxins and mediate inter-cellular competition (Holberger *et al.*, 2012, Jamet *et al.*, 2015, Cao *et al.*, 2016, Ohr *et al.*, 2017, Whitney *et al.*, 2017). Further experimental studies uncovered additional polymorphic toxin delivery systems. *Bacillus* and *Listeria* use large cell-wall associated YD-repeat proteins to exchange polymorphic toxin domains (Koskiniemi *et al.*, 2013). *Xanthomonas* species deliver toxins though a specialized type IV secretion system (T4SS) (Souza *et al.*, 2015), and a type I secretion system (T1SS) mediates contact-dependent competition in *Caulobacter crescentus* (Garcia-Bayona *et al.*, 2017). Most recently, polymorphic lipoprotein toxins were shown to inhibit cell growth when transferred between strains of *Myxococcus xanthus* through outer membrane exchange (Dey *et al.*, 2016, Vassallo *et al.*, 2017). Collectively, these observations indicate that inter-cellular toxin exchange is a fundamental and ubiquitous facet of prokaryotic biology.

Toxin delivery mechanisms are manifold because the architecture of the bacterial cell envelope varies considerably between phyla. Thus, systems capable of breaching the Gram-negative envelope are ineffective against Gram-positive bacteria and vice versa. Nonetheless, divergent secretion systems from unrelated species often deploy closely related toxins. For example, uropathogenic *E. coli* isolates use T5SS/CdiA to deliver Ntox28 RNase domains into target bacteria (Beck *et al.*, 2016). Ntox28 toxins also form the C-terminal domains of

YD-peptide repeat proteins and T7SS effectors from Gram-positive bacteria (Zhang *et al.*, 2012, Diner *et al.*, 2012). The Ntox28 toxins of Gram-negative and Gram-positive bacteria share similar tRNA anticodon nuclease activities, and their cognate immunity proteins are clearly homologous (Johnson *et al.*, 2016). This sporadic distribution across unrelated species strongly suggest that toxin-immunity gene pairs are spread though horizontal gene transfer (Zhang *et al.*, 2012, Poole *et al.*, 2011, Ruhe *et al.*, 2016). In this model, newly acquired toxin-immunity sequences are integrated into resident secretion systems through genetic recombination (Aoki *et al.*, 2010, Poole *et al.*, 2011, Zhang *et al.*, 2012, Arenas *et al.*, 2013, Unterweger *et al.*, 2014). The modular organization of antibacterial effectors facilitates this process. Effector proteins are typically composites in which variable C-terminal toxins are fused to conserved N-terminal domains that guide export through a specific secretion system (Zhang *et al.*, 2012, Koskiniemi *et al.*, 2013). Thus, polymorphic toxins are versatile payloads that can be delivered through many different export pathways.

EndoU RNase domains are among the most commonly deployed toxins in bacterial competition (Zhang *et al.*, 2012). Aravind and colleagues first predicted that these toxins are related in structure and activity to eukaryotic and viral RNA processing enzymes (Zhang *et al.*, 2012, Zhang *et al.*, 2011). XendoU is the founding member of the EndoU superfamily and was isolated from *Xenopus laevis* as a uridylate specific endonuclease that releases small nucleolar RNAs from introns (Caffarelli *et al.*, 1994, Caffarelli *et al.*, 1997, Gioia *et al.*, 2005, Laneve *et al.*, 2003). RNA viruses of the order *Nidovirales* also encode EndoU nucleases. Nidoviral EndoU (NendoU) domains form the C-termini of nonstructural proteins (Nsp) Nsp15 in coronaviruses and Nsp11 in arteriviruses (Snijder *et al.*, 2003, Nedialkova *et al.*, 2009). These endonucleases are essential for viral replication and have recently been

shown to interfere with innate immune responses (Bhardwaj *et al.*, 2004, Ivanov *et al.*, 2004, Kindler *et al.*, 2017, Deng *et al.*, 2017). XendoU and NendoU nucleases have the same core built from two α/β sub-domains and share a common catalytic triad that emanate s from the N-terminal sub-domain (Joseph *et al.*, 2007, Renzi *et al.*, 2006, Ricagno *et al.*, 2006, Xu *et al.*, 2006, Shi *et al.*, 2016, Zhang *et al.*, 2017). Though unrelated in sequence and fold, EndoU and eukaryotic RNase A nucleases have remarkably similar active sites and use the same catalytic mechanism to produce cyclic 2´-3´-phosphodiester and 5´-hydroxyl termini (Laneve *et al.*, 2003, Ivanov *et al.*, 2004, Renzi *et al.*, 2006, Ricagno *et al.*, 2006, Nedialkova *et al.*, 2009). Two bacterial EndoU toxins have been shown to have RNase activity (Holberger *et al.*, 2012, Jamet *et al.*, 2015), but there is currently no high-resolution structural information for the prokaryotic enzymes. Here, we present the first structure of a bacterial EndoU toxin-immunity protein complex from the T5SS/CDI system of *E. coli* STEC_O31. The CdiA-CT[STECO31] nuclease domain adopts the EndoU fold and contains a canonical catalytic triad, but unexpectedly it lacks the characteristic uridylate specificity of the superfamily. Instead, CdiA-CT[STECO31] is an anticodon nuclease that preferentially cleaves after a cytidylate residue in tRNA[Glu]. This unique specificity is correlated with a distinct architecture that differs from previously characterized antibacterial EndoU toxins.

B) Results
Structure of the CdiA-CT/CdiI[STECO31] complex
We co-expressed the CdiA-CT[STECO31] toxin together with His[6]-tagged CdiI[STECO31] immunity protein and purified the complex by Ni[2+]-affinity chromatography. Limited proteolysis with subtilisin was used to generate a complex that is suitable for crystallization. This treatment removed the N-terminal "translocation" domain of CdiA-CT[STECO31], which does not affect toxin function, but is required for transfer of the CdiA-CT into the cytoplasm

14

of target bacteria (Willett *et al.*, 2015a). Truncated CdiA-CT^{STECO31} and full-length CdiI^{STECO31} crystallized in space group $P6_5$ with one complex in the asymmetric unit. The structure was determined at 2.0 Å by the single wavelength anomalous dispersion approach using selenomethionine (SeMet)-labeled proteins (**Table 1**). The final model contains CdiA-CT^{STECO31} residues Lys181 – Lys323 and residues Met1 – Pro120 of CdiI^{STECO31}. The C-terminal domain of CdiA-CT^{STECO31} adopts a V-shaped structure built from two α/β subdomains (**Fig. 1A**). The N-terminal subdomain consists of a β-hairpin (β1 and β2) followed by helix α1 and an antiparallel β-sheet. The sheet is composed of three major strands (β3, β4 and β6) and a shorter β5 strand that together with β6 flanks β4. The C-terminal subdomain begins at helix α2 and continues as an antiparallel β-sheet (β7, β8 and β9). The N-terminal tail crosses over to the adjacent subdomain, bringing the N- and C-termini into close proximity. The CdiI^{STECO31} immunity protein is composed of a core antiparallel β-sheet (β1´, β2´, β3´ and β4´) surrounded by helices α1´, α2´, α3´ and α4´ (**Fig. 1A**). The β-sheet is located topologically between helices α2´ and α3´. The immunity protein also contains three short sections (G1´, G2´ and G3´) that adopt 3_{10}-helical geometry (**Fig. 1A**).

CdiI^{STECO31} binds between the two subdomains of the toxin domain, with its β-sheet inserting into the toxin cleft (**Fig. 1A**). The N-terminus and helix α1´ of CdiI^{STECO31} interact with the C-terminal subdomain of the toxin, and the long G1´ loop contacts the N-terminal subdomain (**Fig. 1A**). The interface buries ~1,600 Å^2, corresponding to ~20% of the total solvent-accessible surface area as determined by PISA (Krissinel & Henrick, 2007). The interaction network is elaborate and features numerous direct contacts and more than 20 water-mediated hydrogen bonds (**Fig. 1B, Tables 2 & 3**). Moreover, the toxin cleft is

strongly electropositive, and several anionic residues from CdiI^{STECO31} form salt-bridges (**Fig. 1C**). Asp48 and Asp78 of CdiI^{STECO31} form salt-bridges with toxin residues Lys261 and Lys181, respectively. Arg307 and Tyr309 from the toxin also converge on the main-chain carbonyl of CdiI^{STECO31} Asp48. CdiI^{STECO31} Tyr75 makes direct H-bond contacts with His187, Gly202 and His204 (**Fig. 1B** & **Table 2**). The side-chains of Trp17, Thr18, Asn28 and Glu29 emanate from CdiI^{STECO31} helix α1´ to make contacts with toxin residues Gly293, Gln295, Asp298 and Arg307, respectively (**Figs. 1B** & **1C**). Additionally, CdiI^{STECO31} Trp17 makes van der Waals contacts in a pocket formed by toxin residues Gln295 and Gln312 (**Fig. 1C**). The N-terminal subdomain of the toxin interacts primarily with CdiI^{STECO31} helix α2´ and the long G1´ loop that links β2´ to β3´. Toxin residue Lys257 is the focal point of interactions with side-chains of Asp34, Glu37 and Asn38 from immunity helix α2´ (**Figs. 1B** & **1C**). Residues Asp61 and Ser63 from the G1´ loop form direct H-bonds with toxin residue Lys197 (**Fig. 1B** and **Table 2**), and CdiI^{STECO31} Tyr60 packs against residues Ala243, Ala244 and Gly245, which form the β-turn connecting toxin strands β4 and β5 (**Fig. 1C**).

CdiA-CT^{STECO31} contains an EndoU ribonuclease domain

Consistent with its annotation as an EndoU toxin, the DALI server returned eukaryotic EndoU domains as the closest structural homologs of CdiA-CT^{STECO31}. Coronaviral NendoU domains were the top hits in the search: coronavirus 229E Nsp15 (PDB:4S1T), $Z = 6.1$, rmsd 3.6 Å over 100 Cα atoms; SARS Nsp15 (PDB:2RHB (Bhardwaj *et al.*, 2008)) $Z = 6.0$, rmsd 3.6 Å over 101 Cα atoms; SARS Nsp15 (PDB:2H85 (Ricagno *et al.*, 2006)), $Z = 6.0$, rmsd 3.5 Å over 100 Cα atoms, and murine hepatitis virus Nsp15 (PDB:2GTH (Xu *et al.*, 2006)), $Z = 5.5$, rmsd 3.8 Å for 101 Cα atoms. The NendoU domains and CdiA-CT^{STECO31} toxin have essentially identical topologies, though some secondary structure elements vary in length and relative position (**Figs. 2A** & **2B**). Notably, NendoU

domains harbor the two catalytic His residues within helix α8 and the adjacent loop, which corresponds to the β1-β2 hairpin of the toxin domain (**Figs. 2A** & **2B**). CdiA-CT^{STECO31} also contains additional residues that form strand β5 and contribute to an elongated helix α2 (**Fig. 2B**). XendoU (PDB:2C1W (Renzi *et al.*, 2006)) is more distantly related to the toxin at $Z = 5.2$, rmsd 3.1 Å over 108 Cα atoms. As with Nsp15, the XendoU active-site loop aligns poorly with the β1-β2 hairpin of the toxin (**Figs. 2C** & **2D**). Close inspection of the XendoU electron density map suggests that the active-site loop may be modeled incorrectly, though the rebuilt loop still does not superimpose well. NendoU domains from arteriviral Nsp11 proteins show the lowest structural similarity to the toxin. Nsp11 from porcine reproductive and respiratory virus (PDB:5DA1) was recovered at $Z = 4.5$ with rmsd 3.7 Å over 84 Cα atoms. Notably, sequence homology between CdiA-CT^{STECO31} and the eukaryotic EndoU domains is low, ranging from ~9% identity with XendoU up to about 15% for the coronavirus NendoU domains (**Figs. 2B** & **2D**). The DALI search also recovered a handful of pseudopilin proteins from T2S systems. Though these latter hits returned Z scores comparable to XendoU and Nsp11 proteins, structural similarity with pseudopilins is limited to the C-terminal subdomain of the toxin.

Two prokaryotic EndoU toxins have been characterized to date. BC0920 from *Bacillus cereus* ATCC 14579 is a putative T7SS effector that degrades rRNA and tRNA when expressed in *E. coli* (Holberger *et al.*, 2012). MafB1 from *Neisseria meningitidis* 8013 cleaves RNA after uridylate residues *in vitro* (Jamet *et al.*, 2015). To determine whether CdiA-CT^{STECO31} possesses similar activity, we expressed the toxin in *E. coli* and examined RNA for degradation. Although the toxin inhibited cell growth (**Fig. S1A**), there was no obvious RNase activity in the intoxicated cells (**Fig. S1B**). Because several CDI toxins

cleave specific tRNAs (Nikolakakis *et al.*, 2012, Willett *et al.*, 2015a, Jones *et al.*, 2017a, Michalska *et al.*, 2017), we used Northern blot analysis to identify possible tRNA substrates. This screen revealed efficient cleavage of $tRNA_{UUC}^{Glu}$ in CdiA-CT^{STECO31} intoxicated cells, but not in cells expressing other toxins (**Fig. 3**). We also detected modest nuclease activity against $tRNA_{GUC}^{Asp}$, $tRNA_{CUG}^{Gln}$, $tRNA_{CCC}^{Gly}$, $tRNA_{UCC}^{Gly}$, $tRNA_{GCC}^{Gly}$, $tRNA_{UUU}^{Lys}$, $tRNA_{GGU}^{Thr}$ and $tRNA_{CCA}^{Trp}$ (**Fig. 3**). These results indicate that CdiA-CT^{STECO31} exhibits a novel tRNase activity for the EndoU superfamily.

CdiI^{STECO31} homologs and the specificity of immunity

A DALI server search for structural homologs of CdiI^{STECO31} returned NMB0503 from *N. meningitidis* MC58 (PDB ID:4Q7O, (Tan *et al.*, 2015)) with $Z = 13.4$ and rmsd 2.6 Å over 111 Cα atoms (**Fig. 4A**). Although it shares only ~16% sequence identity with CdiI^{STECO31} (**Fig. 4B**), NMB0503 is probably an immunity protein because the upstream NMB0502 reading frame encodes a C-terminal EndoU domain that is 39% identical to CdiA-CT^{STECO31} (**Fig. 4C**). Moreover, NMB0502 has the same tRNase activity as CdiA-CT^{STECO31} when expressed in *E. coli* (**Fig. S2**). The structures of the two immunity proteins deviate in three regions: i) NMB0503 contains an N-terminal helix (α0′) that is absent from CdiI^{STECO31}; ii) the loop connecting β2′ to β3′ is extended in NMB0503 and contains an additional 3_{10} helix, and iii) CdiI^{STECO31} contains a 3_{10} helix (G2′) in the loop connecting β3′ to β4′ (**Fig. 4B**). Given that these secondary structure elements interact with the toxin domain in the CdiA-CT/CdiI^{STECO31} structure (**Fig. 4A**), the discrepancies may reflect the toxin-free state of NMB0503 (Tan *et al.*, 2015). However, NMB0502 diverges from CdiA-CT^{STECO31} in regions that interact directly with CdiI^{STECO31} (**Fig. 4C**), suggesting that these differences are critical for binding interactions between toxins and cognate immunity proteins. This conclusion is supported by alignments of closely related EndoU toxins and immunity

proteins, which show that each pair possesses a unique set of residues contributing to the binding interface (**Fig. S3**).

To explore the specificity of immunity proteins, we tested whether EndoU toxins can be neutralized by near-cognate CdiI proteins. We first fused the *cdiA-CT/cdiI*[STECO31] module to the *cdiA*[EC93] gene of *E. coli* EC93, thereby generating a plasmid-borne CDI system capable of delivering CdiA-CT[STECO31] toxin into *E. coli* target cells. Inhibitor cells expressing the chimeric CDI[STECO31] system significantly outcompeted *E. coli* target bacteria in co-culture (**Fig. 5A**). Northern blot analysis revealed cleaved $tRNA_{UUC}^{Glu}$ in the mixed culture (**Fig. 5B**). Because inhibitor cells are immune to the toxin, the latter result suggests that most of the $tRNA_{UUC}^{Glu}$ is degraded in target bacteria. Target cells expressing CdiI[STECO31] were protected from both growth inhibition and tRNase activities (**Figs. 5A & 5B**). By contrast, targets expressing near-cognate NMB0503 were inhibited to the same extent as cells with no immunity gene (**Fig. 5A**), and Northern blot showed that NMB0503 failed to neutralize tRNase activity (**Fig. 5B**). We then tested whether a more closely related immunity protein blocks CdiA-CT[STECO31]. The T5SS/CDI locus of *Yersinia mollaretii* ATCC 43969 encodes an EndoU toxin and immunity protein that are ~55% and ~58% identical to CdiA-CT[STECO31] and CdiI[STECO31], respectively (**Figs**. **4B & 4C**). Despite this homology, *cdiI*[Ymo43969] expression did not protect target cells from CDI[STECO31] mediated growth inhibition (**Figs. 5A & 5B**). To confirm that the *cdiI*[Ymo43969] construct produces functional immunity protein, we tested it against inhibitor cells that express a CdiA[EC93]-CT[Ymo43969] fusion protein. The chimeric CDI[Ymo43969] system is less potent than CDI[STECO31] (**Figs. 5A & 5C**), but toxic nuclease activity was detected in co-cultures with non-immune target cells (**Fig. 5D**). As expected, target cells that express CdiI[Ymo43969] were immune to CDI[Ymo43969] mediated growth inhibition

and nuclease activity, but near-cognate CdiI^{STECO31} failed to protect target bacteria (**Figs. 5C & 5D**). Surprisingly, we found that NMB0503 expressing target cells were not killed in co-culture with CDI^{Ymo43969} inhibitors (**Fig. 5C**). Nevertheless, cleaved tRNA$_{UUC}^{Glu}$ accumulated in this latter co-culture (**Fig. 5D**), demonstrating that NMB0503 immunity protein does not completely block tRNase activity. Presumably, the NMB0503 expressing target cells retained enough full-length tRNA to support growth. Taken together, these results indicate that each EndoU toxin-immunity protein interaction is distinct.

The EndoU nuclease active site
Eukaryotic EndoU nucleases cleave phosphodiester bonds using a His-His-Lys catalytic triad (Ivanov *et al.*, 2004, Guarino *et al.*, 2005, Shi *et al.*, 2016, Kang *et al.*, 2007, Nedialkova *et al.*, 2009). The EndoU catalytic mechanism is similar to that of RNase A, which use two His residues to mediate proton transfers and Lys to stabilize the pentavalent phosphoryl transition-state (Findlay *et al.*, 1962, Deavin *et al.*, 1966, Cuchillo *et al.*, 2011). Structural superimposition suggests that toxin residues His187 and His204 are equivalent to the catalytic His residues in SARS Nsp15 (His234/His249) and XendoU (His162/His178) (**Fig. 6A**). Further, CdiA-CT^{STECO31} Lys261 is positioned near Lys289 of Nsp15 and Lys224 of XendoU (**Fig. 6A**). The EndoU active site also contains conserved residues that contribute to uridylate specificity. Ser293 of SARS Nsp15 (Ser174 in EAV Nsp11, Ser228 in XendoU) is predicted to form a H-bond with O2 of uracil, and Tyr342 (Tyr216 in EAV Nsp11, Tyr280 in XendoU) is thought to stack onto the pyrimidine ring (Bhardwaj *et al.*, 2008, Nedialkova *et al.*, 2009). The structure overlay suggests that CdiA-CT^{STECO31} residues Thr262 and His321 may also function in substrate discrimination (**Fig. 6A**). Finally, toxin residue Pro322 superimposes closely onto conserved Pro residues that are required for Nsp15 nuclease activity (**Fig. 6A**) (Ricagno *et al.*, 2006, Bhardwaj *et al.*, 2008).

We probed the CdiA-CT^{STECO31} active site using site-directed mutagenesis. We first introduced Ala substitutions into full-length CdiA^{EC93}-CT^{STECO31} and tested inhibition activity in competition co-cultures. Mutation of His187, His204, Lys261, Thr262 and His321 abrogated growth inhibition (**Fig. 6B**). In contrast, substitution of Asn319 in the toxin active site had no discernable effect on inhibition activity (**Fig. 6B**). Immunoblot analysis confirmed that each CdiA variant was produced at the same level as the wild-type effector (**Fig. S4**). Furthermore, the mutated effectors were susceptible to degradation with extracellular proteinase K (**Fig. S4**), indicating that each was exported to the cell surface properly. Although target bacteria were not inhibited by cells deploying the Lys261Ala and Thr262Ala toxin variants (**Fig. 6B**), we detected tRNase activity in these co-cultures (**Fig. 6C**). Presumably, these latter toxins are attenuated enough to allow target cell growth. We then compared the activities of mutant toxins to wild-type CdiA-CT^{STECO31} using *in vitro* nuclease assays. Purified wild-type toxin cleaved tRNA$_{UUC}^{Glu}$ efficiently, but its activity was neutralized when purified CdiI^{STECO31} immunity protein was included in the reaction (**Fig. 6D**). By contrast, toxins carrying His187Ala and His204Ala substitutions had little activity *in vitro* (**Fig. 6D**). The Lys261Ala, Thr262Ala and His321Ala variants each exhibited significantly reduced nuclease activity compared to wild-type (**Fig. 6D**). Purified Asn319Ala toxin retained nuclease activity (**Fig. 6D**), consistent with its near wild-type function in the competition co-culture experiments (**Fig. 6C**). Collectively, these results strongly suggest that CdiA-CT^{STECO31} uses same catalytic mechanism as eukaryotic EndoU proteins.

Prokaryotic EndoU toxins have evolved into three clades
    The tRNase activity of CdiA-CT^{STECO31} differs from previously characterized EndoU toxins, suggesting the superfamily may have evolved a broader spectrum of substrate specificities. Analysis of EndoU toxin sequences from predicted T5SS/CdiA effector proteins

shows that the domains segregate into at least three major clades. Clade I domains are related to the previously characterized MafB1 toxin from *N. meningitidis* 8013 (Jamet *et al.*, 2015) (**Fig. 7A**). Clade I toxins are characterized by greatly abbreviated N-terminal subdomains that lack the α1, β3, β4 and β5 elements of CdiA-CT[STECO31] (**Fig**. **S5**). Clade I toxins contain additional residues between α2 and β7 in the C-terminal subdomain and sometimes carry extended C-terminal tails (**Fig. S5**). Clade II toxins are similar to BC0920 from *B. cereus* ATCC 14579 and are predicted to lack helix α1 (**Figs. 7A** & **S5**). CdiA-CT[STECO31], CdiA-CT[Ymo43969] and NMB0502 contain clade III nuclease domains (**Fig. 7A**). All EndoU domains are predicted to carry the two catalytic His residues within a β-hairpin, though residue spacing varies amongst toxins (**Fig. S5**). Clades I and III contain the catalytic Lys residue within a conserved K-(S/T) motif in strand β6 (**Fig. S5**). The equivalent Lys residues are not easily identified in clade II, though many of these toxins contain irregularly spaced K-(S/T) motifs within the β5 region (**Fig. S5**).

To directly compare nucleases from each EndoU clade, we purified representative toxins and examined their RNase activities *in vitro*. For clade I, we chose a predicted CdiA-CT from *Klebsiella aerogenes* GN05224, reasoning that it should be functional when fused to CdiA[EC93] to generate a chimeric effector. Additionally, the CdiA-CT[GN05224] EndoU domain shares 61% sequence identity with the previously characterized *N. meningitidis* MafB1 toxin (Jamet *et al.*, 2015). Purified CdiA-CT[GN05224] degrades 5S rRNA and several tRNA species *in vitro* (**Fig. 7B**). Similar RNase activity was detected in competition co-cultures with inhibitor cells that deploy CdiA-CT[GN05224] (**Fig. S6**), though 5S rRNA degradation was not observed *in vivo*, presumably because it is protected by ribosomal proteins. We also found that CdiA-CT[GN05224] RNase activity is effectively neutralized by CdiI[GN05224] immunity

protein *in vitro* (**Fig. 7B**) and *in vivo* (**Fig. S6**). We chose *B. cereus* BC0920 as a representative clade II toxin and found that its RNase activity is similar to that of CdiA-CT^{GN05224} (**Fig. 7B**). Like the other nucleases, BC0920 activity was suppressed when the reaction was supplemented with purified BC0921 immunity protein (**Fig. 7B**). Finally, we confirmed the substrate specificity of CdiA-CT^{STECO31} toxin *in vitro.* Purified CdiA-CT^{STECO31} cleaves tRNA$_{UUC}$^{Glu} efficiently and shows partial activity on tRNA$_{GCC}$^{Gly}, but has no discernable activity on tRNA$_{GAU}$^{Ile}, tRNA$_{GUA}$^{Tyr} or 5S rRNA (**Fig. 7B**). Moreover, CdiA-CT^{STECO31} is most likely a specific endonuclease because it produces stable 5´ and 3´ fragments of tRNA$_{UUC}$^{Glu} (**Fig. 7B**).

Finally, we mapped the CdiA-CT^{STECO31} cleavage site in tRNA$_{UUC}$^{Glu}. The comparable sizes of the 5´ and 3´ fragments suggest that the toxin cleaves tRNA$_{UUC}$^{Glu} in the anticodon loop. Therefore, we designed an oligonucleotide that anneals downstream of the anticodon for primer extension analysis (**Fig. 8A**). Analysis of 3´-fragments generated from *in vivo* toxin expression revealed cleavage at the 5´-side of m^{2}A38 (**Figs. 8B** & **8C**). This same site was observed in tRNA$_{UUC}$^{Glu} isolated from competition co-cultures. As expected, these cleavages were not detected in cells that express *cdiI*^{STECO31} (**Fig. 8B**). We also determined the same cleavage site in tRNA$_{UUC}$^{Glu} digested purified CdiA-CT^{STECO31} *in vitro* (**Fig. 8B**). This latter result confirms that CdiA-CT^{STECO31} is directly responsible for producing the 3´-fragment, and strongly suggests that the toxin cleaves the phosphodiester linking nucleotides C37 and m^{2}A38.

C) Discussion

Aravind and colleagues first identified the link between eukaryotic EndoU RNases and a broadly distributed family of antibacterial toxins (Zhang *et al.*, 2012, Zhang *et al.*,

2011). The structure of CdiA-CT^{STECO31} confirms those predictions, and mutational analysis suggests that this toxin uses the same catalytic mechanism as eukaryotic EndoU nucleases. However, there are also significant differences between the prokaryotic and eukaryotic nucleases. Purified XendoU and Nsp15 are only active in the presence of millimolar Mn^{2+} or Ca^{2+} (Laneve *et al.*, 2003, Schwarz & Blower, 2014, Ivanov *et al.*, 2004, Bhardwaj *et al.*, 2004), whereas the bacterial toxins examined here have no metal requirement. There is no direct role for divalent metal in the EndoU catalytic mechanism (Ricagno *et al.*, 2006, Renzi *et al.*, 2006), and like the bacterial toxins, arteriviral Nsp11 is active in the absence of metal (Nedialkova *et al.*, 2009). These observations suggest that metals may promote XendoU and Nsp15 nuclease activity through allostery. This model is consistent with data showing that SARS Nsp15 undergoes conformational changes upon titration with Mn^{2+} (Bhardwaj *et al.*, 2004). Allosteric regulation could be mediated by the additional structural elements found in eukaryotic EndoU enzymes. Bacterial EndoU domains are about half the size of XendoU, which contains ~150 additional N-terminal residues that form an extensive α-helical cradle that surrounds and supports the nuclease core (Renzi *et al.*, 2006). In this model, conformational changes in the supporting cradle would reshape the active-site cleft and/or alter the configuration of the catalytic triad. The supporting structures in NendoU domains are abbreviated, but Nsp11 and Nsp15 contain additional N-terminal domains that mediate oligomerization (Bhardwaj *et al.*, 2008, Ricagno *et al.*, 2006). Oligomerization is critical for Nsp15 activity because the active-site loop is stabilized by the β10-β11 "supporting loop" from the neighboring protomer (Joseph *et al.*, 2007, Guarino *et al.*, 2005). In monomeric Nsp15, the active-site loop is untethered and migrates into the substrate-binding groove, where it presumably blocks the interaction with substrate RNA. Similarly, one of the active-

site loops is disordered in crystal structures of homodimeric Nsp11, again raising the possibility that RNase activity is regulated through multimerization (Shi *et al.*, 2016, Zhang *et al.*, 2017, Joseph *et al.*, 2007). These findings suggest that eukaryotic EndoU proteins must be carefully controlled, because unrestrained RNase activity is lethal. Such allosteric control should not be required of prokaryotic EndoU nucleases, because these enzymes are used to inhibit bacterial cell growth. Nevertheless, the potent toxic effects of EndoU nucleases must be suppressed in the producing cell, thus providing the selective pressure to evolve immunity proteins that protect against self-intoxication.

The narrow substrate specificity of CdiA-CT^{STECO31} is surprising given that all previously characterized EndoU enzymes cleave RNAs after uridylate. This substrate selectivity appears to be a common property of clade III EndoU domains, because CDI toxins from *N. meningitidis* and *Y. mollaretii* also preferentially cleave tRNA$_{UUC}^{Glu}$. The determinants governing EndoU nuclease specificity have been examined most extensively for viral NendoU proteins. Ser293 in Nsp15 (Ser228 in XendoU) is predicted to form specific H-bond contacts with uracil, and Ala substitutions at this position relax specificity and increase the rate of cleavage at cytidylate nucleotides (Ricagno *et al.*, 2006, Bhardwaj *et al.*, 2008, Nedialkova *et al.*, 2009). The corresponding residue in CdiA-CT^{STECO31} is Thr262, which is able to form the same H-bond contacts with uracil as Ser. Moreover, Thr residues at this position have been shown to promote uridylate specificity in other NendoU-containing proteins (Bhardwaj *et al.*, 2008, Nedialkova *et al.*, 2009). Nevertheless, CdiA-CT^{STECO31} does not recognize uridylate residues and instead cleaves tRNA$_{UUC}^{Glu}$ between nucleotides C37 and m^{2}A38. The *in vivo* toxin activity screen suggests that the C37-A38 motif is a positive recognition determinant, because tRNA$_{QUC}^{Asp}$ and tRNA^{Gly} isoacceptors share this element

and are subject to partial cleavage. However, the toxin probably monitors the overall structure of the anticodon loop, because it clearly discriminates against $tRNA_{QUC}^{Asp}$, which differs from $tRNA_{UUC}^{Glu}$ only at position 34. All of the clade III toxins examined here have anticodon nuclease activity, suggesting that their shared architecture determines substrate specificity. Clade III domains are characterized by helix α1, which is apparently absent from the other clades. Helix α1 supports the β1-β2-hairpin in the CdiA-CT^{STECO31} structure, and therefore it is conceivable that this element shapes the active-site cleft. We note that clade III also contains toxins from *Pseudomonas* and *Acinetobacter* species that appear to lack β5 in the N-terminal sub-domain (see **Fig. S5**). It will be of interest to determine whether these enzymes share anticodon nuclease activity with CdiA-CT^{STECO31} or perhaps exhibit distinct substrate specificities.

The EndoU superfamily is one of the most widely used toxins in inter-bacterial competition. Aravind and colleagues initially identified EndoU effectors associated with T2SS/MafB, T5SS, T6SS and T7SS delivery platforms (Zhang *et al.*, 2012, Holberger *et al.*, 2012, Jamet *et al.*, 2015). More recent database searches indicate that EndoU domains are deployed by several other systems – some of which have been previously described and others that appear to be novel. We identified bacterial EndoU domains fused to the C-termini of PrsW, SpvB and MuF-like phage-head morphogenesis proteins (**Table S3**), all of which have been proposed to deliver polymorphic toxins (Zhang *et al.*, 2012). Intriguingly, EndoU nuclease are also fused to VgrG-like proteins in Gram-positive *Paenibacillus* species (**Table S3**). VgrG is a critical structural component of T6SS and often carries C-terminal toxin domains (Russell *et al.*, 2014, Ho *et al.*, 2014). However, T6SSs have only been characterized in Gram-negative bacteria, and *Paenibacillus* genomes contain no recognizable

T6SS core genes. Nevertheless, these VgrG proteins are encoded together with DUF4280 (PAAR-like) and pentapeptide repeat adaptor proteins that are similar to those found in T6SS effector operons in Gram-negative bacteria. We also found a PAAR-like protein with a C-terminal EndoU domain in *Clostridium* sp. ASBs410 (**Table S3**). Together, these observations strongly suggest that Gram-positive bacteria possess phage-like secretion systems that are functionally analogous to T6SS. In addition to these recognizable toxin delivery systems, we recovered EndoU-containing proteins with unusual N-terminal domains and several others that contain no other conserved domains whatsoever (**Table S3**). Some of these proteins carry N-terminal secretion signal sequences, and nearly all are encoded together with probable immunity proteins. We speculate that these latter uncharacterized proteins are also used for inter-bacterial toxin delivery. Finally, we identified a handful of extraordinarily large (1.5 to 2.9 MDa) proteins from Actinobacteria that carry internal EndoU domains. These proteins typically contain one or more APH aminoglycoside phosphotransferase domains together with a variety of RNA/DNA-binding modules (**Table S4**). The function of these proteins is not known, but they appear unlikely to mediate inter-cellular competition given the central location of the EndoU domain and the absence of linked immunity genes. The unusual domain assemblage suggests that these enormous proteins could process nucleic acids, or perhaps synthesize secondary metabolites or antibiotics in a manner similar to polyketide synthases. Regardless of their precise activities, these latter proteins suggest that EndoU domains have also been coopted to perform non-inhibitory functions in bacteria.

D) Experimental Procedures

Plasmid constructions

Plasmid constructs and oligonucleotides used in this study are listed in **Tables 4** and **S1**, respectively. DNA fragments encoding CdiA-CT/CdiI proteins from *E. coli* STEC_O31 and *Y. mollaretii* ATCC 43969 were synthesized by Genscript (Piscataway, NJ) and supplied in plasmid pUC57. The fragment from *E. coli* STEC_O31 (Genbank: AFEX01000038) encodes residues Val2931 – Lys3253 of CdiA^{STECO31} (locus tag: ECSTECO31_4009) and CdiI^{STECO31} (ECSTECO31_4008). The initiation codon for ECSTECO31_4008 is not annotated correctly, and CdiI^{STECO31} contains 121 residues. The fragment from *Y. mollaretii* ATCC 43969 (NCBI reference: NZ_AALD02000043.1) encodes residues Val2694 – Asn2963 of CdiA^{Ymo43969} (YMOLL0001_RS03095) and its putative CdiI^{Ymo43969} immunity protein (YMOLL0001_RS0219640). A fragment encoding Val1 – Lys360 of CdiA-CT^{GN05224} (YA39_RS24095) and CdiI^{GN05224} (YA39_RS16570) from *K. aerogenes* GN05224 (NCBI reference: NZ_LDBZ01000036.1) was synthesized by GenArt Gene Synthesis (ThermoFisher) and provided in plasmid pMA-T. The *cdiA-CT/cdiI*^{STECO31} module was amplified with primers 204F40 and 204R49 and introduced into plasmid pMCSG58 using ligation-independent cloning (LIC) (Eschenfeldt *et al.*, 2013, Eschenfeldt *et al.*, 2010, Eschenfeldt *et al.*, 2009). All of the resulting clones contained frame-shift mutations in the toxin coding region, suggesting that the toxin was not completely neutralized by CdiI^{STECO31}. Therefore, *cdiI*^{STECO31} was amplified with primer pair 204CdiIF/04R49 and introduced into plasmid pMCSG88 as described above. Vector pMCSG88 was generated by digesting pMCSG58 with HindIII/NcoI, followed by ligation to plasmid pMCSG76 (Eschenfeldt *et al.*, 2013). The resulting pMCSG88 construct contains the LIC site with the Clo DF13 origin of replication and a spectinomycin-resistance marker. The pMCSG58-CdiA-CT/CdiI^{STECO31} and

pMCSG88-CdiI^{STECO31} expression vectors were then introduced into *E. coli* BL21(DE3) to generate an over-expression strain for large-scale protein purification and crystallography.

The *cdiA-CT/cdiI*^{STECO31} (CH3538/CH3539) and NMB0502/NMB0503 (CH2839/CH2785) modules were amplified and ligated to NcoI/SpeI-digested plasmid pCH12599 to tag the immunity proteins with C-terminal ssrA(DAS) degrons. The *cdiA-CT/cdiI*^{STECO31} product was also ligated to pCH6505 to generate plasmid pCH11884 for the purification of CdiA-CT/CdiI^{STECO31} complex for biochemical analyses. This latter plasmid was also used as a template for PCR-based site-directed mutagenesis using primers CH4016, CH4017, CH4018, CH4019, CH4269 and CH4270. The *cdiA-CT/cdiI*^{GN05224} module was amplified with primers SK724/DL3986 and ligated to pET28a via NcoI/XhoI restrictions sites to generate plasmid pDAL8956. The *cdiI*^{STECO31} (CH3854/CH3539), *cdiI*^{GN05224} (SK733/DL3986) and BC0921 (CH4398/CH1706) immunity genes were also cloned into pET21S for the purification of His$_6$-tagged immunity proteins under non-denaturing conditions. The *cdiI*^{STECO31} (CH3854/CH3624), *cdiI*^{GN05224} (SK733/SK734), NMB0503 (CH2784/CH2785) and *cdiI*^{Ymol43969} (CH3976/CH3977) genes were also cloned into pTrc99a derivatives to test for immunity function in competition co-cultures.

Plasmid-borne chimeric CDI systems were constructed by allelic exchange of the counter- selectable *pheS** marker from plasmid pCH10163 (Morse *et al.*, 2012). Wild-type and mutant *cdiA-CT/cdiI*^{STECO31} sequences were amplified with primers CH3172/CH3569 and fused to fragments amplified from regions upstream and downstream of the *cdiA*^{EC93} gene. The upstream homology fragment was amplified using primer pair CH4100/CH4101 and the downstream fragment with primers CH4102/CH4103. The three products were then fused to each other using overlap-extension PCR with primers CH4100/CH4103. The final

DNA product (100 ng) was electroporated together with plasmid pCH10163 (300 ng) into *E. coli* DY378 cells (Thomason *et al.*, 2007). Recombinant plasmid clones were selected on yeast extract glucose-agar supplemented with 33 µg/mL chloramphenicol and 10 mM D/L-*p*-chlorophenylalanine. The same procedure was used to fuse the *cdiA-CT/cdiI*[Ymol43969] and *cdiA-CT/cdiI*[GN05224] modules (amplified primers CH3747/CH3748 and SK693/SK694, respectively) to the *E. coli* EC93 CDI system.

Protein expression and purification for crystallization
      E. coli BL21(DE3) cells carrying pMCSG58-CdiA-CT/CdiI[STECO31] and pMCSG88-CdiI[STECO31] were grown at 37 °C for 6 - 8 h in LB medium supplemented with ampicillin (100 µg/mL) and spectinomycin (50 µg/mL). A portion of the culture (0.5 mL) was diluted into 50 mL of M9 minimal medium supplemented with 0.5% glycerol, 100 µg/mL of ampicillin, 50 µg/mL of spectinomycin, trace minerals and vitamins for overnight culture at 37 °C. Large-scale cultures were grown to an optical density at 600 nm (OD_{600}) of 0.8 then cooled to 18 °C. SeMet was added to a final concentration of 60 µg/mL together with L-isoleucine, L-leucine, L-lysine, L-phenylalanine L-threonine and L-valine to a final concentration of 100 µg/mL and incubated for 20 min. Protein expression was induced with 0.5 mM isopropyl-D-thiogalactopyranoside (IPTG) for overnight culture. Cells were harvested by centrifugation the next day, and the pellets washed and resuspended in 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 mM 2-mercaptoethanol (2-ME), 10% glycerol. Cell were broken with Fast Break reagent (Promega) supplemented with 10 µg/mL lysozyme, 500 U Benzonase Nuclease HC (Novagen) and Complete Protease Inhibitor Cocktail (Roche, Mannheim, Germany). The lysate was clarified by centrifugation at 10,000 rpm for 1 h, followed by passage through a 0.22 µm filter prior to loading onto a 5 mL Nickel (II) Sepharose HisTrap column (GE Healthcare Biosciences). The column was washed with 5

volumes of 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole, 10% glycerol, followed by elution with 250 mM imidazole in the same buffer. Fractions were pooled and loaded onto a Hiload 26/60 Superdex200 size exclusion column equilibrated with 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 2 mM dithiothreitol (DTT). Fractions containing two proteins were pooled and concentrated to 14 mg/mL using an Amicon Ultracel 20K concentrator (Millipore).

Crystallization, data collection, structure solution and refinement
The protein sample was treated with subtilis (20 ng/µL) on ice overnight prior to crystallization trials. Crystallization conditions were screened using the Pi-PEG Screen HTS (Jena Biosciences GmBh). The complex was crystallized in 10.7% PEG 4000, 50 mM Tris-HCl (pH 8.0), 8.6% PEG 3000 at 4 °C by sitting-drop vapor diffusion in 96-well Crystal Quick plates (Greiner Bio-one). Prior to flash cooling in liquid nitrogen, the crystals were cryo-protected in mother liquor supplemented with 17% glycerol. The single-wavelength anomalous diffraction (SAD) dataset was collected at 100 K near the selenium K-absorption edge on beamline 19-ID at the Advanced Photon Source, Argonne National Laboratory. Diffraction images were processed with the HKL3000 suite (Minor *et al.*, 2006). Intensities were converted to structure factor amplitudes in Ctruncate (French & Wilson, 1978, Padilla & Yeates, 2003) from the CCP4 package (Winn *et al.*, 2011). Data collection and processing statistics are presented in **Table 1**. The structure was solved using the HKL3000 software pipeline (Minor *et al.*, 2006) by the SAD method with Se peak data. The pipeline applied SHELXD for the search of heavy atom sites and SHELXE for initial phases calculations (Sheldrick, 2008). The phases were improved through iterations of MLPHARE (Otwinowski, 1991) and DM (Cowtan, 1994). The initial protein model was built in HKL-Builder utilizing Buccaneer (Cowtan, 2006). The final model was obtained through alternating manual

rebuilding in COOT (Emsley & Cowtan, 2004) and crystallographic refinement in Refmac (Murshudov *et al.*, 1997, Winn *et al.*, 2011). The protocol refinement included optimization of TLS parameters with 5 and 2 groups defined for chains A (CdiA-CT^{STECO31}) and I (CdiI^{STECO31}), respectively. The refinement statistics are presented in **Table 1**. The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 5HKQ.

Protein purification and in vitro nuclease assays

Cultures of *E. coli* CH2016 harboring expression plasmids were grown to OD$_{600}$ ~ 0.7, and expression was induced with 1.5 mM IPTG. After incubation for 2 h, cells were harvested and frozen at –80 °C. Cell pellets were re-suspended in lysis buffer [20 mM sodium phosphate buffer (pH 7.0), 150 mM NaCl, 10 mM 2-ME, 0.05% Triton X-100, and broken by two French press passages at 20,000 psi. Cell debris was removed by centrifugation at 16,000 x g at 4 °C. His$_6$-tagged immunity proteins were purified by Ni^{2+}-affinity chromatography in lysis buffer and eluted with 250 mM imidazole as described (Nikolakakis *et al.*, 2012). EndoU toxins were first purified in complex with His$_6$-tagged immunity proteins, then eluted from the column by denaturation in 6 M guanidine-HCl, 20 mM sodium phosphate buffer (pH 7.0). Purified proteins were dialyzed against 20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 10 mM 2-ME. Proteins were quantified by absorbance at 280 nm using the following extinction coefficients: CdiA-CT^{STECO31}, 25,330 cm^{-1} M^{-1}; CdiI^{STECO31}, 20,400 cm^{-1} M^{-1}; CdiA-CT^{GN05224}, 47,330 cm^{-1} M^{-1}; CdiI^{GN05224}, 7,450 cm^{-1} M^{-1}; BC0920, 22,920 cm^{-1} M^{-1} and BC0921, 15,470 cm^{-1} M^{-1}.

In vitro nuclease assays were performed in 20mM Tris-HCl (pH 7.5), 150 mM NaCl, using *E. coli* total RNA (0.5 mg mL^{-1}) as a substrate. Reactions were initiated by addition of

2 µM purified toxins, followed by incubation for 30 min at 37 °C. Where indicated, reactions were supplemented with 2 to 6 µM purified immunity protein. Reactions were quenched with denaturing gel-loading buffer and run on 50% urea-10% polyacrylamide gels buffered with 1 x Tris-borate EDTA. Gels were stained with ethidium bromide or electroblotted to positively charged nylon membranes for subsequent Northern blot analyses.

In vivo toxin activity and competition co-cultures

Nuclease activity screens were performed by activating toxins from *Enterobacter cloacae* ATCC 13047 (Whitney *et al.*, 2014) and *E. coli* isolates STEC_O31, 3006 and 96.154 inside *E. coli* X90 cells. Each cognate immunity protein was tagged with a C-terminal ssrA(DAS) degron as described (McGinness *et al.*, 2006, Poole *et al.*, 2011). *E. coli* X90 cells carrying these expression constructs were seeded at OD_{600} ~ 0.05 in LB media supplemented with 25 µg/mL tetracycline and incubated with shaking at 37 °C. After 60 min, the cultures were adjusted to 0.2% L-arabinose and cultured for an additional 4 h. Cell growth was monitored by measuring the OD_{600} every 30 min. Culture samples were harvested into an equal volume of ice-cold methanol after 2 h of induction, and the cells frozen at –80 °C for subsequent RNA extraction and analysis.

Bacterial competitions were conducted in shaking LB broth at 37 °C. Media were seeded with inhibitor and target strains at a 1:1 ratio, and the co-culture incubated for 1 to 3 h. *E. coli* EPI100 cells that express chimeric CDI systems were used as inhibitors. Inhibitor strains carried plasmids pCH2408 (wild-type CDI[STECO31]), pCH13871 (His187Ala), pCH13306 (His204Ala), pCH13881 (Lys261Ala), pCH13309 (Thr262Ala), pCH13307 (Asn319Ala), pCH13215 (His321Ala), pDAL8914 (CDI[GN05224]) or pCH12847 (CDI[Ymo43969]). *E. coli* MC4100 target strains carried plasmids pTrc99A (*cdiI⁻*), pCH4496 (*cdiI*[STECO31]),

pCH4066 (*cdiI*[Ymo43969]), pDAL8924 (*cdiI*[GN05224]) or pCH13638 (*cdiI*[MC58]). Viable inhibitor and target cells were enumerated as colony forming units per mL (CFU mL^{-1}) on selective media at the beginning and end of co-culture. Competitive indices were calculated as the ratio of target cells to inhibitor cells at the end of co-culture divided by the initial target to inhibitor cell ratio. Competitive indices are reported for each independent experiment together with the average ± standard error of the mean. Co-culture samples were also harvested into equal volumes of ice-cold methanol for RNA extraction.

RNA isolation and analyses
        Frozen cell pellets were resuspended in guanidinium isothiocyanate (GITC)-phenol and total RNA extracted as described (Garza-Sánchez *et al.*, 2006). RNAs (5 µg) were run on 50% urea-10% polyacrylamide gels buffered with 1 x Tris-borate EDTA and electroblotted to positively charged nylon membranes for Northern blot analysis. Blots were hybridized with [^{32}P]-labeled oligonucleotide probes that are specific for individual *E. coli* tRNAs (see **Table S2**) (Garza-Sánchez *et al.*, 2006, Hayes & Sauer, 2003). Blots were visualized by phosphorimaging using Bio-Rad Quantity One software. Primer extension analysis was performed as described (Beck *et al.*, 2014) using oligonucleotide CH4275 to map the cleavage site in tRNA$_{UUC}^{Glu}$. The primer and marker oligonucleotides were 5´-radiolabeled with [^{32}P] using T4 polynucleotide kinase. The radiolabeled primer was hybridized with RNA samples for 5 min at 50 °C, then extended with Superscript reverse transcriptase at 37 °C for 30 min. Reactions were quenched with denaturing gel-loading buffer and heated to 95 °C. Primer extension reactions were run on a 50% urea, 15% polyacrylamide gel buffered with 1 × Tris-borate-EDTA and visualized on a Bio-Rad phosphorimager using Quantity One software.

Acknowledgements
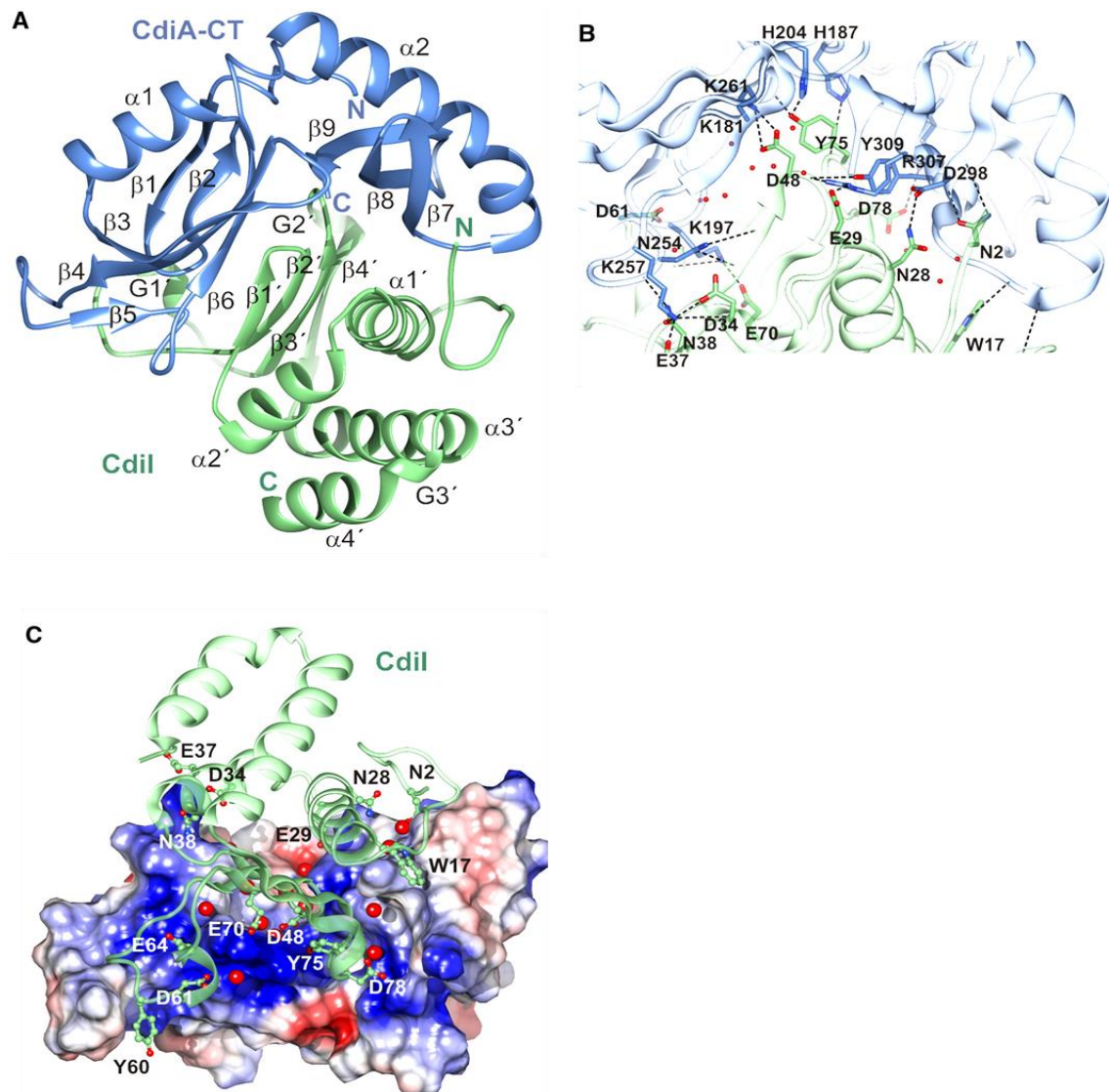
Figure 1: Structure of the CdiA-CT/CdiI[STECO31] complex.

A. The CdiA-CT/CdiI[STECO31] complex is depicted in cartoon with the toxin domain colored blue and the immunity protein colored green. Secondary structure elements are labeled with CdiI[STECO31] elements denoted by a prime (′) symbol.

B. The toxin-immunity protein interface is depicted with selected side-chains forming direct hydrogen bonds (black dashed lines) shown in a stick representation. Water molecules that mediate interactions are shown as red spheres.

C. Charge complementarity at the toxin-immunity protein interface. The electrostatic potential of the toxin surface was calculated using Coulomb's law with Chimera (Pettersen et al., 2004). Potentials range from –10 kcal/mol*e (red) to +10 kcal/mol*e (blue). Water molecules that mediate interactions are shown as red spheres.

[Colour figure can be viewed at wileyonlinelibrary.com]
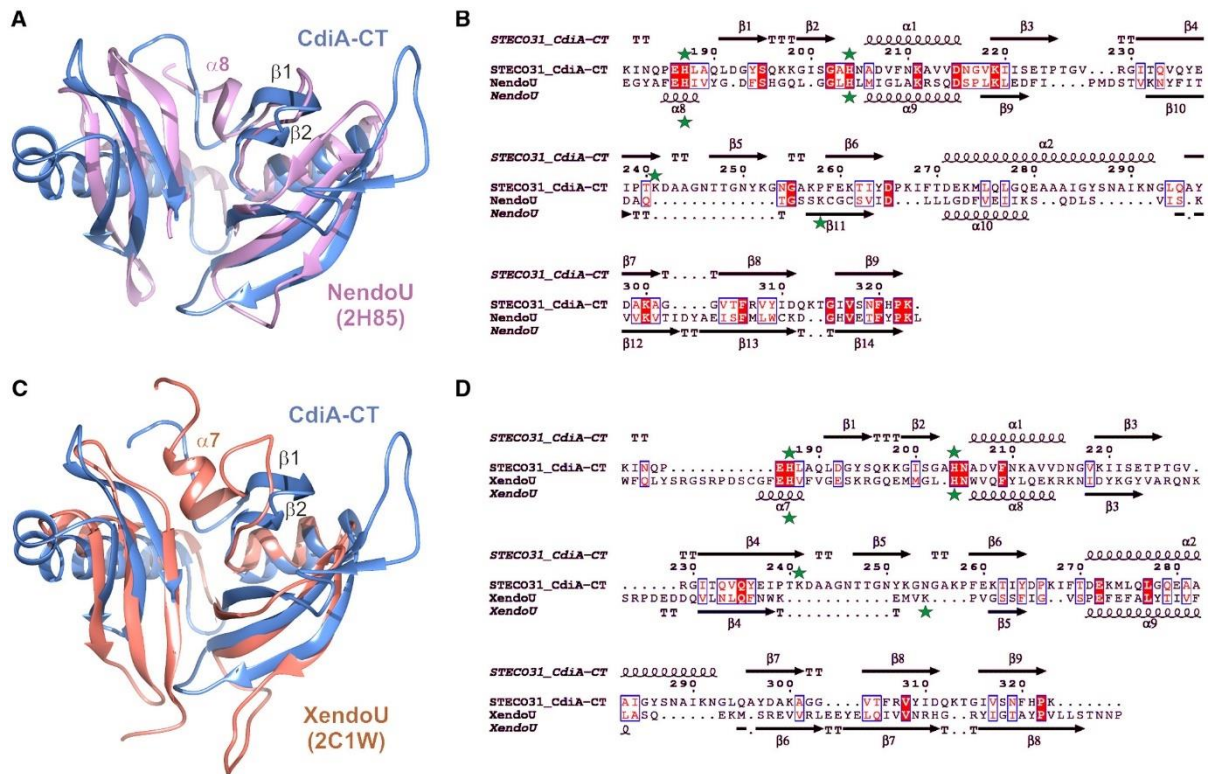
Figure 2: The CdiA-CT^STECO31 toxin is a prokaryotic EndoU domain.

A. CdiA-CT^STECO31 (blue) and the C-terminal domain of SARS Nsp15 (PDB: 2H85, pink) were aligned using secondary structure matching (SSM) superposition calculated in Coot (Emsley and Cowtan, 2004). The β1-β2 hairpin of CdiA-CT^STECO31 and helix α8 in the NendoU active-site loop are indicated.

B. The sequences of the CdiA-CT^STECO31 and SARS Nsp15 C-terminal domains were aligned based on structure using DALI (Holm and Rosenstrom, 2010). The resulting sequence alignment was rendered using Esprit (Robert and Gouet, 2014). Identical residues are highlighted in red, and similar residues are shown in red font. The predicted catalytic triad residues are marked with green stars.

C. CdiA-CT^STECO31 (blue) and the C-terminal domain of XendoU (PDB: 2C1W, coral) were aligned by SSM superposition. The β1-β2 hairpin of CdiA-CT^STECO31 and helix α7 in the XendoU active-site loop are indicated.

D. CdiA-CT^STECO31 and XendoU sequences were aligned as described in panel B.

[Colour figure can be viewed at wileyonlinelibrary.com]
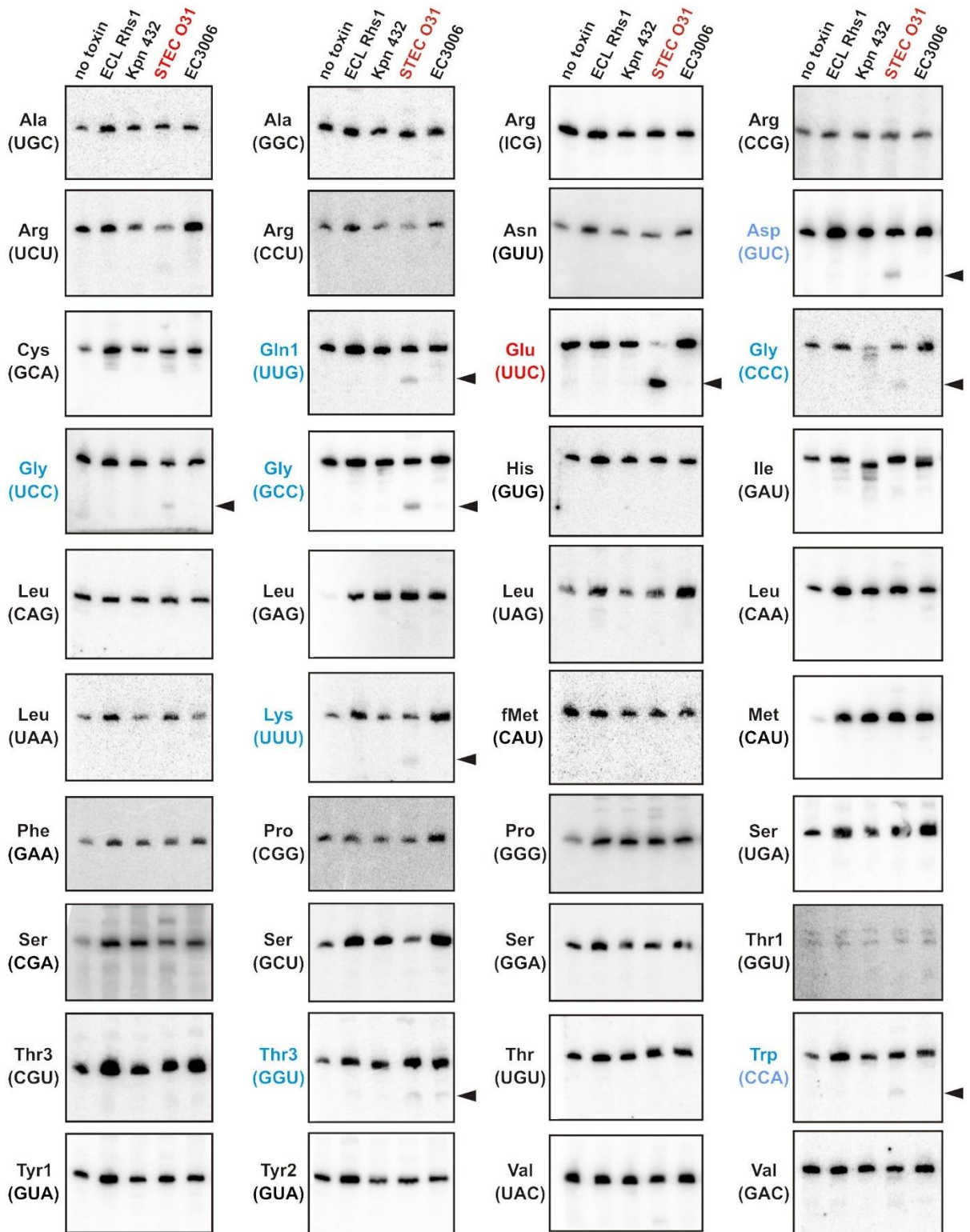
Figure 3: CdiA-CT^{STECO31} is a specific tRNase.

Northern blot analyses of RNA isolated from intoxicated *E. coli* cells. The C-terminal toxin domains of Rhs1 from *E. cloacae* ATCC 13047 (**ECL Rhs1**: YP_003612075.1), CdiA from *K. pneumoniae* 342 (**Kpn 342**: ACI08381.1), CdiA from *E. coli* STEC_O31 (**STEC O31**: EJK94116.1) and CdiA *E. coli* 3006 (**EC3006**: EKI34460.1) were expressed in *E. coli*, and total RNA was isolated for Northern blot hybridization using radiolabeled probes to the indicated tRNAs. Anticodon sequences are shown in parentheses for specific isoacceptors. Arrows to the left of the blot indicate cleavage products detected in CdiA-CT^{STECO31} intoxicated cells.

[Colour figure can be viewed at <u>wileyonlinelibrary.com</u>]
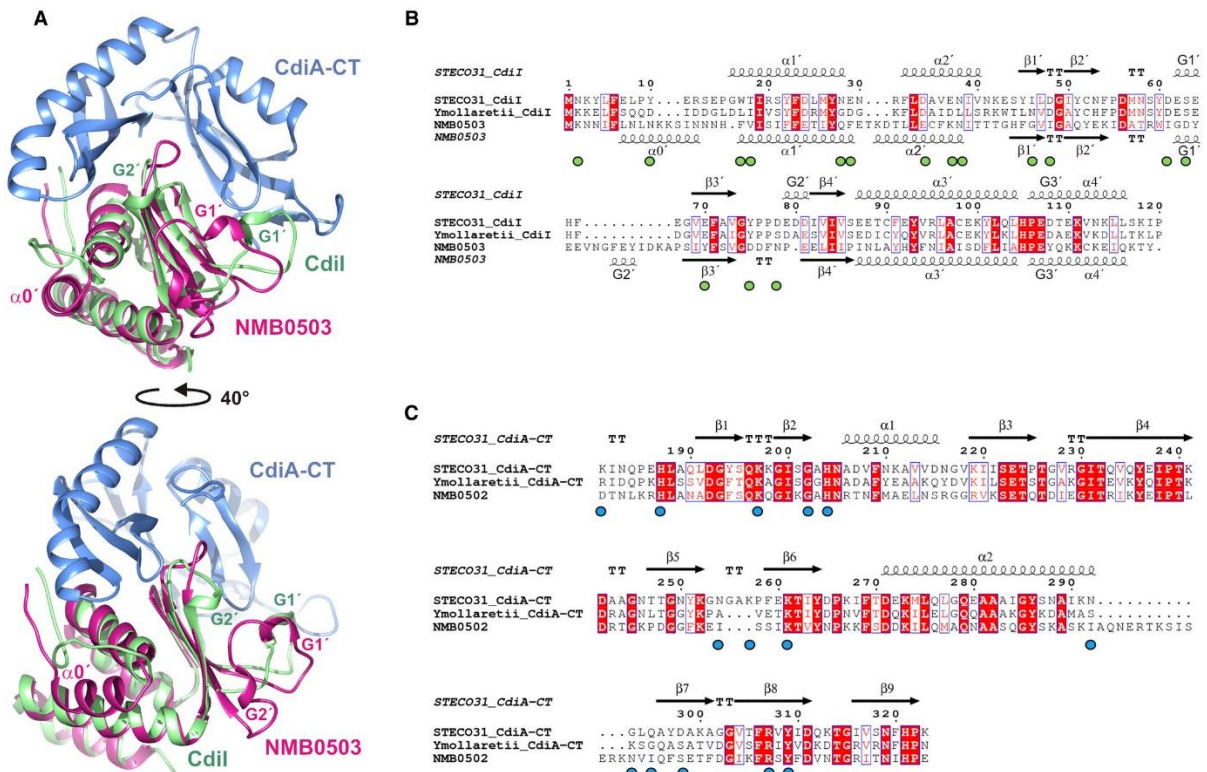
Figure 4: Structure-based alignment of CdiI[STECO31] and the NMB0503 immunity protein of *N. meningitidis* MC58.

A. CdiI[STECO31] (green) and NMB0503 (PDB: 4Q7O, magenta) were aligned by SSM superposition. The CdiA-CT[STECO31] toxin domain (blue) is included to illustrate the toxin-immunity protein binding interface. The bottom view is rotated 40° with respect to the top view.

B. CdiI[STECO31] and NMB0503 sequences were aligned based on structure using DALI. The CdiI[Ymo43969] immunity protein from *Y. mollaretii* ATCC 43968 was aligned using Clustal-Omega. The resulting alignment was rendered using Espript (Robert and Gouet, 2014) with identical residues highlighted in red, and similar residues shown in red font. Green circles below the alignment indicate CdiI[STECO31] residues that form direct H-bonds with the EndoU domain of CdiA-CT[STECO31].

C. EndoU toxins from *E. coli* STEC_O31, *Y. mollaretii* ATCC 43969 and *N. meningitidis* MC58 were aligned using Clustal-Omega. The resulting alignment was rendered using Espript (Robert and Gouet, 2014) with identical residues highlighted in red, and similar residues shown in red font. Blue circles below the alignment indicate CdiA-CT[STECO31] residues that form direct H-bonds with CdiI[STECO31] immunity protein.

[Colour figure can be viewed at wileyonlinelibrary.com]
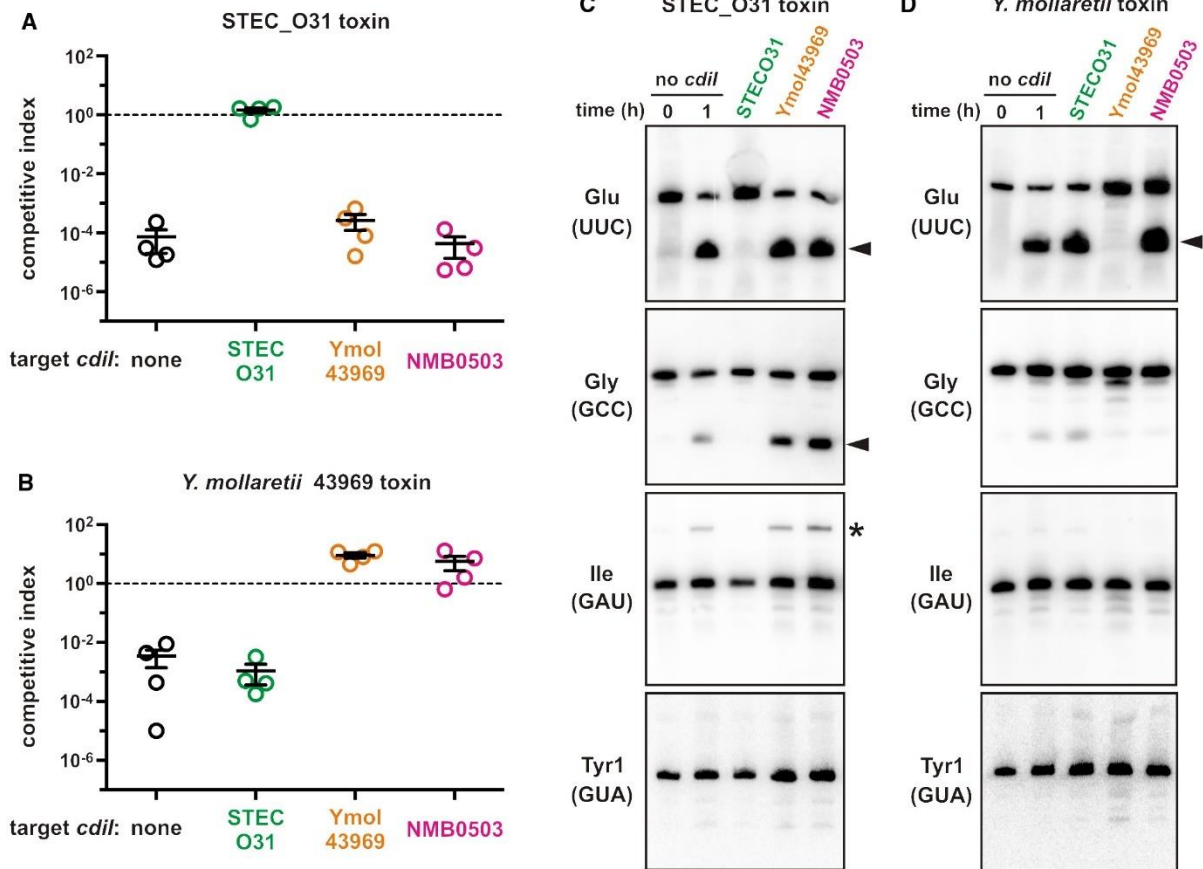
Figure 5: Specificity of CdiI immunity function.
E. coli inhibitor strains that deploy CdiA-CT^{STECO31} (panel A) or CdiA-CT^{Ymo43969} (panel B) were cultured at a 1:1 ratio with *E. coli* target cells that express the indicated cdiI immunity genes. Viable inhibitor and target bacteria were enumerated as colony forming units (cfu) at t = 0 and after 1 h of co-culture CDI^{STECO31} competitions (panel A) and 3 h for CDI^{Ymo43969} competitions (panel B). The competitive index = (cfu$_{t = final}^{targets}$/cfu$_{t = final}^{inhibitors}$)/(cfu$_{t = 0}^{targets}$/cfu$_{t = 0}^{inhibitors}$). Competitive indices for four independent experiments are reported together with mean ± standard error. Northern blot analysis of total RNA isolated from CDI^{STECO31} (panel C) and CDI^{Ymo43969} (panel D) competition co-cultures. One sample was collected immediately after mixing (t = 0 h) and all other samples were collected after 1 h of co-culture. Carets indicate cleavage products, and the asterisk indicates an incompletely processed tRNA$_{GAU}^{Ile}$ transcript that accumulates in CdiA-CT^{STECO31} intoxicated cells. [Colour figure can be viewed at wileyonlinelibrary.com]
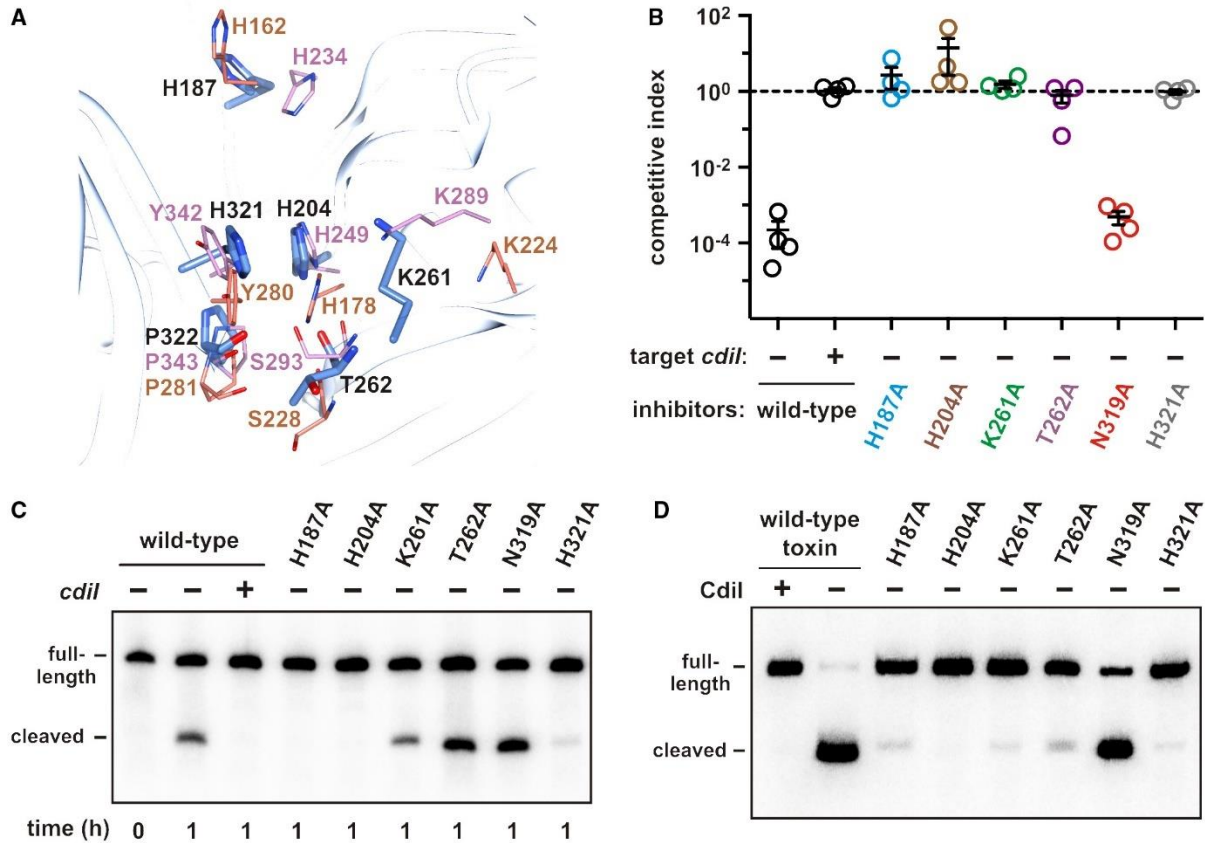
Figure 6: The CdiA-CT^{STECO31} active site.

A. Superimposition of XendoU and SARS Nsp15 active-site residues onto CdiA-CT^{STECO31}. Side-chains from SARS Nsp15 are shown in pink, and those from XendoU are in coral. CdiA-CT^{STECO31} residues are rendered in blue and labeled with black font.

B. *E. coli* inhibitors that deploy the indicated CdiA-CT^{STECO31} variants were cultured at 1:1 ratio with *E. coli* target cells. Viable inhibitor and target bacteria were enumerated as colony forming units (cfu) at t = 0 and t = 1 h, and the competitive index calculated. Competitive indices for four independent experiments are reported together with mean ± standard error.

C. Northern blot analysis of tRNA^{Glu} isolated from the competition co-cultures in panel B.

D. In vitro activities of CdiA-CT^{STECO31} variants. The indicated toxins were purified and incubated with *E. coli* total RNA for 30 min at 37 °C. Reactions were analyzed by Northern blot using a probe to tRNA$_{UUC}^{Glu}$.

[Colour figure can be viewed at wileyonlinelibrary.com]
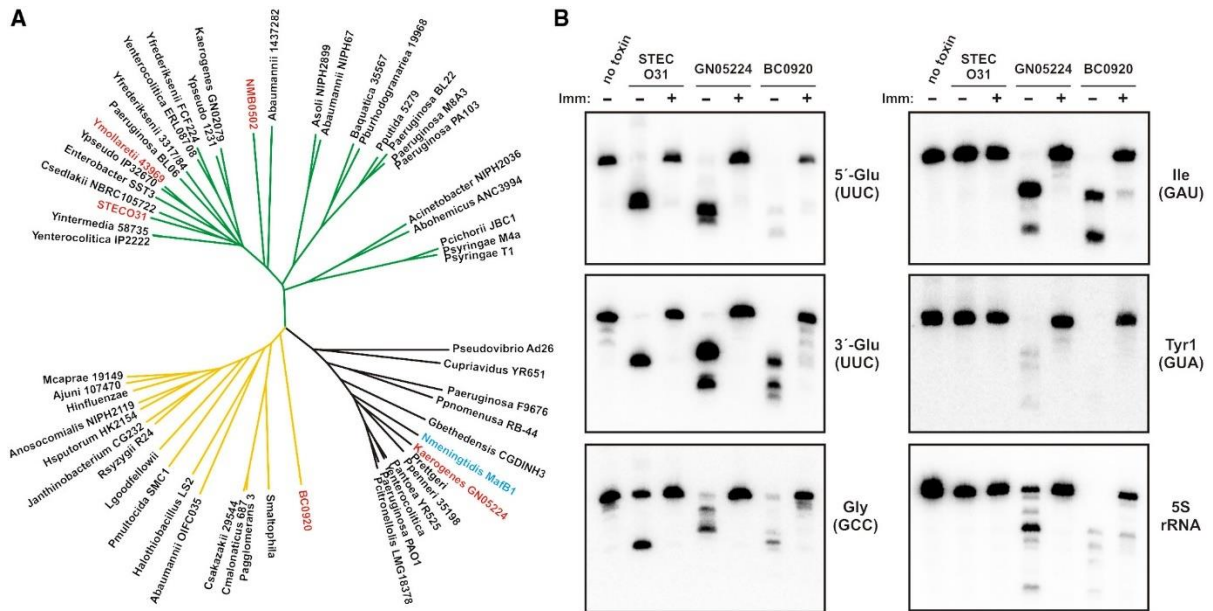
Figure 7: Diversification of prokaryotic EndoU toxin domains.
A. EndoU domains from predicted CdiA proteins were aligned using Clustal-Omega, and the associated unrooted tree rendered using the Tree of Life website. The source multiple sequence alignment is presented in Fig. S7, which also lists the NCBI Refseq accession numbers for each protein. EndoU domains from *N. meningitidis* MafB1 and *B. cereus* BC0920 are included for comparison.
B. In vitro activities of EndoU toxins. The indicated toxins were purified and incubated with *E. coli* total RNA in the absence or presence of cognate immunity protein. Reactions were analyzed by Northern blot hybridization with probes to the indicated RNAs.
[Colour figure can be viewed at wileyonlinelibrary.com]

A

tRNA$_{UUC}^{Glu}$ 5´ GUCCCCUUCGUCΨAGAGGCCCAGGACACCGCCCUS**UCAC**GGCGGUAACAGGGGTΨCGAAUCCCCUAGGGGACGCCA 3´

```
                                                           |||||||||||||||||||||||||||
                                    RT primer              ccattgtccccaagcttagg-5´☢

                                    C39 marker             ccgccattgtccccaagcttagg-5´☢

                                    U36 marker             gtgccgccattgtccccaagcttagg-5´☢

                                    C33 marker             aaagtgccgccattgtccccaagcttagg-5´☢
```

B

intra- | CDI | *in vitro*
cellular | co-culture | reaction | no RNA | no RT

CdiI + − + − + −

C33 −
U36 −
C39 −
RT primer −

C

tRNA$_{UUC}^{Glu}$

Figure 8: CdiA-CT^{STECO31} is an anticodon nuclease.
A. tRNA$_{UUC}^{Glu}$ sequence showing the hybridized reverse transcriptase (RT) primer and oligonucleotide standards used to map the CdiA-CT^{STECO31} cleavage site.
B. Primer extension analysis. RNA was isolated competition co-cultures and cells intoxicated by intracellular CdiA-CT^{STECO31} expression. Samples from in vitro nuclease assays were also analyzed. The neutralizing effect of CdiI^{STECO31} was tested where indicated. The reverse transcriptase (RT) primer was radiolabeled, hybridized to tRNA$_{UUC}^{Glu}$, and extended with reverse transcriptase. Reactions and radiolabeled marker oligonucleotides were run on a 50% urea, 15% polyacrylamide gel and visualized by phosphorimaging.
C. Position of CdiA-CT^{STECO31} cleavage within the tRNA$_{UUC}^{Glu}$ anticodon loop. Modified nucleotides are highlighted in red.
[Colour figure can be viewed at wileyonlinelibrary.com]

Table 1: Data processing and refinement statistics

Processing	
Wavelength (Å)	0.9793
Resolution range (Å)a	30.0 – 2.00 (2.03 – 2.00)
Space group	$P6_5$
Unit cell parameters (Å)	a = 89.43 b = 89.43 c = 76.15
Unique reflections	23,397 (1,192)
Multiplicity	6.3 (5.7)
Completeness (%)	99.2 (99.8)
$<I>/<\sigma I>$	17.03 (2.21)
Wilson B factor ($Å^2$)	21
Rmergeb	0.117 (0.839)
CC1/2	0.689
CC*	0.903
Refinement	
Resolution (Å)	30.00 - 2.00
Reflections work/test set	22,180/1,140
R_{work}/R_{free}	0.1674/0.2076
Average B factor ($Å^2$) (No of atoms)	
macromolecules	30.4 (2,125)
solvent	34.3 (194)
Rmsd bond lengths (Å)	0.014
Rmsd bond angles (°)	1.487
Ramachandran favoredd (%)	98.1
Ramachandran outliers (%)	0
Clashscored	0.72

- [a] Values in parentheses correspond to the highest resolution shell.
- [b] $R_{merge} = \Sigma h \Sigma j |Ihj – <Ih>|/\Sigma h \Sigma j Ihj$, where Ihj is the intensity of observation j of reflection h.
- [c] $R = \Sigma h |Fo|–|Fc|/\Sigma h |Fo|$ for all reflections, where Fo and Fc are observed and calculated structure factors respectively. R_{free} is calculated analogously for the test reflections, randomly selected and excluded from the refinement.
- [d] As defined by Molprobity (Davis et al., 2004).

Table 2: Direct H-bonds and salt-bridges at the CdiA-CT/CdiI^{STECO31} interface

CdiA-CT atom	CdiI atom	distance (Å)
Lys197 (NZ)	Ser63 (O)	3.23
Lys257 (NZ)	Glu37 (OE1)	2.81
Lys197 (NZ)	Glu70 (OE2)	2.54
Lys257 (NZ)	Asp34 (O)	3.16
Lys257 (NZ)	Asp34 (OD1)	2.85
Lys261 (NZ)	Asp48 (OD1)	3.3
Lys261 (NZ)	Asp48 (OD2)	2.76
Lys197 (N)	Asp61 (OD2)	3.26
Lys181 (NZ)	Asp78 (OD2)	3.49
Lys257 (NZ)	Asn38 (OD1)	2.77
His187 (NE2)	Tyr75 (O)	2.84
His204 (NE2)	Tyr75 (OH)	2.69
Gly202 (N)	Tyr75 (OH)	2.7
Gln295 (NE2)	Thr18 (OG1)	3.25
Asp298 (N)	Asn2 (OD1)	2.8
Asn254 (ND2)	Ile46 (O)	2.96
Asp298 (O)	Asn2 (ND2)	3.11
Asp298 (OD2)	Asn28 (ND2)	2.85
Asn254 (O)	Asn38 (ND2)	3.01
Arg307 (NH1)	Glu29 (OE1)	2.96
Arg307 (NE)	Asp48 (O)	3.38
Arg307 (NH2)	Asp48 (O)	2.74
Tyr309 (OH)	Asp48 (O)	3.01
Asn292 (O)	Tyr10 (N)	3.1
Gly293 (O)	Trp17 (NE1)	2.92
Asn254 (OD1)	Ile46 (N)	2.75

Table 3: Plasmids

Plasmid	Description[a]	Reference
pTrc99a	IPTG-inducible expression plasmid, Amp^R	GE Healthcare
pTrc99KX	Derivative of pTrc99a with additional 5′ KpnI, and 3′ SpeI and XhoI restriction sites, Amp^R	Koskiniemi et al. (2014)
pMCSG58	LIC overexpression vector containing *argU* and *ileX* tRNA genes, Amp^R	Eschenfeldt et al. (2013)
pMCSG76	Over-expression vector with Clo DF13 origin of replication, Spc^R Str^R	Eschenfeldt et al. (2013)
pMCSG88	LIC overexpression vector containing *argU* and *ileX* tRNA genes and Clo DF13 origin of replication, Spc^R Str^R	This study
pDAL8914	Constitutive expression of chimeric *cdiA*^{EC93}-*CT*^{GN05224} and *cdiI*^{GN05224} genes, Cm^R	This study
pDAL8924	pTrc99KX::*cdiI*^{GN05224}, Amp^R	This study
pDAL8956	pET28A(+)::*cdiA-CT*//*cdiI*^{GN05224}, Kan^R	This study
pMCSG58-APC200204	pMCSG58::*cdiA-CT/cdiI*^{STECO31}, Amp^R	This study
pMCSG88-APC111471	pMCSG88::*cdiI*^{STECO31}, Spc^R Str^R	This study
pCH450	pACYC184 derivative with *E. coli araBAD* promoter for arabinose-inducible expression, Tet^R	Hayes and Sauer (2003)
pCH2408	Constitutive expression of chimeric *cdiA*^{EC93}-*CT*^{STECO31} and *cdiI*^{STECO31} genes, Cm^R	This study
pCH4066	pTrc99KX::*cdiI*^{Ymo43969}, Amp^R	This study
pCH4496	pTrc99KX::*cdiI*^{STECO31}, Amp^R	This study
pCH6278	pUC57::*cdiA-CT/cdiI*^{STECO31}, Amp^R	Genscript
pCH6290	pUC57::*cdiA-CT/cdiI*^{Ymo43969}, Amp^R	Genscript
pCH6505	pET21S::*cdiA-CT/cdiI*^{Dd3937}, Amp^R	Aoki et al. (2010)
pCH8102	pET21::*BC0921*, Amp^R	This study
pCH9273	pET21::*BC0920-CT/BC0921*, Amp^R	Holberger et al. (2012)
pCH10163	Cosmid pCdiA-CT/*pheS** that carries a *kan-pheS** cassette in place of the *E. coli* EC93 *cdiA-CT/cdiI* coding sequence. Used for allelic exchange and counter-selection. Cm^R Kan^R	Morse et al. (2012)
pCH11140	pCH450::*rhsA-CT/rhsI*^{ECL}-*DAS*, Tet^R	This study
pCH11884	pET21::*cdiA-CT/cdiI*^{STECO31}, Amp^R	This study
pCH12158	pCH450::*cdiA-CT/cdiI*^{Kpn342}-*DAS*, Tet^R	This study
pCH12599	pCH450::*cdiA-CT/cdiI*^{EC3006}-*DAS*, Tet^R	Willett et al. (2015)
pCH12847	Constitutive expression of chimeric *cdiA*^{EC93}-*CT*^{Ymo43969} and *cdiI*^{Ymo43969} genes, Cm^R	This study
pCH12964	pET21::*cdiA-CT(N319A)/cdiI*^{STECO31}, Amp^R	This study

a Abbreviations: Amp^R, ampicillin-resistance; Cm^R, chloramphenicol-resistance; Kan^R, kanamycin-resistance; Spc^R, spectinomycin-resistance; Str^R, streptomycin-resistance; and Tet^R, tetracycline-resistance.

Table 4: Oligonucleotides

Identifier	Descriptive name	Sequence
204F40		5′-GTC TCT CCC ATG GTT GAT AAT AAC TAC CTG AGC GTG TCT G-3'
204R49		5′-TGG TGG TGC CCA GCG GAT GGG ATT TTA GAC AGT AAT TTG TTT ACT TTT T-3'
204CdiIF		5′-GTC TCT CCC ATG AAT AAA TAT TTA TTT GAA TTG CCA TAT GAA CGT TCT G-3'
CH367	Gly-GCC-probe	5′-CTT GGC AAG GTC GTG CT-3'
CH368	Pro-GGG-probe	5′-CAC CCC ATG ACG GTG CG-3'
CH374	Ala-UGC-probe	5′-TCC TGC GTG CAA AGC AG-3'
CH379	Arg-CCU-probe	5′-CCT GCA ATT AGC CCT TAG G-3'
CH380	Gly-UCC-probe	5′-CCC GCA TCA TCA GCT TGG AAG GC-3'
CH381	Leu-CAA-probe	5′-CCC GCA CAG CGC GAA CGC CG-3'
CH382	Phe-GAA-probe	5′-TGC TCT ACC GAC TGA GCT A-3'
CH383	Pro-CGG-probe	5′- CTT CGT CCC GAA CGA AGT G-3'
CH385	Thr-CGU-probe	5′-CCT ACG ACC TTC GCA TT-3'
CH406	5S probe	5′-ATG CCT GGC AGT TCC CTA CTC TC-3'
CH407	16S probe	5'-TGC GCT TTA CGC CCA GTA ATT CC-3'
CH408	23S probe	5′-GTT TAG CCC CGT TAC ATC TTC CG-3'
CH436	Arg-CCG-probe	5′-CCT GAG ACC TCT GCC TCC GGA-3'
CH451	Arg-UCU-probe	5′-CCT GCG GCC CAC GAC TTA G-3'
CH452	Arg-ICG-probe	5′-CCT CCG ACC GCT CGG TTC G-3'
CH487	Cys-GCA-probe	5′-GGA CTA GAC GGA TTT GCA A-3'
CH577	Ile-GAU-probe	5′-ACC GAC CTC ACC CTT ATC AG-3'
CH618	Trp-CCA-probe	5′-CCC AAC ACC CGG TTT TGG-3'
CH791	Gln-CUG-probe	5′-TCG AAT GCC GGA ATC AGA A-3'
CH798	Tyr1-GUA-probe	5′-CTT CGA AGT CTG TGA CGG CAG-3'
CH799	Tyr2-GUA-probe	5′-CTT CGA AGT CGA TGA CGG CAG-3'
CH800	Asp-QUC-probe	5′-TCG AAC CCG CGA CCC CCT GCG-3'
CH801	Asn-QUU probe	5′-CTC GAA CCA GTG ACA TAC GG-3'
CH837	Ser-GCU-probe	5′-CCC CGG ATG CAG CTT TTG ACC-3'
CH1046	Leu-GAG-probe	5′-CCC GTA AGC CCT ATT GGG CA-3'
CH1047	Ser-UGA-probe	5′-AAC CCT TTC GGG TCG CCG GTT TTC-3'
CH1248	Val-UAG-probe	5′-CGC CGA CCC CCT CCT TGT AAG-3'
CH1249	Lys-UUU-probe	5′- CCT GCG ACC AAT TGA TTA AA-3'
CH1417	Glu-UUC-probe	5′- CCC CTG TTA CCG CCG TG-3'
CH1706	BC0921-Nhe-rev	5′-CTC TAT <u>GCT AGC</u> TTT CAA TTC TTT ATT TTT TCC-3'
CH2032	Gly-CCC-probe	5′-CCC TCG TAT AGA GCT TGG GAA-3'
CH2034	Leu-CAG-probe	5 -CCC CCA CGT CCG TAA GGA CA-3'
CH2035	Leu-UAG-probe	5′ - CAC CTT GCG GCG CCA GAA
CH2036	Leu-UAA-probe	5′ - CCC GCA CAG CGC GAA CGC CG
CH2037	fMet-CAU-probe	5′ - CGG GTT ATG AGC CCG ACG A
CH2038	Met-CAU-probe	5′ - CCT GTG ACC CCA TCA TTA TGA
CH2040	Ser-CGA-probe	5′ - GTA GAG TTG CCC CTA CTC CGG
CH2042	Thr-GGU-probe	5′ - CTG GGG ACC CCA CCC CT
CH2784	NMB0503-Kpn-for	5′ - GGT GGT ACC ATG AAA AAT AAT ATT TTT C
CH2785	NMB0503-Spe-rev	5′ - TTC ACT AGT GGT TTC ATG CAG GCT AC
CH2839	NMB0502-Nco-for	5′ - CTC CCA TGG TGA AAA ATA ATC AGC
CH3172	EC3006-CT(OE)-for	5′ - CAG GTA GGA ACT CGG TTG AGA ATA ATT ATC TTA GCG TGT CTG AAA AGA CAG AGC
CH3538	STECO31-CT-Nco-for	5'- TTT CCA TGG TTG ATA ATA ACT ACC
CH3539	STECO31-cdiI-Spe-rev	5′ - TTT ACT AGT GGA TGG GAT TTT AGA C
CH3569	STECO31-mid-rev	5′ - GGT CTG GTG TCT AAC CTT TGG TTA GGA TGG GAT TTT AGA CAG TAA TT
CH3624	STECO31-cdiI-Xho-rev	5′ - TTT CTC GAG TTA GGA TGG GAT TTT AG

48

CH3747	Ymol-mid-for	5´ - CAG GTA GGA ACT CGG TTG AGA ATA ATA ACC TTA GTT TTG GCA AAG G
CH3748	Ymol-mid-rev	5´ - GGT CTG GTG TCT AAC CTT TGG GTT AAG CAG GTA ATT TAG TCA GTA AAT C
CH3854	STECO31-cdiI-Kpn-for	5'- TTT GGT ACC ATG AAT AAA TAT TTA TTT GAA TTG C
CH3976	Ymol-cdiI-Kpn-for	5´ - TTT GGT ACC ATG AAA AAG GAA TTA TTT AGT CAA C
CH3977	Ymol-cdiI-Spe-rev	5´ - TTT ACT AGT AGC AGG TAA TTT AGT CAG
CH4016	STECO31-H187A-rev	5'-CAA GTT GAG CTA AGG CTT CCG GCT G
CH4017	STECO31-H204A-rev	5´ - GAA TAC ATC AGC ATT AGC TGC ACC GCT AAT TCC
CH4018	STECO31-N319A-rev	5´ - CCT ATT TCG GGT GGA AGG CAC TTA CAA TTC C
CH4019	STECO31-H321A-rev	5´ - CCT ATT TCG GGG CGA AGT TAC TTA CAA TTC C
CH4100	EC93-cdiA-8220-for	5´ - GAA GCG ATG AAA GCA GCC AGG
CH4101	EC93-cdiA-8718-rev	5´ - CTC AAC CGA GTT CCT ACC TGC CTG
CH4102	EC93-cdiI-296-for	5´ - AAC CCA AAG GTT AGA CAC CAG ACC
CH4103	EC93-cdiI-+486-rev	5´ - AAG TAG GCA TTC TCG ACC CTG
CH4269	STECO31-K261A-rev	5´ - ATC ATA AAT AGT GGC CTC AAA AGG CTT
CH4270	STECO31-T262A-rev	5´ - AGG ATC ATA AAT GGC TTT CTC AAA AGG
CH4272	gltW-C33-marker	5´ - GAT CGA ACC CCT GT TAC CGC CGT GAA A
CH4273	gltW-U36-marker	5´ - GGA TTC GAA CCC CTG TTA CCG CCG TG
CH4274	gltW-C39-marker	5´ - GGA TTC GAA CCC CTG TTA CCG CC
CH4275	gltW-RT-primer	5´ - GGA TTC GAA CCC CTG TTA CC
CH4398	BC0921-Kpn-for	5´ - ACA GGT ACC ATG AAA TAT CCA TAT AGC TTT GAA G
CH4649	tRNA-Glu 5´ probe	5´ - GAA AGG GCG GTG TCC TG
SK693	GN05224-CT-for	5´ - TTC TTG GGG CCG GAG AGC CGA GAG
SK694	GN05224-CT-rev	5´ - CCG GCG ATC TGC GGG TGT ACA AGC
SK733	GN05224-cdiI-Kpn-for	5´ - GAG CGG TAC CAT GAG TAT TAA TAA TAG
SK734	GN05224-cdiI-Xho-rev	5´ - GAG GCT CGA GTT ATC TCT TAT TTT CTG G
SK724	GN05224-CT-Nco-for	5´ - CAG ACC ATG GTG GAG AAT AAC TTC TTG GG
DL3986	GN05224-cdiI-Xho-rev	5´ - CCA GCT CGA GTC TCT TAT TTT CTG GAT TAA TAT ATA AAT ACC TAT TAT CTA TTT C

Chapter III: Discovery of another EF-Tu dependent RNase toxin in *Escherichia coli* O32:H37

A) Introduction

Bacteria employ multiple strategies to constantly compete for resources to survive. In a close contact environment, the contact-dependent inhibition (CDI) system is one of the weapons used among Gram-negative pathogens to compete for space and resources with other bacteria (Aoki et al., 2010; Hayes et al., 2014; Ruhe, Low, et al., 2013). The CDI system resides within the inhibitor (CDI+) cells, is comprised of three genes within its locus; a beta-barrel transporter CdiB allows its cognate CdiA filament to translocate through its lumen and deliver a polymorphic toxin derived from its C-terminal domain (CdiA-CT), which possesses a cytoplasmic-entry domain required for inner-membrane protein recognition (Aoki et al., 2008; Willett et al., 2015). Most importantly, an immunity factor is always encoded at the end of the locus to safeguard the CDI+ from being affected (or killed by) self- and kin-intoxications.

Our previous studies have shown that the CDI ribonuclease toxins from different systems share structural features similar to the C-terminal domains of colicins, which are diffusible effector proteins released by many colicinogenic *E. coli* (Cascales et al., 2007). CdiA-CT[EC16] from *Erwinia chrysanthemi* EC16, CdiA-CT[BAA-461] from *Bordetella petrii* BAA-461, CdiA-CT[K96243] from *Burkholderia pseudomallei* K96243, and CdiA-CT[McGHS1] from *Morxella catarrhalis* McGHS1 resemble similar folds to the nuclease found in colicins E3, D, E5, and E2, respectively (Aoki et al., 2010; Walker et al., 2004). Common substrates of these toxins include anticodon loops of certain tRNAs (CdiA-CT[K96243] and colicin E5) and the 16s rRNA (CdiA-CT[EC16] and colicin E3) (Beck et al., 2014; Nikolakakis et al., 2012).

Nevertheless, CdiA-CTs from some strains can exploit normal cytoplasmic proteins as their permissive factors to either stabilize or activate the toxin activities. The CdiA-CT[UPEC536] domain deployed by *Escherichia coli* strain 536 (UPEC536) is a latent tRNase until activated by the biosynthetic enzyme *O*-acetylserine sulfhydrylase A, CysK. Commonly, *E. coli* cells have CysK and CysM, *O*-acetylserine sulfhydrylase isoenzymes to facilitate the second and final step of cysteine synthesis from serine. It was found that the CysK/CdiA-CT[UPEC536] interaction mimics the one found in CysK/CysE (serine *O*-acetyltransferase). Remarkably, target cells in the absence of CysK, are completely resistant to the toxin (Diner et al., 2012). Furthermore, two other common translation factor proteins, EF-Tu (elongation factor thermos unstable) and EF-Ts (elongation factor thermos stable), were found to associate with several CdiA-CTs, CdiA-CT[NC101] binds to domain 2 of EF-Tu to cleave the single stranded 3'- end of tRNAs containing the guanine discriminator nucleotides (Michalska et al., 2017), CdiA-CT[EC869] binds strongly with EF-Tu, follows by the help of EF-Ts for stabilizing the GTP·EF-Tu·tRNA ternary complexes, which facilitate the tRNA^{Gln} and tRNA^{Asn} cleavages (Jones et al., 2017), and CdiA-CT[Kp342] requires both EF-Tu and EF-Ts to cleave the acceptor stem of deacylated tRNA$_{GAU}^{Ile}$(Gucinski et al., 2019).

In this study, we report that CdiA-CT isolated from *Escherichia coli* strain O32:H37 is a general RNase when delivered inside *E. coli* cells. Interestingly, CdiA-CT[O32:H37] shares structural homology with the nuclease domain colicin D despite a lack of primary sequence homology. CdiA-CT[O32:H37] has a similar positively charged active central groove as colicin D's but lacks substrate specificity for the tRNA^{Arg} isoacceptor. Furthermore, we found that Trp52 and His67 residues are critical to stabilize the N-terminal sequence of the toxin and allow it to bind with EF-Tu to facilitate the nuclease activity. We propose that the toxin is

generally unstable after delivery to the target cells and the presence of either EF-Tu or EF-Ts enhances the toxin stability and promote the nuclease activity towards RNAs.

B) Results

CdiA-CT/I$^{O32:H37}$ complex structure and EF-Tu co-purification

We co-expressed the CdiA-CT$^{O32:H37}$ toxin along with the His$_6$- tagged CdiI$^{O32:H37}$ immunity protein and purified the complex by Ni^{2+} affinity chromatography. Dr. Karolina M. Michalska from Argonne National Laboratory crystallized the purified structure. The crystal structure of CdiA-CT$^{O32:H37}$ adopts a fold consisting of five α-helices (α1-α5) and four beta-strands (β1-β4). The active residues are located within the α1 helix and β4 strand, forming a catalytic pocket. In contrast, the immunity protein has a relatively non-structural fold with two small alpha helices (α'1 and α'2) (**Fig. 1A**). These two proteins form a stable complex with several key residues involved in hydrogen bonds and salt-bridges, as identified by PDBePISA at the interface between the toxin and immunity protein (**Table 1**).

Among these residues, Glu32 of CdiI$^{O32:H37}$ directly contacts two putative catalytic residues of CdiA-CT$^{O32:H37}$, His71 and Thr130, within the catalytic pocket, with distances of 2.65 and 2.67 Å, respectively (**Table 1**). Outside of the pocket, Glu19 and Tyr23 of the immunity factor form hydrogen bonds and salt-bridges with Tyr140 and Arg129 from the toxin (**Fig. 1B**). These direct contact points allow the immunity protein to fully cover the toxin from accessing its substrates. However, the immunity protein may have a short half-life due to its unique fold, with four cysteine residues (Cys9, Cys12, Cys26, and Cys29) chelating an iron atom to stabilize its non-structural fold (**Fig. 1C**).

Upon further analysis of the complex, we observed that the complex co-purified with an elongation factor-Tu (EF-Tu) migrating around the size of 43 kDa. We purified the

complex using two different approaches, as indicated in **Fig. 1D**. The first approach involved purifying the full complex under native conditions, where the protein-expressing cells were lysed in a native buffer, and proteins were eluted from Ni^{2+}-NTA agarose. The eluant contains a mixture of three species, EF-Tu (43 kDa), CdiA-CT$^{O32:H37}$ (15.7 kDa), and CdiI$^{O32:H37}$ (6.5 kDa) (**Fig. 1E,** *sample 1*). Notably, the immunity protein did not migrate at its predicted size (6.5 kDA) and was difficult to stain with Coomassie blue. The second approach involved first eluting the EF-Tu/CdiA-CT$^{O32:H37}$ complex from the CdiI$^{O32:H37}$ (**Fig 1E,** *sample 2*) under denatured condition and then natively eluting the CdiI$^{O32:H37}$ off the Ni^{2+}-NTA agarose (**Fig 1E,** *sample 3*). This purification method revealed that the EF-Tu strongly binds to the CdiA-CT/CdiI$^{O32:H37}$complex, and the denaturing buffer could not remove the EF-Tu from the mixture without disrupting the toxin and immunity interaction.

CdiA-CT$^{O32:H37}$ is a general RNase

To investigate the enzymatic activity of CdiA-CT$^{O32:H37}$, we performed a Phyre2 search for homologous structure. The Phyre2 search results suggest that CdiA-CT$^{O32:H37}$ closely resembles the folding of colicin D, a known tRNase, belonging to a colicin D/E5 superfamily. Although the sequences of CdiA-CT$^{O32:H37}$and colicin D were poorly aligned, the crystal structures super-impose well (**Fig. 2**). The CdiA-CT$^{O32:H37}$ structure consists of four beta-strands, in contrast to colicin D, which has three-beta-strands. Notably, the very last beta-strand of both structures (β3-strand of colicin D and β4-strand of CdiA-CT) forms the catalytic pocket with their α1 helix (**Fig. 2A**). Additionally, the last helix of colicin D, α4, is split into two helices, α4 and α5 in CdiA-CT$^{O32:H37}$ (**Fig. 2B**).

Minor differences in the structure of the CdiA-CT$^{O32:H37}$ could lead to more diversified substrate recognition compared to its counterparts, colicin D. Internal expression

assays were performed to demonstrate the toxicity of CdiA-CT$^{O32:H37}$. It is known that the toxin is too lethal to be solely expressed by the host, and an induced immunity protein is required to facilitate the cell growth (Michalska et al., 2018). Hence, the toxin was cloned under an arabinose inducible promoter and co-transformed with its cognate immunity protein regulated by a *trc (trp-lac)* promoter under the induction of isopropyl ß-D-1-thiogalactopyranoside (IPTG). *E. coli* cells harboring both toxin and immunity plasmids were seeded at OD$_{600}$ of 0.05 in the presence of IPTG to ensure immunity proteins were expressed prior to the induction of the toxin. At OD$_{600}$ of 0.2, the culture was split into two independent flasks, one with D-glucose repression and the other with L-arabinose induction. The expression of toxin halted bacterial growth at OD$_{600}$ of 0.3 whereas the toxin repression allowed cells to grow normally until OD$_{600}$ of 1.7 at the final timepoint (**Fig. 3A**).

The colicin D nuclease domain is found to be less potent than the CdiA-CT$^{O32:H37}$ and *E. coli* cells with only an arabinose induction tRNase domain plasmid managed to survive without acquiring mutations. We performed a similar approach that included a low concentration of D-glucose when seeding the culture to suppress the toxin expression. When the cells reached an OD$_{600}$ of 0.2, the cultures were also split into two for further D-glucose suppression and L-arabinose induction. *E. coli* cell growth was halted at OD$_{600}$ of 0.5 when colicin D nuclease domain was expressed (**Fig. 3A**). RNA samples from both systems were collected at the indicated times. Total RNA samples from the CdiA-CT$^{O32:H37}$ reveals a striking pattern of RNA degradation, especially of the high molecular weight rRNAs (**Fig. 3B** & **S1**).

Northern blot analyses were used to investigate the substrate specificity of CdiA-CT$^{O32:H37}$. This toxin appears to be a general RNase, capable of targeting rRNAs, including

23s rRNA (**Fig. 3B** & **S1**) and the 5' end of 16s rRNA (**Fig. 3C** & **S1**). Two representative tRNA species, tRNA$^{Glu}_{UUC}$ and tRNA$^{Arg}_{CCU}$, were examined. In contrast to colicin D, which can only recognize tRNA^{Arg} isoacceptors, CdiA-CT$^{O32:H37}$ can also recognize other tRNA species (**Fig. 3C** & **S1**). The CdiA-CT$^{O32:H37}$ can indiscriminately target variable regions of tRNAs, either at the tRNA anticodons (e.g., tRNA^{Arg}, tRNA^{Asp}, tRNA^{Cys}, tRNA^{Glu}, tRNA^{Gly2}, tRNA^{His}, tRNA^{Ile}, and tRNA^{Val}) or destabilize tRNA structures beyond detection level (e.g., tRNA^{Ala}, tRNA^{Pro}, tRNA^{Sec}, and tRNA^{Ser}) (**Fig. S1**). Furthermore, the 5s rRNA substrate specificity appears to vary depending on the dosage of toxin and the availability of the cleavage site during ribosome translation; it is either cleaved poorly or remains inert (**Fig. 3C** & **S1**).

CdiA-CT$^{O32:H37}$ active site

In general, most RNases use a catalytic mechanism similar to that of RNase A, which involves two critical His residues to facilitate the nuclease activity (Findlay et al., 1962; Gilliland, 1997). In the case of CdiA-CT$^{O32:H37}$, two catalytic His residues, His67 and His71, can be found in α1 helix, along with two other residues, Thr130 and Phe132 in the β4-strand. These four residues correspond to the catalytic residues found in the colicin D nuclease domain, namely Lys608, His611, Ser677, and Trp679, respectively (**Fig. 4A**) (Graille et al., 2004). To test the hypothesis that specific residues were required for enzymatic activity, a co-culture assay was performed with inhibitor cells carrying either wild-type or a mutant *cdiBAI* locus mixed with the target cells carrying the plasmid-borne tRNA$_{CCU}^{Arg}$ gene (pArgW), with or without the known outer-membrane receptor, Tsx. The rationale behind having a pArgW plasmid in the target cells is that tRNA$_{CCU}^{Arg}$ has low copy numbers in the cells and could not be detected at the basal level in the inhibitor cells; hence, the signal detected from the northern blot directly comes from the target cells.

First, we analyzed the CDI expression level in all the inhibitor cells by conducting immunoblots probing for the TPS-domain. We found that the filaments were expressed properly, allowing us to conclude that the defects in activity are caused by impaired catalytic residues, not by expression level or delivery issues (**Fig. 4C**). Subsequently, we observed that target cells without Tsx showed resistance to wild-type toxin delivery, resulting in a competitive index equal to one, whereas target cells expressing plasmid borne Tsx were fully susceptible to the toxin (**Fig. 4B**). Mutations at the residues responsible for the catalytic activity, His67Ala, His71Ala, and Phe132Ala mutants exhibited inert toxin activity, with target cells remaining viable, and $tRNA_{CCU}^{Arg}$ was not cleaved (**Fig. 4B** & **D**). We also found that the Lys70Ala and Arg104Ala mutations had attenuated activity, which was observed in both targe cell viability counts and $tRNA_{CCU}^{Arg}$ blot (**Fig. 4B** & **D**). These observations suggest that these two residues play an important role in stabilizing the RNA substrates in the catalytic pocket due to their positive charges. As expected, the Lys73 residue faces away from the catalytic side and does not participate in the nuclease activity, carrying out an identical activity as the wild-type counterpart (**Fig. 4B** & **D**).

Remarkably, when the proteins were purified natively to retain the protein complex, we found that all wild-type and mutant variants were able to secure EF-Tu, except for the mutation at His67. This provides another critical piece of evidence for the importance of this residue in the nuclease activity (**Fig. 4E**). Finally, we decided to conduct *in vitro* reactions using purified wild-type or indicated mutant variant of CdiA-CT$^{O32:H37}$, CdiI$^{O32:H37}$, and EF-Tu, incubated with the extracted total RNA to confirm the observations made during *in vivo* delivery. Under *in vitro* conditions, we were able to probe for more RNA species to confirm the activity, including $tRNA_{UUC}^{Glu}$ and 5s rRNA. The blots for these RNA also support the

observation made in *in vivo* delivery. Mutations in His67, His71, Thr130, and Phe132 fully abolished the nuclease activity, whereas the mutation at Arg104 appears to be less active under *in vitro* conditions (**Fig. 4F**).

CdiA-CT$^{O32:H37}$is required co-factor for activity

As demonstrated in previous studies, a co-factor is always required for some toxins to facilitate their activities. This includes CysK for CdiA-CT^{UPEC536} (Hayes et al., 2014), EF-Tu in CdiA-CT^{NC101}(Michalska et al., 2017), and EF-Tu/ EF-Ts in CdiA-CT^{EC869} (Jones et al., 2017) and CdiA-CT^{Kp342}(Gucinski et al., 2019). In the case of CdiA-CT$^{O32:H37}$, the co-purification of EF-Tu demonstrates its important role in the nuclease activity (**Fig. 4E**). To validate the necessity of EF-Tu in the nuclease activity, a dosage titration assay was conducted (**Fig. 5A**). The presence of EF-Tu enhanced the toxin activity. For example, when comparing the *in vitro* nuclease activity at the 100 nM of CdiA-CT$^{O32:H37}$, we found that in the absence of EF-Tu, the toxin is inert to its substrates, including 5s rRNA and tRNA$_{UUC}^{Glu}$. However, when EF-Tu was supplemented into the same conditions, the 5s rRNA and tRNA$_{UUC}^{Glu}$ were slowly cleaved. We speculate that the toxin may recognize the tRNA substrate better than the rRNA as indicated by the full conversion of tRNA$_{UUC}^{Glu}$ at 100 nM compared to the pseudo-cleavage in 5s rRNA (**Fig. 5A**). From our data, we found that 600 nM of toxin and 1 µM of EF-Tu are the optimal concentrations fully cleave both RNA species.

Since CdiA-CT$^{O32:H37}$ and the colicin D nuclease domain shared structural homology, we decided to investigate the requirement of EF-Tu for colicin D activity. As expected, only CdiA-CT$^{O32:H37}$ required EF-Tu to facilitate its activity. In all the cases tested, 5s, tRNA$_{UUC}^{Glu}$, and tRNA$_{CCU}^{Arg}$, the toxin was only able to initiate the cleavage when EF-Tu

was introduced. In contrast, the colicin D nuclease domain was able to cleave its cognate substrate $tRNA_{CCU}^{Arg}$ without the assistance of EF-Tu (**Fig. 5B**). Notably, the off-target effect observed in the colicin D aligns with data previously described in Graille et al., 2004.

Subsequently, we investigated the requirement of EF-Ts through *in vitro* nuclease activity and co-cultured assay using previous described viable *tsf* mutant, Ala202Glu and Δcoiled-coli (Jones et al., 2017). Our data indicate that EF-Ts may play a role in stabilizing the toxin in the absence of EF-Tu by promoting an attenuated nuclease activity, where the probed RNA species, 5s rRNA and $tRNA_{UUC}^{Glu}$, were incompletely cleaved (**Fig. 6A,** lane 5). The toxin's activity returned to its full potential when EF-Tu was supplemented (**Fig. 6A,** lanes 3**,** 6, & 7). Importantly, the toxin activity in the presence of EF-Tu can be neutralized with its cognate immunity protein (**Fig. 6A,** lane 8). Furthermore, we also examined *tsf(*Ala202Glu) and *tsf(*Δcoiled-coil), which encode EF-Ts with an Ala202Glu substitution at the tip of the coiled-coil domain and a deletion of the coiled-coil domain. Unlike the reported CdiA-CTs from Kp342 (Gucinski et al., 2019) and EC869 (Jones et al., 2017), these *tsf* mutants did not influence the toxin activity. Competitive indices of the co-cultures revealed that the target cells with either the wild-type or mutant *tsf* were inhibited at the same rate (**Fig. 6B**), which can also be further confirmed by the analysis of RNAs isolated from the same co-cultured assay (**Fig. 6C**). In summary, we found that CdiA-CT$^{O32:H37}$ required a co-factor to facilitate its nuclease activity and EF-Tu is the preferred partner to have for an optimal activity.

The importance of Trp52 in CdiA-CT$^{O32:H37}$

We hypothesized that the structure of CdiA-CT$^{O32:H37}$ may lack features that contribute to stability in the absence of a co-factor. Based on the unexpected result of the

His67Ala point mutant (**Fig. 4C**), we used AlphaFold to model the EF-Tu residues that interact with His67, in the hope of pinpointing the crucial residues that secure the fold of CdiA-CT$^{O32:H37}$ and its interaction with EF-Tu. AlphaFold predicted that the His67 can form hydrogen bond with another aromatic residue, Trp52, which secures a non-structural turn with a short α0 helix. Remarkably, this short fragment was not resolved during the crystallization process (**Fig. 1A**).

AlphaFold provided putative protein conformations for the wild-type and Trp52Phe, His67Ala, Trp52Ala, and ΔIle51-Trp52 mutant proteins. The structures of Trp52Phe and His67Ala resemble the wild-type conformation (**Fig. 7A**), suggesting that the aromatic residue is crucial in this location to secure the hydrogen bonding. Surprisingly, the model of His67Ala mutant did not reveal a disruption of the structure. Notably, however, the loss of the Trp52 and its neighboring residue, Ile51, disrupted the structure and unfolded the N-terminal sequence along with the α0 helix, similar to the effect of the single aromatic residue replacement, Trp52Ala (**Fig. 7B**). To further validate these predictions, we purified the toxin complex of these wild-type and mutant variants. The purified proteins revealed that the loss of the aromatic residue at position 52 can disrupt the interaction with EF-Tu (**Fig. 7C**). This means that the turn secured by Trp52 and His67 is critical for EF-Tu docking.

The loss of structural features at the N-terminus not only disrupted EF-Tu binding, but also toxin activity. Both the ΔIle51-Trp52 and Trp52Ala mutants completely abolished toxin activity, as demonstrated by *in vitro* reactions and *in vivo* delivery assays. In both cases, there was no RNA cleavage in the *in vitro* assays, and the target cells remained viable after a 15-minute mixing period (**Fig. 7D & E**). However, the Trp52Phe mutant was still able to inhibit the growth of the target cells, even though it could not target the RNA in the *in vitro*

condition (**Fig. 7D** & **E**). This strongly indicates that the 52nd residue in CdiA-CT$^{O32:H37}$ must have an aromatic side chain in order to secure the turn and α0 helix for EF-Tu binding and nuclease activity.

The interface between EF-Tu and CdiA-CT$^{O32:H37}$

To further investigate the relationship between EF-Tu and CdiA-CT$^{O32:H37}$, we utilized AlphaFold to predict the trimeric complex of EF-Tu, CdiA-CT, and CdiI. The model suggested that the complex formed with CdiI immunity protein occupied the catalytic pocket of CdiA-CT$^{O32:H37}$, whereas the Domain 1 of EF-Tu interacts with the turn formed by Trp52 and His67 (**Fig. 8A**). In this configuration, the nuclease domain is closer to Domain 3 of EF-Tu, which also serves as the tRNA binding site (Wang et al., 1997). This suggests that once the toxin is delivered in the intoxicated cells, EF-Tu not only stabilizes the toxin conformation, but also facilitates the access of substrates to the nuclease domain. We found that the turn made by Trp52 and His67 in the nuclease domain allows Lys45 and Ile51 to interact with Asp143 and Glu145 in EF-Tu, respectively (**Fig. 8B** & **Table 2**). Furthermore, we also discovered that the toxin can only form a stable complex with EF-Tu in the presence of immunity. When a catalytic mutant (His71Ala toxin mutant) was purified natively in the absence of immunity, it no longer was bound to EF-Tu, whereas the one with immunity can bind tightly to EF-Tu (**Fig. 8C**).

Additionally, we introduced mutations at several residues in EF-Tu, Asp142, Asp143, Glu145, and Leu149 based on the AlphaFold model and tested whether these mutants could disrupt the toxin activity. We discovered that Asp142Ala and Glu145Ala did not disrupt the toxin activity, as they exhibited similar toxin activity to wild-type EF-Tu after 120-minute of incubation (**Fig. 8D**). Mutations at Asp145 and Leu149 interrupted the nuclease activity, as

demonstrated with a slower rate of RNA degradation when comparing the 30-minute and 60-minute timepoints (**Fig. 8D**). However, when these mutants were incubated with RNA substrates for a longer time (120 minutes), all RNA species were ultimately degraded (**Fig. 8D**). Collectively, these data reveal that the presence of EF-Tu enhances the stability of toxin and increases the rate of RNA degradation.

C) Discussions

The delivery of CdiA toxin requires an intricate design since it must be unfolded and refolded to pass through two layers of the inhibitor cell and then those of the target cell. This design must ensure that the toxin is flexible enough to be unfolded when passing through its cognate CdiB upon binding to its outer-membrane receptor in the target cell and, ultimately, to be imported through the target inner-membrane transporter (Aoki et al., 2005; Ruhe et al., 2018; Ruhe, Wallace, et al., 2013; Willett et al., 2015). It is crucial for the toxin to refold properly in the target cell's cytoplasm. Most of the documented CdiA-CT systems from various sources have harnessed different cellular factors to either enhance stability or trigger activation (Gucinski et al., 2019; Hayes et al., 2014; Jones et al., 2017; Michalska et al., 2017). We hypothesize that these toxins may have evolved from a common ancestor that can utilize one of the most abundant proteins in bacteria. It is widely recognized that EF-Tu ranks among the most prevalent proteins in *Escherichia coli*, constituting up to 6% of the total expressed protein (Furano, 1975). The likelihood of encountering EF-Tu exceeds that of other factors in the cytoplasm. Hence, upon release from the inner-membrane transporter, the toxins can readily associate with EF-Tu, ensuring their proper folding and optimal function on their substrates.

Furthermore, this hypothesis is supported by the moonlighting function of EF-Tu. It has been reported that EF-Tu acts as a protein chaperone, facilitating the refolding of citrate synthase, α-glucosidase, and rhodanese after denaturation (Caldas et al., 1998; Kudlicki et al., 1997). EF-Tu can form stable complex with unfolded proteins, including carboxymethyl alpha-lactalbulmin and unfolded bovine pancreatic trypsin inhibitor (Caldas et al., 1998). Similarly, our data also suggest that EF-Ts plays a minor role in toxin activity under *in vitro* conditions, as EF-Ts can act as a steric chaperone, promoting proper folding and solubility in mutant EF-Tu (Krab et al., 2001).

The His67 residue of CdiA-CT$^{O32:H37}$ is a critical feature, required for RNA hydrolysis and ensuring the hydrogen bonding with Trp52 for securing the turn at the N-terminus. This non-structural turn contains Lys45 and Ile51, facilitating direct interaction with residues in Domain 1 of EF-Tu. Our data suggest that when the toxin reaches a high concentration, it does not require the assistance of EF-Tu for nuclease activity. This indicates that EF-Tu acts as an enhancing factor during contact-dependent toxin delivery, especially when the toxin concentration cannot reach the µM range. Although CdiA-CT$^{O32:H37}$ shares structural homology with the nuclease domain of colicin D, colicin D's nuclease activity does not require a co-factor, and it also has a more selective bias towards its substrate. The turn formed by Trp52 and His67 could not be found in the colicin D nuclease domain. Instead of His67, colicin D uses Lys608, which does not make contact with other residues in the non-structured N-terminal domain (**Fig. S2**). Additionally, CdiA-CT$^{O32:H37}$ utilizes the standard mechanism similar to RNaseA (Raines, 1998), where the His67 and His71 can act either as an acid or base for donating and accepting electrons. The positively charged Arg104 and

phenyl ring of Phe132 help to stabilize the phosphate group and the nucleotide ring, respectively.

In conclusion, we have successfully characterized CdiA-CT$^{O32:H37}$ as a general ribonuclease. The structural importance of Trp52 and His67 residues is evident in their role in facilitating EF-Tu binding and RNA hydrolysis. EF-Tu serves as an enhancer, significantly increasing the catalytic rate of the toxin.

D) Experimental Procedures

Strain and plasmid constructions

Plasmid pCH6281 with the synthetic locus of CdiA-CT/CdiI$^{O32:H37}$, acquired from Argonne national laboratory, was the template to build the CdiA-CT/CdiI variants in this study. The chimeric plasmid, pCH2816 was constructed by amplifying pCH6281 with oligos CH4262/CH4716 and cloned into pCH226, CdiA^{STEC3} stick with NheI/XhoI and selected on LB-Amp plate. Similarly, the mutant variants of CdiA-CT$^{O32:H37}$ was constructed with the megaprimer approach by amplifying off pCH2816 with the forward oligos CH4262 paired with the indicated site-directed mutagenesis oligos, including CH6209 (Trp52Phe), CH6210 (ΔIle51-Trp52), CH6211 (Trp52Ala), CH4794 (His67A), CH4795 (Lys70Ala), CH3897 (His71Ala), CH4796 (Lys73Ala), CH4797 (Arg104Ala), CH4798 (Thr130Ala), and CH4799 (Phe132Ala). These megaprimers were later combined with oligo CH4716 to amplify the full-length fragment from pCH2816 and cloned into pCH226 via NheI/XhoI to have the final chimeric toxin plasmids, including pCH7029 (Trp52Ala), pCH10085 (ΔIle51-Trp52), pCH10086 (Trp52Phe), pCH5506 (His67Ala), pCH5507 (Lys70Ala), pCH5508 (His71Ala), pCH5509 (Lys73Ala), pCH5510 (Arg104Ala), pCH5511 (Thr130Ala), and pCH5512 (Phe132Ala).

The N-terminal his-tagged wild-type and mutant toxins expression plasmids were constructed by amplifying from corresponding templates with oligos CH5152/CH4716 and cloned into pMCSG63 with KpnI/XhoI, including pCH6442 (wild-type), pCH7030 (Trp52Ala), pCH4112 (ΔIle51-Trp52), pCH4113 (Trp52Phe), pCH1746 (His67Ala), pCH1748 (Lys70Ala), pCH1750 (His71Ala), pCH1752 (Lys73Ala), pCH1754 (Arg104Ala), pCH1756 (Thr130Ala), and pCH1758 (Phe132Ala). Similarly, the C-terminal his-tagged wild-type toxin expression (pCH12716) and immunity DAS-tagged (pCH13459) constructs were built from pCH2816 with oligos CH3896/CH3898 and cloned into pET21P and pCH450 with NcoI/SpeI, respectively. The other immunity DAS-tagged constructs for CdiA-CT$^{EC97.0246}$ and CdiA-CT^{NMB0520} were amplified off pCH6276 and pCH1712 with oligos CH3534/CH3535 and CH839/CH2785, respectively. Then, the fragments were cloned into pCH450 with NcoI/SpeI.

Additionally, the C-terminal his-tagged toxin-immunity locus of ColD-ImmD (pCH2485) was cloned into pET21b with oligos CH714/CH560 via NdeI/XhoI. Furthermore, the wild-type CdiA-CT$^{O32:H37}$ (pCH3782) and colicin D nuclease domain (V585) (pCH1554) were cloned into pCH450 with oligos CH3896/CH4003 and oligos CH5283/5824 to have arabinose inducible plasmid. Notably, the wild-type CdiA-CT$^{O32:H37}$ must be cloned into a cell line carrying its cognate immunity plasmid, which was constructed by amplifying the immunity from pCH2816 with oligos CH3604/CH4716 and cloned into pTrc99a with KpnI/XhoI.

All C-terminal his-tagged EF-Tu, both wild-type and mutant variants were built by amplifying *tufA* locus from X90 cells with oligos the forward oligo CH4748, reverse oligo CH4749, and megaprimer oligos for indicated mutants, including CH6114 (Glu145Ala),

CH6115 (Asp142Ala), CH6116 (Asp132Ala), CH6224 (Leu149Gln), and CH6226 (Leu149Arg). All fragments were later cloned into pET21b with XbaI/XhoI.

In vivo expression and competition co-culture assays

For two-plasmid system O32:H37 expression, fresh wild-type X90 *E. coli* transformants with the pCH450::CdiA-CT$^{O32:H37}$ and pTrc99aKX::CdiI$^{O32:H37}$ were selected on LB-agar plates supplemented with 15 μg/mL tetracycline, 150 μg/mL ampicillin, and 0.4% D-glucose. Transformants were harvested off the agar plates and resuspended in 1 x M9 salts. Cells were seeded at $OD_{600}= 0.05$ in the fresh LB-media supplemented with 15 μg/mL tetracycline, 150 μg/mL ampicillin, and 1.5mM IPTG. Cultures were measured every 30 minutes at OD_{600} and a final concentration of 0.2% L-arabinose was to each culture when OD_{600} reaching 0.2. The growth curve measurement was accomplished when it reached the 5-hour timepoint. The full panel RNA substrate screen was conducted by using C-terminal ssrA(DAS) degron described (McGinness et al., 2006; Poole et al., 2011) on the cognate immunity proteins of each CdiA-CTs. *E. coli* X90 cells carrying these expression constructs were seeded at $OD_{600} \sim 0.05$ in LB media supplemented with 15 µg/mL tetracycline and incubated with shaking at 37 °C. After 60 min, the cultures were adjusted to 0.2% L-arabinose and cultured for an additional 4 hours.

A similar approach was taken for expressing colicin D tRNase domain (pColD), fresh X90 *E. coli* transformants with pColD were selected on LB-agar plates supplemented with 15 µg/mL tetracycline and 0.4% D-glucose. Transformants were harvested off the agar plates and resuspended in 1 x M9 salts. Cells were seeded at $OD_{600}= 0.05$ in the fresh LB-media supplemented with 15 µg/mL tetracycline and 0.1% D-glucose. Cultures were measured every 30 minutes at OD_{600} and a final concentration of 0.2% L-arabinose was to each culture

when OD_{600} reaching 0.2. The growth curve measurement was accomplished when it reached the 5-hour timepoint. Finally, all culture samples were harvested into an equal volume of ice-cold methanol after 2-3 hours of induction, and the cell pellets frozen at –80 °C for subsequent RNA extraction and analysis.

Inhibitor and target cells were grown to mid-log phase and mixed in test tube with 1: 1 ratio to a final volume of 10 mL. Allowed cells to mix on shaker for 15 minutes at 37°C, the cells were spotted on LB agar plate supplemented with 150 μg/mL ampicillin for inhibitor cell selection and LB agar plate supplemented with 15 μg/mL tetracycline for target cell selection. Colonies from both plates were enumerated and calculated as competitive indices (CI). CI is a ratio of target: inhibitor cells (CFU/mL) at the final timepoint divided by the ratio of target: inhibitor cells (CFU/mL) at the initial timepoint. Co-culture samples were harvested into an equal volume of ice-cold methanol at the final timepoint, and the cell pellets frozen at –80 °C for subsequent RNA extraction and analysis.

Protein purifications
All protein expressing *E. coli* cells (CH2016) were grown to OD_{600}=0.6 and induced expression by addition of isopropyl β-d-1-thiogalactopyranoside (IPTG) to 1.5 mM for 1 hour and stored in respective buffer plus 50% glycerol at -20 °C. The C-terminal His-affinity tagged immunity-toxin construct was purified in two independent approaches. Cell pellets were incubated in buffer A [20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 25mM imidazole] in the presence of 1 mg/mL lysozyme for 30 minutes, then cell lysates were sonicated for two rounds to achieve complete lysis. Ni^{2+}-NTA resins were added to the cleared supernatants (obtained by centrifugation for 15 minutes at 15,000 RPM). First, buffer B [20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 250mM imidazole] was used to elute

the whole intact complex. The alternative way was to separate the CdiA-CT/EF-Tu from the CdiI-His$_6$ by eluting the native lysed extract with denatured buffer [6 M guanidine-HCl, 20 mM Tris-HCl (pH 7.0)] and dialyzed once for 4 hours in buffer C [20 mM Tris-HCl (pH 7.0), 150 mM NaCl]. Consequently, the remaining CdiI-His$_6$ on the resins were washed thrice with buffer A before eluting in buffer C. Similarly, the C-terminal His-affinity tagged immunity was purified in the same native conditions, with two rounds of 4 hours dialysis in buffer C. All eluted proteins were resolved on SDS-PAGE (10 % acrylamide, 110V for 1 hour).

The N-terminal His-affinity tagged toxins were purified in two different conditions. The native purification approach was carried out to demonstrate the EF-Tu can bind to the toxin. Cell pellets were incubated in buffer A in the presence of 1mg/mL lysozyme followed by sonication. Ni^{2+}-NTA resins were added to the cleared supernatants (obtained by centrifugation for 15 minutes at 15,000 RPM). Then, the resins were washed thrice with buffer A. The CdiA-CT/CdiI/EF-Tu complex from wild-type and mutant CdiA-CTs were collected by using buffer B.

The denatured condition was conducted to study the toxin activity of CdiA-CT and its variants without the presence of EF-Tu. Cell pellets of CdiA-CT variants and colicin D nuclease domain were resuspended in denatured buffer. Cell debris was removed by centrifugation for 15 minutes at 15, 000 RPM. Ni^{2+}-NTA resins were added to the cleared supernatants to purify the N-terminal His-tagged toxin only. The resins were washed thrice with denatured lysis buffer before eluting with the denatured elution buffer [6 M guanidine-HCl, 20 mM Tris-HCl (pH 7.0), 250 mM Imidazole]. The unfolded toxins were refolded by

dialyzing twice against buffer D [20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 10 mM β-ME]. Finally, proteins were resolved on SDS-PAGE (10 % acrylamide, 110V for 1 hour).

EF-Tu and EF-Ts were purified as described in (Jones et al., 2017). Briefly, His_6-tagged proteins were eluted with buffer B and dialyzed against buffer D. After dialysis, His_6-from CdiA-CT variants and His_6-TrxA from EF-Tu and EF-Ts were cleaved using TEV protease and removed by Ni^{2+}-NTA resin prior to the *in vitro* nuclease reaction. Alternatively, C-terminal his-tagged EF-Tu variants were purified with the native approach with buffer A, eluted in buffer B, and stored in buffer C.

In vitro nuclease reaction

In vitro nuclease assays were performed in 20mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM β-ME using *E. coli* total RNA from cells carrying pArgW (0.5 mg mL-1) as a substrate. The titration reactions were initiated by addition of purified toxins to the final concentration of 1 mM either with or without 1 µM EF-Tu, followed by incubation for 30 minutes at 37 °C. Other *in vitro* reactions were also conducted at 37 °C for 30 minutes by addition of 0.6 µM of toxins (either CdiA-CT or colicin D nuclease domains). Where indicated, reactions were supplemented with 1.2 µM purified immunity protein. All reactions were quenched with denaturing gel-loading buffer and run on 50% urea-10% polyacrylamide gels buffered with 1 x Tris-borate EDTA. Gels were stained with ethidium bromide or electroblotted to positively charged nylon membranes for subsequent Northern blot analyses.

RNA isolation and analyses

Frozen cell pellets were resuspended in guanidinium isothiocyanate (GITC)-phenol and total RNA extracted as described (Garza-Sánchez *et al.*, 2006). RNAs (5 µg) were run on 50% urea-10% polyacrylamide gels buffered with 1 x Tris-borate EDTA and electroblotted to positively charged nylon membranes for Northern blot analysis. Blots were hybridized

with [^{32}P]-labeled oligonucleotide probes that are specific for individual *E. coli* tRNAs (Garza-Sánchez *et al.*, 2006, Hayes & Sauer, 2003). Blots were visualized by phosphorimaging using Bio-Rad Quantity One software.
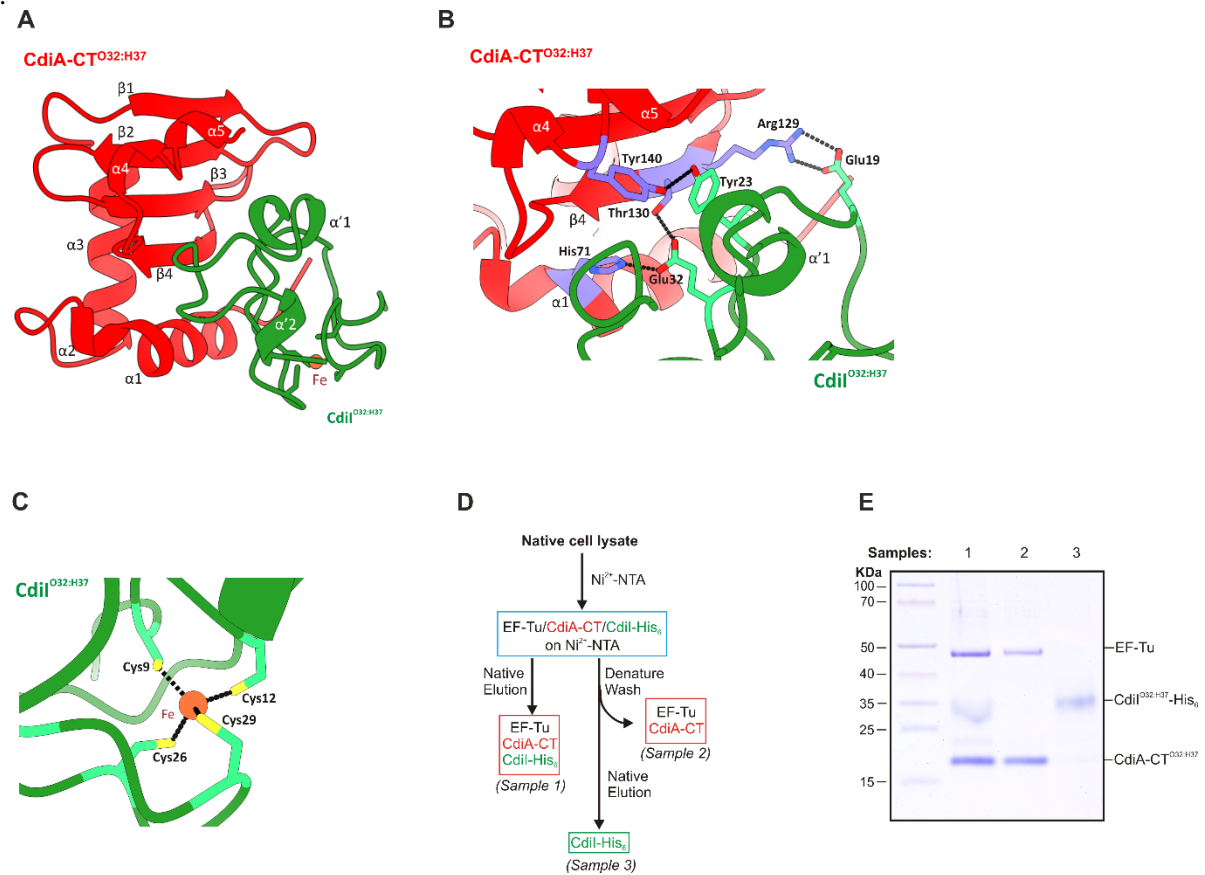
Figure 1: CdiA-CT/CdiI$^{O32:H37}$ structure and purification.

A. The X-ray crystal structure of the CdiA-CT/CdiI$^{O32:H37}$ complex was acquired by Dr. Karolina M. Michalska at the Argonne National Laboratory. Within this structure, the crystal representation of the CdiA-CT (depicted in red) with its catalytic active site completely covered by its corresponding immunity protein, CdiI (presented in green). Importantly, CdiI is found to be intricately associated with an iron atom.

B. The toxin-immunity protein interface is depicted with selected side-chains forming direct hydrogen bonds (black dashed lines) shown in a stick representation.

C. Immunity protein of O32:H37 co-purified with its iron co-factor, secured by four cysteines residues.

D. Illustration of the purification workflow to isolate CdiA-CT/CdiI$^{O32:H37}$ under both native and denatured conditions. Sample 1 was procured through native elution of the complex off Ni^{2+}-NTA agarose. This elution captured three protein species, CdiA-CT/CdiI$^{O32:H37}$/EF-Tu. Sample 2 was derived from a denaturing wash aimed at breaking CdiA-CT$^{O32:H37}$/EF-Tu interaction from CdiI$^{O32:H327}$. Finally, sample 3 resulted from a concluding native elution step, obtained by removing CdiI$^{O32:H37}$ from Ni^{2+}-NTA agarose following denaturing wash.

E. Samples 1-3 were resolved on a SDS-PAGE gel and stained with Coomassie Blue. MW markers are indicated on the left.
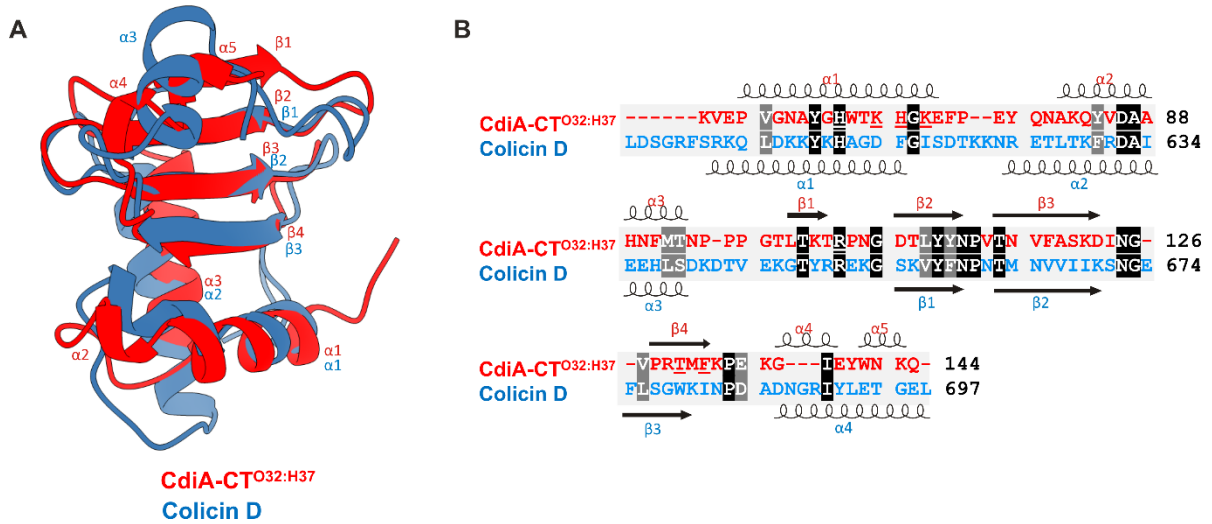
Figure 2: **CdiA-CT^O32:H37 shares structural homology with colicin D.**
A. Superimposed structures of CdiA-CT^O32:H37 (red) and colicin D (PDB_ID IV74, blue).
B. Sequence alignment of CdiA-CT^O32:H37 and colicin D nuclease domain with conserved residues labeled in black. Note the lack of sequence similarity in contrast to substantial structural similarity.
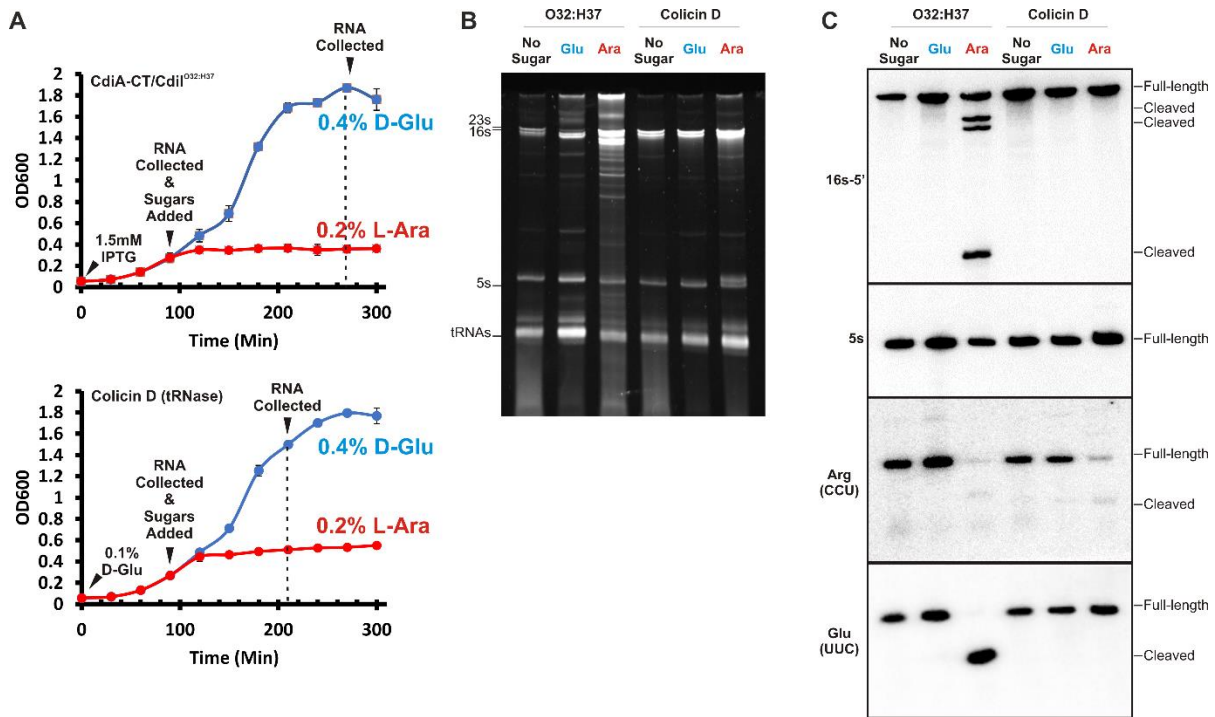
Figure 3: RNase activity of CdiA-CT$^{O32:H37}$.
A. The growth curves present a comparison of growth behaviors of *E. coli* cells harboring pCdiA-CT$^{O32:H37}$/pCdiI$^{O32:H37}$ (top) and *E. coli* containing pColD (nuclease domain; bottom). These comparisons were made under the conditions of D-glucose repression and L-arabinose induction. In the O32:H37 system, 1.5mM IPTG was introduced at the start of the growth curves to induce immunity. In the case of the colicin D system, 0.1% D-glucose was also included at time 0 to prevent leaky expression of the toxin.
B. The RNA samples obtained from the growth curves in panel A for both the O32:H37 and colicin D systems are separated using a urea-acrylamide gel. Additionally, RNA samples without added sugar were included as a quality control measure.
C. Northern blots of the identical RNA samples shown in panel B were incubated with the inidicated radiolabeled probes.
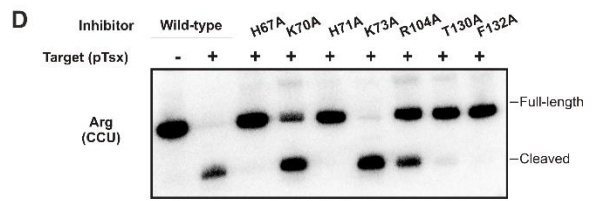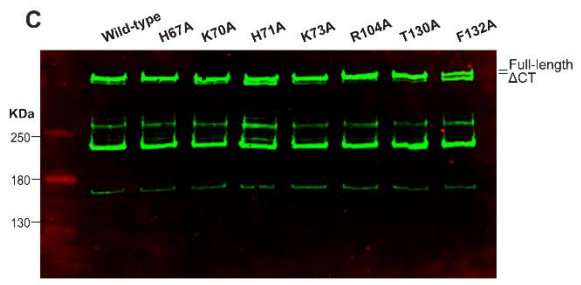
A

β1

Arg104

β2
β1

Phe132
Trp679

β3
β2

Thr130
Ser677

β4
β3

His71
His611

His67
Lys608

α1
α1

CdiA-CT^O32:H37

Colicin D

Lys73

B

Inhibitor	Wild-type		H67A	K70A	H71A	K73A	R104A	T130A	F132A
Target (pTsx)	−	+	+	+	+	+	+	+	+

competitive index

10^1

10^-1

10^-3

10^-5

C

Wild-type H67A K70A H71A K73A R104A T130A F132A

=Full-length
=ΔCT

KDa

250

180

130

D

Inhibitor	Wild-type		H67A	K70A	H71A	K73A	R104A	T130A	F132A
Target (pTsx)	−	+	+	+	+	+	+	+	+

Arg
(CCU)

Full-length

Cleaved

E

KDa   Wild-type H67A K70A H71A K73A R104A T130A F132A

100
70

50
40

35

25

EF-Tu

His₆-CdiA-CT^O32:H37

F

0.6µM CdiA-CT^O32:H37   Wild-type   H67A K70A H71A K73A R104A T130A F132A
1.2µM CdiI^O32:H37   −   +   −   −   −   −   −   −   −

Arg
(CCU)
Full-length
Cleaved

Glu
(UUC)
Full-length
Cleaved

5s
Full-length
Cleaved

73

Figure 4: Mutagenic analyses of the CdiA-CT$^{O32:H37}$ active site.
A. The highlighted catalytic residues of CdiA-CT$^{O32:H37}$ predicted to facilitate nuclease activity are located on the α1-helix and β4-strand. These are superimposed with the colicin D nuclease domain and its homologous residues found on α1-helix and β3-strand.
B. *E. coli* inhibitor cells that deploy the indicated CdiA-CT$^{O32:H37}$ variants were cultured at 1:1 ratio with *E. coli* target cells expressing pArgW. Viable inhibitor and target bacteria were enumerated as colony forming units (cfu) at t = 0 and t = 15 min, and the competitive index calculated. Competitive indices for three independent experiments are reported together with mean ± standard error.
C. Anit-CdiA immunoblot for the inhibitor cells used in panel B.
D. Northern blot analysis of tRNA^{Arg} isolated from the competition co-cultures in panel B.
E. Purified CdiA-CT$^{O32:H37}$ variants under native condition revealing the binding of EF-Tu with toxin.
F. Northern blot analysis of tRNA^{Arg}, tRNA^{Glu}, and 5s rRNA from *in vitro* nuclease activity of denatured purified CdiA-CT$^{O32:H37}$ and native purified EF-Tu with extracted *E. coli* RNA for 30 min.
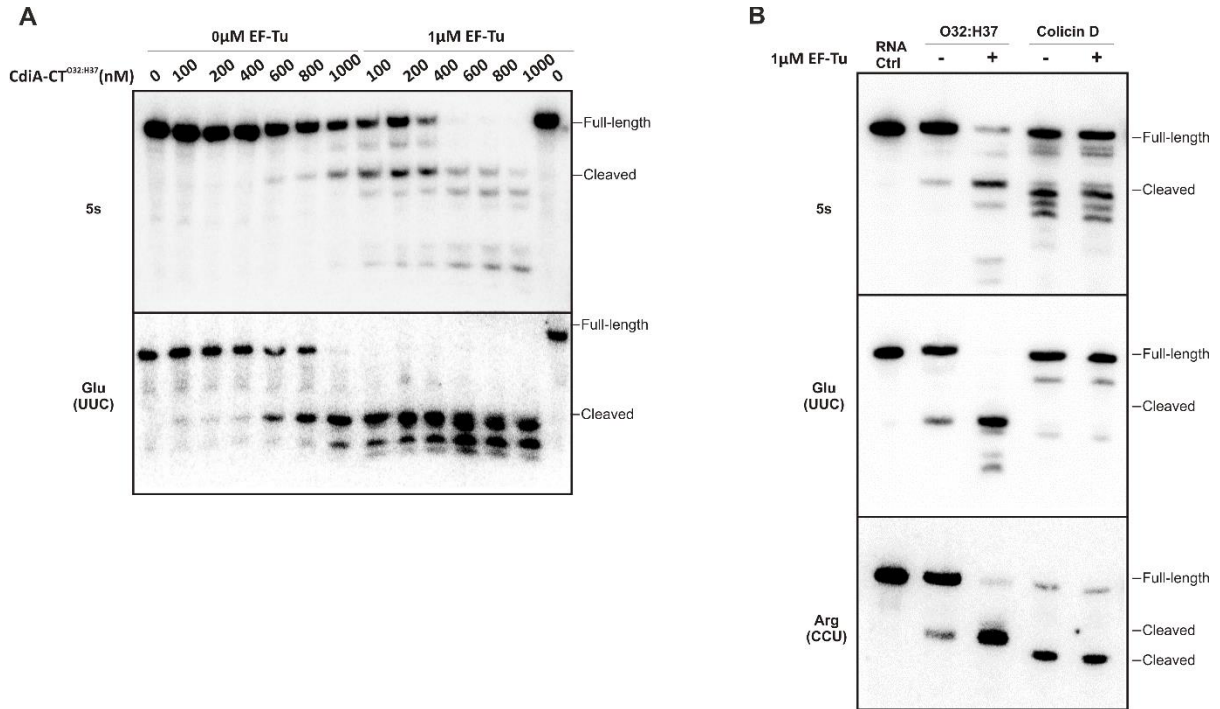
Figure 5: EF-Tu enhances the CdiA-CT$^{O32:H37}$ nuclease activity.
A. Northern blot analysis of 5s rRNA and tRNA^{Glu} from a dosage determined *in vitro* nuclease assays. The *in vitro* reactions were performed by titrating different concentrations of CdiA-CT$^{O32:H37}$ with a constant concentration of EF-Tu.
B. Northern blot analysis of 5s rRNA, tRNA^{Glu}, and tRNA^{Arg} from *in vitro* nuclease assays to determine that EF-Tu is only required for CdiA-CT$^{O32:H37}$ for enhancing nuclease activity.
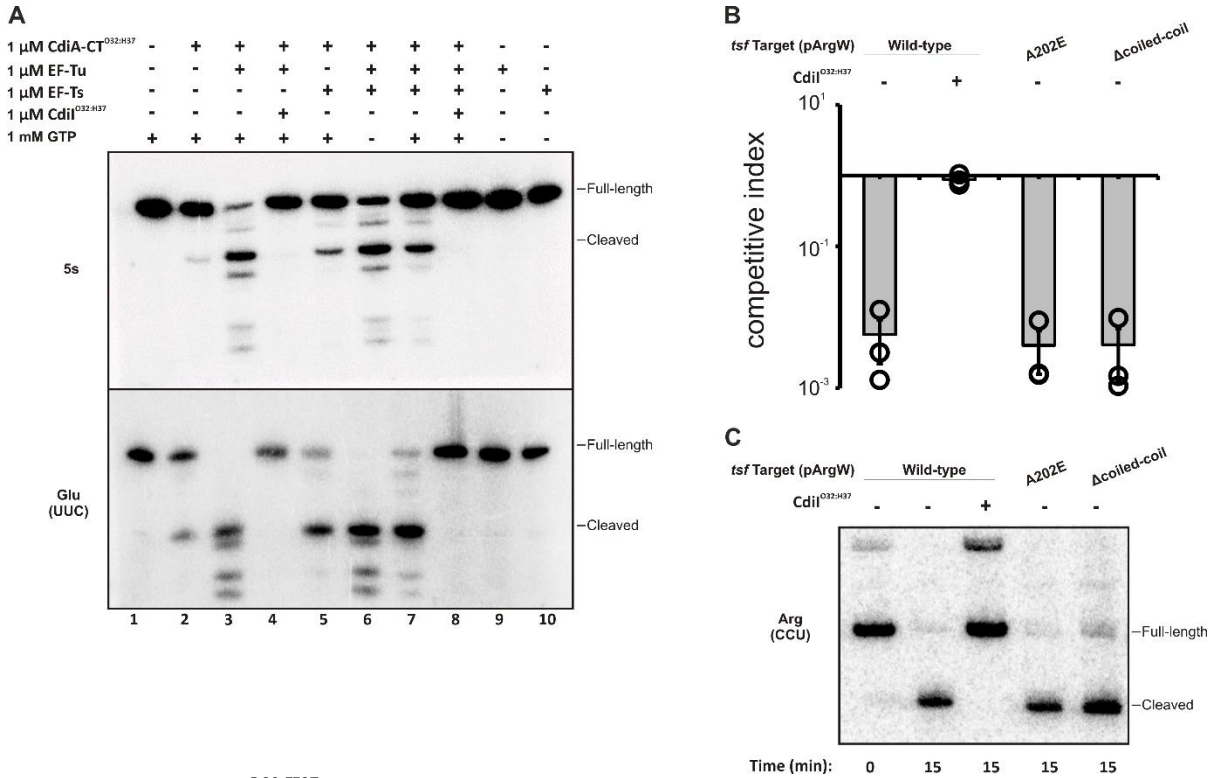
Figure 6: CdiA-CT^(O32:H37) does not require EF-Ts for nuclease activity.
A. Northern blot analysis of 5s rRNA and tRNA^(Glu) from *in vitro* nuclease assays. The reactions were added with indicated components, including CdiA-CT^(O32:H37), CdiI^(O32:H37), EF-Tu, EF-Ts, and GTP.
B. *E. coli* inhibitors that deploy wild-tu[eCdiA-CT^(O32:H37) were cultured at 1:1 ratio with wild-type or mutant variant (Ala202Glu and Δcoiled-coli) of *tsf* target cells with pArgW. Viable inhibitor and target bacteria were enumerated as colony forming units (cfu) at t = 0 and t = 15 min, and the competitive index calculated. Competitive indices for three independent experiments are reported together with mean ± standard error.
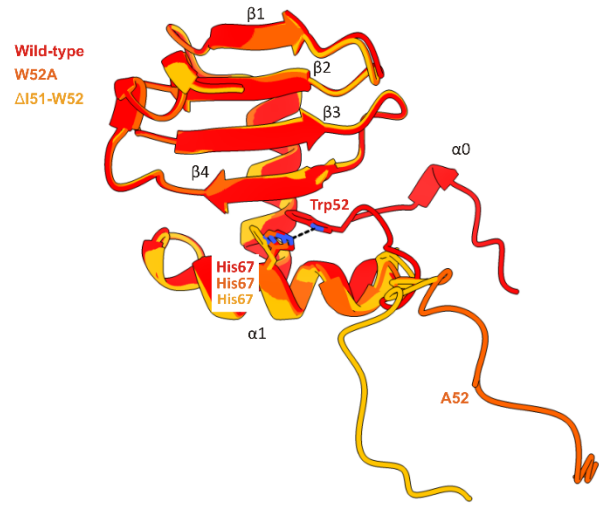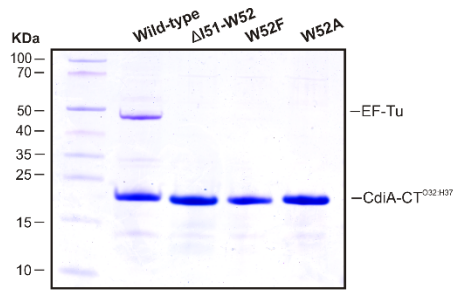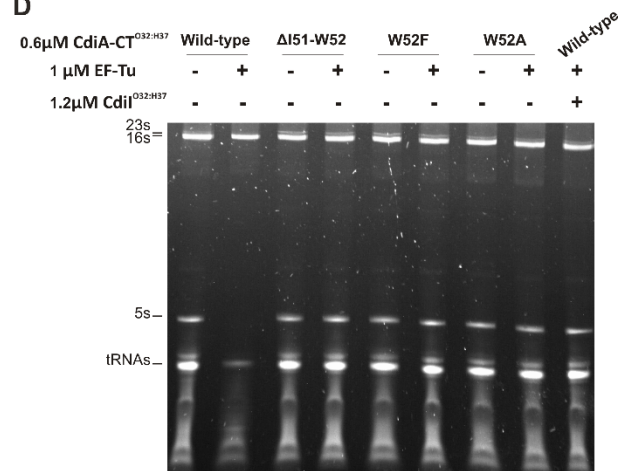C. Northern blot analysis of tRNA^(Arg) isolated from the competition co-cultures in panel B.

A

Wild-type
W52F
H67A

β1
β2
β3
β4
α0

Trp52
Phe52
Trp52

His67
His67
Ala67

α1

B

Wild-type
W52A
ΔI51-W52

β1
β2
β3
β4
α0

Trp52

His67
His67
His67

α1

A52

C

KDa	Wild-type	ΔI51-W52	W52F	W52A
100 —				
70 —				
50 —				—EF-Tu
40 —				
35 —				
25 —				
				—CdiA-CT$^{O32:H37}$
15 —				
10 —				

D

0.6μM CdiA-CT$^{O32:H37}$	Wild-type		ΔI51-W52		W52F		W52A		Wild-type
1 μM EF-Tu	-	+	-	+	-	+	-	+	+
1.2μM CdiI$^{O32:H37}$	-	-	-	-	-	-	-	-	+

23s
16s

5s

tRNAs

E

Inhibitor	Wild-type		ΔI51-W52	W52F	W52A
Target (pTsx)	-	+	+	+	+

competitive index

10^1

10^{-1}

10^{-3}

10^{-5}

Figure 7: The unique Trp52 and His67 interaction within CdiA-CT$^{O32:H37}$.
A. The superimposed AlphaFold model compares the wild-type (red), Trp52Phe mutant (cyan), and His67A mutant (green), demonstrating a more condensed structure of CdiA-CT$^{O32:H37}$.
B. The superimposed AlphaFold model compares the wild-type (red), Trp52Ala mutant (dark orange), and ΔIle51-Trp52 mutant (orange), revealing a less condensed structure of CdiA-CT$^{O32:H37}$, with the N-terminal sequence unfolded, exposing the Ala52 point mutant, and losing one small α-helix (α0).
C. The SDS-PAGE gel resolved purified wild-type, ΔIle51-Trp52, Trp52Phe, and Trp52Ala variants of CdiA-CT$^{O32:H37}$, showing that these mutant variants lost the ability to bind to EF-Tu.
D. In the *in vitro* nuclease assay resolved RNA with a urea-acrylamide gel, EF-Tu was supplemented to the reaction with the indicated toxin variants. Only wild-type could target the RNAs.
E. *E. coli* inhibitors that deploy the indicated CdiA-CT$^{O32:H37}$ variants were cultured at 1:1 ratio with *E. coli* target cells with pArgW. Viable inhibitor and target bacteria were enumerated as colony forming units (cfu) at t = 0 and t = 15 min, and the competitive index calculated. Competitive indices for two independent experiments are reported together with mean ± standard error.
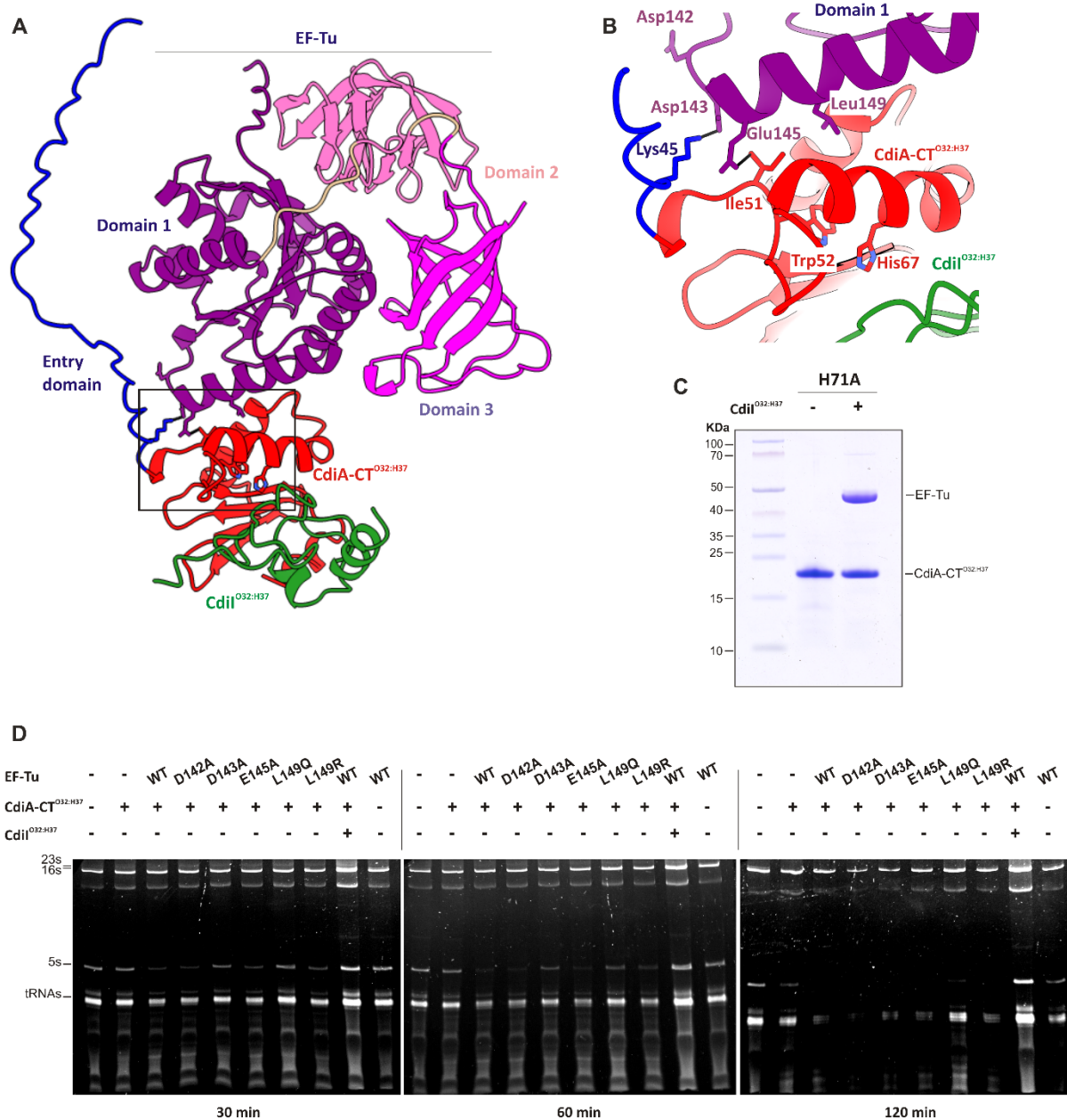
Figure 8: The interaction between EF-Tu and CdiA-CT$^{O32:H37}$.

A. The AlphaFold model of a CdiA-CT/CdiI$^{O32:H37}$/EF-Tu complex with the Cdi-CT entry domain colored in blue, nuclease domain in red, CdiI in green, EF-Tu with three domains in different shades of purple.

B. A closer look at the residues within interface between EF-Tu and CdiA-CT$^{O32:H37}$

C. The SDS-PAGE gel resolved natively purified wild-type and His71Ala variant either with or without the presence of CdiI.

D. A time-course urea-acrylamide gel analysis on the RNA samples incubated with wild-type toxin and indicated EF-Tu variants

79

Table 1: Direct H-bonds and salt-bridges at the CdiA-CT/CdiI$^{O32:H37}$ interface

CdiA-CT atom	CdiI atom	Distance (Å)
LYS 57[N]	ASP 18[OD2]	2.69
VAL 58[N]	ASP 18[OD2]	3.51
ARG 129[NH2]	GLU 19[OE1]	3.15
ARG 129[NH1]	GLU 19[OE2]	2.69
ASN 106[ND2]	CYS 22[SG]	3.65
HIS 67[NE2]	TYR 23[O]	2.78
TYR 140[OH]	TYR 23[OH]	2.84
HIS 71[NE2]	GLU 32[OE1]	2.65
LYS 70[NZ]	GLU 32[OE1]	3.40
THR 130[OG1]	GLU 32[OE2]	2.67
TYR 140[OH]	ASP 34[OD2]	2.52
LYS 143[NZ]	ASP 40[OD2]	2.81
LYS 136[NZ]	SER 44[O]	2.71
LYS 136[NZ]	SER 44[OG]	3.55
GLU 135[N]	GLY 45[O]	3.34
LYS 136[N]	GLY 45[O]	3.14
LYS 70[NZ]	ALA 47[O]	2.91
ASP 108[OD2]	CYS 22[SG]	3.89
GLU 135[OE2]	SER 49[OG]	2.79

Table 2: Direct H-bonds and salt-bridges at the CdiA-CT$^{O32:H37}$/EF-Tu interface

CdiA-CT atom	EF-Tu atom	Distance (Å)
ILE 51[N]	GLU 145[OE2]	2.85
TRP 68[NE1]	GLU 153[OE1]	3.57
ASN 80[ND2]	MET 113[SD]	3.59
ASN 80[ND2]	PRO 112[O]	3.26
ALA 81[N]	GLU 153[OE2]	2.94
LYS 82[NZ]	THR 109[O]	3.05
LYS 42[NZ]	GLU 148[OE2]	3.05
LYS 42[NZ]	GLU 148[OE1]	2.02
LYS 45[NZ]	ASP 143[OD1]	2.46
LYS 45[NZ]	ASP 143[OD2]	2.15
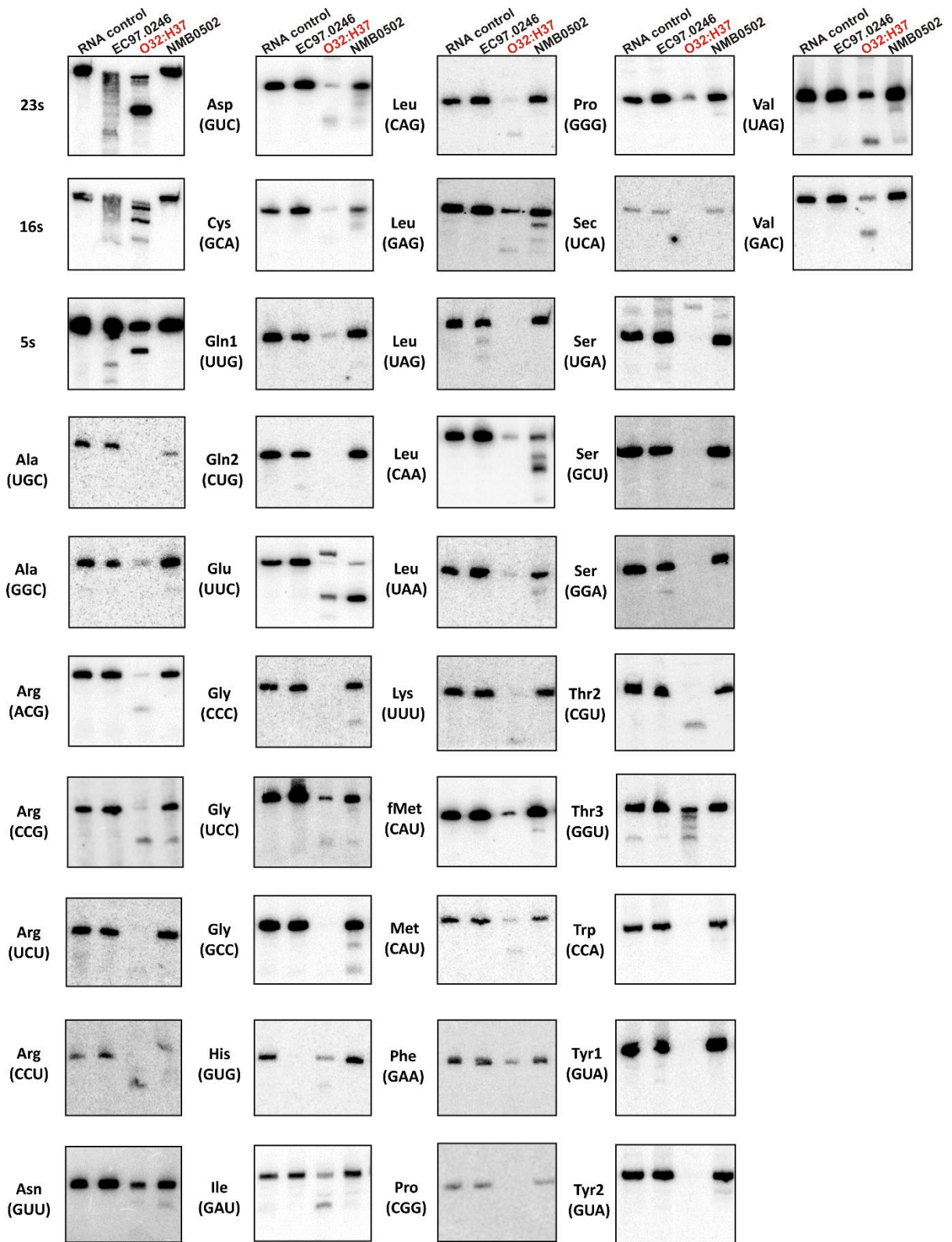
Figure S1: CdiA-CT$^{O32:H37}$ is a general RNase.
Northern blot analyses of RNA isolated from intoxicated *E. coli* cells. The C-terminal toxin domains of from *E. coli* 97.0246, CdiA from *Escherichia coli* O32:H37, and CdiA from *N. meningitidis* NMB0502 were expressed in *E. coli*. Total RNA was isolated for Northern blot hybridization using radiolabeled probes to the indicated rRNAs and tRNAs. Anticodon sequences are shown in parentheses for specific isoacceptors.
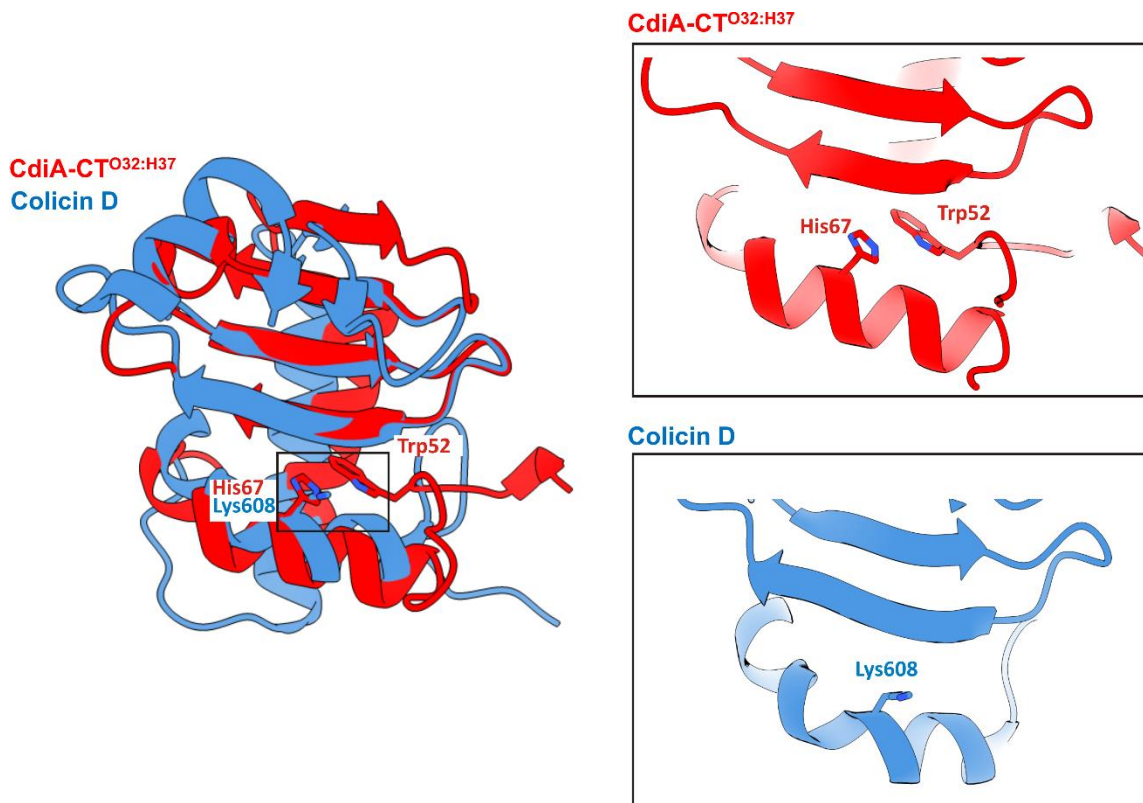
Figure S2: The absence of the Trp52-His67 turn in the colicin D nuclease domain.
Superimposed structures of CdiA-CT$^{O32:H37}$ predicted by AlphaFold (highlighted in red) and colicin D (PDB_ID IV74, shown in blue). In place of His67, colicin D has Lys608, which does not form a bond with another aromatic residue as CdiA-CT$^{O32:H37}$.

Table 3: Bacterial strains and plasmids

Strains	Description[a]	References
X90	F' *lacI^q lac' pro' /ara Δ(lac-pro) nal1 argE(amb) rif'thi-1*, Rif[R]	(Zhang et al., 2012)
CH2016	X90 (DE3) *Δrna ΔslyD*::kan, Kan[R]	(Aoki et al., 2008)
CH7157	X90 *ΔclpX ΔclpA*::kan, Kan[R]	(Poole et al., 2011)
CH14016	MG1655 *Δwzb Δtsx*	(Ruhe et al., 2018)
CH14387	MG1655 ara::spec, Spc[R]	(Jones et al., 2017)
CH14388	MG1655 ara::spec *tsf* A202E, Spc[R]	(Jones et al., 2017)
CH14389	MG1655 ara::spec *tsf* Δcoiled-coil, Spc[R]	(Jones et al., 2017)

Plasmids	Description[a]	References
pCP20	plasmid that shows temperature-sensitive replication and thermal induction of FLP synthesis, Amp[R] and Cm[R]	(Cherepanov & Wackernagel, 1995)
pTrc99a	IPTG-inducible expression plasmid, Amp[R]	GE Healthcare
pTrc99KX	Derivative of pTrc99a with additional 5′ KpnI, and 3′ SpeI and XhoI restriction sites, Amp[R]	(Koskiniemi et al., 2014)
pCH1554	pCH450::Colicin D(V585), Tet[R]	This study
pCH1712	pMCSG58-NMB0502/503-His6, Amp[R]	(Michalska et al., 2018)
pCH1746	pMCSG63::H6-TEV-CdiA-CT(H67A)-cdiI (O32:H37), Amp[R]	This study
pCH1748	pMCSG63::H6-TEV-CdiA-CT(K70A)-cdiI (O32:H37), Amp[R]	This study
pCH1750	pMCSG63::H6-TEV-CdiA-CT(H71A)-cdiI (O32:H37), Amp[R]	This study
pCH1752	pMCSG63::H6-TEV-CdiA-CT(K73A)-cdiI (O32:H37), Amp[R]	This study
pCH1754	pMCSG63::H6-TEV-CdiA-CT(R104A)-cdiI (O32:H37), Amp[R]	This study
pCH1756	pMCSG63::H6-TEV-CdiA-CT(T130A)-cdiI (O32:H37), Amp[R]	This study
pCH1758	pMCSG63::H6-TEV-CdiA-CT(F132A)-cdiI (O32:H37), Amp[R]	This study
pCH2485	pET21b::colD-immD-His6, Amp[R]	This study
pCH2675	pET21b::tufA-His6(*E.coli*), Amp[R]	This study, CSH
pCH2684	pET21b::tufA (D142A)-His6(*E.coli*), Amp[R]	This study, CSH
pCH2685	pET21b::tufA (D143A)-His6(*E.coli*), Amp[R]	This study, CSH
pCH2686	pET21b::tufA (E145A)-His6(*E.coli*), Amp[R]	This study, CSH
pCH2713	pET21b::tufA (L149Q)-His6(*E.coli*), Amp[R]	This study, CSH
pCH2714	pET21b::tufA (L149R)-His6(*E.coli*), Amp[R]	This study, CSH
pCH2816	pET21b::cdiBA STEC- O32:H37 cdiA(CT)-cdiI Chimera, Amp[R]	This study
pCH3781	pCH450/ pTrc99a, Tet[R], Amp[R]	This study
pCH3782	pCH450::CT (O32::H37) / pTrc99aKX::CdiI (O32::H37), Tet[R], Amp[R]	This study
pCH4112	pMCSG63::H6-TEV-O32-CT(ΔI51-W52)-CdiI, Amp[R]	This study
pCH4113	pMCSG63::H6-TEV-O32-CT(W52F)-CdiI, Amp[R]	This study
pCH5506	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (H67A)-cdiI Chimera, Amp[R]	This study
pCH5507	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (K70A)-cdiI Chimera, Amp[R]	This study
pCH5508	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (H71A)-cdiI Chimera, Amp[R]	This study
pCH5509	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (K73A)-cdiI Chimera, Amp[R]	This study
pCH5510	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (R104A)-cdiI Chimera, Amp[R]	This study
pCH5511	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (T130A)-cdiI Chimera, Amp[R]	This study
pCH5512	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (F132A)-cdiI Chimera, Amp[R]	This study
pCH6276	pUC57::cdiA-CT/cdiI(*E. coli* 97.0246), Amp[R]	Argonne National Lab
pCH6281	pUC57::cdiA-CT/cdiI(*E. coli* O32:H37), Amp[R]	Argonne National Lab
pCH6603	pMCSG63::H6-TEV-CdiA-CT(H71A) (O32:H37), Amp[R]	This study
pCH6442	pMCSG63::H6-TEV-O32-CT-I (VENN), Amp[R]	This study
pCH7029	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (W52A)-cdiI Chimera, Amp[R]	This study
pCH7030	pMCSG63::H6-TEV-O32-CT(W52A)-CdiI, Amp[R]	This study
pCH10085	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (ΔI51-W52)-cdiI Chimera, Amp[R]	This study
pCH10086	pET21b::cdiBA STEC- O32:H37 cdiA(CT) (W52F)-cdiI Chimera, Amp[R]	This study
pCH12602	pSH21P::trxA-TEV-tsf, Amp[R]	(Jones et al., 2017)
pCH12603	pSH21P::trxA-TEV-tufA, Amp[R]	(Jones et al., 2017)
pCH12716	pET21P::O32:H37 CT-imm-H6, Amp[R]	This study
pCH13458	pCH450::cdiA-CT-cdiI (EC97.0246)-DAS, Tet[R]	This study
pCH13459	pCH450::cdiA-CT-cdiI (O32:H37)-DAS, Tet[R]	This study
pCH13527	pCH450::NMB0502-NMB0503-DAS, Tet[R]	This study
pCH15124	pZS21/pArgW, Kan[R], Tet[R]	This study
pCH15125	pZS21:tsx/pArgW Kan[R], Tet[R]	This study

Table 4: Oligonucleotides

Identifier	Descriptive name	Sequence	References
CH367	Gly-GCC-probe	5´-CTT GGC AAG GTC GTG CT-3'	Chapter II
CH368	Pro-GGG-probe	5´-CAC CCC ATG ACG GTG CG-3'	Chapter II
CH374	Ala-UGC-probe	5´-TCC TGC GTG CAA AGC AG-3'	Chapter II
CH379	Arg-CCU-probe	5´-CCT GCA ATT AGC CCT TAG G-3'	Chapter II
CH380	Gly-UCC-probe	5´-CCC GCA TCA TCA GCT TGG AAG GC-3'	Chapter II
CH381	Leu-CAA-probe	5´-CCC GCA CAG CGC GAA CGC CG-3'	Chapter II
CH382	Phe-GAA-probe	5´-TGC TCT ACC GAC TGA GCT A-3'	Chapter II
CH383	Pro-CGG-probe	5´- CTT CGT CCC GAA CGA AGT G-3'	Chapter II
CH385	Thr-CGU-probe	5´-CCT ACG ACC TTC GCA TT-3'	Chapter II
CH406	5S probe	5´-ATG CCT GGC AGT TCC CTA CTC TC-3'	Chapter II
CH407	16S probe	5'-TGC GCT TTA CGC CCA GTA ATT CC-3'	Chapter II
CH408	23S probe	5´-GTT TAG CCC CGT TAC ATC TTC CG-3'	Chapter II
CH436	Arg-CCG-probe	5´-CCT GAG ACC TCT GCC TCC GGA-3'	Chapter II
CH451	Arg-UCU-probe	5´-CCT GCG GCC CAC GAC TTA G-3'	Chapter II
CH452	Arg-ICG-probe	5´-CCT CCG ACC GCT CGG TTC G-3'	Chapter II
CH487	Cys-GCA-probe	5´-GGA CTA GAC GGA TTT GCA A-3'	Chapter II
CH560	immD-Xho	5' - TGG ACT CGA GTA ATT TAA ATT TTT CCA AG-3'	This study
CH577	Ile-GAU-probe	5´-ACC GAC CTC ACC CTT ATC AG-3'	Chapter II
CH618	Trp-CCA-probe	5´-CCC AAC ACC GGG TTT TGG-3'	Chapter II
CH714	colD(M590)-Nde	5' - GAA GCT GTT TAT CAT ATG CTT AAT GAT CCT C-3'	This study
CH791	Gln-CUG-probe	5´-TCG GAA TGC CGG AAT CAG A-3'	Chapter II
CH798	Tyr1-GUA-probe	5´-CTT CGA GTC TG TGA CGG CAG-3'	Chapter II
CH799	Tyr2-GUA-probe	5´-CTT CGA GTC GA TGA CGG CAG-3'	Chapter II
CH800	Asp-QUC-probe	5´-TCG AAC CCG CGA CCC CCT GCG-3'	Chapter II
CH801	Asn-QUU probe	5´-CTC GAA CCA GTG ACA TAC GG-3'	Chapter II
CH837	Ser-GCU-probe	5´-CCC CGG ATG CAG CTT TTG ACC-3'	Chapter II
CH1046	Leu-GAG-probe	5´-CCC GTA AGC CCT ATT GGG CA-3'	Chapter II
CH1047	Ser-UGA-probe	5´-AAC CCT TTC GGG TCG CCG GTT TTC-3'	Chapter II
CH1248	Val-UAG-probe	5´-CGC CGA CCC CCT CCT TGT AAG-3'	Chapter II
CH1249	Lys-UUU-probe	5´- CCT GCG ACC AAT TGA TTA AA-3'	Chapter II
CH1417	Glu-UUC-probe	5´- CCC CTG TTA CCG CCG TG-3'	Chapter II
CH1706	BC0921-Nhe-rev	5´-CTC TAT GCT AGC TTT CAA TTC TTT ATT TTT TCC-3'	Chapter II
CH2032	Gly-CCC-probe	5´-CCC TCG TAT AGA GCT TGG GAA-3'	Chapter II
CH2034	Leu-CAG-probe	5 -CCC CCA CGT CCG TAA GGA CA-3'	Chapter II
CH2035	Leu-UAG-probe	5´ - CAC CTT GCG GCG CCA GAA-3;	Chapter II
CH2036	Leu-UAA-probe	5´ - CCC GCA CAG CGC GAA CGC CG-3'	Chapter II
CH2037	fMet-CAU-probe	5´ - CGG GTT ATG AGC CCG ACG A-3'	Chapter II
CH2038	Met-CAU-probe	5´ - CCT GTG ACC CCA TCA TTA TGA-3'	Chapter II
CH2040	Ser-CGA-probe	5´ - GTA GAG TTG CCC CTA CTC CGG-3'	Chapter II
CH2042	Thr-GGU-probe	5´ - CTG GGG ACC CCA CCC CT-3'	Chapter II
CH2785	NMB0503-Spe-rev	5´ - TTC ACT AGT GGT TTC ATG CAG GCT AC-3'	Chapter II
CH2839	NMB0502-Nco-for	5´ - CTC CCA TGG TGA AAA ATA ATC AGC-3'	Chapter II
CH3534	EC97.0246-CT-Nco-for	5' - TTT CCA TGG CTG AAA ACA ATA TGC - 3'	This study
CH3535	EC97.0246-cdiI-Spe-rev	5' - TTT ACT AGT GGC AAG TAG CTC TAA TT - 3'	This study
CH3603	O32:H37-cdiI-Kpn-for	5' - AAA GGT ACC ATG AAT AAT GGT TC - 3'	This study
CH3896	O32:H37-CT-Kpn/Nco-for	5' - TTT GGT ACC ATG GTT GAG AAT AAT GCG C - 3'	This study
CH3897	O32:H37-H71A-rev	5' - GGA AAC TCT TTT CCA GCT TTT GTC CAA TGC C - 3'	This study
CH3898	O32:H37-cdiI-Spe-rev	5' - TTT ACT AGT CTG TTC GTT AAA TGC TCG - 3'	This study
CH4003	O32-CT-Xho-rev	5' - TTT CTC GAG TCA TTG CTT GTT CCA ATA TTC G - 3'	This study
CH4262	EC93-A2894-Nhe-for	5' - AAA GCT AGC GCC GGT ACG GGG GC - 3'	This study
CH4716	O32-cdiI-Xho-rev	5'- TTT CTC GAG TTA CTG TTC GTT AAA TGC TCG TTT C-3'	This study
CH4748	Ec-tufA-Xba-for	5' - CGG TTC TAG ACA TGG GTA TTC GTC TGC AC - 3'	This study
CH4749	Ec-tufA-Xho-rev	5' - ACA GCT CGA GTA AGA TAT GCC GTC AAC AAA TGC-3'	This study
CH4794	O32:H37-H67A-rev	5'- ATG TTT TGT CCA AGC CCC ATA GGC ATT-3'	This study
CH4795	O32:H37-K70A-rev	5'- CTC TTT TCC ATG TGC TGT CCA ATG CCC ATA-3'	This study
CH4796	O32:H37-K73A-rev	5'- CTC TGG AAA CTC TGC TCC ATG TTT TGT CCA-3'	This study
CH4797	O32:H37-R104A-rev	5'- GTC TCC GTT AGG GGC GGT TTT AGT TAA-3'	This study

CH4798	O32:H37-T130A-rev	5'- CTC TGG TTT AAA CAT TGC ACG CGG TAC ACC-3'	This study
CH4799	O32:H37-F132A-rev	5'- TTT CTC TGG TTT AGC CAT TGT ACG CGG TAC-3'	This study
CH5152	O32H37-Kpn-TEV-VENN-for	5'- TTT GGT ACC GAG AAC CTG TAC TTC CAA GTT GAG AAT AAT GCG CTG GG-3'	This study
CH5283	colD-V585-Kpn/Nco-for	5'- GCT GGT ACC ATG GGT GAA GCT GTT TAT GTG ATG-3'	This study
CH5824	colD-Xho-rev	5'- CAA CTC GAG CAT TGC CAT CTT ATT CAT AG-3'	This study
CH6114	tufA-E145A-Sac-rev	5' - CAT TTC AAC GAG CTC CAG CAG CGC TTC GTC ATC-3'	This study
CH6115	tufA-D142A-Sac-rev	5' - AAC GAG CTC CAG CAG CTC TTC GTC AGC AAC CAT GTC-3'	This study
CH6116	tufA-D143A-Sac-rev	5' - AAC GAG CTC CAG CAG CTC TTC GGC ATC AAC CAT-3'	This study
CH6209	O32-W52F-rev	5'- CTT GGT TTC TGT AAA AAT TTG AGT ATC TTT-3'	This study
CH6210	O32-del-IW-rev	5'- ACT TTC TTG GTT TCT GTT TGA GTA TCT TT-3'	This study
CH6211	O32-W52A-rev	5'- CTT GGT TTC TGT CGC AAT TTG AGT ATC-3'	This study
CH6224	tufA-L149Q-rev	5' - CTT CCA TTT CAA CCT GTT CCA GCA GCT CTT C-3'	This study
CH6226	tufA-L149R-rev	5' - CTT CCA TTT CAA CCC GTT CCA GCA GCT CTT C-3'	This study

Chapter IV: Exploitation of the Peptide Antibiotic Transporter SbmA by CdiA-CT$^{O32:H37}$ for Translocation

A) Introduction

Gram-negative bacteria, sometimes referred to as diderm bacteria, possess a more complex membrane system compared to their counterparts, Gram-positive or monoderm bacteria. First, their peptidoglycan cell walls are strong enough to withstand approximately 3 atm of turgor pressure (Koch, 1998) and elastic enough to expand several times their initial surface area (Koch & Woeste, 1992). In addition to the robust and elastic cell wall, diderms consist of two layers of membranes, the outer- and inner- membranes that sandwich the cell wall, encapsulating their vital cellular components. To avoid confusion, it is important to note that different Gram-negative could have different cellular components at the most external layer, namely capsules, S-layers, or sheaths (Beveridge, 1981). In most cases, the outer-membrane layer is considered to serve as an external barrier for the bacteria, composed of proteins and lipopolysaccharides (LPSs). This is followed by the inner-membrane layer, which also contains embedded with membrane proteins and forms a boundary between the periplasmic and cytoplasmic spaces. While the design of the diderm structure might be thought of as impenetrable defense in the harsh natural environment, various bacteriocins and secretion systems have managed to exploit pre-existing import and export proteins to transport toxic effectors across the intricate layers of Gram-negative bacteria.

Bacteriocins, such as colicins, are specialized in recognizing outer-membrane proteins in *Escherichia coli*. For instance, in the case of group A colicins, they recognize proteins like OmpA and OmpF, while some group B colicins recognize FepA or BtuB. These bacteriocins bind and then translocate to the periplasmic space and subsequently utilize inner-membrane proteins to enter the cytoplasmic domain. In the case of group A colicins,

they use the Tol systems, whereas group B colicins utilize the TonB-ExbB-ExbD system. Another example found among Gram-negative is the type V secretion system known as contact-dependent growth inhibition (CDI). This system involves the expression of a surface filamentous CdiA protein with the help of its two-partner secretion protein, CdiB. This CdiA protein binds to specific outer-membrane receptor on neighboring target cells. For instance, CdiA[EC93] binds to the BamA protein (Aoki et al., 2008; Ruhe, Wallace, et al., 2013) and CdiA[STEC3] binds to the Tsx protein (Ruhe et al., 2018). Once binding occurs, CdiA delivers a polymorphic toxin derived from its C-terminal domain (CdiA-CT). This CdiA-CT possesses a cytoplasmic-entry domain that is essential for recognizing specific inner-membrane proteins. These inner membrane proteins facilitate the translocation of CdiA-CT to the cytoplasm of the target cell. Each CdiA-CT variant recognizes a specific host cell inner membrane protein: CdiA-CT[MHI813] via MetI, CdiA-CT[TTO1] via GltJ/K, CdiA-CT[NC101]/CdiA-CT[EC3006] via PtsG, CdiA-CT[Dd3937] via RbsC, and CdiA-CT[EC869-o11] via YciB (Willett et al., 2015). Additionally, the d*Burkholderia* CDI system, BcpA-CT employs GltJ/K for its recognition (Myers-Morales et al., 2021).

In this study, we sought to identify the inner membrane protein used by CdiA-CT[O32:H37]. We used genetic approaches and discovered that mutations in an inner-membrane protein, SbmA, confer resistance to CdiA-CT[O32:H37]. Although SbmA lacks a known physiological function, it has been identified as a crucial transporter for a wide range of antibiotics. It was first identified in the genetic screen for the azole-modified antibiotic microcin B17 (MccB17) (Laviña et al., 1986). Subsequently, it was reported to be an importer for other unique antibiotic peptide, such as microcin J25 (MccJ25) (Salomón & Farías, 1995), klebsazolicin (KLB) (Metelev et al., 2017), and azole containing antibiotic

bleomycin (Yorgey et al., 1994). Additionally, SbmA can import translation-inhibiting proline-rich antimicrobial peptides (PrAMPs) (Corbalan et al., 2013; Mattiuzzo et al., 2007). Consistent with this, a proline motif, PPxxPxxPLP, has been identified on the CdiA-CT$^{O32:H37}$ cytoplasmic-entry domain. Site-directed mutagenesis and deletion of this motif in CDI+ cells resulted in loss of target cell inhibition. Collectively, these findings suggest that CdiA-CT$^{O32:H37}$ may utilize its proline repeat for SbmA recognition and cell entry, akin to the mechanism used by PrAMPs. In the future, a bioinformatic analysis can be applied to identify other CDI toxins that share a homologous proline motif for SbmA recognition and CdiA-CT$^{O32:H37}$ can now be used to study the transport mechanism of SbmA.

B) Results

Mutation within the sbmA-yaiW operon confers resistance to CdiA-CT$^{O32:H37}$

CdiA-CT$^{O37:H37}$ requires EF-Tu to facilitate its ribonuclease activity as described in chapter III and yet, mutations in EF-Tu that confer resistant to CdiA-CT$^{O32:H37}$ is challenging to generate since *E. coli* cells carry two copies of *tuf* genes encoded for EF-Tu, *tufA* and *tufB*. To achieve this goal, a copy of *tufB* was removed in *E. coli* cells and sequentially mutagenized with UV-radiation and TnSC189 transposon. The resulting UV-mutagenized and transposon inserted libraires competed against *E. coli* with a plasmid-borne CdiBA^{STEC3}-CT$^{O32:H37}$ chimeric system. Independent UV and transposon mutants surviving the competition were harvested and enriched with sequential competition. Two enrichments were performed for ten transposon libraries and five UV-libraries (**Fig. 1A** & **1B**) and yielded complete CdiA-CT$^{O32:H37}$ resistant mutants. A single colony was isolated from each library for identifying the mutated gene(s).

Rescue cloning of a single transposon library was performed and revealed that the TnSC189 was inserted in the *sbmA* gene within the *sbmA-yaiW* locus rather than the *tufA*

locus. The remaining libraries and pools were screened by PCR with oligos specific to *sbmA,* revealed that all the transposon libraries have an insertion at the *sbmA* gene (**Fig. 1C**). Further PCR analysis was conducted with the downstream gene, *yaiW* from these resistant libraries and pools and showed that *yaiW* is not being mutagenized. Then, DNA sequencing analysis was performed on the PCR products (shown in **Fig. 1C**), which revealed the precise insertions of TnSC189 from transposon libraries and point mutations of UV pools (**Fig. 1D**). In the UV mutagenized experiment, the non-irradiated control (Sp) acquired an IS5 element insertion at the *sbmA-yaiW* promoter when it was enriched twice with the inhibitor cells. Although there are many mutations represented from the UV pools, there were only two mutants that can be used for further analysis in the role of *sbmA*, Pool 1 (Gly186Arg) and Pool 4 (Arg363Cys) (**Fig. 1D**).

SbmA is required for toxin translocation to the target cytoplasm

To establish the essential role of SbmA in CdiA-CT$^{O32:H37}$ translocation, a comprehensive knockout of SbmA (ΔsbmA) was introduced in the wild-type *E. coli* target cells. This knockout rendered the cells resistant to the toxic effects of CdiA-CT$^{O32:H37}$ in a competition assay. Moreover, the susceptibility of these target cells to the toxin was effectively restored upon the introduction of a plasmid borne SbmA (pSbmA). Indeed, cells harboring pSbmA exhibit twofold increase in susceptibility to the toxin compared to the wild-type background (**Fig. 2A**). Despite our efforts with the CDI-resistant libraries and pools, no mutations in the *yaiW* gene were identified. As a next step, we opted to perform a *yaiW* knockout and assess its impact on toxin susceptibility. In the co-culture assay, it was observed that the target cells lacking *yaiW* (ΔyaiW) remained vulnerable to the toxin. Intriguingly, the toxin exhibited an even greater lethality towards this population compared to the wild-type, as evidenced by a lower competitive index (**Fig. 2A**).

To validate the specific involvement of SbmA in toxin translocation exclusively within the target cells, we conducted an experiment involving the expression of the toxin within Δ*sbmA E. coli* cells, as detailed in chapter III. In a concise summary, the inclusion of pCdiI was crucial to ensure the basal expression of pCdiA-CT. Notably, the pCdiA-CT plasmid's activity was repressed by glucose, permitting Δ*sbmA E. coli* cells able to sustain comparable growth rates to those observed in the empty vector controls (**Fig. 2B**). However, distinct behavior was evident upon induction of the pCdiA-CT plasmid using arabinose. Under this condition, the growth of Δ*sbmA E. coli* cells experienced a complete halt, mirroring the growth inhibition observed when pCdiA-CT was expressed in the wild-type background (**Fig. 2B**) (as discussed in Chapter III). This observation strongly suggests that SbmA plays a pivotal role exclusively within the target cells, specifically facilitating toxin translocation.

Important SbmA residues for toxin translocation
The import mechanism of SbmA has undergone extensive investigation, leading to the identification of crucial residues (Corbalan et al., 2013; Ghilarov et al., 2021). In our study, we evaluated these significant residues and probed whether the toxin employs a mechanism akin to that of other microcins for import. The AlphaFold models predictions eluciate the dimer configuration of SbmA in the inner membrane, whereas the monomer configuration is to unveil the key residues of interest, namely Trp52, Phe60, Val102. Tyr116, Gly186, Phe219, Glu276, Asn308, Arg363 (**Fig. 3A**). To assess the involvement of these residues, we tested *E. coli* cells lacking *sbmA* (Δ*sbmA*), which were then complemented with a plasmid-borne *sbmA* in the wild-type and mutants forms. Our experiment encompassed nine mutations, Trp52Gly, Phe60Gly, Val102Gly, Tyr116Gly, Gly186Arg, Phe219Gly, Glu276Gly, Asn308Gly, and Arg363Cys, in the co-culture assays. Intriguingly, solely

mutations in Gly186, Glu276, and Arg363 exhibited the capacity to entirely hinder toxin translocation, substantiated by competitive indices analogous to the empty vector control. Conversely, mutations in Trp53, Phe60, Tyr116, and Asn308 facilitated the restoration of SbmA function, permitting toxin translocation with approximately five-fold inhibition as seen in the wild-type SbmA control (**Fig. 3B**). Notably, the Val102 mutation conferred partial susceptibility to the toxin translocation, failing to produce a fully resistant profile (**Fig. 3B**).

To validate our findings, we examined the expression levels of both wild-type and mutant SbmA proteins via immunoblotting the total cell lysates with anti-SbmA antiserum and assessed their susceptibility to zeocin by monitoring their growth rates. All mutant protein exhibited identical expression levels compared to the wild-type SbmA control. Furthermore, the absence of signal from the empty vector further underscores the high specificity of the antiserum towards SbmA (**Fig. 3C**).

While our ΔsbmA cell line expressing pSbmA had a slower growth rate than the one expressing empty vector, cells with wild-type and mutant variants of SbmA were still able to uptake zeocin in their growth cycles, resulting in the inhibition of bacterial growth (**Fig. 4**). This suggests that the mutations introduced to SbmA only prevented the translocation of CdiA-CT$^{O32:H37}$ without affecting other SbmA-dependent small molecules passing through SbmA. Importantly, our immunoblot analysis establishes that the variation in toxin translocation efficacy does not arise from differences in protein expression levels but rather from the intrinsic nature of the residues themselves.

To further understand the import of toxin via translocation, we delved into the protein sequence of the cytoplasmic domain of CdiA-CT$^{O32:H37}$. We found that the N-terminus of CdiA-CT$^{O32:H37}$, features an arrangement of six consecutive proline residues. This peculiar proline arrangement imparts a unique conformational kink at the entry domain, closely resembling the analogous conformation observed in microcin B17 (MccB17) (Arnison et al., 2013) (**Fig. 5A**). Remarkably, the five prolines follow almost in a contiguous order (Pro14, Pro15, Pro18, Pro21, and Pro23), while the sixth proline (Pro33) is separated by a span of ten residues (**Fig. 5B**). To investigate the implications of these proline residues, we devised a variety of mutations. This involved both single-point mutations at each proline position (Pro14Ala, Pro15Ala, Pro18Ala, Pro23Ala, and Pro33Ala) and more substantial replacements or deletions (Pro18Ala/Pro33Ala, Penta PA, Hexa PA, and ΔPro), outlined in **Fig. 5B** & **5C**. Furthermore, to establish a comparative control, we introduced a set of four alanine replacements for the glycine residues (Gly36-Gly39), the GA mutant.

The co-cultured assays involved combining inhibitor cells carrying mutated entry domains mixed with the target cells with the cognate receptor, Tsx. Among these assays, the single point mutants (Pro14Ala, Pro15Ala, Pro18Ala, Pro23Ala, and Pro33Ala), as well as the Pro18Ala/Pro33Ala double mutant, demonstrated remarkable potency in eradicating the target cells as wild-type. Conversely, the ΔPro mutant displayed a complete lack of activity. Notably, the Penta PA, Hexa Pa, and GA mutants exhibited decreased cytotoxicity towards the target cells compared to the wild-type toxin (**Fig. 5C**). An additional internal expression experiment was conducted to confirm that the observed inactivity of Penta PA, Hexa PA, and ΔPro toxins was not attributed to mutations in the RNase domain. To achieve this, the inducible toxin plasmids of wild-type and mutant variants exhibited cytotoxicity with the

presence of arabinose as evidence of the stall of growth after 5 hours. It is worth noting that the solely viable culture emerged when *E. coli* was cultivated without the inclusion of toxin plasmids (**Fig. 5D**).

Penta PA, Hexa PA, and ΔPro toxins were purified as described in chapter III to characterize the structural integrity. The purified ΔPro protein displayed an anticipated phenotype, closely resembling the wild-type protein but with a slightly accelerated mobility, attributed to the truncation of amino acids Pro14-Pro23. However, the purified Penta PA and Hexa PA variants exhibited a similar profile, featuring an additional degraded band of smaller size than the anticipated molecular weight (**Fig. 5E**). It is noteworthy that despite these structural variations, all purified toxins consistently retained the ability to bind EF-Tu from the cytoplasm. This compellingly suggests the functional competence of these proteins and supports the conclusion that the loss of toxicity is due to an inability to translocate via interaction with SbmA.

Putative model of CdiA-CT$^{O32:H37}$ and SbmA interaction
The AlphaFold model provided a detailed prediction of CdiA-CT$^{O32:H37}$ translocation through the lumen of SbmA (**Fig. 6A**). All five highly ranked models from the AlphaFold model agree that the entry domain of CdiA-CT$^{O32:H37}$ threads through the lumen of SbmA with similar interactions, while the RNase domain resides either in the periplasmic or cytoplasmic domain (data not shown). The long, non-structured entry domain of CdiA-CT$^{O32:H37}$ enables the flexibility to snake through the dimer and interact with distinct residues from both monomers. As predicted by PDBePISA, the SbmA monomer has four key residues that contact the proline-rich region of CdiA-CT$^{O32:H37}$, including Tyr112, Tyr116, Gln328, and Asn331(**Fig. 6B** & **D**). Similarly, Tyr112', Tyr116', Gln328' and Asn331', from the

other monomer, SbmA', interact with the entry domain at the same region with additional two interactions at the glycine repeats, Gly105' and Val106' (**Fig. 6C** & **D**). Notably, the residues from SbmA do not make direct contact with any proline residues but interact with Ile17, Ile19, Leu22, and Ile24 at the proline-rich hotspot. This observation strongly suggests that the presence of proline residues is important for structural recognition, as described in **Fig. 5A** & **B**.

C) Discussion

Previously, CDI inner-membrane transporters have been identified, but a molecular interaction mechanism has been lacking (Willett et al., 2015). Here, we identify SbmA, an inner-membrane protein that interacts with CdiA-CT$^{O32:H37}$ and propose a mechanism. In-depth molecular studies of microcin and other antimicrobial peptide uptake by SbmA provided us a more detailed insight into the toxin entry domain and its interaction with SbmA. In this study, we report that the *sbmA-yaiW* locus is required for target cells to translocate the CdiA-CT$^{O32:H37}$. Interestingly, although two genes were found within the locus, only mutations in *sbmA* conferred resistance to the toxin, Surprisingly, we found that when *yaiW* was knocked out from the target cells, they became more susceptible to the toxin compared to the wild-type background. YaiW is co-transcribed with *sbmA* under the sbmAp promoter, which is regulated by the CpxR regulator (Raivio et al., 2013). Given that YaiW a surface-exposed outer-membrane lipoprotein (Arnold et al., 2014), we speculate that the removal of such a protein results in outer-membrane destabilization. Paradoxically, this membrane destabilization may induce the expression of the CpxR regulator, resulting in the upregulation of *sbmA* expression (Raivio et al., 2013).

A mechanism called SbmA-like peptide transporter (SLiPT) was previously proposed by Ghilarov et al., 2021, and may provide insights into the translocation of CdiA-CT$^{O32:H37}$. In this model, SbmA remains in an outward-open conformation in its resting state, allowing access to both microcin peptides and protons. The two residues, Tyr116 and Tyr116', serve as the periplasmic gate for translocating the microcin peptides. When binding to the microcin, it is believed that the Tyr116 residues stack against the azole groups of microcin, inducing a conformational change in SbmA to a transient state. As microcin translocates, SbmA undergoes structural rearrangement and shifts to an inward-open conformation. Microcin must pass through the internal gate, formed between Tyr368 and the cytoplasmic gate (Tyr285 and Asp288'). This translocation event is coupled with the uptake of protons, which move through SbmA via a series of conserved Glu residues, known as glutamate ladder, including Glu203, Glu378, Glu193, Glu276, and Glu287. Subsequently, the release of protons into the cytoplasm is also coupled with the release of microcin. After the release of both protons and microcin, SbmA is reset to its resting state.

Our data illustrate that Tyr116 may not play a crucial role for translocating the toxin through periplasmic gate. This is indicated by the fact that the targe cell with Tyr116Gly mutation can still be inhibited by the CdiA-CT$^{O32:H37}$ toxin. Furthermore, we discovered two novel mutations in our mutagenized pools, Gly186Arg and Arg363Cys, which also exhibited a CdiA-CT$^{O32:H37}$ resistance phenotype. Gly186 is located at the end of the cytoplasmic gate, and mutating this residue to a positively charged residue may trap the toxin, preventing its release from the lumen of SbmA. On the other hand, Arg363 resides within the lumen of SbmA, close to the periplasmic gate. Changing this residue to a cysteine may introduce an unexpected disulfide bond, which could disrupt the lumen of SbmA and prevent the toxin

from passing through. Otherwise, CdiA-CT$^{O32:H37}$ utilizes similar residues from SbmA to pass through the inner membrane, including the glutamate ladder, as seen in other reported microcins (Ghilarov et al., 2021).

The AlphaFold model predicted the possible translocating conformation of SbmA dimer and CdiA-CT$^{O32:H37}$with the entry domain threads into the lumen of SbmA with unique kinks. The presence of these kinks is due to the arrangement of proline residues repeats throughout the sequence, which mimics the oxazole and thiazole found in microcins (Arnison et al., 2013). We found that replacing the proline from Pro14-Pro23 to alanine residues lead to protein instability and premature processing. Furthermore, missing residues from Pro14 to Pro23 completely abolish the toxin delivery mechanism. Combining with the AlphaFold model, we conclude that the proline-repeats play a critical role in introducing the unique kinks to allow the Ile17, Ile19, Leu22, and Ile24 to may direct contact with Gln328'/Asn331', Tyr116', and Tyr112' on SbmA dimer, respectively. Collectively, we were able to propose a model for the import pathway of CdiA-CT$^{O32:H37}$ with the reported toxin activity described in Chapter III (**Fig. 7**).

D) Experimental Procedures
Transposon libraries construction and selection for CDI^R mutants
Construction of transposon libraries was performed as described in Willett et al., 2015 with minor modifications. Donor *E. coli* donor cells carrying plasmid pSC189 with *mariner* transposon, were mated into *E. coli* CH10229 recipient cells with conjugation. Donor and recipient cells were grown to mid-log phase in lysogeny broth (LB) medium. Donor cells were supplemented with 150 μg/mL ampicillin and 60 μM of diaminopimelic acid (DAP) and recipients with 33 μg/mL chloramphenicol. Donors and recipients were harvested in the same microcentrifuge tube with the ratio of 2:1 respectively based on OD$_{600}$. Supernatants

from both cultures were removed and the pellet was resuspended in 100 μL of 1 x M9 salts. 20 μL of cell suspension was then spotted on plain LB agar and allowed incubation at 37°C for 5 hours. The cells were then harvested off the plate with sterile swabs and suspended in 2 mL of 1 x M9 salts. Transposon insertion mutants were selected by plating 10-fold serial dilution on LB-agar supplemented with 50 μg/mL of kanamycin.

More than 20, 000 colonies from each overnight transposon library were collected from the agar plates in 1 x M9 salts and inoculated into fresh 25 mL of LB in a 250 mL baffled flask. Transposon inserted libraries were then mixed with CDI+ inhibitor strains at mid-log phase with a 1:1 ratio and cultured for 5 hours with shaking at 37°C. Viable cells from the transposon library were enumerated as CFU/mL on LB agar plate supplemented with 50 μg/mL of kanamycin. The survivors from the first round of co-culture were harvested and subjected to the next round of co-culture until the third round to acquire fully resistant mutants. Individual clones from the resistant mutants were isolated and confirmed their CDI resistance (CDI^R) phenotype by another co-culture assay. The transposon mutants were then transduced into wild-type *E. coli* with P1 phage, and the resulting transductants were once confirmed its CDI^R phenotype with co-culture assay.

Rescue cloning and sbmA identification
The genomic DNA (gDNA) from a transposon linked CDI^R mutant was extracted from the overnight culture with phenol/chloroform and ethanol precipitated. Then 1 μg of gDNA was digested with 1 μL of each AgeI and XmaI restriction endonucleases for 5 hours at 37°C, followed by heat inactivation at 75°C for 5 minutes. Reactions were cooled down to room temperature before adding ATP to the mixture with the final concentration of 1.5 mM along 2 μL of purified T4 ligase and allowed DNA to ligate overnight at 16 °C. The ligated

DNA was electroporated into *E. coli* EPI100 *pir+* (CH2552) and recovered in LB at 37°C for 5 hours. Cells were pelleted and selected on LB agar plates supplemented with 50 μg/mL kanamycin. Selected transformants were grown in LB media with 50 μg/mL kanamycin, followed by plasmid extraction. Isolated plasmids were sequenced by using primer CH2260 to identify the transposon insertion junctions. Once the transposon insertion junctions were identified in *sbmA* from one of the libraries, the remaining resistant libraries were screened by using the oligos upstream and downstream of *sbmA* (CH5000/ CH5001).

UV mutagenized libraries and selection for CDI^R mutants

UV mutagenized libraries construction protocol was performed as described in Jones et al., 2017 with minor modifications. *E. coli* MG1655 cells were grown to mid-log phase, harvested by centrifugation, and resuspended in 0.1 M $MgSO_4$ at the $OD_{600}= 0.4$. The suspension was irradiated at 64 J/m^2 in s Startalinker 1800 on a sterile petri dish. Once the cells were irradiated, all subsequent procedures must be maintained in the dark environment. Irradiated cells were collected by centrifugation and resuspended in LB medium, followed by overnight recovery.

The selection of CDI^R mutants was performed with the same protocol as described in the transposon libraries section. The surviving target cells were isolated for colony PCR by using oligos CH5000/CH5001 for *sbmA* mutants. Furthermore, the downstream gene of *sbmA*, *yaiW* was also being amplified from the CDI^R mutants with oligos CH5002/CH5003 for confirming *sbmA* was the only target.

Plasmid constructions and gene knockout

The wild-type and mutant variants of *sbmA* were amplified with the forward oligo CH4717, reverse oligo CH4718, selected point mutation oligos, CH5101 (Trp53Gly), CH5102 (Phe60Gly), CH5103 (Val102Gly), CH5104 (Tyr116Gly), CH5105 (Phe219Gly),

CH5106 (Glu276Gly), and CH5107 (Asn308Gly). The other two mutant variants, Gly185Arg and Arg363Cys were amplified from the UV resistant mutants. Subsequently, the wild-type and all variants of *sbmA* were cloned into pCH450 with KpnI/XhoI and selected on LB-Tet15. These plasmids were kept in the Δ*sbmA:;Kan* strain (CH6402), which was transduced with P1 phage from Δ*sbmA:;Kan* (JW0368) to the CH15206. Furthermore, the Δ*yaiW::Kan* (JW0369) was also introduced to CH15206 to have the final strain CH6403. Notably, strain CH15206 was derived from CH14910. Initially, Δ*tufB::Kan* (JW3943) was introduced to X90. Subsequently, the Kan cassette was cured with pCP20 before introducing the second gene knockout, Δ*wzb::*Kan (JW2046). Finally, the Kan cassette was again cured with pCP20 to achieve the desired background.

The chimeric CdiA-CT$^{O32:H37}$ fused with CdiA^{STEC3} plasmid generated from Chapter III, pCH2816, was used as a template for introducing the indicated proline/glycine mutants, by using forward oligo CH4262, reverse oligo CH4716, and indicated mutant oligos, CH5036 (Pro14Ala), CH5037 (Pro15Ala), CH5038 (Pro18Ala), CH5039 (Pro21Ala), CH5040 (Pro23Ala), CH5041 (Pro33Ala), CH5038 (PP18,33AA), CH5099 (PentaPA), CH5100 (ΔPro), and CH5604 (GA). All fragments were cloned into pCH2816 with NheI/XhoI and selected on LB-Amp150. Additionally, HexaPA was built by using pCH6414 as a template with oligos CH5041 to introduce the last P33A mutation and later also cloned into pCH2816 with NheI/XhoI.

The site-directed mutants were later amplified with oligos CH3896/CH3898 and cloned into pET21P with NcoI/SpeI, as described in Chapter III, to create the expression constructs with a C-terminal his-tagged immunity protein. These constructs, pCH6444 (ΔPro), pCH6445 (PentaPA), and pCH6446 (HexaPA), were stored in CH2016 strain.

Additionally, other indicated site-directed mutants were also cloned into pCH450, using oligos CH3896/CH4003 to obtain the toxin domain only. The final toxin plasmids were constructed in the presence of pTrcKX::CdiI (O32:H37), as outlined in Chapter III, and were kept in CH6402, including pCH3783 (ΔPro), pCH3784 (PentaPA), and pCH3785 (HexaPA).

In vivo expression and competition co-culture assays
Fresh *E. coli* transformants (wild-type or Δ*sbmA*) with the pCH450::CdiA-CT[O32:H37] (wild-type or entry domain mutants) and pTrc99aKX::CdiI[O32:H37] were selected on LB-agar plates supplemented with 15 μg/mL tetracycline, 150 μg/mL ampicillin, and 0.4% D-glucose. Transformants were harvested off the agar plates and resuspended in 1 x M9 salts. Cells were seeded at OD_{600}= 0.05 in the fresh LB-media supplemented with 15 μg/mL tetracycline, 150 μg/mL ampicillin, and 1.5mM IPTG. Cultures were measured every 30 minutes at OD_{600} and a final concentration of 0.2% L-arabinose was to each culture when OD_{600} reaching 0.2. The growth curve measurement was accomplished when it reached the 5-hour timepoint.

Inhibitor and target cells were grown to mid-log phase and mixed in test tube with 1:1 ratio to a final volume of 2 mL Allowed cells to mix on shaker for 1 hour at 37°C, the cells were spotted on LB agar plate supplemented with 150 μg/mL ampicillin for inhibitor cell selection and LB agar plate supplemented with 15 μg/mL tetracycline for target cell selection. Colonies from both plates were enumerated and calculated as competitive indices (CI). CI is a ratio of target: inhibitor cells (CFU/mL) at the final timepoint divided by the ratio of target: inhibitor cells (CFU/mL) at the initial timepoint.

Protein purifications
All cells carrying pCH12716 (wild-type), pCH6444 (ΔPro), pCH6445 (PentaPA), and pCH6446 (HexaPA) were purified as described in Chapter III. Briefly, all protein expressing *E. coli* cells (CH2016) were grown to OD_{600}=0.6 and induced expression by addition of

isopropyl β-d-1-thiogalactopyranoside (IPTG) to 1.5 mM for 1 hour and stored in respective buffer plus 50% glycerol at -20 °C. All proteins were lysed in native buffer A [20mM Tris-HCl (pH 7.5), 150 mM NaCl, 25mM imidazole] in the presence of 1 mg/mL lysozyme for 30 minutes, then cell lysates were sonicated for two rounds to lyse cells completely before 15-minute centrifugation at 15,000 RPM to remove cell debris. Then, Ni^{2+}-NTA resins were added to the cleared supernatants to bind His6-tagged proteins for 1 hour at 4 °C. All protein bound resins were washed thrice with native buffer A before stripping the toxin domains and EF-Tu from the immunity by denature buffer [6 M Guanidine-HCl, 20 mM Tris-HCl (pH 7.5)]. Finally, all proteins were dialyzed twice for 4 hours in the storage buffer [20mM Tris-HCl (pH 7.5), 150 mM NaCl] and analyzed on SDS-PAGE (10% acrylamide at 110V for 1 hour).

SbmA detection with immunoblot

 E. coli strains carrying plasmids with different point mutations of SbmA were grown to mid-log phase and harvested by centrifugation. Each cell pellet from 1 mL culture was resuspended in 100 μL of urea lysis buffer (8 M urea, 10 mM Tris-HCl (pH 8.0), 150 mM NaCl) supplemented with 1 % SDS. Cells were broken with a freeze-thaw cycle at -80°C. Cell lysates were cleared by centrifugation and normalized with Bradford solution before resolving on SDS-PAGE (10% acrylamide at 110V for 1 hour), and then electrotransferred onto Immunobilon®-P PVDF transfer membrane (0.45 μm pore size) at 17 V (constant) for 1 hour. The membrane was blocked with 4 % non-fat milk in 1 x PBS for 30 minutes at room temperature and incubated overnight with primary anti-SbmA antibody in 0.1% non-fat milk PBS (1:10,000 dilution, Scocchi lab from University of Trieste). On the next day, the membrane was washed thrice with 1 x PBS at room temperature before incubating with 800CW-conjugated goat anti-rabbit IgG (1:25,000 dilution, LICOR) in 1 x PBS. Finally, the

membrane was washed thrice with 1 x PBS prior to visualizing with a LI-COR Odyssey infrared imager.

SbmA mutant functions assessment with zeocin

Overnight strains carrying the empty vector (pCH6405), wild-type (pCH6404), and mutant variants of SbmA (pCH6406-pCH6412) were grown in LB-Tet15 liquid cultures. They were then seeded at an $OD_{600} \approx 0.1$ in low-salt LB media (5g/L NaCl) in a 96-well plate with a final volume of 100 μL per well. Each strain was cultured in two independent wells, one containing low-salt LB-Tet15 and the other with low salt LB-Tet15-Zeo50. Subsequently, the 96 well-plate was placed in the Victor 3V^{TM} multilabel plate reader and incubated at 37°C, with OD_{600} measurements taken every 10 minutes for 5 hours. All growth curves were independently repeated three times using the same setup.
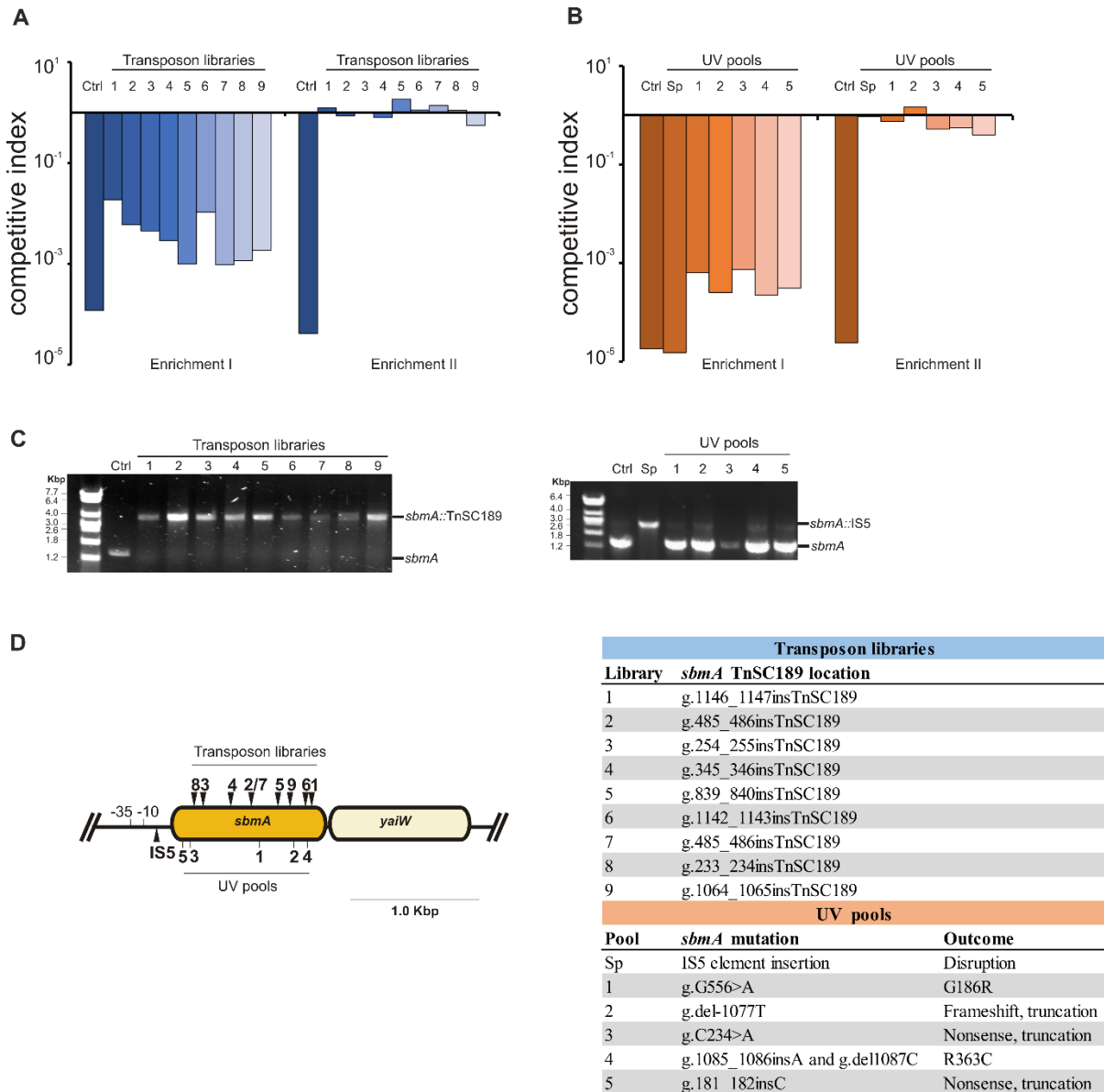
Figure 1: Selection for CdiA-CT$^{O32:H37}$ resistant mutants reveals the requirement of SbmA

A. Competitive indices derived from different transposon libraries cultured against CDI+ cells and acquired resistant phenotype after second round of enrichment.

B. Competitive indices of independent UV mutagenized pools against CDI+ cells and acquired resistant phenotype after second round of enrichment, including a control with spontaneous mutation (Sp).

C. Amplification of *sbmA* was conducted across all resistant libraries and pools to validate the presence of transposon insertions in each of the samples.

D. Transposon-insertion sites are shown for each resistant mutant along with the UV pool mutated sites. A comprehensive table is of the precise location for each insertion and mutation event is also shown.
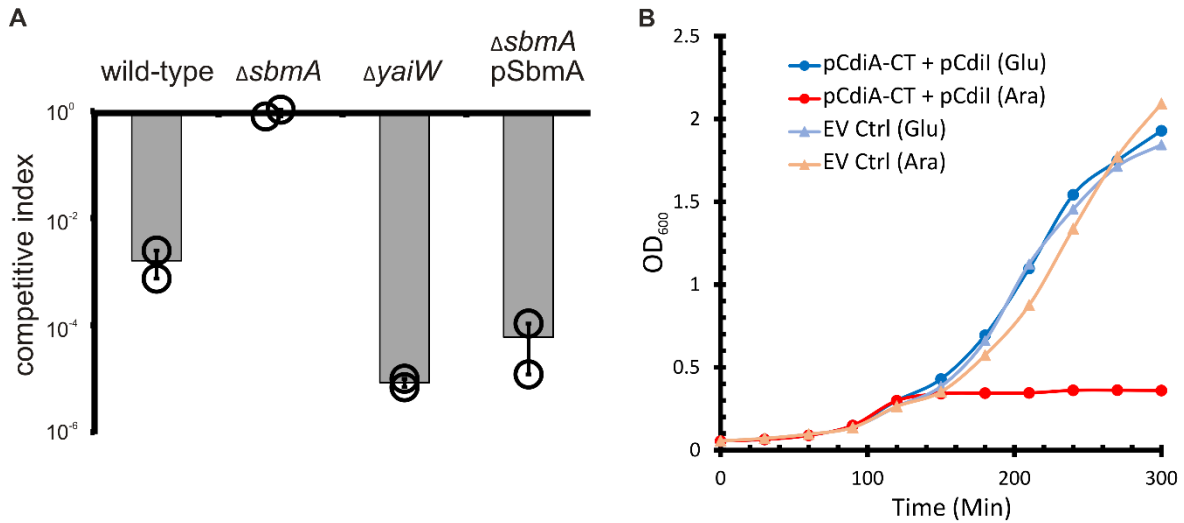
Figure 2: SbmA is required for toxin translocation in target cells.
A. *E. coli* target cells of the indicated genotypes were incubated with CDI+ cells at a 1:1 ratio. Notably, the CDI resistant background can be complemented by introducing plasmid-borne *sbmA* (pSbmA). Average competitive indices (±SEM) are presented for two independent experiments.
B. Growth curves detailing the behavior of *E. coli* cells with Δ*sbmA* background carrying the toxin plasmid (pCdiA-CT) driven by an arabinose inducible promoter along with the immunity plasmid (pCdiI), juxtaposed with the control empty vector counterparts for the respective plasmids. Average competitive indices (±SEM) are presented for two independent experiments.
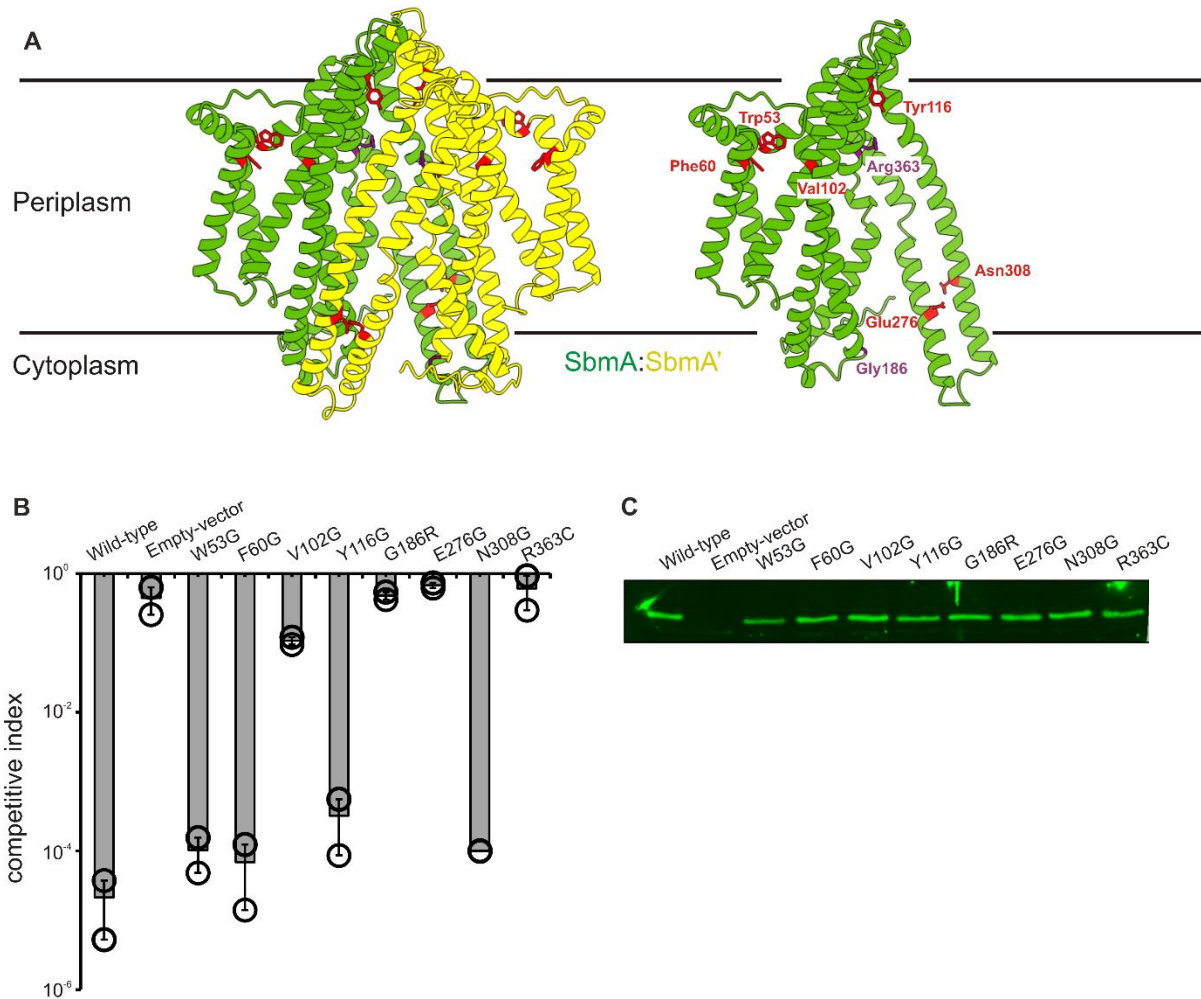
Figure 3: Important SbmA residues for translocating CdiA-CT[O32:H37].

A. The AlphaFold models depict SbmA in both dimeric and monomeric forms. The residues of particular significance have been annotated. Notably, the residues highlighted in green correspond to those reported in Corbalan et al., 2013, while the purple residues have been derived through cloning from UV pools 1 and 4.

B. To assess the importance of SbmA residues, *E. coli* target cells lacking the Δ*sbmA* gene, but harboring plasmid-borne SbmA in various forms including wild-type and selected mutants, were combined with CDI+ cells at a 1:1 ratio. This mixture was incubated for one hour to allow interactions. The average competitive indices (±SEM) arising from two distinct experiments are presented. As a baseline, an empty vector was introduced and examined as a control.

C. In order to validate the expression levels of SbmA in the forms of wild-type and mutants, an SbmA immunoblot was performed on the *E. coli* target cells from the aforementioned experiment in panel B.
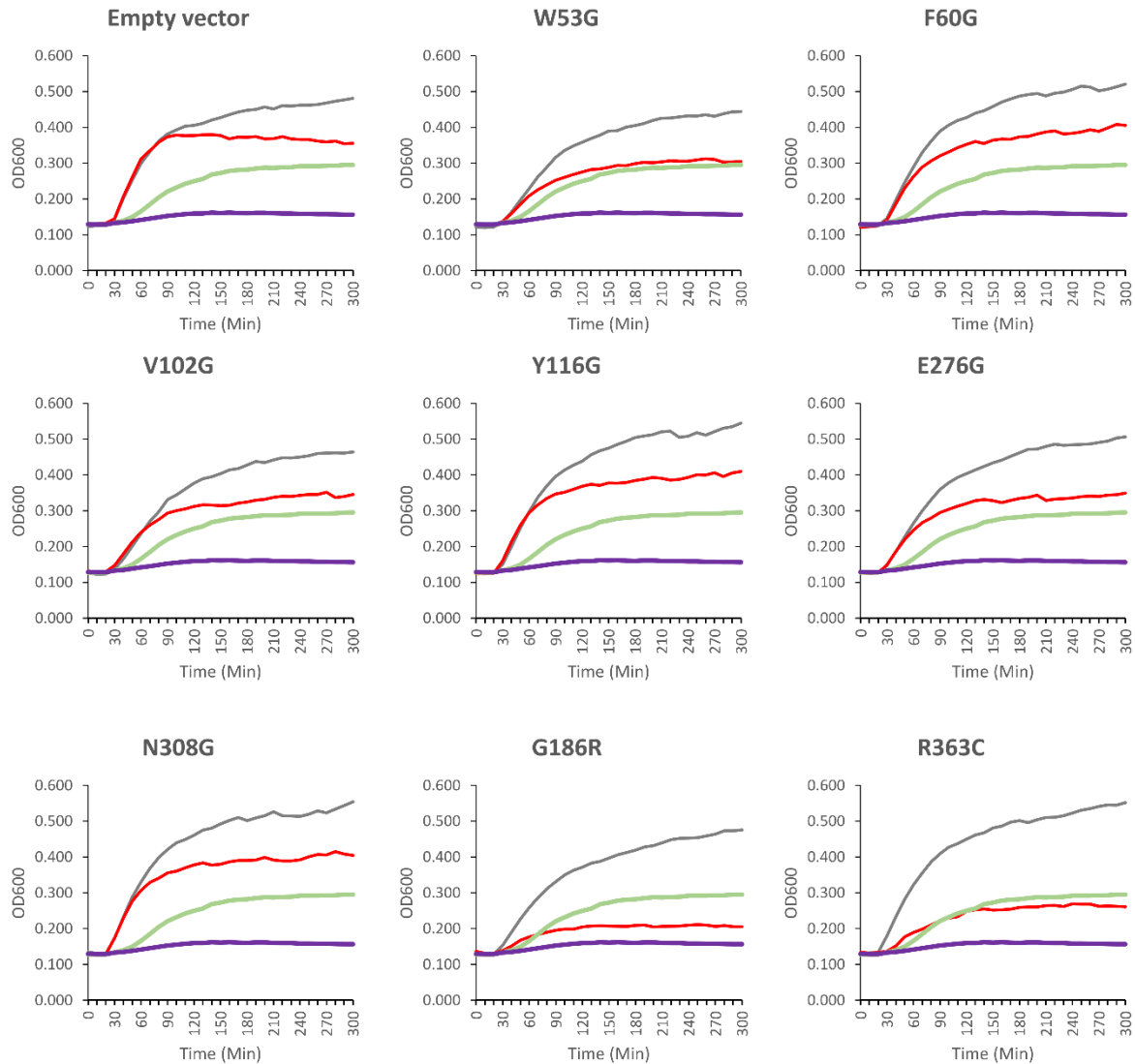
Figure 4: Zeocin susceptibility tests of wild-type and mutant variants of SbmA.
The five-hour growth curves were observed in the Δ*sbmA* cell line, which contained either an empty vector or pSbmA, including both wild-type and specified mutant variants (n=3). The green and purple curves represent wild-type cells, illustrating their growth in the absence or presence of zeocin, respectively. The gray curves represent indicated variants without zeocin, while the red curves illustrate their growth with zeocin. The wild-type growth curves were overlayed onto all indicated empty vector and mutant variant curves as a reference.

A Microcin B17

VGIGGGGGGGGG

CdiA-CT^O32:H37

VENNALGNKNDKL ... LNGGGG

B

```
            1415 18  21  23              33        39
            PPIIPINPLPIGVEGADGEPLNGGGG
P14A        APIIPINPLPIGVEGADGEPLNGGGG
P15A        PAIIPINPLPIGVEGADGEPLNGGGG
P18A        PPIIAINPLPIGVEGADGEPLNGGGG
P21A        PPIIPINALPIGVEGADGEPLNGGGG
P23A        PPIIPINPLAIGVEGADGEPLNGGGG
P33A        PPIIPINPLPIGVEGADGEALNGGGG
P18A/P33A   PPIIAINPLPIGVEGADGEALNGGGG
Penta PA    AAIIAINALAIGVEGADGEPLNGGGG
Hexa PA     AAIIAINALAIGVEGADGEALNGGGG
ΔPro        ---------------IGVEGADGEPLNGGGG
GA          PPIIPINPLPIGVEGADGEPLNAAAA
```

C

D

- EV + pCdiI
- pCdiA-CT + pCdiI
- pCdiA-CT (ΔPro)+ pCdiI
- pCdiA-CT (Penta PA) + pCdiI
- pCdiA-CT (HexaPA) + pCdiI

E

Figure 5: The proline-rich cytoplasmic entry domain of CdiA-CT$^{O32:H37}$.

A. The cytoplasmic entry domain of CdiA-CT$^{O32:H37}$ exhibits a structural resemblance to microcin B17, evident in their analogous configurations. The proline residues within this domain induce characteristic kinks that mirror the structural motifs seen in the thiazole and oxazole turns of microcin B17.

B. Schematic illustration of the mutations introduced to the cytoplasmic entry domain of CdiA-CT$^{O32:H37}$. Single alanine replacements were introduced to individual proline from N-terminus (Pro14, Pro14, Pro18, Pro21. Pro23, and Pro33). Multiple alanine substitutions were also included (Pro18-22, Penta Pro, Hexa Pro). A mutant with a deletion from residue Pro14 to Pro23 was annotated as ΔPro. Nevertheless, the glycine residues (35-39) were mutated to alanine as GA mutant.

C. Competitive indices of *E. coli* inhibitor cells carrying CDI plasmids with entry domain mutations cultured against target *E. coli* cells either with or without the cognate outer membrane receptor at the ratio of 1: 1 for 1 hour. Average competitive indices (±SEM) are presented for two independent experiments.

D. Growth curves of *E. coli* cells expressing CdiA-CT$^{O32:H37}$ wild-type, ΔPro, PentaPA, and HexaPA under the control of arabinose inducible promoter to demonstrate the mutations at the entry domain do not affect the toxin activity of the RNase domain. Average competitive indices (±SEM) are presented for two independent experiments.

E. Purified proteins from CdiA-CT$^{O32:H37}$ wild-type, ΔPro, PentaPA, and HexaPA to ensure the toxins can still be expressed. Notably, the PentaPA and HexaPA toxins appear to be unstable, and a smaller molecular weight band was observed.
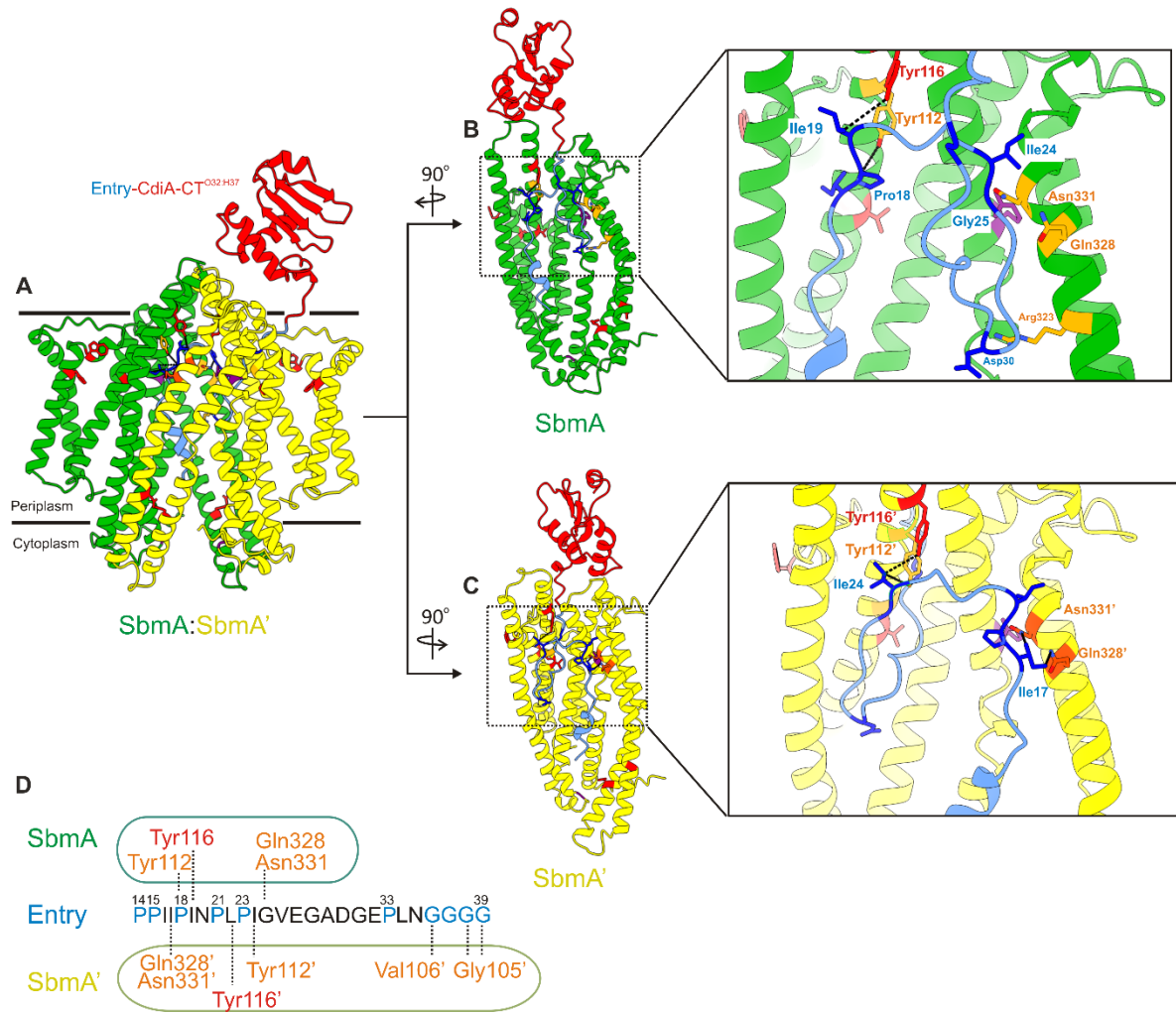
Figure 6: Predicted interaction of SbmA with CdiA-CT$^{O32:H37}$.

A. The AlphaFold model predicts the potential interaction between SbmA and CdiA-CT$^{O32:}$H37, with the entry domain threading into the lumen of the SbmA dimer.

B. A 90-degreen turn to the left of the complex, with SbmA' concealed, illustrates the predicted hydrogen bonding between selected residues on CdiA-CT$^{O32:H37}$ entry domain and those on SbmA. In this study, residues in red were tested, while orange represents the predicted residues from PDBePISA (**Table 1**).

C. A 90-degreen turn to the right of the complex, with SbmA hidden, depicts the remaining hydrogen boding of the entry domain with SbmA. In this study, residues in red were tested, while orange represents the predicted residues from PDBePISA (**Table 1**).

D. A schematic illustration depicts the hydrogen bonding between the proline-rich and glycine-rich regions of CdiA-CT$^{O32:H37}$ and the key residues of both SbmA monomers.

112
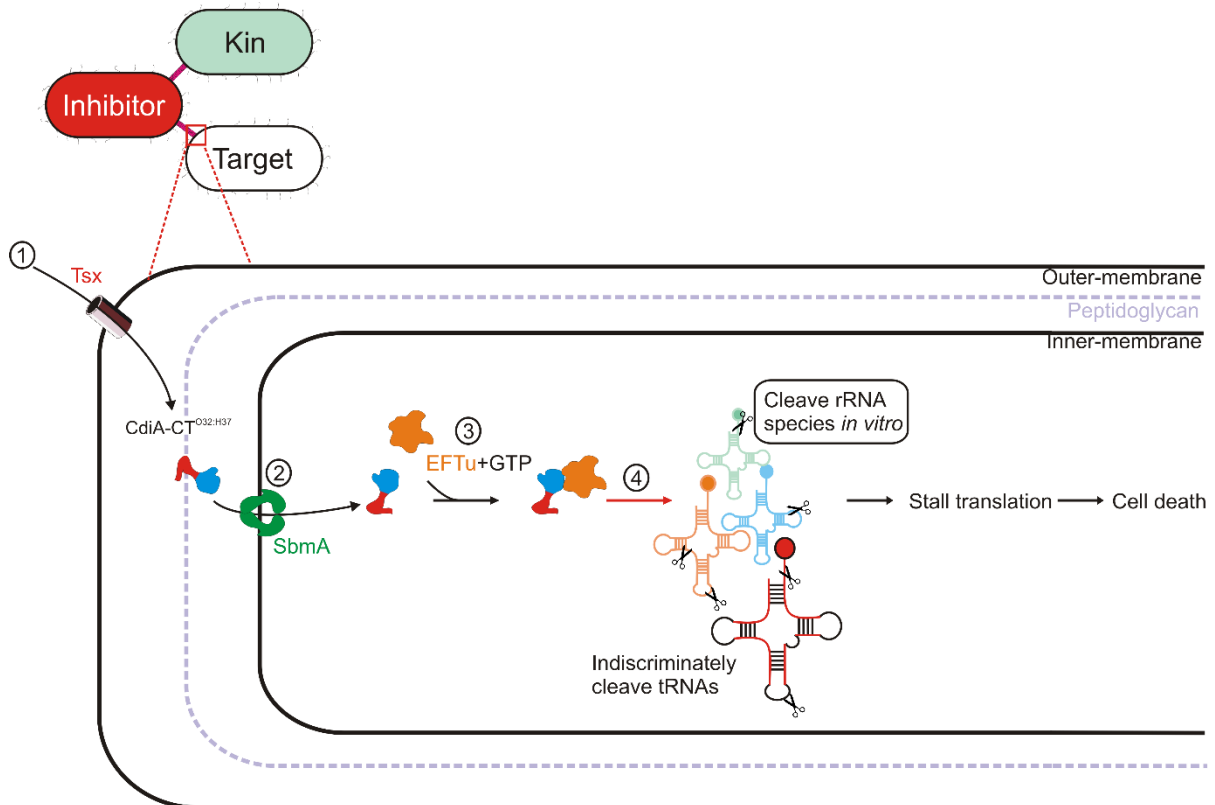
Figure 7: A proposed mechanism of CdiA-CT$^{O32:H37}$ intoxication.

In a proposed mechanism within a co-culture environment, the CDI+ inhibitor cells deploy the toxin to both kin and target bacterial species. Only the targeted species is being affected by the toxin, **(1)** CdiA$^{O32:H37}$ recognizes its receptor, in this experimental setup, we used Tsx, facilitating toxin deployment in the periplasm. **(2)** In the periplasmic compartment, the entry domain of CdiA-CT$^{O32:H37}$ is translocated to the cytoplasm via the SbmA via proline-rich region. **(3)** Once CdiA-CT$^{O32:H37}$ enters the cytoplasm, it requires EF-Tu in the GTP bound form to stabilize the conformation and enhance its toxin activity as an RNase. **(4)** The activated CdiA-CT$^{O32:H37}$ can now locate its substates, rRNAs and tRNAs, and cleave indiscriminately. Consequently, damage to the RNA species impairs the protein translation process Ultimately, this cascade of events results in cell death.

Table 1: Direct H-bonds and salt-bridges at the SbmA and CdiA-CT$^{O32:H37}$ interface

CdiA-CT atom	SbmA atom	Distance (Å)
LYS 12[NZ]	GLU 203[OE2]	3.45
GLU 27[N]	GLN 328[OE1]	2.79
LYS 12[O]	ASN 156[ND2]	3.51
PRO 18[O]	TYR 112[OH]	2.22
ILE 19[O]	TYR 116[OH]	2.78
GLY 25[O]	GLN 328[NE2]	2.95
GLY 25[O]	ASN 331[ND2]	2.55
ASP 30[O]	ARG 323[NH1]	3.82

CdiA-CT atom	SbmA' atom	Distance (Å)
LYS 9[NZ]	GLU 269[OE1]	2.62
ASN 10[ND2]	THR 374[OG1]	3.58
ILE 17[N]	GLN 328[OE1]	3.71
ILE 24[N]	TYR 116[OH]	3.46
GLY 38[N]	GLY 105[O]	2.85
GLY 39[N]	GLY 105[O]	3.18
ILE 17[O]	GLN 328[NE2]	3.27
ILE 17[O]	ASN 331[ND2]	2.42
LEU 22[O]	TYR 116[OH]	3.29
ILE 24[O]	TYR 112[OH]	2.44
ASP 30[OD1]	ASN 156[ND2]	3.77
GLY 36[O]	VAL 106[N]	3.16
GLY 37[O]	ALA 110[N]	3.77

Table 2: Ability of target cells to transport CdiA-CT$^{O32:H37}$

O32 mutation

Mutation	Result		Legend
P14A	🟩 Transport	🟩	Transport
P15A	🟩 Transport		
P18A	🟩 Transport	🟧	Slow transport
P21A	🟩 Transport		
P23A	🟩 Transport	🟥	No transport
P33A	🟩 Transport		
P18A/P33A	🟩 Transport		
Penta PA	🟧 Slow transport		
Hexa PA	🟧 Slow transport		
ΔPro	🟥 No transport		
GA	🟧 Slow transport		

SbmA mutation

Mutation	Result
W53G	🟩 Transport
F60G	🟩 Transport
V102G	🟧 Slow transport
Y116G	🟩 Transport
G186R	🟥 No transport
E276G	🟥 No transport
N308G	🟩 Transport
R363C	🟥 No transport

Table 3: Bacterial strains and plasmids

Strains	Description[a]	References
X90	F' *lacI*^{q} *lac' pro' /ara Δ(lac-pro) nal1 argE(amb) rif'thi-1*, Rif^R	(Zhang et al., 2012)
MFD*pir*	MG1655 RP4-2-Tc::[Δ*Mu1::aac(3)IV-*Δ*aphA-*Δ*nic35-*Δ*Mu2*::zeo] Δ*dapA*::(erm-pir) Δ*recA,* Amp^R, Zeo^R, Erm^R	(Ferrières et al., 2010)
CH2016	X90 (DE3) Δ*rna* Δ*slyD*::kan, Kan^R	(Aoki et al., 2008)
CH2552	EPI100 *pir+*	(Willett et al., 2015)
CH6402	X90 Δ*tufB* Δ*wzb* Δ*sbmA*::Kan, Kan^R	This study
CH6403	X90 Δ*tufB* Δ*wzb* Δ*yaiW*::Kan, Kan^R	This study
CH10229	JCM158 *rif*^{R} Δwzb::cat, Rif^R, Cm^R	(Ruhe et al., 2013)
CH14016	MG1655 Δ*wzb* Δ*tsx*	(Ruhe et al., 2018)
CH14910	X90 Δ*tufB*::Kan, Kan^R	This study
CH14914	X90 Δ*tufB*	This study
CH15176	X90 Δ*tufB* Δ*wzb*::Kan, Kan^R	This study
CH15206	X90 Δ*tufB* Δ*wzb*	This study
JW0368	BW25113 F-, *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), ΔsbmA742::kan, λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514*, Kan^R	Keio collection
JW0369	BW25113 F-, *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), ΔyaiW743::kan, λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514*, Kan^R	Keio collection
JW2046	BW25113 F-, *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, Δwzb-759::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514*, Kan^R	Keio collection
JW3943	BW25113 F-, Δ(*araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, ΔtufB756::kan, hsdR514*, Kan^R	Keio collection

Plasmids	Description[a]	References
pCH48	pCP20, plasmid that shows temperature-sensitive replication and thermal induction of FLP synthesis, Amp^R and Cm^R	(Cherepanov & Wackernagel, 1995)
pCH450	pACYC184 derivative with *E. coli araBAD* promoter for arabinose-inducible expression, Tet^R	(Hayes & Sauer, 2003)
pCH2816	pET21b::cdiBA STEC- O32:H37 cdiA(CT)-cdiI Chimera, Amp^R	Chapter III
pCH3781	pCH450/ pTrc99a, Tet^R, Amp^R	Chapter III
pCH3782	pCH450::CT (O32::H37)/ pTrc99aKX::CdiI (O32::H37), Tet^R, Amp^R	Chapter III
pCH3783	pCH450::CT (O32::H37)-(Del-P14-P23) / pTrc99aKX::CdiI (O32::H37), Tet^R, Amp^R	This study
pCH3784	pCH450::CT (O32::H37)-(P14-P23 all P to A)/ pTrc99aKX::CdiI (O32::H37), Tet^R, Amp^R	This study
pCH3785	pCH450::CT (O32::H37)-(P14-P33 all P to A)/ pTrc99aKX::CdiI (O32::H37). Tet^R, Amp^R	This study
pCH6404	pCH450::sbmA., Tet^R	This study
pCH6405	pCH450, Tet^R	This study
pCH6406	pCH450::sbmA (W53G), Tet^R	This study
pCH6407	pCH450::sbmA (F60G), Tet^R	This study
pCH6408	pCH450::sbmA (V102G), Tet^R	This study
pCH6409	pCH450::sbmA (Y116G), Tet^R	This study
pCH6410	pCH450::sbmA (F219G), Tet^R	This study
pCH6411	pCH450::sbmA (E276G), Tet^R	This study
pCH6412	pCH450::sbmA (N308G), Tet^R	This study
pCH6413	pET21b::cdiBA STEC- O32:H37 cdiA-CT(Del-P14-P23)-cdiI Chimera, Amp^R	This study
pCH6414	pET21b::cdiBA STEC-O32:H37 cdiA-CT(P14-P23 all P to A)-cdiI Chimera, Amp^R	This study
pCH6430	pET21b::cdiBA STEC-O32:H37 cdiA-CT(P14-P33 all P to A)-cdiI Chimera, Amp^R	This study
pCH6431	pET21b::cdiBA STEC- O32:H37 cdiA-CT(PP18,33AA)-cdiI Chimera, Amp^R	This study
pCH6444	pET21::O32:H37 CT(Del-P14-23)-imm-H6, Amp^R	This study
pCH6446	pET21::O32:H37 CT(P14-23A all Pro to Ala)-imm-H6, Amp^R	This study
pCH7133	pET21::O32:H37 CT(P14-33A all Pro to Ala)-imm-H6, Amp^R	This study
pCH9582	pSC189-Kan, Amp^R, Kan^R	(Aoki et al., 2008)
pCH12716	pET21::O32:H37 CT-imm-H6, Amp^R	Chapter III
pCH15124	pZS21/pArgW, Kan^R, Tet^R	Chapter III
pCH15125	pZS21:tsx/pArgW Kan^R, Tet^R	Chapter III
pCH15270	pET21b::cdiBA STEC- O32:H37 cdiA-CT(P14A)-cdiI Chimera, Amp^R	This study

pCH15271	pET21b::cdiBA STEC- O32:H37 cdiA-CT(P15A)-cdiI Chimera, Amp^R	This study
pCH15272	pET21b::cdiBA STEC- O32:H37 cdiA-CT(P21A)-cdiI Chimera, Amp^R	This study
pCH15273	pET21b::cdiBA STEC- O32:H37 cdiA-CT(P23A)-cdiI Chimera, Amp^R	This study
pCH15274	pET21b::cdiBA STEC- O32:H37 cdiA-CT(P33A)-cdiI Chimera, Amp^R	This study
pCH15275	pCH450::sbmA (G186R), Tet^R	This study
pCH15276	pCH450::sbmA (R363C), Tet^R	This study

a Abbreviations: Amp^R, ampicillin-resistance; Cm^R, chloramphenicol-resistance; Kan^R, kanamycin-resistance; Spc^R, spectinomycin-resistance; Erm^R, erythromycin-resistance; and Tet^R, tetracycline-resistance.

Table 4: Oligonucleotides

Identifier	Descriptive name	Sequence	References
CH3896	O32:H37-CT-Kpn/Nco-for	5' - TTT GGT ACC ATG GTT GAG AAT AAT GCG C - 3'	Chapter III
CH3898	O32:H37-cdiI-Spe-rev	5' - TTT ACT AGT CTG TTC GTT AAA TGC TCG - 3'	Chapter III
CH4003	O32-CT-Xho-rev	5' - TTT CTC GAG TCA TTG CTT GTT CCA ATA TTC G - 3'	Chapter III
CH4262	EC93-A2894-Nhe-for	5' - AAA GCT AGC GCC GGT ACG GGG GC - 3'	Chapter III
CH4716	O32-cdiI-Xho-rev	5'- TTT CTC GAG TTA CTG TTC GTT AAA TGC TCG TTT C-3'	Chapter III
CH4717	sbmA-Eco-for	5'- TAA GAA TTC AGC AGG AGT GCA TAT GTT TAA GTC-3'	This study
CH4718	sbmA-XhoI-rev	5'- CTT CTC GAG TTT AGC TCA AGG TAT GGG-3'	This study
CH5000	sbmA screen-for	5'- TGG GTC ACA ATG GGC GGT- 3'	This study
CH5001	sbmA screen-rev	5'- TCA ACG GTT GCG GAG CCT G-3'	This study
CH5002	yaiW screen-for	5'- GTT GAT GTC TAT CTA CAA ACG T-3'	This study
CH5003	yaiW screen-rev	5'- ATC GCC GCC AGT AAG TTT T-3'	This study
CH5036	O32-H37-P14A-rev	5'- ATA ATG GGT GCC AGC TTG TCA T-3'	This study
CH5037	O32-H37-P15A-rev	5'- CGG AAT AAT GGC TGG CAG CTT GT-3'	This study
CH5038	O32-H37-P18A-rev	5'- CAG CGG ATT TAT CGC AAT AAT GGG-3'	This study
CH5039	O32-H37-P21A-rev	5'- GAT TGG CAG CGC ATT TAT CGG-3'	This study
CH5040	O32-H37-P23A-rev	5'- CAA CCC CGA TTG CCA GCG GAT T-3'	This study
CH5041	O32-H37-P33A-rev	5'- CGT TCA GCG CCT CAC CAT CT- 3'	This study
CH5099	O32-H37-P13-23-rev	5'- ACC TTC AAC CCC GAT CAG CTT GTC ATT CTT GTT AC-3'	This study
CH5100	O32-H37-Pro-less rev	5'- CAA CCC CGA TTG CCA GCG CAT TTA TCG CAA TAA TGG CTG CCA GCT TGT CA-3'	This study
CH5101	sbmA-W53G-for	5'- GCG CGT TTC GGG TCG TTG GAT-3'	This study
CH5102	sbmA-F60G-for	5'- TTG GAT TTC CTG ATT GGT TAC GCT TAC-3'	This study
CH5103	sbmA-V102G-rev	5'- CCC GAC TTC CCC CAA AAA CCA G-3'	This study
CH5104	sbmA- Y116G-rev	5'- GAA TCA GAT CAC CGA ACG GCG CAT-3'	This study
CH5105	sbmA- F219G-rev	5'- ACC GGC AGG CCG GCG ATC AA-3'	This study
CH5106	sbmA-E276G-for	5'- CAG CGT GTA GGG GCT GCC TAC C-3'	This study
CH5107	sbmA-N308G-for	5'- GTA CGG AAA GGC TAT TTC CGC CT-3'	This study
CH5604	O32-G36-39A-rev	5'- TTT AGC TAT CGC AGC AGC AGC GTT CAG CTC-3'	This study

Chapter V: A C39 family protease in *Citrobacter rodentium* DBS100 targets gyrase and topoisomerase IV

A) Introduction

Bacterial growth inhibition has been intensively studied over the years, with numerous effector proteins identified in various systems, particularly in the contact-dependent growth inhibition (CDI) system. These proteins all share a common mechanism, targeting essential molecular components of intoxicated cells to either halt or abolish bacterial growth. These targeted components include DNA, RNA, and cell membranes. The nuclease family is one of the well-described effectors in the CDI system, which includes DNase from *Escherichia. coli* EC869 o11 (Morse et al., 2012), responsible for degrading bacterial chromosomal DNA. Similarly, RNases have been isolated from various sources, such as *E. coli* NC101(Michalska et al., 2017) , *E. coli* EC3006 and *Klebsiella pneumoniae* 342 (Gucinski et al., 2019) , *E. coli* EC869 (Jones et al., 2017), and *E. coli* STEC_O31 (Michalska et al., 2018). These RNases primarily target the RNAs of intoxicated cells or, in some cases, specifically go after tRNA's anticodon and 3'-acceptor stems. Another family of effectors includes pore-forming toxins isolated from *E. coli* EC93, which dissipate the proton motive force and reduce ATP synthesis (Aoki et al., 2009).

In recent years, many more putative toxin families have slowly been revealed in different delivery systems. An example is the, DNA deaminase from DddA in the type VI secretion system (T6SS) of *Burkholderia cenocepacia*. This deaminase directly induces lethal mutations in the target genomes through a process of cytosine deamination (de Moraes et al., 2021). Using bioinformatic analysis, we also identified a putative peptidase CDI toxin domain in the rodent pathogen *Citrobacter rodentium* DBS100, CdiA-CT[DBS100]. This domain belongs to the MEROPS peptidase C39, clan CA (MER0002036) (Rawlings et al., 2014).

The Clan CA/C39 peptidase is a cysteine protease family that removes the N-terminal leader peptides of bacteriocins, including those from non-lantibiotic and lantibiotic bacteriocins. This is required for the ATP-binding cassette transporters to export these antimicrobial peptides or bacteriocins from Gram-negative and Gram-positive bacteria (Rice et al., 2014). In most cases, the C39 peptidase targets the two conserved glycine or double-glycine motifs in positions -1 and -2 of the leader peptide (Rawlings et al., 2014). The active site of this clan is well conserved with a catalytic triad composed of Cys-His-Asn/Asp, followed by a Gln in some cases (Barrett et al., 1982).

Research on the C39 peptidase as an effector is limited because it is typically found only in ABC transporter systems. However, there has been a recent discovery of a C39-like effector, IpaJ from *Shigella felxneri*. IpaJ is a T3SS pathogenic virulent factor that cleaves the peptide bond between the N-myristyolated glycine-2 and asparagine-3 of human ARF1 and promote host secretory inhibition (Burnaevskiy et al., 2013), which could use as framework for our study. We found that CdiA-CT^{DBS100} is located at the C-terminal end of the CDI-like filament, known as StbD, which shares a similar topology to a typical CDI filament. It features FHA-1 repeats, a receptor binding domain (RBD), FHA-2 repeats, and a VENN motif prior to the toxin domain (Ruhe et al., 2018), located upstream of a known T3SS in *C. rodentium*. Notably, the *cdiB* gene and its TPS, responsible for recognizing CdiB, were precisely replaced by the type I usher and a type I chaperone genes. These replacements may suggest that the filament in *C. rodentium* may play a role in pathogenesis.

For the scope of this study, our primary interest lay in the function of the C-terminal effector protein. Consequently, we decided to re-engineer the StbD filament back to a CDI topology by actively replacing these two domains with CdiB^{STEC3} and TPS^{STEC3} from *E. coli*

STEC_O31 (Ruhe et al., 2018). This modification provided us with a functional CDI filament for bacterial inhibition. We demonstrated that CdiA-CT^{DBS100} becomes toxic to *E. coli* after activation by an autoproteolytic event occurring between Gly158-Ala159 under reducing conditions. This activation releases the toxin domain into the cytoplasmic compartment and induces cell filamentation with non-partitioning chromosome phenotypes. This evidence provides important insights for narrowing down the substrate identification for this toxin. Specifically, the ATPase domains found in GyrB and ParE subunits of gyrase and topoisomerase IV, respectively, are identified as substrates. Notably, the cleavage event occurs precisely at the ATP binding pockets between the two glycine residues: Gly101-Gly102 in GyrB and Gly97-Gly98 in ParE. Additionally, we have observed that the first thirteen residues, Ser2-Lys14, of GyrB play a critical role in binding the toxin domain and orienting its active pocket towards the cleavage motif LHAGGK.

B) Results

CdiA-CT/CdiI^{DBS100} structure

The AlphaFold model provides insight into the structural composition of CdiA-CT^{DBS100}, revealing a total of fifteen alpha-helices and seven beta-strands. The cytoplasmic entry domain encompasses ten alpha-helices (α1-α10), while the C39 protease domain initiates with three alpha-helices (α11-α13). Additionally, α14 is situated between two beta-strands (β1, β2). The protease domain further extends with three consecutive beta-strands β4, β5, and β6, and culminates in the final helix, α15, before concluding with β7. These two domains are connected by the 44-residue long extended linker region (**Fig. 1A & B**). The entry domain potentially obstructs the active pocket, where the putative active residues are labeled in yellow (**Fig. 1A**). This led us to an unexpected observation involving the immunity protein, CdiI. The CdiI^{DBS100} consists mainly of beta-strands (β'1-β'5) with two alpha helices

(α'1-α'2) positioned between β'4 and β'5 (**Fig. 1B**). These two alpha helices make direct contact with the backside of CdiA-CT^{DBS100}, facing away from the active pocket. Specifically, Met62 of α'1 forms a hydrogen bond with Trp250 of CdiA-CT^{DBS100}, Asn70 of α'2 forms hydrogen bond with Phe233 of α14., Asp52 of α'2 forms bond with Arg216 of α13, and Val49 of β'4 forms bond with Phe230 of β1 (**Figs. 1B** & **C** and **Table 1**).

Following this structural analysis, we sought to confirm the autoproteolysis of the toxin. The toxin domain was purified and analyzed in the absence and presence of a reducing agent, dithiothreitol (DTT). We observed that the toxin underwent self-processing in the presence of DTT. The full-length protein running at approximately 35 KDa, was converted into a processed band, migrating around 15 KDa (**Fig. 1D,** lanes 1 & 4). This autoproteolytic event can be prevented in the presence of its cognate immunity protein (**Fig. 1D,** lane 6).

CdiA-CT^{DBS100} auto-cleaves between Gly158 and Ala159
N-terminal His6-tagged full-length wild-type and an active residue mutant CdiA-CT were created to precisely analyze the cleavage site of the toxin. Both the full-length (His6-Full) and auto-processed product (His6-Entry) were purified and isolated by using high performance liquid chromatography (HPLC) (**Fig. 2A**) and then analyzed by electrospray ionization mass spectrometry (ESI-MS). The processed His6-Entry product was found to have a size of 18, 094.0±1 Da (**Fig. 2B**), closely matching the predicted size of entry domain of CdiA-CT^{DBS100} up to Gly158, which is 18, 097 Da (**Fig. 2C**).

To confirm the ESI-MS analysis, serine point mutations were introduced at the potential cleavage site, specifically at both Gly157 and Gly158, or a single Gly158. Wild-type and mutant proteins were then purified and analyzed in the absence and presence of DTT. It was observed that serine mutations at both Gly157 and Gly158, as well as a single

Gly158, prevented the auto-cleavage event from the toxin domain, resulting in the retention of the full-length toxin, similar to the wild-type toxin, in the presence of its cognate immunity protein (**Fig. 2C,** lanes 3, 5, & 8).

In terms of spatial arrangement, AlphaFold models provided insights into the potential regulation of this auto-proteolytic event. In the presence of the immunity protein Gly158, located on the linker region, faces away from the active pocket. When the immunity protein is removed from the model, the linker region reorients to face toward the active pocket, bringing Gly157 and Gly158 into proximity with the putative active residues (**Fig. 2D**).

CdiA-CT^{DBS100} targets GyrB and ParE

Two fluorescent *E. coli* cell lines were introduced for a co-culture assay to observe the toxin effect on intoxicated cells as a first step toward identifying targets of the toxin. A super folded YFP (sfYFP) inhibitor strain carries a CDI-plasmid to deliver the CdiA-CT^{DBS100} to the mKate target strain. After co-culturing, the mKate target cells becomes filamentous and the nucleoids labeled with Hoescht 33342, are unable to separate to the cell poles. The intoxicated cells progressively become filamentous after a long period of co-culturing. These intoxicated cells share a similar phenotype with those treated with nalidixic acid and novobiocin (**Fig. 3A**). This shared phenotype, particularly the presence of unsegregated nucleoids suggests that the DNA replication machinery has been affected an may provide us with better insights on the targets of the toxin.

A second approach involved a specific analysis of the DNA topoisomers from the co-cultured cells. Target cells carrying a high-copy plasmid species [pBluescript II SK(+)], were intoxicated by the inhibitor cells. After 15 minutes of cell mixing, plasmid DNAs from the

cell mixture were collected and resolved on a chloroquine agarose gel. Inhibitor cells carrying a larger plasmid species (16 Kbp) and pBluscript II SK(+) in the target cells (approximately 3 Kbp). It was observed that the plasmid DNA species from target cells lost the ability to form supercoiled DNA species compared to the mock control, and different species of DNA topoisomers migrated according to their linker numbers (**Fig. 3B**). Together with the phenotype observed (**Fig. 3A**), this suggests that the toxin is targeting DNA replication and/or the cell division machinery.

To investigate this further, we independently introduced VSV-g epitope tags in the chromosome downstream to five possible candidates, GyrA, GyrB, ParC, ParE, and FtsZ. A co-culture assay was performed between the CDI-plasmid inhibitor and the VSV-g tagged target cells. The cell mixtures were collected for immunoblot analysis with anti-VSV-g antibody, revealing that only GyrB and ParE were cleaved by the toxin whereas GyrA, ParC, and FtsZ remain untouched by the toxin (**Fig. 3C**).

CdiA-CT^{DBS100} specifically affects the ATPase activity of GyrB and ParE

GyrB and ParE exhibit ATPase activity. Thus, we tested whether this domain was cleaved by the toxin. The ATPase domain of GyrB was tagged at C-terminus with His6 (ATPase^{GyrB}). This was overexpressed in the target cells, which were then mixed with the inhibitor cells. The ATPase^{GyrB} domain was purified from the co-culture with Ni^{2+} -NTA affinity chromatography followed by SDS-PAGE analysis. The data confirmed previous observations: the ATPase^{GyrB} domain is cleaved by the toxin (**Fig. 4A**). Furthermore, the full-length (Full-His6) and cleaved (Processed-His6) ATPase^{GyrB} were separated by HPLC (**Fig. 4B**) and analyzed by ESI-MS. Full-His6 was found to be 25, 210.00±1 Da (**Fig. 4C**) and Processed-His6 was 14, 609.00±1 Da (**Fig. 4D**), which closely matches the predicted size of

the full-length ATPase^{GyrB} domain (25, 219.23 Da). The processed product strongly suggests that the cleavage site is between Gly101 and Gly102, generating the predicted product with the size of 14613.38 Da (**Fig. 4E**).

To further confirm the ESI-MS analyses, serine point mutations were introduced either to both Gly101 and Gly102 or individual at Gly101 or Gly102 of GyrB. The wild-type and mutant ATPase^{GyrB} domains were overexpressed in the target cells and mixed with inhibitor cells, and the proteins were further purified for analysis. It was found that only a double-serine mutation (GG101, 101SS) fully prevented the toxin from cleaving, whereas individual mutation at Gly101 and Gly102 were still recognized and cleaved by the toxin. However, serine mutations at another pair of double glycine (Gly113 and Gly114) downstream to Gly101 and Gly102 did not prevent toxin cleavage, revealing that the toxin specifically targets the Gly101 and Gly102 pair (**Fig. 4F**). Based on the sequence homology between GyrB and ParE ATPase domains, it is implied that the toxin would cleave ATPase^{ParE} domain between Gly97 and Gly98 with the conserve motif LHAGGKF (**Fig. 4G**).

The specificity of a double-glycine motif for protease recognition prompted us to investigate several other ATPase domains found in *E. coli* identified by the PEC database within the protein family of ATPase$^{GyrB/ParE}$, including HtpG, YedV, CheA, UhpB, RcsD, BaeS, PhoR, PhoQ, EnvZ, CpxA, RstB, AcrB, TorS, BarA, RcsC, GlnL, AtoS, and ZraS (Yamazaki et al., 2008). Sequence alignment revealed that other ATPase domains also contained double-glycine motifs scattered among the sequences, but only GyrB and ParE had the motif of LHAGGKF (**Fig. 5A**). The phylogenetic tree suggested that the ATPase$^{GyrB/ParE}$ is closely related to CheA, HtpG, and RcsD with its own branch (**Fig. 5B**). We decided to

select two other ATPase domains from EnvZ and HtpG to test whether CdiA-CT[DBS100] could recognize other ATPases. A similar cell mixing experiment was conducted to purify the proteins of interest after intoxication. It was found that only ATPase[GyrB] can be cleaved by the toxin, while other ATPases are unaffected by the toxin (**Fig. 5C**).

The C39 protease active site
The C39 peptidase family can be found in many membrane transporter systems for processing lantibiotic or other secreted peptides. Among the C39 peptidase from different species categorized by the MEROPS database, such as SunT from both *Bacillus cereus* and *B. subtilis*; PedD (*Pediococcus acidilactici*); McnB, CvaB, MchF (*Escherichia coli*); ScnT (*Streptococcus pyogenes*); MutT (*Streptococcus mutans*); NukT (*Staphylococcus warneri*); lcnDR3 (*Lactococcus lactis*), BT_4288 (*Bacteroides thetaiotaomicron*); and MrsT (*Bacillus* sp. HIL-Y85/54728), all share the three crucial catalytic residues: Cys-His-Asp (Rawlings et al., 2014). The only exception is NukT, which replaces Asp with Asn. Most of these C39 peptidases also employ another Cys residue downstream of the catalytic one to form a disulfide bond, except in the case of PedD (**Fig. 6A**). CdiA-CT[DBS100] is equipped with all the essential C39 peptidase elements: the three catalytic residues, Cys183, His258, and Asp275, along with a secondary Cys187 for disulfide bonding. The AlphaFold model presents the putative model of the active pocket, comprised of the catalytic residues, targeting the known ATPase[GryB] cleavage site, Gly101 and Gly102 (**Figs. 6A** & **B**). We conducted a thorough point mutation screening to evaluate the physiological importance of the motif for the autoproteolytic event (Gly157 and Gly158), active site residues (Cys183, His258, and Asp275), and the disulfide bonding residue, Cys187.

A co-culture assay was performed with inhibitor cells carrying either wild-type or specific *cdiBAI* locus mutants mixed with the target cells, with or without the known outer-membrane receptor, OmpF (data not shown). Target cells without OmpF showed resistance to the wild-type toxin delivery, resulting in a competitive index equal to one, whereas target cells expressing plasmid borne OmpF are fully susceptible to the toxin (**Fig. 6C**). Mutations at the residues responsible for the autoproteolysis, GG157,158SS and Gly158Ser mutants, expressed the CdiA filament normally (**Fig. 6D,** lanes 3-6) and exhibited attenuated toxin activity with target cells remaining viable (**Fig. 6C**). Based on the evidence of the competitive index and immunoblot, the GG157, 158SS mutant had a slower rate of cleaving GyrB compared to the single mutation, Gly158Ser (**Fig. 6D,** lanes 4 & 6), which is also evident in the competitive indices where target cells intoxicated by the double mutant showed slightly higher viability compared to Gly158Ser (**Fig. 6C**). Furthermore, the fully expressed CdiA filaments of three catalytic mutants completely abolished the toxin activity, as indicated by competitive indices precisely equal to one (**Figs. 6C** & **D**) and the immunoblot revealed that these mutant variants were not able to cleave GyrB-VSV in the target cells (**Fig. 6D,** lanes 8, 12 & 14). Surprisingly, Cys187Ser mutant was still able to kill target cells and fully cleave GyrB like the wild-type toxin (**Figs. 6C** & **D,** lane 10), but it was prematurely processed in the absence of receptor (**Fig. 6D,** lane 9).

Finally, *in vitro* reactions were conducted using purified GyrB incubated with the purified wild-type and mutant toxins to confirm the observations made during *in vivo* delivery. Similar to the co-culture assay, mutations at Cys183, His258, and Asp275 resulted in the inability to convert full-length GyrB into the cleaved products (**Fig. 6E**). Unlike the *in vivo* delivery, GG157, 158SS and Gly158Ser were able to cleave the GyrB in the rate similar

to that of the wild-type toxin and the Cys187Ser variant (**Fig. 6E**). These findings strongly suggest that CdiA-CT^{DBS100} utilizes the Cys-His-Asp catalytic triad to facilitate enzymatic activity, and the double-glycine residues (Gly157, Gly158) along with the Cys187 are important for the toxin delivery process.

CdiA-CT^{DBS100} interacts with the GyrB ATP binding site

Gyrase exists as a heterotetramer in bacterial cells, comprising two homodimeric subunits of GyrA and two of GyrB (Brino et al., 2000). GyrB plays a critical role in ATP binding, enabling gyrase to negatively supercoil closed circular DNA to maintain chromosomes in an underwound state (Gibson et al., 2020). Extensive structural studies have been conducted on GyrB to investigate gyrase-targeting compounds. The ATPase domains reside within GyrB can bind to two ATP molecules for a closed conformation (Maxwell, 1993). A high-resolution crystal structure reveals that the first thirteen residues (Ser2-Lys14) of the GyrB is critical for securing the ATP analog adenylyl-imidodiphosphate (ANP) within its binding pocket (Hearnshaw et al., 2015). These residues from one subunit form an extended arm that covers the ATP pocket of the other subunit, ensuring ATP binding (**Fig. 7A**).

Furthermore, the crystal structure highlights the presence of a double-glycine motif, Gly101 and Gly102 within the ATP binding pocket. This observation led us to explore ways of blocking the toxin from accessing the substrate (**Fig. 7B**). *In vitro* reactions were conducted to test whether the toxin could access the substrate in the presence of ATP or ANP. Our data indicate that the presence of 0.5 mM of the ATP and ANP were able to prevent 50 % of the toxin's activity compared to the absence of both nucleotides (**Fig. 7C**). However, higher concentrations of these nucleotides did not provide additional protection to

the substrates, suggesting that the toxin targets the substrates after ATP hydrolysis or in their open conformation.

AlphaFold modeling provides additional insights into the interaction between the toxin and ATPase^{GyrB} domains. It suggests that the first thirteen residues (Ser2-Lys14) make direct contact with the backside of CdiA-CT^{DBS100}, facing away from the active site of the toxin (**Fig. 8A**). To test this interaction, we constructed a deletion mutant of ATPase^{GyrB}, ATPase (Δ1-14) ^{GyrB}, and assessed its binding affinity with the inactive toxin domain, A159-CT(C183A)^{DBS100}. Our data strongly indicate that the loss of the first thirteen residues impairs binding affinity, as the His6-ATPase(Δ1-14)^{GyrB} did not retain A159-CT(C183A)^{DBS100} after co-incubation, while the wild-type His6-ATPase^{GyrB} retained a portion of the toxin domain (**Fig. 8B,** lanes 6 & 9).

Furthermore, we conducted a time-course and concentration titration experiments to evaluate the processing rate of the toxin. First, a time-course experiment revealed that the wild-type ATPase^{GyrB} was fully processed by the toxin after 10-minute incubation, whereas ATPase (Δ1-14) ^{GyrB} required 30 minutes for complete processing (**Fig 8C,** lanes 4 & 11). Then, a concentration titration experiment was also included, and we found that ATPase (Δ1-14) ^{GyrB} required 100 times more toxin (1 µM toxin) to be fully processed compared to the wild-type ATPase (0.01 µM toxin) (**Fig 8D,** lanes 4 & 10). This underscores the essential role of the first thirteen residues (Ser2-Lys14) in the ATPase domain, not only in securing ATP in the binding pocket but also in toxin binding affinity.

C) Discussion

These findings provide valuable insights into the novel discovery of antimicrobial mechanisms. Gyrase is a common target to develop antimicrobial agents; many antibiotic

drugs and toxins primarily target both subunits of gyrase. For example, aminocoumarin binds to the GyrB subunit and prevent ATP from binding (Maxwell, 1993). Quinolone inhibits both gyrase and topoisomerase IV, inducing the formation of chromosomal breakage (Pham et al., 2019). The CcdB protein, a type II toxin-antitoxin (TA) that maintains plasmids, poisons the gyrase-DNA complex and blocks the passage of polymerase (Bahassi et al., 1999). FicT toxin inactivates both gyrase and topoisomerase IV by adenylylation (Harms et al., 2015).

Here, we report a toxin that efficiently cleaves the gyrase and topoisomerase IV subunits, completely halting cell growth. These results reveal that the CdiA-CT/CdiI[DBS100] complex has a unique structural arrangement. Most of the reported CDI toxin commonly encode the immunity protein to fully bind to the active pocket of the toxin (Gucinski et al., 2019; Jones et al., 2017; Michalska et al., 2017; Nikolakakis et al., 2012), whereas in the case of DBS100, the immunity pwas designed to interact with the interface that makes direct contact with the first thirteen residues (Ser2-Lys14) of ATPase domain. When this surface is occupied by the immunity protein, the toxin is unable to access its extended linker region, which is essential for recognizing the double-glycine motifs on its own linker region and in the GyrB ATPase domain to remove the entry domain and deactivate GyrB, respectively.

With intensive structural studies on gyrase and topoisomerase IV, we have acquired the data to identify the spatial and temporal aspects of toxin cleavage. As a member of the C39 peptidase family, CdiA-CT[DBS100] utilizes the Cys-His-Asp catalytic triad to target the double-glycine motif (Rawlings et al., 2014). Our findings suggest that the toxin adheres strictly to a recognition motif, LHAGGK, which can only be found in the ATPase domains of topoisomerases and not found in other ATPase domains. This motif is a non-structural

peptide that aligns perfectly to the active groove when the ATPase domain orients itself toward CdiA-CT^{DBS100} using the anchor of Ser2-Lys14 arm.

It is known that when gyrase binds to ATP, the two Ser2-Lys14 arms cross each other to secure ATP binding, completely burying the LHAGGK motif (Brino et al., 2000; Hearnshaw et al., 2015). The arms only open after ATP hydrolysis, creating an opportunity for the toxin to impose its catalytic groove onto the LHAGGK region. Our nucleotide blockage data support this hypothesis by demonstrating that the substrate is less susceptible to toxin recognition. The lack of complete protection may indicate that a portion of gyrase in the purified mixture may not be fully functional and could not bind to the nucleotides. Consequently, those non-functional gyrase molecules may have remained as the only available substrates for the toxin.

Furthermore, this toxin system provides a new perspective for studying the CDI delivery mechanism. CDI delivery has been extensively described by Ruhe et al., 2018; Willett et al., 2015, where the toxin must recognize its cognate outer-membrane and inner-membrane proteins to enable toxin secretion into target cells. In our study, we demonstrate that an additional processing can occur when a cysteine protease is invovled. Upon toxin delivery in the target cell periplasmic compartment, the entry domain utilizes the inner-membrane partner, YajC, to translocate into the cytoplasm (unpublished data) (**Fig. 9**). We could speculate that the entry domain may interact with the inner-membrane protein translocator after the toxin domain is activated by auto-processing under reducing conditions, facilitated by the Cys-His-Asp triad. It is worth noting that the presence of Gly157-Gly158 and Cys187 play vital roles in ensuring toxin deployment and preventing premature toxin activation.

In the case of replacing Gly157-Gly158 to serine residues, the rate of cleaving of GyrB *in vivo* is slower compared to the wild-type, indicating that the toxin domain may be associated with the inner-membrane transporter and may not have full access to the substrate (**Fig. 6D,** lanes 2 & 4). On the other hand, when Cys187 is replaced, the safeguard of the toxin is now removed, which means it could prematurely remove the toxin within the inhibitor cell's periplasmic oxidizing compartment before delivery. This prediction was confirmed by observations in the immunoblot, where the CdiA filament was processed either with or without the presence of the cognate receptor (**Fig. 6D,** lane 9) and the attenuated toxin activity in the target cells (**Fig 6C**). However, there was another unexpected premature processing event that occurred in the Asp275Ala filament before binding to its receptor. This might be due to the instability of the filament caused by this point mutation (**Fig 6D,** lane 13), but the catalytic activity could be confirmed by the *in vitro* reaction.

Finally, the observed phenotypes in the target cells suggest that the toxin targets both GyrB and ParE simultaneously. However, considering the high level of catenated DNA in the intoxicated cells, we suspect that the most lethal effect of this process is the cleavage in ParE. A future study could investigate the targeting of the toxin at different stages of cell growth. It is known that topoisomerase IV is responsible for decatenating DNA after replication, aiding chromosome segregation (Zechiedrich et al., 1997). We hypothesize that when the toxin is delivered into actively growing cells, the target cells should retain the filamentous phenotypes due to the high expression level of topoisomerase IV required for segregation. However, if toxin is delivered at the stationary phase or later, the cells may not exhibit filamentation due to the plateau of growth, and gyrase becomes the sole target within the cells.

D) Experimental Procedures

Strain and plasmid constructions for CdiA-CT^{DBS100}

  Citrobacter rodentium DBS100 (CH87) served as the parental strain for constructing multiple downstream strains. To construct a *cdiB*^{STEC3}-TPS^{STEC3} substitution upstream of the StbD filament, a fragment of *cdiB*^{STEC3}-TPS^{STEC3} (S260) from pCH3779, amplified by oligos CH5561/CH5562, was cloned into pCH501 using NcoI/SalI and stored in EPI100*pir* strain. Subsequently, a fragment carrying the fragment of StbD (I585-S388), amplified with oligos CH5559/CH5560, was cloned into pCH3779 to create a final integrative plasmid pCH3780, which was stored in MFD*pir* strain. Finally, the pCH3780 was integrated into CH87 in the absence of diaminopimelic acid (DAP) and selected with LB-Kan50. Final integrant, CH4201 was confirmed by PCR screening. Subsequently, the full CdiB-TPS-StbD chimeric locus was completely excised from CH87 by rescue cloning via SphI digestion for 5 hours and re-ligated with T4 ligase for overnight at 16°C. The confirmed L-rhamnose inducible rescue clone, pCH8906, was selected with LB-Kan50 and stored in CH8896.

  CH8906 served as a framework to construct the expression plasmids. First, the wild-type CdiA-CT/CdiI^{DBS100} genes were amplified using oligos CH4346/CH4347 and cloned into pET21P with NcoI/SpeI to have the first toxin expression plasmid with C-terminal his tagged immunity protein, pCH13984. Later, mutant variants were introduced to pCH13984 using megaprimer PCR method. The mutant fragments generated from forward oligos CH4346, reverse oligo CH4347, and indicated mutant oligos, CH5425 (GG157-158SS), CH5453 (Gly158Ser), CH5336 (Cys183Ala), CH5458 (Cys187Ser), CH5381 (His258Ala), and CH5402 (D275A), were also cloned into pET21P using NcoI/SpeI to make the final expression plasmids, pCH6007, pCH6009, pCH998, pCH5997, pCH9108, and pCH9109. respectively. Additionally, several N-terminal His tagged constructs were created by cloning

into pMCSG63 with KpnI/XhoI. First, the CdiA-CT/CdiI^{DBS100} from CH13984 was amplified using CH5356/CH5338 to have N-terminal His tagged wild-type toxin (pCH1762). Finally, the full-length Cys183Ala toxin domain construct, pCH1764 was cloned by itself by using oligos CH5356/CH5357 from pCH998 and the truncated Cys183Ala domain starting from Ala159, pCH9103 was cloned by using oligos CH5765/CH5357 from pCH998.

These point mutations were also introduced to the CdiA-CT^{DBS100} domain by using overlap extension (OE)-PCR. The universal FH2-PT fragment was amplifying off pCH8906 with oligos CH5899/CH5950 and combined with the OE-PCR fragments using oligos CH5951/CH5952 amplified from indicated strains, pCH6007 (GG157-158SS), pCH6009 (Gly158Ser), pCH998 (Cys183Ala), pCH5997 (Cys187Ser), pCH9108 (His258Ala), pCH9109 (Asp275Ala). Finally, those full-length PCRs would be cloned into pCH8906 using KpnI/XbaI and selected with LB-Kan50 to generate the final chimeric plasmids, pCH8908 (GG157-158SS), pCH8909 (Gly158Ser), pCH8910 (Cys183Ala), pCH8911(Cys187Ser), pCH8912 (His258Ala), and pCH8913 (Asp275Ala), stored in CH8896.

Furthermore, some other constructs were built for different assays. First the constitutive expression plasmid for both wild-type and C183A CdiA-CT/CdiI^{DBS100} fused with the CdiA^{EC93} filament (pCH1004) using NheI/XhoI. Finally, the immunity gene was also cloned into both expression plasmid pET21P and pCH405Δ using KpnI/XhoI to have pCH9110 (oligos CH4497/CH4347) and pCH9102 (oligos CH4497/CH5338), respectively.

Strain and plasmid constructions for the substrates
Several steps were taken to build the chromosomal VSV-g tagged strains for FtsZ, GyrA, GyrB, ParC, and ParE. First, the VgG2-VSV fragment from pCH14397 was

subcloned into pCH902 to generate the integrative plasmid with EcoRI/KpnI, pCH5955. This plasmid subsequently served as a cloning vector using EcoRI/SpeI to clone fragments of FtsZ (oligos CH5429/CH5452), GyrA (oligos CH5694/CH5696), GyrB (oligos CH 5698/CH5700), ParC (oligos CH5767/CH5768), and ParE (oligos CH5702/CH5766), resulting in the final integrative plasmids: pCH5958, pCH7987, pCH7988, pCH7990, and pCH7991, respectively. These plasmids were ultimately integrated into MG1655 Δ*wzb* (CH7367) and selected on Cm66 to prevent the integration from reforming the plasmid, resulting in the final VSV-tagged strains: CH5960 (FtsZ), CH7993 (GyrA), CH7994 (GyrB), CH7992 (ParC), and CH7995 (ParE). Alternatively, the VSV-tagged was also introduced to a MG1655 Δ*wzb* Δ*ompF* strain (CH8900) to serve as a target cell line for a co-culture assay (CH8901), which later transformed with pTrc99a and pTrc99aKX::ompF(K12) to have final target strains, CH8914 and CH8915. Notably, MG1655 Δ*wzb* Δ*ompF* was generated through a P1-transduction to introduce Δ*ompF*::Kan (JW0912) to CH7367 and later being cured with pCP20 (pCH48).

All gyrase expression plasmids were constructed by amplifying each subunit from the X90 strain and cloning into pET21b with NheI/XhoI. Oligos CH5695/CH5697 were used to amplify GyrA subunit, and CH5699/CH5701 were used for GyrB subunit, resulting in pCH8187 and pCH8189, respectively. Subsequently, pCH8193 was created using oligos CH5695/CH5714 to amplify the ATPase domain of GyrB from pCH8189. Finally, pCH8193 served as a template for introducing site-directed mutagenesis with forward oligos CH5699, reverse oligos CH5714, and indicated site-directed mutation oligos Gly101Ser (CH5824), Gly102Ser (CH5825), GG101-102SS (CH5748), and GG113-114SS (CH5749) to construct pCH8511, pCH8512, pCH5358, and pCH5359, respectively. Additionally, the ATPase

135

domains of EnvZ (oligos CH5742/CH5743) and HtpG (oligos CH5744/5745) were also cloned into pET21P with NcoI/XhoI. Alternatively, the first 14 residues deletion construct, pCH4032, was created with oligos CH6080/CH5714 and cloned into pET21P using Kpn/XhoI.

Co-culture protein purification

Target cells carrying the proteins of interest, including ATPase^{GyrB}-His6 from pCH8193, ATPase^{GyrB} (GG101/102SS)-His6 from pCH5358, ATPase^{GyrB} (GG113/114SS)-His6 from pCH5359, ATPase^{GyrB} (G101S)-His6 from pCH8511, ATPase^{GyrB} (G102S)-His6 from p8512 ATPase^{EnvZ}-His6 from pCH5355, and ATPase^{HtpG}-His6 from pCH5356 were OD600=0.6 and induced expression by addition of isopropyl β-d-1-thiogalactopyranoside (IPTG) to 1.5 mM for 1 hour. Subsequently, equal density of induced target cells and inhibitor cells that carrying CDI locus, pCH9099 were mixed in LB media in a 10 mL final volume and grown for additional 30-minutes before collecting the co-cultured pellet from 5-minute centrifugation at 6,000RPM. The cell pellets were lysed in urea lysis buffer [20 mM Tris-HCl (pH 8.0), 8 M Urea, 150 mM NaCl] and 50 µL of Ni-NTA resin was added to the cleared supernatant. Bound proteins were washed thrice in urea lysis buffer supplemented with 25 mM of imidazole before elution in urea lysis buffer supplemented with 250 mM of imidazole. Purified proteins were analyzed on analyzed on SDS-PAGE (10 % acrylamide, 110V for 1 hour). The fragments of ATPase^{GyrB} were further separated with RP-HPLC and analyzed with ESI-MS.

In vitro proteolytic and binding reactions

All protein expressing *E. coli* cells (CH2016) were grown to OD$_{600}$=0.6 and induced expression by addition of isopropyl β-d-1-thiogalactopyranoside (IPTG) to 1.5 mM for 1 hour and stored in respective buffer plus 50% glycerol at -20 °C. All proteins of interest,

G15-ATPase$^{\text{GyrB}}$-His6 expressed from pCH4032, GyrB-His6 from pCH8189, and, wild-type toxin complex CdiA-CT/CdiI$^{\text{DBS100}}$-His6 from pCH13984, mutant variants from pCH998 (Cys183Ala), pCH5997 (Cys187Ala), pCH6007 (GG157-158SS), pCH6009 (Gly158Ser), pCH9108 (His258Ala), pCH9109 (Asp275Ala), and CdiI$^{\text{DBS100}}$ from pCH9110 were lysed in native buffer A [20mM Tris-HCl (pH 7.5), 150 mM NaCl, 25mM imidazole] in the presence of 1 mg/mL lysozyme for 30 minutes, then cell lysates were sonicated for two rounds to lyse cells completely before 15-minute centrifugation at 15,000 RPM to remove cell debris. Then, Ni^{2+}-NTA resins were added to the cleared supernatants to bind His6-tagged proteins for 1 hour at 4 $^{\text{o}}$C. All protein bound resins except for the toxin complex (pCH13984) were washed thrice with native buffer A before eluting with buffer A supplemented with 250 mM imidazole. For the toxin complex (pCH13984), the toxin domain was stripped from the immunity by denature buffer [6 M Guanidine-HCl, 20 mM Tris-HCl (pH 7.5)]. Finally, all proteins were dialyzed twice for 4 hours in the storage buffer [20mM Tris-HCl (pH 7.5), 150 mM NaCl]. Protein concentrations were determined using an extinction coefficient at 280 nm of 61, 210 M^{-1} cm^{-1} (wild-type CdiA-CT$^{\text{DBS100}}$, GG157-158SS, G158S, H258A, and D275A), 61, 085 M^{-1} cm^{-1} (C183A & C187A), 14, 440 M^{-1} cm^{-1} (CdiI$^{\text{DBS100}}$).

All *in vitro* reactions for conducted in 20 µL final volume. For autoproteolysis analysis, equal concentration (1 µM) of purified wild-type toxin, the indicated mutant variants, along with the cognate immunity were mixed in 1 x reaction buffer [20mM Tris-HCl (pH 7.5), 150 mM NaCl]. The reactions initiated in the presence of the reducing reagent DTT and incubated for 30 minutes at 37 $^{\text{o}}$C. Similarly, *in vitro* reactions for cleaving gyrase were conducted in the same reaction buffer with the presence of either DTT or β-ME, including GyrB-His6, ATPase$^{\text{GyrB}}$-His6, G15-ATPase$^{\text{GyrB}}$-His6, were incubated with the

wild-type and indicated mutant toxin variants for 30 minutes (or at the indicated timepoint) at 37 °C. Additionally, the nucleotide blockage reactions were conducted with wild-type toxin and GyrB-His6 protein in 1x reaction for 30 minutes at 37 °C by titrating different concentrations of ATP and ANP accordingly. Ultimately, all final reactions were analyzed on SDS-PAGE (10 % acrylamide, 110V for 1 hour).

Proteins for *in vitro* binding studies, including His6-A159-CT(Cys183Ala) from pCH8858, G15-ATPase^{GyrB}-His6 from pCH4032, and ATPase^{GyrB}-His6 from pCH8193, were lysed and purified in a denaturing buffer [6 M Guanidine-HCl, 20 mM Tris-HCl (pH 7.5)] with three washes in buffer A before elution in buffer A supplemented with 250 mM imidazole. All proteins were refolded through two rounds of 4-hour dialysis in a storage buffer. Subsequently, the His6-tag was removed from the refolded A159-CT(Cys183Ala) using TEV protease. Protein concentrations were determined using an extinction coefficient at 280 nm of 60, 960 M^{-1} cm^{-1} [A159-CT(Cys183Ala)], 14, 440 M^{-1} cm^{-1} (ATPase^{GyrB}-His6), and 12, 950 M^{-1} cm^{-1} (G15-ATPase^{GyrB}-His6). Equal concentrations (5 µM) of confirmed untagged A159-CT(Cys183Ala) and His6-tagged bait proteins were incubated for 1-hour at 4 °C in 1x reaction buffer with a final volume of 600 µL. Afterward, 15 µL of the mixtures were collected as the Input fraction before adding 50 µL of Ni-NTA resin for 30 minutes binding at 4 °C. A brief at max speed was used to collect Ni-NTA. Another 15 µL of the supernatant was collected as the Unbound fraction before removal. The resin was then washed once by 600 µL of buffer A before the final elution in 60 µL of buffer A supplemented with 250 mM imidazole. Finally, 15 µL of the eluted proteins were also collected as the Bound fraction. Subsequently, all fractions were analyzed on SDS-PAGE (10 % acrylamide, 110V for 1 hour).

Protein purification and RP-HPLC analyses

To determine the mass of the autoproteolytic products, two N-terminal His-tagged constructs were used, pCH1762 (wild-type toxin) and pCH1764 (Cys183Ala mutant). Expressing cells (CH2016) with the plasmids were grown to OD_{600}=0.6 and induced expression by IPTG to 1.5 mM for 1 hour. The collected cell pellets were lysed completely in denature buffer [6 M Guanidine-HCl, 20 mM Tris-HCl (pH 7.5)] before 15-minute centrifugation at 15,000 RPM to remove cell debris. Then, Ni^{2+}-NTA resins were added to the cleared supernatants to bind His6-tagged proteins for 1 hour at 4 °C. Finally, all proteins were washed twice in denature buffer supplemented with 25 mM imidazole before eluting in denature buffer supplemented with 50 mM EDTA.

On the other hand, the mass of cleaved ATPase domain was determined by purifying the C-terminal His-tagged ATPase^{GyrB} domain from a co-culture assay between cells carrying pCH9099 (DBS100 inhibitor cells) and pCH8193 (ATPase^{GyrB} expressing cells). Cells with pCH8193 were induced with IPTG to 1.5mM for 30 minutes before mixing in the inhibitor cells. After 30 minutes of co-culture, the cell pellet was collected and lysed in urea lysis buffer. Then, a 15-minute centrifugation at 15,000 RPM cycle, was applied to obtain cleared supernatant. Ni^{2+}-NTA was added to supernatant to bind the His-tagged protein. Finally, all proteins were washed twice in urea lysis buffer supplemented with 25 mM imidazole before eluting in urea lysis buffer supplemented with 50 mM EDTA.

Collectively, all purified proteins analyzed by RP-HPLC as described in Halvorsen et al., 2021. Briefly, samples were passed through a 0.22-μm cellulose acetate spin filter (Costar) and then injected onto a Vydac 15- by 300-mm C4 column in buffer A (0.06% trifluoroacetic acid) at a flow rate of 1 ml/min. After 5 min, the column was developed with a

0 to 100% linear gradient of buffer B (0.052% trifluoroacetic acid in 80% acetonitrile) over 60 min, and eluted proteins were detected by absorbance at 214 nm using a Waters UV spectrophotometer. All HPLC-purified proteins were dried by SpeedVac and redissolved in formic acid for electrospray ionization-mass spectrometry (ESI-MS).

Co-culture assay and immunoblot
Inhibitor (CH8906, CH8908-8913) and target (CH8914, 8915) cells were grown to mid-log phase. Then, 0.2% L-rhamnose was added to the inhibitor cells for CDI induction prior to cell mixing. Induced inhibitor cells and target cells were mixed in test tube with 1: 1 ratio to a final volume of 3 mL in LB media supplemented with 0.2% L-rhamnose. Allowed cells to mix on shaker for 3 hours at 37°C, the cells were spotted on LB agar plate supplemented with 50 μg/mL kanamycin for inhibitor cell selection and LB agar plate supplemented with 150 μg/mL ampicillin for target cell selection. Colonies from both plates were enumerated and calculated as competitive indices (CI). CI is a ratio of target: inhibitor cells (CFU/mL) at the final timepoint divided by the ratio of target: inhibitor cells (CFU/mL) at the initial timepoint.

1-mL of each co-culture sample was collected, and the cell pellets frozen at –80 °C before lysing in 100 µL of 8 M urea lysis buffer [20 mM Tris-HCl (pH 8.0), 8 M Urea, 150 mM NaCl]. Lysates were normalized and analyzed by Tris-tricine SDS-PAGE on a two-tier 5%/15% polyacrylamide gels run at 110 V (constant) for 5 hours. Gels were soaked for 5 min in 25 mM Tris, 192 mM glycine (pH 8.6), 10% methanol, then electrotransferred to PVDF membranes using a semi-dry transfer apparatus at 17 V (constant) for 1 hour. Transferred membrane was blocked in 4 % non-fat milk in 1 x PBS for 30 minutes at room temperature and incubated with two primary antibodies [1:150,000 dilution of mouse anti-

VSV-g (Sigma-Aldrich) and 1:10,000 dilution of rabbit anti-TPS] in 0.1 % not-fat milk in 1 x PBS overnight at 4 °C. Blot was then incubated with 800CW-conjugated goat anti-rabbit IgG (1:25,000 dilution, LICOR) and 680LT-conjugated goat anti- mouse IgG (1:25,000 dilution, LICOR) in 1× PBS. Immunoblots were visualized with a LI-COR Odyssey infrared imager.

Supercoiling DNA analysis

Inhibitor (CH8906 and CH8910) and target (CH69) cells were grown to mid-log phase. Subsequently, 0.2% L-rhamnose was added to the inhibitor cells for CDI induction before cell mixing. Induced inhibitor cells and target cells were mixed in test tube with 1: 1 ratio, resulting in a final volume of 10 mL in LB media supplemented with 0.2% L-rhamnose for 30 minutes. Co-cultured cells were collected, and plasmid DNA samples were isolated with standard miniprep procedure.

Next, plasmid DNA samples from both co-cultures were analyzed with chloroquine agarose gel, following the protocol described in Gibson et al., 2020, with slight modifications. A 1 % agarose gel supplemented with 1 µg/mL of chloroquine diphosphate was used to resolve the plasmid DNA samples. Ethidium bromide was omitted from the gel, and the electrophoresis was conducted in the dark using 1 x TAE (40 mM Tris-acetate, 1 mM EDTA) running buffer for 4 hours at 100V or until bromophenol blue had migrated off the gel. Subsequently, the gel was soaked in water containing 1 µg/mL of ethidium bromide for 30 minutes. It was then washed thrice with distilled water before imaging with a UV lightbox. It is worth noting that all the parts of the electrophoresis apparatus were soaked in DI water overnight before conducting the chloroquine gel experiment to prevent unexpected DNA intercalation caused ethidium bromide.

Fluorescent microscopy

CDI+ cells with sfYFP (CH9099) were co-cultured with the mKate target cells with an empty vector (CH9101) and pCdiI^{DBS100} (CH9102) with OD_{600} = 0.2 and 0.4 in 2 mL of LB media, respectively. The cells were mixed for 3 hours on shaker at 37°C. Then, 1 mL of each mixture was collected, and the cell pellets were washed once with 1 mL of 1 x PBS. The washed pellets were resuspended in 100 μL of 1 x PBS supplemented with 1 μg/mL of Hoechst 33342 for 5 minutes. The stained cell mixtures were then washed once with 1 mL 1 x PBS before having the cell pellet resuspended in 100 μL of 1 x PBS. Finally, 5 μL of each cell suspension was added and air dried on agarose pad (1 % agarose in 1 X PBS) before placing the cover slips. The prepared samples were visualized on ECHO Revloved- R4 microscope under 100x objective lens with immersion oil.
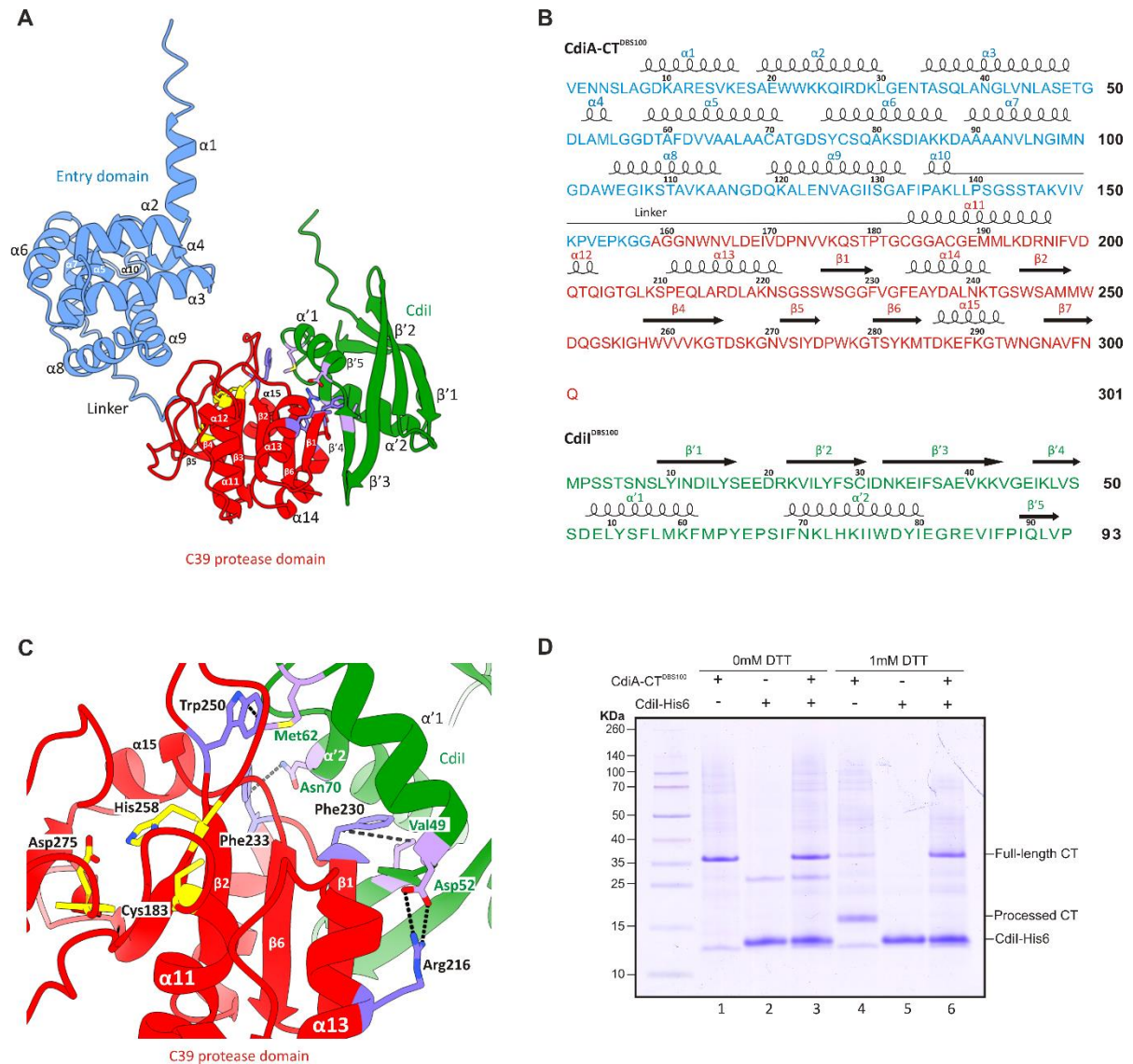
Figure 1: CdiA-CT/CdiI [DBS100] unique binding.

A. AlphaFold model of the CdiA-CT/CdiI[DBS100] complex with the CdiA-CT cytoplasmic entry domain depicted in blue, the C39 protease domain in red, and the CdiI in green.

B. The sequences of the corresponding proteins from panel A are displayed with alpha-helices and beta-strands appropriately annotated for clarity.

C. The interface between CdiA-CT and CdiI. The key interactive residues are shown, Arg216, Phe230, Phe233, and Trp250 of CdiA-CT make contacts with Asp52, Val49, Asn70, and Met62 of CdiI, respectively.

D. Unique auto-processing of CdiA-CT[DBS100] under reducing condition.

143
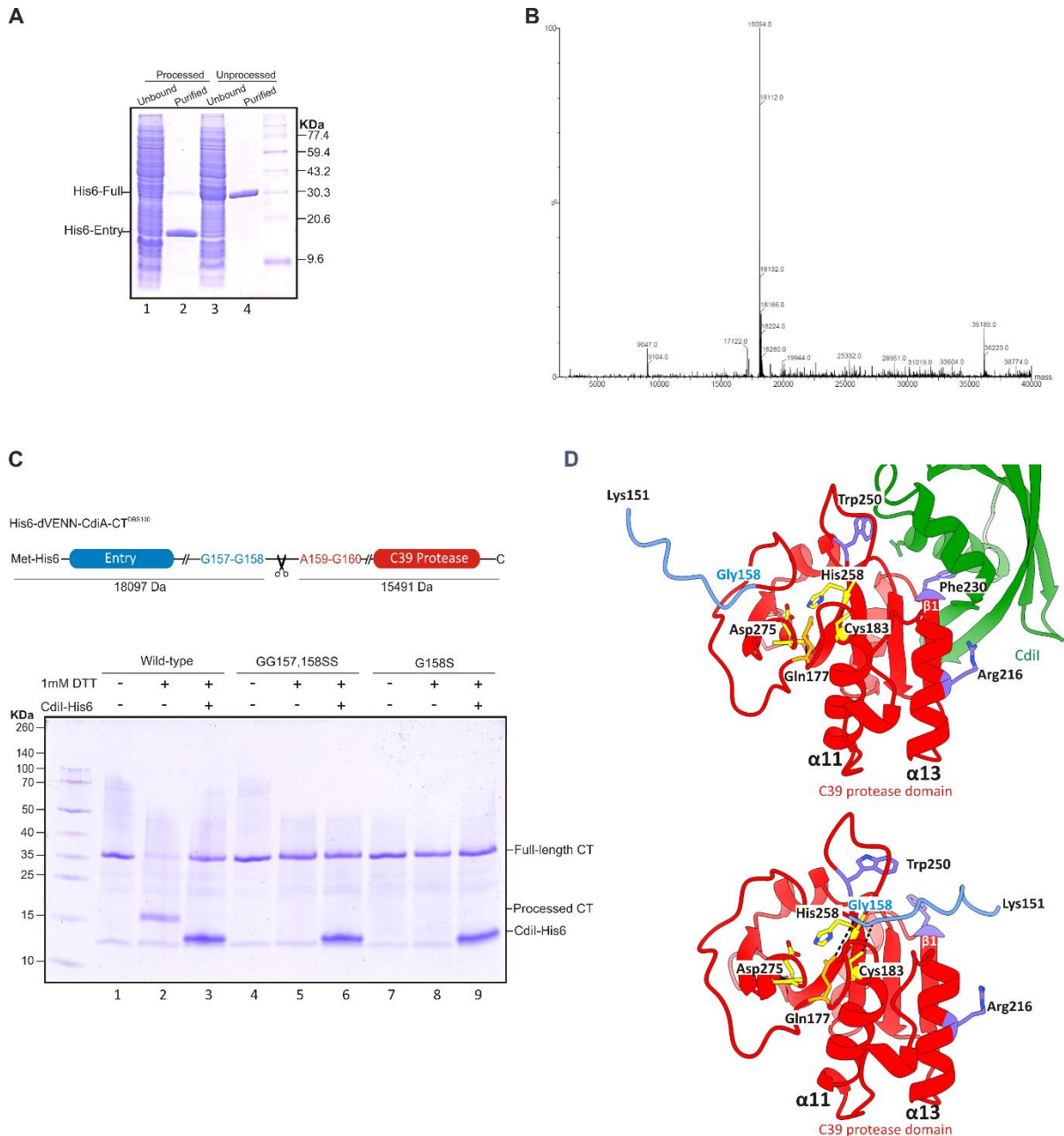
Figure 2: CdiA-CT^DBS100 undergoes auto-proteolytic at Gly158.
A. Purified N-terminal His6 tagged proteins as either a processed or a full-length product.
B. ESI-MS analysis on the His6-processed product.
C. A schematic illustration of the auto-proteolytic site at the linker region between Gly158 and Ala159. Mutations at Gly158 prevent the C39 protease from cleaving itself.
D. AlphaFold models reveal the subtle conformational shifts within the CdiA-CT protein as it transitions from its bound state with CdiI to its free form, enabling the linker region to become exposed to the active binding pocket where the catalytic residue, Cys183 directly interacting with the cleavage site, Gly158.
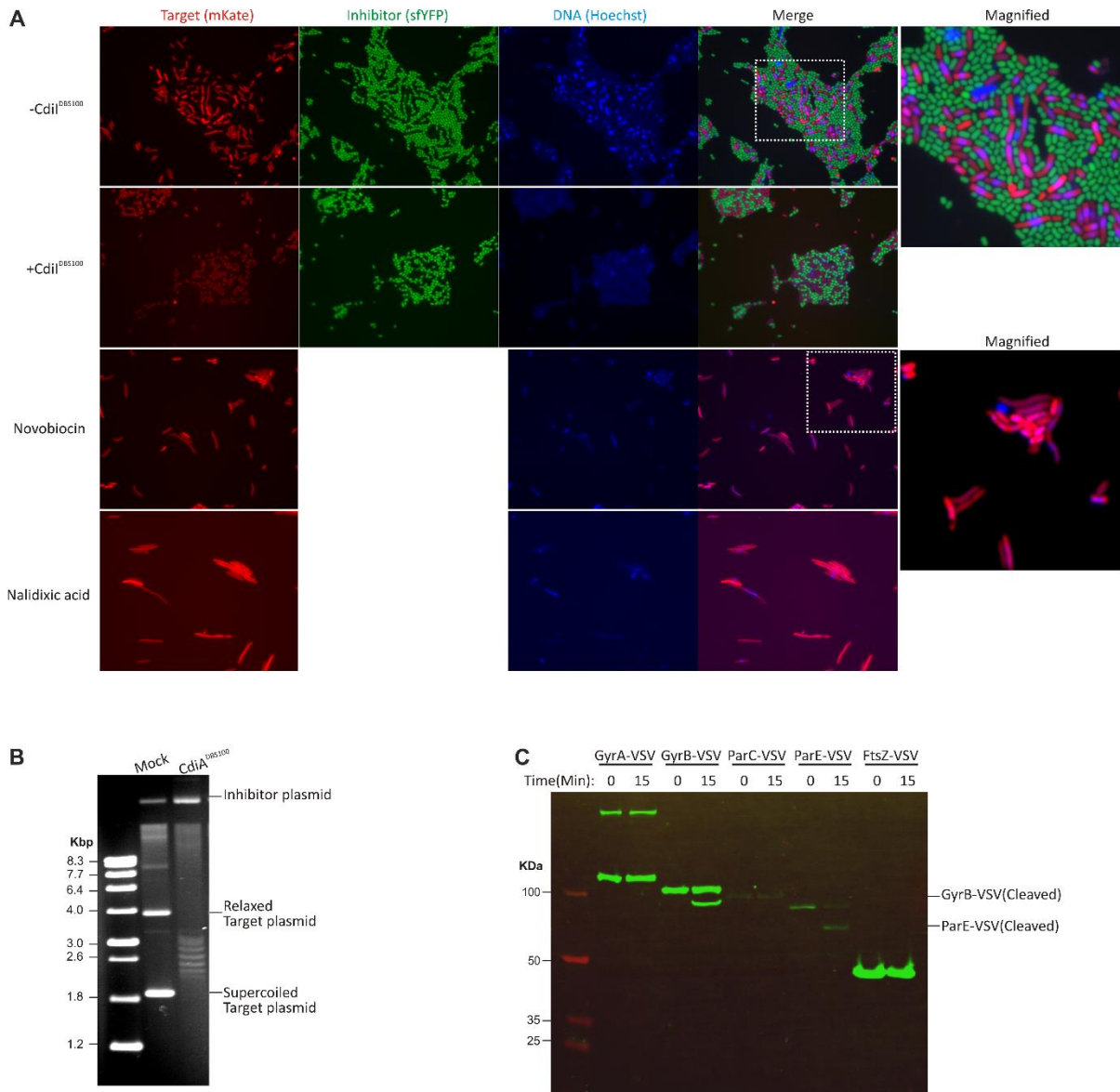
Figure 3: Identification of protease substrates in target cells.

A. Fluorescent micrographs of *E. coli* cells from co-culture assays with mKate-expressing inhibitor cells and green-labeled sfYFP target cells either carrying the empty vector or the cognate immunity protein. The chromosomal DNA was stained by Hoechst 33342. Target cells intoxicated by novobiocin and nalidixic acid were also included for comparison.

B. Chloroquine agarose gel resolution of the plasmid DNA from the co-culture assays revealing different DNA topoisomers of the intoxicated target cells.

C. Anti-VSV immunoblot was performed to reveal the tagged candidate proteins of interest from a 15-min co-culture assay.
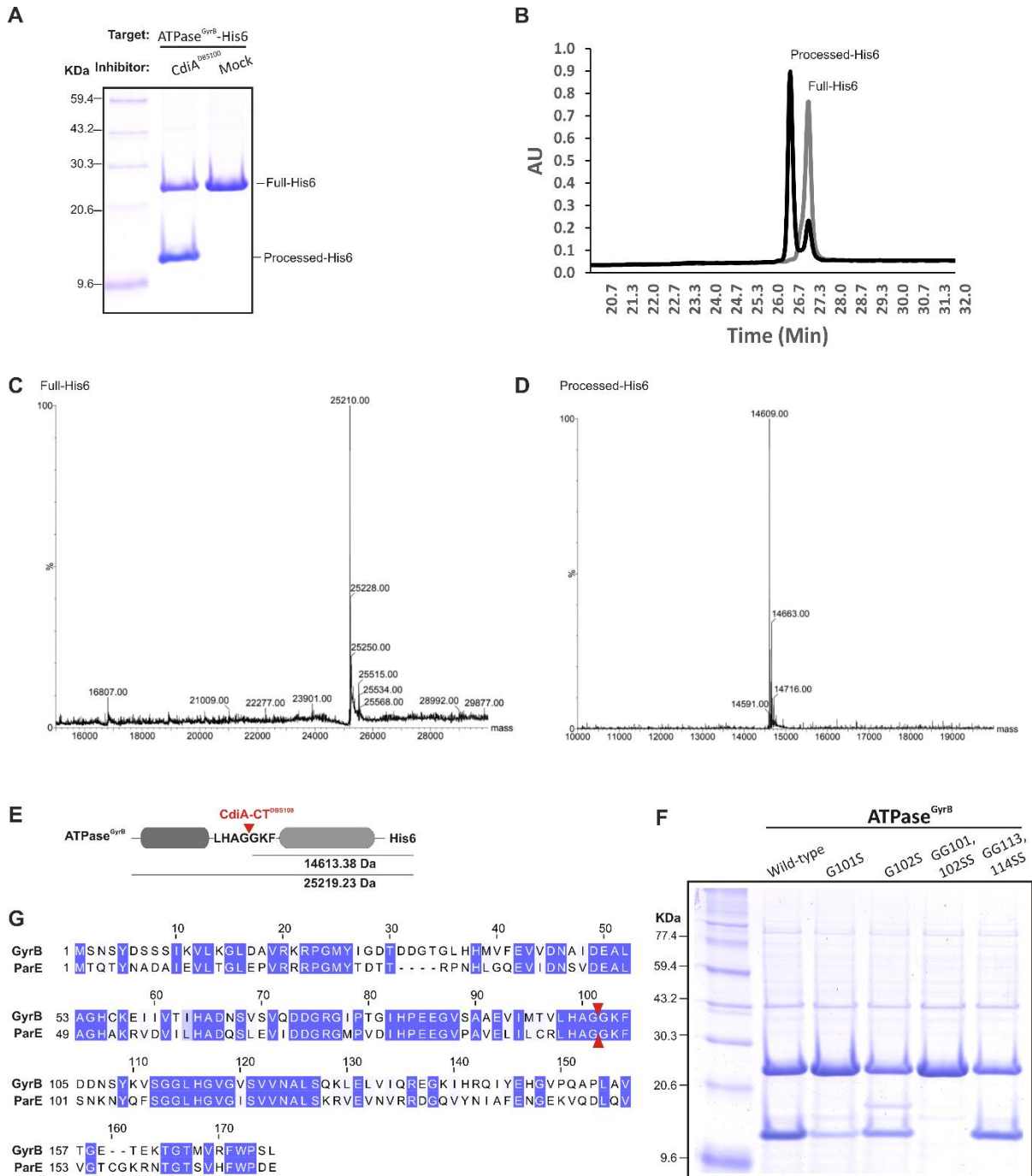
Figure 4: Protease cleaves between Gly101 and Gly102 of the ATPase domains of GyrB.

A. The purified ATPase^{GyrB} domain from a co-culture assay undergoes processing, resulting in a cleaved band running above 9.5 KDa.
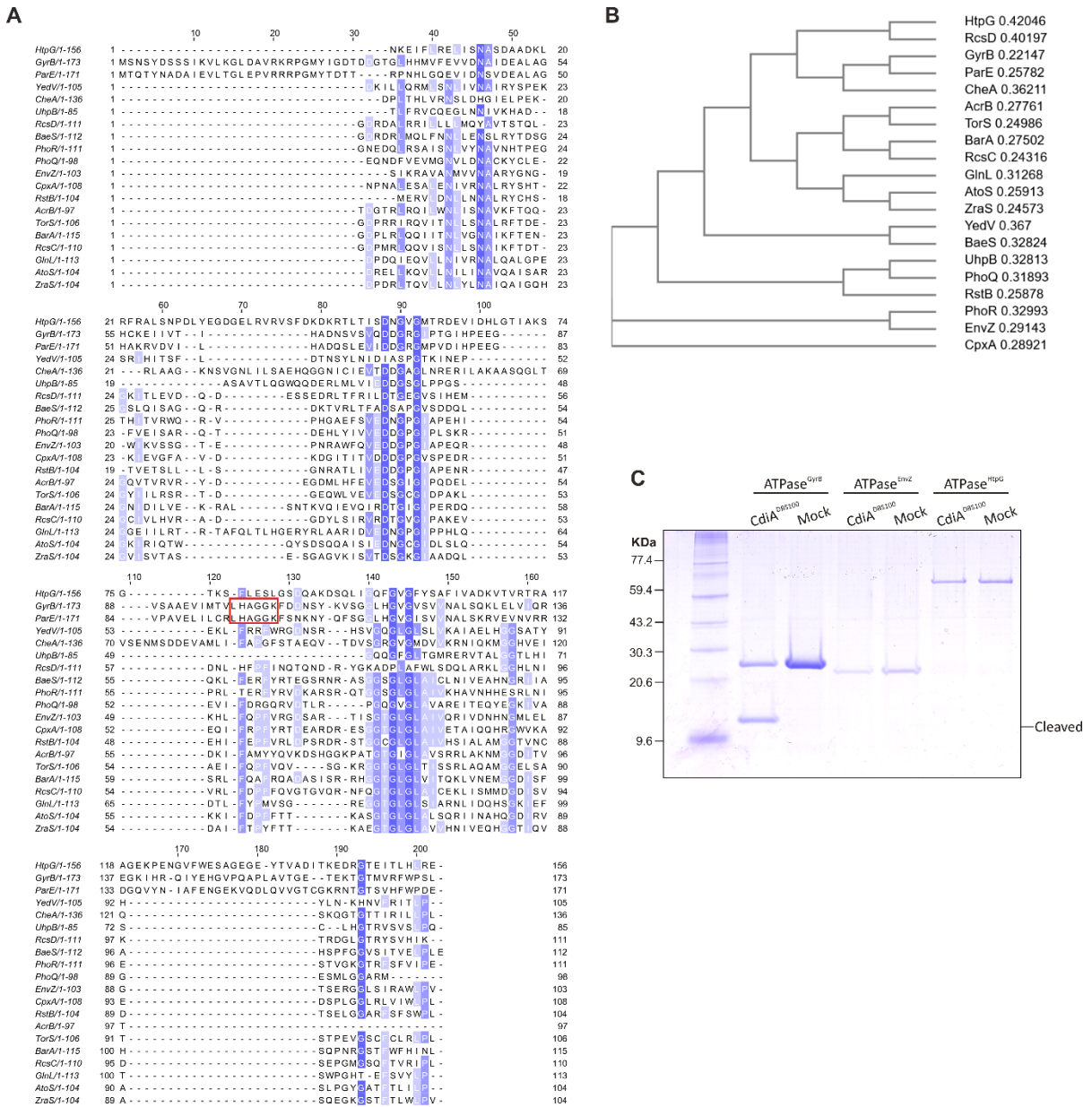
B. High-performance liquid chromatography (HPLC) is used to separate the mixture of full-length and cleaved bands from samples shown in panel A.

C. and D. ESI-MS analyses of both full-length and cleaved bands.

E. A schematic illustration of the cleavage site of the CdiA-CT^{DBS100} on the ATPase domain of GyrB.

F. Purified wild-type and mutants of ATPase^{GyrB} from a co-culture assay.

G. A sequence alignment between the GyrB and ParE ATPase domains with indicated cleavage site.

146

Figure 5: Protease targets specifically at the ATPase domains of GyrB and ParE.

A. A sequence alignment of GyrB, ParE, and other selected ATPase domains found in *E. coli* proteins. The specific cleavage motif LHAGGKF is boxed in red.

B. A phylogenetic tree of the ATPase domains generated from panel A.

C. The purified ATPase domains of GyrB, EnvZ, and HtpG from co-culture assays revealed that the toxin can only cleave ATPase^{GyrB}.

A

Cys183 Cys187

```
DBS100(159-293)  AGGNWNVLDEIVDPNVVKQSTPTGCGGACGEVMLKDRNIFVDQTQIGTGLKSPEQLARDLAKNSGSLWGGFV-----GF---EAYDA---------
SunT-Bs(1-148)   --MN--KKKKY--VHTKQFNSHDCGLACISSLTKFHNLNYGIDFLLDLIG-D---------KEGYILRDLIVIFKKMGI-KTRPLELQENK-TFEA
SunT-Bc(8-150)   ------------QEEYDCGIACVASLFKLYHLHHGINYLDQVT-Y---------KNGYDLKDLLALFSQVNHFSCKLVEINKED-IEEA
PedD(1-150)      --M--WTQKWIHKYYTAQVDENDCGLAALNMILKLYGSDYMLAILPQLAKTT---------ADGTTVLGLVKAAKHLNL-NAELVRADMDAL---T
McnB(17-158)     --IG--LRRRI--PVVHQTLSSLCGLACLSMICGHYGRHIDLSTLRQFNLS---------ALGTTLAGLITELGSQLCM-ETRLFSLDLNL-----
CvaB(14-155)     -DLR--WQRRV--PVIHQTETAECGLACLAMICGHFGKNIDLIYLRKFNLS---------ARGATLAGLINGIAEQLCM-ATRLLSLELDE-----
MchF(20-161)     ------RV--PVIHQTETAFCGLACLAMICGHFGKNIDLISLRKFNLS---------ARGANLAGLINGIAEQLCM-VTRLLSLELDE-----
ScnT(1-139)      ----------M--NIILQNNEEDCLLACYTMLLNDLGHKVPLYEIYDKDTLP---------ADGLNVLYLLSLSDRFSV-KINAYHASFDE-LLKV
MutT(1-139)      ----------M--KLVLQNNEQDCLIACYAMILSSFGCNVSLEDLYLDEYIP---------PDGLLVSYIHSLNTKFRL-KMQVLKGDGQQ-VLEY
NukT(1-140)      ------L--KIALQNSDQDCLLACYSMILSLFGKNVSINSLYKREMIP---------PDGLLVLYIKELNIKYEL-NMKVYRIKDKEKTFRV
IcnDR3(1-140)    ----------M--KIVLQNNEQDCLLACYSMLGLFGRDVAIHELYSGEMIP---------PDGLLVLYLKNINMKHQV-SMHVYKTDKKNSPNKI
BT_4288(1-147)   --M--QTRQF--PWEYQMDAKDCGPACIKMIAKLYGKYYSFQYLLDLCGIT---------REGVLFLDISYAAFKIGL-RTVAVKATMEN--LTN
MrsT(1-142)      ------MRRRV--PLVRQMGQYECGPACLTMILSLYGSTISLNKISEQCDAQ---------RNGVLVSILKSVSEYYGL-NCKVYQVSFKD--LKK
```

```
DBS100(159-293)  LNKTGSWSAMMWDQ--GSKIGHWVVVKGTDSLGNVSLYDP-WKGTSYLMTDKLLKGTW---------
SunT-Bs(1-148)   KQIKLPLIALLEGE----EYGHYITIYEI-RNLYLLVSDFD-KDK-ITLKKLDFESKLTNFILEIDKELIPEKNL------
SunT-Bc(8-150)   KHIEGGLIALINKQVNESYEGHYIIIYKR-LKNKLVSDFD-NDK-ISTLTIDLFKKHTLGIFLLIESKNSNSDKLT-----
PedD(1-150)      ASQLPLIPVIVLVFKK---NKLPLYYVLYQV-TELDLILGDFDPTVK-TTLLSKSQFAKEWLQIAIIIALTVKYKPIL-----
McnB(17-158)     SVLLKLPLILIW------EFSHFVVLVSV-RLNHFVLHDPLF-RGLRRTVGLAMSQCLGVALEVWLGTELFVQET------
CvaB(14-155)     RVLLTPLFILIW------DFSHFVVLVSV-RLNRYVLHDPLF-RGLIRYLSRLEMSRYLGVALEVWLGTELQSET------
MchF(20-161)     GALLMLPLILIW------DFSHFVVLVSV-RLNRYVLHDPLF-RGLRRYLGRLEMSRYLGIALEVWLGLFLAFTQQIRIS
ScnT(1-139)      YGEKLQRMLLILIW-------NNDHFVVLWDKC-TLKSVNLMDPL-SLGLTITLPISSLLDNLSSYILIFSKQLNYRPVL-----
MutT(1-139)      IYQHLCKAIAYW------KNSLHFIVVLDKC-TLKSVNLMDPL-SLGLTITLPISSLLDNLSSYILIFSKQLNYRPVL-----
NukT(1-140)      ISKILKLPLILIW------DLNHFVILKNV-LKHIELVNPL-EILGKVLSKLMLLEHLSNVLLMFDLKLDLTKD-------
IcnDR3(1-140)    FYPKMLPVIIQW------NDNHFVVLTKI-YRKNVTLIDPL-AIGLKVLYNYNDLMKKLSGYIITLSLPLSLTKKL-----
BT_4288(1-147)   --RIPLPLTILIW------DQHFIVLNYKT-LKKGKIYVSDPL-KGLLLSYPELDFKDRWYKEGEEFGMLMVLEPMANFKQIE
MrsT(1-142)      YINSYLPCTIFW------DERHFVVLLEQI-LKKGLFHLIDPL-NRGLKLLLSELFKRHYSKVILTFKKSD-----KFKEMM
```

His258   Asp275

B

C

D

E

148
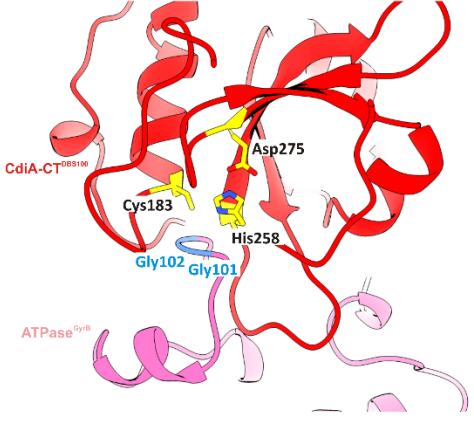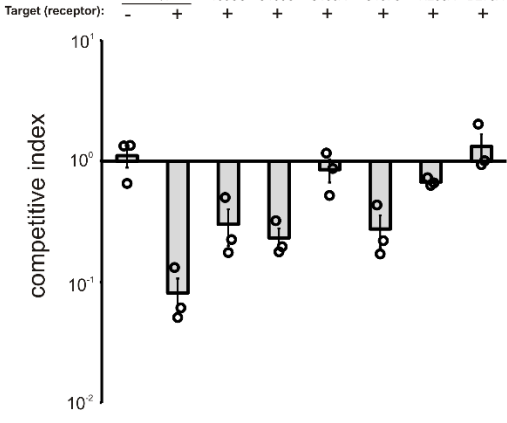
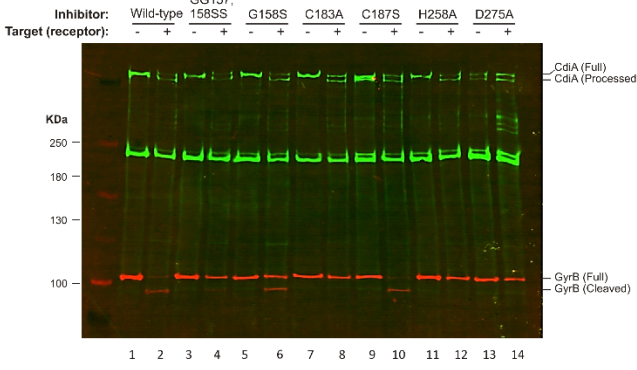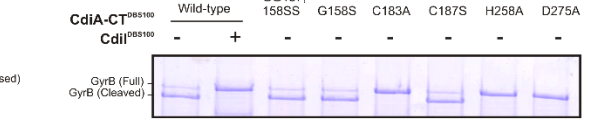**Figure 6: CdiA-CT^{DBS100} active site.**

A. A sequence alignment between CdiA-CT^{DBS100} and selected C39 peptidase domains, including SunT_Bs (*Bacillus subtilis*); SunT_Bc (*Bacillus cereus*); PedD (*Pediococcus acidilactici*); McnB, CvaB, MchF (*Escherichia coli*); ScnT (*Streptococcus pyogenes*); MutT (*Streptococcus mutans*); NukT (*Staphylococcus warneri*); lcnDR3 (*Lactococcus lactis*), BT_4288 (*Bacteroides thetaiotaomicron*); and MrsT (*Bacillus sp.* HIL-Y85/54728).

B. An AlphaFold model of the active residues, Cys183, His258, and Asp275 orients towards the cleavage site, Gly101 and Gly102 of GyrB.

C. Competitive index scores $[(CFU_{t=final}^{targets}/ CFU_{t=final}^{inhibitors})/(CFU_{t=0}^{targets}/ CFU_{t=0}^{inhibitors})]$ of wild-type CdiA-CT^{DBS100} and indicated mutants in liquid co-culture assay (n=3).

D. A dual channel immunoblot of lysates from a co-culture assay. The green channel distinguishes CdiA processing pattern of wild-type and indicated mutants in the absence and presence of its cognate receptor using anti-TPS antibodies. The red channel detects the level of cleavage of GyrB-VSV in target cells with anti-VSV antibodies.

E. The purified GyrB-His6 was incubated with indicated toxins for 30min at 37°C. Reactions were analyzed on SDS-PAGE to visualize cleavage.
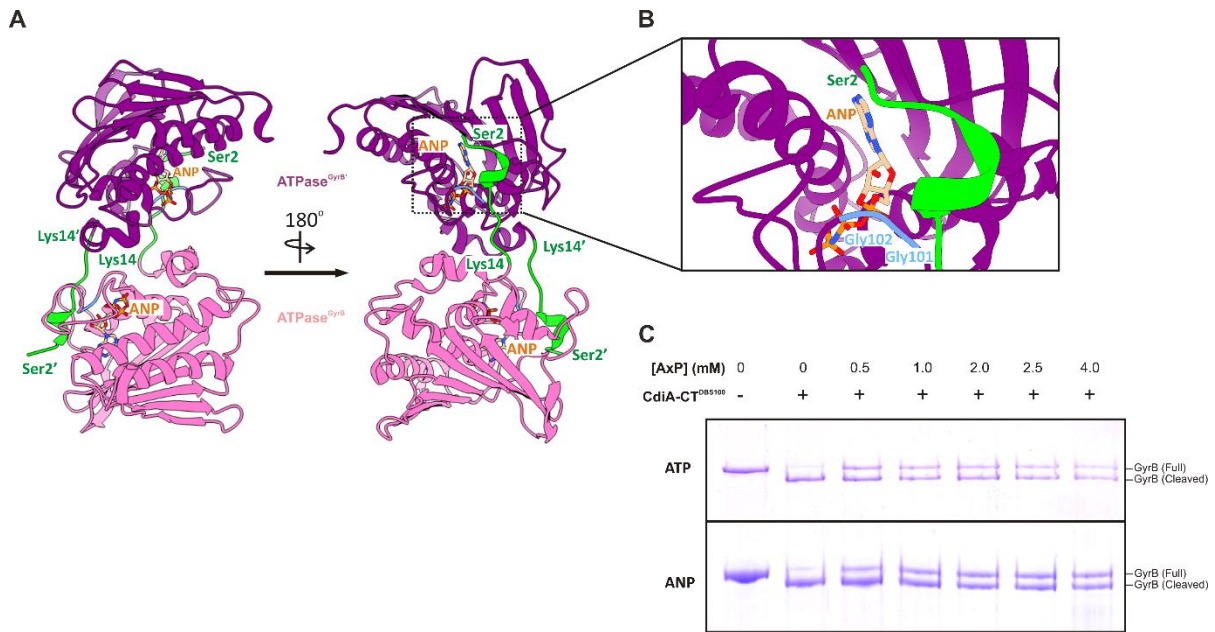
Figure 7: ATP bound form of GyrB prevents toxin recognition.

A. A crystal structure of ATPase domains of GyrB from *E. coli* (PDB_ID 4WUD) reveals a homodimeric structure with one monomer in pink, ATPase^{GyrB} and its cyan counterpart, ATPase$^{GyrB'}$. The first fourteen residues (Ser2-Lys14) from both monomers extend over each other to cover the ATP binding pocket. The structure was obtained through co-crystallization with Mg^{2+} and the ATP analog, adenylyl-imidodiphosphate (ANP) (Hearnshaw et al., 2015).

B. A closer look at the ATP binding pocket with highlighted Gly101 and Gly102 interacting with ANP.

C. *In vitro* proteolysis reactions were conducted using purified toxin and GyrB with the presence of varying concentrations of ATP and ANP.
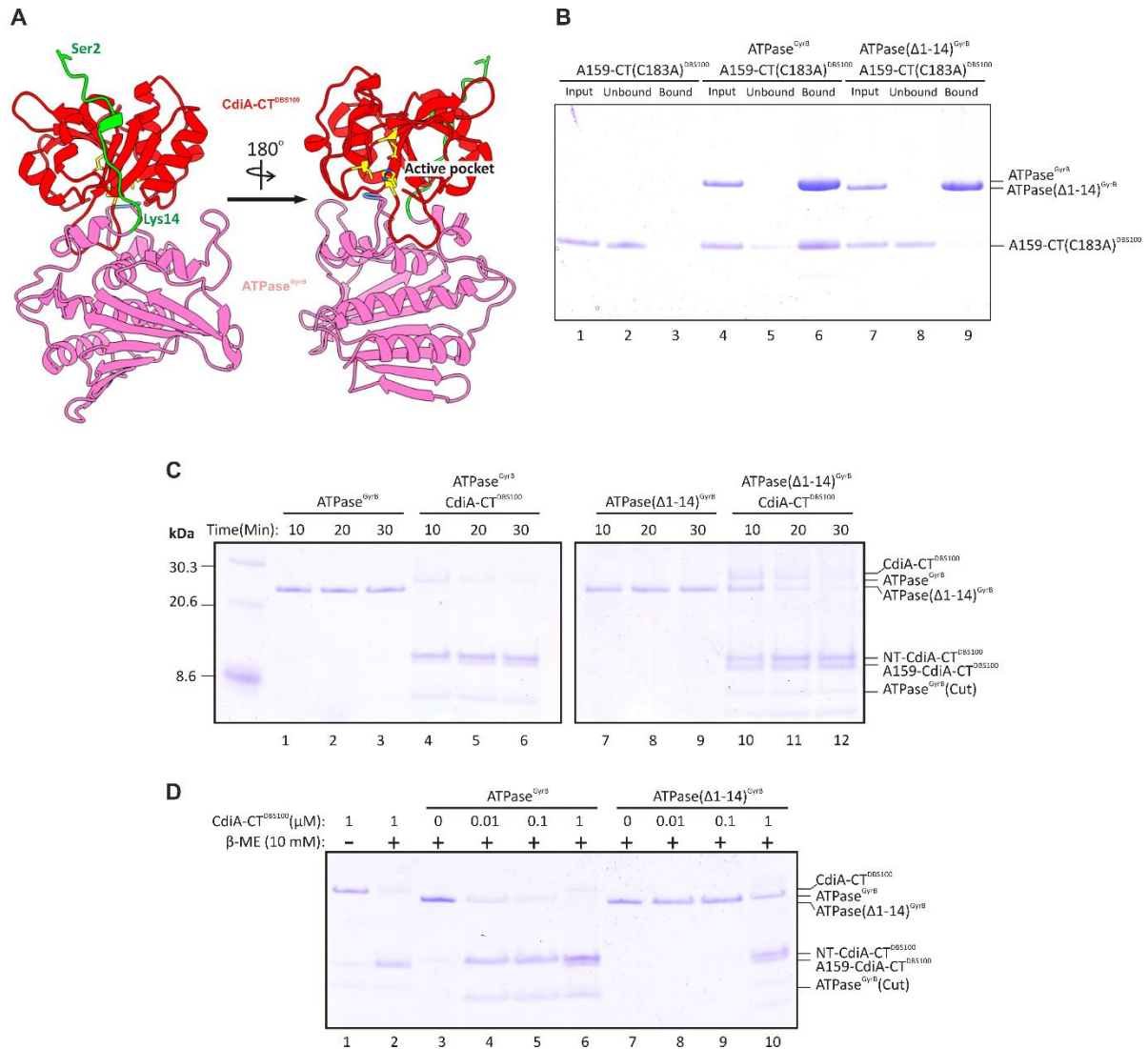
Figure 8: ATPase^GyrB (Ser2-Lys14) is important for toxin binding.

A. An AlphaFold model predicts a potential interaction between ATPase^GyrB with CdiA-CT^DBS100. The toxin domain utilizes the first fourteen residues from GyrB (Ser2-Lys14), labeled in green, to anchor itself onto the ATPase domain, allowing the active pocket to access to the Gly101 and Gly102.

B. *In vitro* binding reactions were conducted using purified His6-tagged ATPase^GyrB domain as bait to bind untagged inactive toxin domain.

C. A time-course experiment was performed to determine the rate at which the toxin recognizes the wild-type and (Δ1-14) ATPase^GyrB domain.

D. A toxin titration experiment was carried out to determine the optimal concentration for the toxin to cleave the wild-type and (Δ1-14) ATPase^GyrB domain.

Figure 9: A proposed mechanism of CdiA-CT^{DBS100} intoxication.
In a proposed mechanism within a co-culture environment, the CDI+ inhibitor cells deploy the toxin to both kin and target bacterial species. Only the targeted species is being affected by t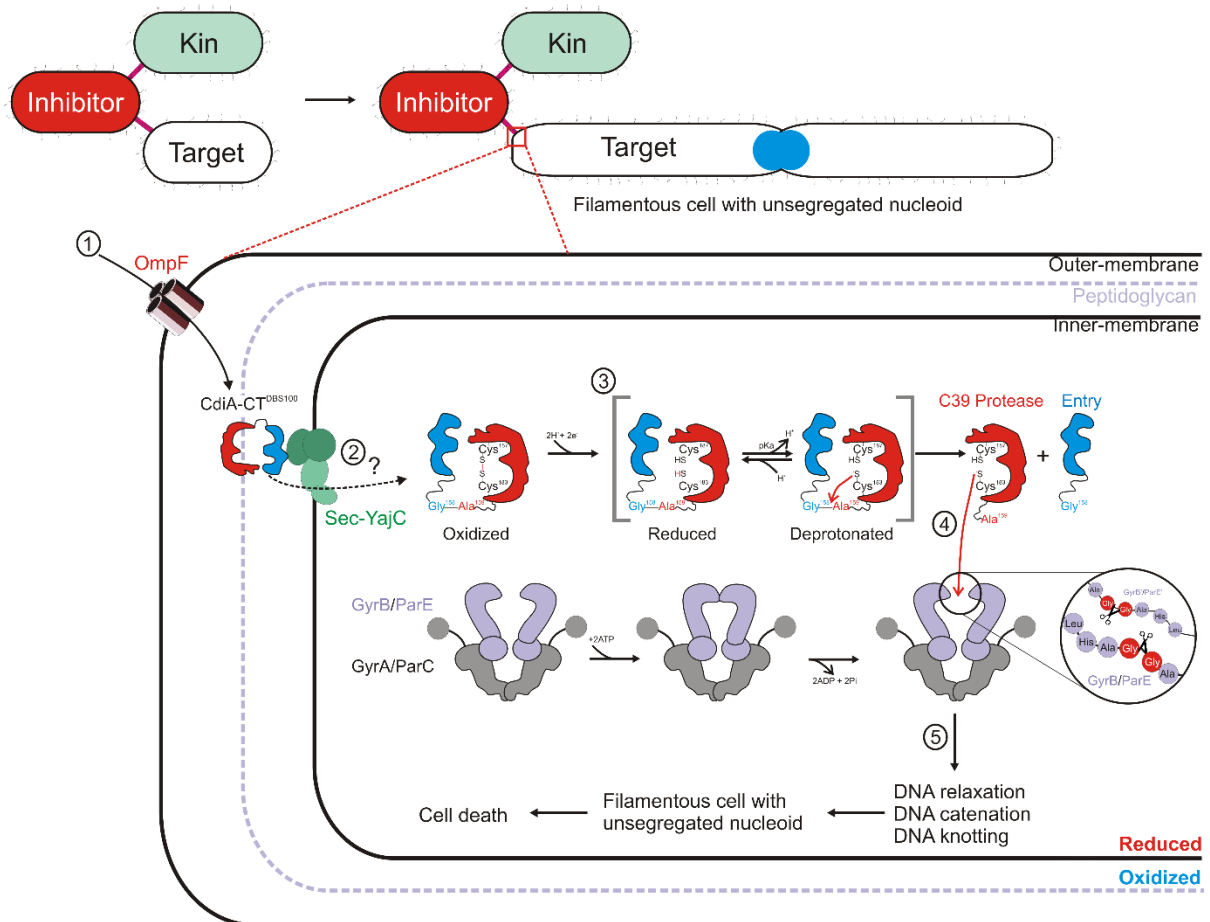he toxin and undergoes filamentation with unsegregated nucleoid. **(1)** CdiA^{DBS100} recognizes its receptor OmpF, facilitating toxin deployment in the periplasm. In the periplasmic compartment, toxin remains full-length and inactive due to an oxidizing environment. **(2)** The entry domain of CdiA-CT^{DBS100} is translocated to the cytoplasm via the Sec-YajC with unknown machinery. **(3)** Once CdiA-CT^{DBS100} enters the cytoplasm, the disulfide bond between Cys183 and Cys187 is reduced, activating the C39 protease domain, which then facilitates the autoproteolytic event between Gly158-Ala159. **(4)** The activated C39 protease can now locate its substates, GyrB and ParE, and cleave between Gly101-Gly102 and Gly97-Gly98, respectively. **(5)** Consequently, damage to the gyrase and topoisomerase IV complexes impairs the DNA replication and separation process, leading to filamentation and unsegregated nucleoids. Ultimately, this cascade of events results in cell death.

Table 1: Direct H-bonds and salt-bridges at the CdiA-CT/CdiI^{DBS100} interface

CdiA-CT atom	CdiI atom	Distance (Å)
PHE 230[N]	VAL 49[O]	2.94
GLN 213[N]	ASP 52[OD1]	3.74
ARG 216[NH2]	ASP 52[OD1]	3.23
SER 210[OG]	ASP 52[OD1]	2.49
ARG 216[NH1]	ASP 52[OD2]	2.85
ARG 216[NH2]	ASP 52[OD2]	3.15
TRP 250[NE1]	MET 62[O]	3.83
PHE 233[N]	ASN 70[OD1]	3.34
PHE 230[O]	VAL 49[N]	3.05
GLU 212[OE1]	SER 50[OG]	3.40
GLY 228[O]	SER 50[OG]	2.77
GLU 212[OE1]	SER 51[N]	2.80

Table 2: Bacterial strains and plasmids

Strains	Description[a]	References
X90	F' *lacI^q lac' pro' /ara Δ(lac-pro) nal1 argE(amb) rif'thi-1, Rif^R*	(Zhang et al., 2012)
CH87	*Citrobacter rodentium* DBS100	Endy lab
MFD*pir*	MG1655 RP4-2-Tc::[Δ*Mu1*::*aac(3)IV*-Δ*aphA*-Δ*nic35*-Δ*Mu2*::zeo] Δ*dapA*::(erm-pir) Δ*recA*, Amp^R, Zeo^R, Erm^R	(Ferrières et al., 2010)
CH2016	X90 (DE3) Δ*rna* Δ*slyD::kan,* Kan^R	(Aoki et al., 2008)
CH2445	CH8251 *galK*::sYFP2opt-cat, Cm^R	(Morse et al., 2015)
CH2567	CH8251 mKate2::cat, Cm^R	(Morse et al., 2015)
CH4201	*Citrobacter rodentium* DBS100 *Prha-CdiBA'(STEC3)-StbD-sphI,* Kan^R	This study
CH5960	MG1655 Δ*wzb* FtsZ-VSV-pCat, Cm^R	This study
CH7367	MG1655 Δ*wzb*	(Ruhe et al., 2018)
CH7992	MG1655 Δ*wzb parC*-VSV-pCat, Cm^R	This study
CH7993	MG1655 Δ*wzb gyrA*-VSV-pCat, Cm^R	This study
CH7994	MG1655 Δ*wzb gyrB*-VSV-pCat, Cm^R	This study
CH7995	MG1655 Δ*wzb parE*-linker-VSV-pCat	This study
CH8896	MG1655 pir-116(DHFR) Δ*ompF* Δ*wzb*::*cat,* Cm^R	This study
CH8897	MG1655 Δ*wzb* Δ*ompF*::Kan, Kan^R	This study
CH8900	MG1655 Δ*wzb* Δ*ompF*	This study
CH8901	MG1655 Δ*wzb* Δ*ompF* gyrB-vsv::pCat, Cm^R	This study
JW0912	BW25113 F-, *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, ΔompF746::kan, rph-1, Δ(rhaD-rhaB)568, hsdR51,* Kan^R	Keio collection

Plasmids	Description[a]	References
pCH48	pCP20, plasmid that shows temperature-sensitive replication and thermal induction of FLP synthesis, Amp^R and Cm^R	(Cherepanov & Wackernagel, 1995)
pCH69	pBluescript II SK(+), Amp^R	Addgene
pCH501	pRE121, Kan^R	This study
pCH902	pCAT-waaL(EAE), Cm^R	(Halvorsen et al., 2021)
pCH998	pET21p::cdiA-CT C183A + cdiI-His6 (Citrobacter rodentium DBS100)	This study, FGS
pCH1004	pET21(DHpa/Nhe)::EC93-EC869CT/cdiIo11, AmpR	(Jones et al., 2017)
pCH1762	pMCSG63::H6-TEV-dVENN-CdiA-CT-cdiI (C.rod DBS100), Amp^R	This study
pCH1764	pMCSG63::H6-TEV-dVENN-CdiA-CT (C183A) (C.rod DBS100), Amp^R	This study
pCH3779	pRE121::STEC3-CdiB-CdiA(S260)	This study
pCH3780	pRE121::STEC3-CdiB-CdiA(S260)-StbD(I585)	This study
pCH3998	pCH501::cdiBA'(STEC3)-sbtD(Crod), Kan^R	This study
pCH4032	pET21P-K::gyrB (G15)-His6 (ATPase domain), Amp^R	This study
pCH5355	pET21P::envZ (M238)-ATPase-His6, Amp^R	This study
pCH5356	pET21P::htpG-His6, Amp^R	This study
pCH5358	pET21b::gyrB-G221 (ATPase) (GG101/102SS)-His6, Amp^R	This study
pCH5359	pET21b::gyrB-G221 (ATPase) (GG113/114SS)-His6, Amp^R	This study
pCH5955	pCAT::VgrG2-VSV(EcoRI-SpeI)	This study
pCH5958	pCAT::FtsZ(G5357)-VSV(EcoRI-SpeI), Cm^R	This study
pCH5997	pET21P::CdiA CT (C187S)-I-his6(C.rod DBS100), Amp^R	This study
pCH6007	pET21P::CdiA CT (GG157-158SS)-I-his6(C.rod DBS100), Amp^R	This study
pCH6009	pET21P::CdiA CT (G158S)-I-his6(C.rod DBS100), Amp^R	This study
pCH7987	pCAT-VSV-gyrA(F557), Cm^R	This study
pCH7988	pCAT-VSV-gyrB(E482), Cm^R	This study
pCH7990	pCAT-VSV-parC(M489), Cm^R	This study
pCH7991	pCAT-VSV-parE(W313-extra linker), Cm^R	This study
pCH8187	pET21b::gyrA-His6, Amp^R	This study
pCH8189	pET21b::gyrB-His6, Amp^R	This study, CSH
pCH8193	pET21b::gyrB-G221 (ATPase)-His6, Amp^R	This study
pCH8511	pET21b::gyrB-G221 (ATPase) (G101S)-His6, Amp^R	This study
pCH8512	pET21b::gyrB-G221 (ATPase) (G102S)-His6, Amp^R	This study
pCH8906	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD-I(Crod DBS100), Kan^R	This study, CSH
pCH8908	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (GG157,158SS)-I(Crod DBS100), Kan^R	This study, CSH
pCH8909	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (G158S)-I(Crod DBS100), Kan^R	This study, CSH
pCH8910	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (C183A)-I(Crod DBS100), Kan^R	This study, CSH
pCH8911	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (C187S)-I(Crod DBS100), Kan^R	This study, CSH

pCH8912	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (H258A)-I(Crod DBS100), Kan^R	This study, CSH
pCH8913	pCH501::cdiB (STEC3)-TPS(STEC3)-stbD (D275A)-I(Crod DBS100), Kan^R	This study, CSH
pCH8914	pTrc99a, Amp^R	GE Healthcare
pCH8915	pTrc99aKX::ompF (K12), Amp^R	(Beck et al., 2016)
pCH9099	pET21::EC93-Crod(wt)-cdiI, Amp^R	This study
pCH9101	pCH405Δ, Tet^R	This study
pCH9102	pCH405::CdiI (C. rodentium DBS100), Tet^R	This study
pCH9103	pMCSG63-H6-TEV-A159-CT (C183A) (toxin domain only) , Amp^R	This study
pCH9108	pET21P::CdiA CT (H258A)-I-his6(C.rod DBS100), Amp^R	This study
pCH9109	pET21P::CdiA CT (D275A)-I-his6(C.rod DBS100), Amp^R	This study
pCH9110	pET21P::CdiI-his6(C.rod DBS100), Amp^R	This study
pCH13984	pET21p-K::CdiA' CT-I-his6(C.rod DBS100), Amp^R	This study, GDE
pCH14397	pBAD24::VgrG2-VSV, Amp^R	(Donato et al., 2020)

a Abbreviations: Amp^R, ampicillin-resistance; Cm^R, chloramphenicol-resistance; Kan^R, kanamycin-resistance; Spc^R, spectinomycin-resistance; Erm^R, erythromycin-resistance; Zeo^R, zeocin-resistance; and Tet^R, tetracycline-resistance.

CSH-plasmids were built by Dr. Christopher S. Hayes
FGS-plasmid was built by Dr. Fernando Garza-Sánchez
GDE-plasmid was built by Dr. Gregory D. Ekberg

Table 3: Oligonucleotides

Identifier	Descriptive name	Sequence	Sources
CH4346	Crod-CT-Nco-for	5'- TTT CCA TGG TGG AGA ATA ACT CGC TGG-3'	This study
CH4347	Crod-CdiI-Spe-rev	5'- TTT ACT AGT GGG AAC TAA CTG GAT CGG G-3'	This study
CH4497	Crod-cdiI-Kpn-for	5' - TCA GGT ACC ATG GAA ATG CCG TC - 3'	This study
CH5336	Crod-C183A-rev	5'-CCG CAA GCA CCG CCA GCA CCT GTT GGT GTT G- 3'	This study
CH5337	Crod-cdiI-Kpn-for	5'- TTT GGT ACC ATG GAA ATG CCG TCT TCA ACC- 3'	This study
CH5338	Crod-cdiI-Xho-rev	5'- TTT CTC GAG TTA GGG AAC TAA CTG GAT CGG GA-3'	This study
CH5356	Crod-Kpn-TEV-delVENN-for	5'-CTG GGT ACC GAG AAC CTG TAC TTC CAA TCG CTG GCC GGG GAT AAG-3'	This study
CH5357	Crod-CT-Xho-rev	5'-TAT CTC GAG TTA CTG GTT GAA GAC GGC ATT TC-3'	This study
CH5381	Crod-H258A-for	5'- AAA ATT GGA GCT TGG GTT GTT-3'	This study
CH5402	Crod-D275A-for	5'- ATC TAT GCT CCG TGG AAA-3'	This study
CH5423	Crod-G158S-rev	5'- GTT TCC ACC GGC GCT GCC CTT AGG CTC-3'	This study
CH5425	Crod-GG157-158SS-rev	5'- GTT TCC ACC GGC GCT GCT CTT AGG CTC-3'	This study
CH5454	Crod-CT-Kpn-For	5' GCA GGT ACC GGA GAC GAA GG-3'	This study
CH5559	Crod-StbD (I585)-Xba-rev	5'- GGG TCT AGA TGT CGA GGG TAG T-3'	This study
CH5560	Crod-StbD (S388)-Sal-for	5'- GAC GTC GAC CGA GAG CGG CG-3'	This study
CH5561	STEC3-CdiA(S260)-Sal-rev	5'- TCA GTC GAC GTC AGA TGA ATA CG-3'	This study
CH5562	STEC3-CdiB-Nco-for	5'- TTA CCA TGG CCA GAA TAC GGG-3'	This study
CH5694	gyrA-F557-Eco-for	5'- GAA GAA TTC ATC GAC CGA CTG CTG G-3'	This study
CH5695	gyrA-Nhe-for	5' - TAG GCT AGC GAC CTT GCG AGA GAA ATT ACA CC -3'	This study
CH5696	gyrA-Spe-rev	5'- GTA ACT AGT TTC TTC TTC TGG CTC GTC GTC AAC G-3'	This study
CH5697	gyrA-Xho-rev	5'- GTA CTC GAG TTC TTC TTC TGG CTC GTC GTC AAC G-3'	This study
CH5698	gyrB-E482-Eco-for	5'- GAC GAA TTC AAC CCG GAC AAA CTG CG-3'	This study
CH5699	gyrB-Nhe-for	5'- TTG GCT AGC AAT TCT TAT GAC TCC TCC AGT ATC AAA GTC C-3'	This study
CH5700	gyrB-SpeI-rev	5'- GCG ACT AGT AAT ATC GAT ATT CGC CGC TTT CAG GG-3'	This study
CH5701	gyrB-Xho-rev	5'-GCG CTC GAC AAT ATC GAT ATT CGC CGC TTT CAG GG-3'	This study
CH5712	gyrB-Spe-for	5'- TTT ACT AGT TCG AAT CTT ATG ACT CCT CCA GT AT-3'	This study
CH5713	gyrB-native-Xho-rev	5'- TTA CTC GAG TTA AAT ATC GAT ATTT CGC CGC TTT CAG-3'	This study
CH5714	gyrB-G221-Xho-rev	5'- CGC CTC GAG GCC GCC TTC ATA GTG G-3'	This study
CH5715	parE-Nhe-for	5'- TTT GCT AGC ACG CAA ACT TAT AAC GCT GAT GCC ATT G-3'	This study
CH5716	parE-Xho-rev	5'- TTT CTC GAG AAC CTC AAT CTC CGC CAT GTC GCC-3'	This study
CH5736	parC-Nhe-for	5'- TTA GCT AGC GAT ATG GCA GAG CGC C-3'	This study
CH5737	parC-Xho-Rev	5'- GGC CTC GCG CTC TTC GCT ATC ACC GCT GC-3'	This study
CH5742	envZ-M238-Nco-for	5'- TGC CCA TGG CGG GGG TAA GTC ACG-3'	This study
CH5743	envZ-Xho-rev	5'- GTT CTC GAG CCC TTC TTT TGT CGT GCC CTG-3'	This study
CH5744	htpG-Nco-for	5'- CCT CCA TGG AAG GAC AAG AAA CTC GTG G-3'	This study
CH5745	htpG-Xho-rev	5'- ATT CTC GAG GGA AAC CAG CAG CTG GTT CAT ACG-3'	This study
CH5748	gyrB-GG101/102SS-rev	5'-GAG TTA TCG TCA ATT TTA CTG CTT GCG TGC AGA ACG G-3'	This study
CH5749	gyrB-GG113/114SS-for	5'- CTC CTA TAA AGT GTC CAG CAG TCT GCA CGG CGT TGG-3'	This study
CH5765	Crod-TEV-A159-Kpn-For	5'- CTC GGT ACC GAG AAC CTG TAC TTC CAA GCC GGT GGA AAC TGG AAT GTG C-3'	This study
CH5766	parE-linker-Spe-rev	5'-TTT ACT AGT AGA AGA AGT ACC ACT ACC TCC AGA ACT AAC CTC AAT CTC CGC CAT GTC GC-3'	This study
CH5767	parC-M489-Eco-for	5'- AGC GAA TTC GAT GAG CGA GCA CGA C-3'	This study
CH5768	parC-Spe-rev	5'- GGC ACT AGT CTC TTC GCT ATC ACC GCT GCT GG-3'	This study
CH5899	stbD-G3267-Kpn-for	5'- ATC GGT ACC GGC CAC GCT GAA GAG AAC-3	This study

CH5950	Crod-VENN-rev	5' - CCG GCC AGC GAG TTA TTC TCC AC-3'	This study
CH5951	Crod-VENN-for	5' - GTG GAG AAT AAC TCG CTG GCC GG-3'	This study
CH5952	Crod-cdiI-Xba-rev	5' - TTT TCT AGA TTA GGG AAC TAA CTG GAT CGG GAA G-3'	This study

Chapter VI: Conclusion

Research has explored both contact-independent and contact-dependent bacterial competition systems within the context of two distinct populations, wherein an inhibitor population consistently gains the upper hand in its confrontation with a vulnerable target population. However, bacteria demonstrate remarkable adaptability in their strategies. Gram-negative bacteria, for instance, employ diverse mechanisms to defend against other bacterial species. Some of them may even employ multiple systems simultaneously to gain a growth advantage in competitive environments. It is worth considering that certain toxin systems may work synchronously to achieve maximal effectiveness in eliminating the surrounding microbes. This thesis examines effector proteins from contact-dependent growth inhibition (CDI) systems in two cell populations. The primary focus of this thesis is to understand and characterize the molecular mechanisms of RNase and protease found in CDI loci.

In Chapter II, we demonstrated a simple mechanism of an EndoU ribonuclease toxin found in *E. coli* STEC_O31, which primarily cleaves the tRNA^{Glu} with the His-His-Lys catalytic triad between nucleotides C37 and m^{2}A38 of the anticodon. Chapter III presented a more complex system of a general ribonuclease toxin from *E. coli* strain O32:H37, where it required a co-factor EF-Tu to have a faster and higher rate of substrate cleavage. In Chapter IV, we examined the transport mechanism of the ribonuclease toxin described in Chapter III and revealed the role of inner-membrane SbmA protein in toxin uptake. Finally, in Chapter V, we ventured beyond the typical toxin activity zone found in CDI systems and successfully characterized a novel C39 protease isolated from *Citrobacter rodentium* that targets gyrase and topoisomerase IV.

The overarching goal of these Chapters is to propose a clear molecular interaction between substrates and enzymes along with the physiological role of these systems in nature. However, there are limitations in this thesis that prevent us from fully demonstrating certain aspects of the systems. In Chapter III, it is evident that EF-Tu plays a crucial role in facilitating the nuclease activity. Nevertheless, further fine-tuning is needed to gain a more concrete understanding of the molecular interaction between EF-Tu and CdiA-CT$^{O32:H37}$. For instance, introducing more mutations = in the CdiA-CT$^{O32:H37}$ at the EF-Tu binding interface could help confirm the model predicted by AlphaFold. Additionally, the requirement of EF-Ts can be further demonstrated in a more refined study by using AlphaFold to predict potential interactions between EF-Ts and CdiA-CT.

Unfortunately, in Chapter IV, we could not fully demonstrate the mechanism of protein translocation. There are mainly questions to be resolved in future studies. First, in our zeocin suppressibility test, we observed that our strain with wild-type SbmA plasmid grew at a slower rate compared to the empty vector and other mutant variants. This outcome may be attributed to the expression levels of SbmA under the low-salt conditions, which could lead to a defect in bacterial growth. Further, it is imperative to establish the ability of SbmA to unfold proteins. While the entry domain of CdiA-CT$^{O32:H37}$ is nonstructural and presumably easily translocated, similar to other microcin peptides, the toxin domain is well-structured. This implies that the toxin domain must be unfolded to pass through the cytoplasmic gate of SbmA and then refolded in the cytoplasm in the presence of EF-Tu. To test this model in the future, a chimeric protein carrying the entry domain and a well-structured mNeon protein can be mixed with liposomes carrying SbmA to measure the fluorescent signal. Additionally,

more point mutations can be introduced to CdiA-CT$^{O32:H37}$ (Ile19 and Leu22) and SbmA (Gly105, Val106, Tyr112, Gln328, and Asn331) to validate the AlphaFold prediction.

Chapter V provided a glimpse of a novel protease that has a great potential to further development of antimicrobial substances. First, we must demonstrate that CdiA-CT^{DBS100} is inactive toward the mammalian topoisomerase, TOP2A from mouse and human. This will give us a green light to proceed with this protein as a potential bacterial-specific drug candidate. We firmly believe that this toxin does not target eukaryotic topoisomerase due to the absence of LHAGGK motif, but this should be demonstrated directly. Furthermore, an outstanding question in the pseudo-CDI system of *C. rodentium* concerns the function of the fimbrial-CDI chimeric system in host colonization and interbacterial interactions. To gain a better understanding, introducing an inducible promoter upstream to the fimbrial-CDI chimeric system is a promising approach. In future studies, we plan to employ mammalian tissue culture systems to investigate the interaction of this chimeric filament and the impact of CdiA-CT^{DBS100} on eukaryotic cells.

In conclusion, our findings in this thesis have raised questions about the evolutionary aspect of CDI systems. Many CDI systems have the potential for horizontal gene transfer, and gene recombination may lead to the emergence of new, unique features within these systems. Future research will not only unveil the intricate mechanistic details of toxin activity but also delve into the realms of bacterial gene evolution and bacterial physiology.

References

Aoki, S. K., Malinverni, J. C., Jacoby, K., Thomas, B., Pamma, R., Trinh, B. N., Remers, S., Webb, J., Braaten, B. A., Silhavy, T. J., & Low, D. A. (2008). Contact-dependent growth inhibition requires the essential outer membrane protein BamA (YaeT) as the receptor and the inner membrane transport protein AcrB. *Mol. Microbiol.*, *70*, 323–340.

Aoki, S. K., Pamma, R., Hernday, A. D., Bickham, J. E., Braaten, B. A., & Low, D. A. (2005). Contact-dependent inhibition of growth in Escherichia coli. *Science*, *309*, 1245–1248.

Aoki, S. K., Webb, J. S., Braaten, B. A., & Low, D. A. (2009). Contact-dependent growth inhibition causes reversible metabolic downregulation in Escherichia coli. *Journal of Bacteriology*, *191*(6), 1777–1786. https://doi.org/10.1128/JB.01437-08

Aoki, S. K., Diner, E. J., de Roodenbeke, C. T., Burgess, B. R., Poole, S. J., Braaten, B. A., Jones, A. M., Webb, J. S., Hayes, C. S., Cotter, P. A., & Low, D. A. (2010). A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature*, *468*(7322), 439–442. https://doi.org/10.1038/nature09490

Arenas, J., Schipper, K., van Ulsen, P., van der Ende, A., & Tommassen, J. (2013). Domain exchange at the 3´ end of the gene encoding the fratricide meningococcal two-partner secretion protein A. *BMC Genomics*, *14*, 622.

Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K.-D., Fischbach, M. A., Garavelli, J. S., … van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports*, *30*(1), 108–160. https://doi.org/10.1039/c2np20085f

Arnold, M. F. F., Caro-Hernandez, P., Tan, K., Runti, G., Wehmeier, S., Scocchi, M., Doerrler, W. T., Walker, G. C., & Ferguson, G. P. (2014). Enteric YaiW is a surface-exposed outer membrane lipoprotein that affects sensitivity to an antimicrobial peptide. *Journal of Bacteriology*, *196*(2), 436–444. https://doi.org/10.1128/JB.01179-13

Bahassi, E. M., O'Dea, M. H., Allali, N., Messens, J., Gellert, M., & Couturier, M. (1999). Interactions of CcdB with DNA gyrase. Inactivation of Gyra, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. *The Journal of Biological Chemistry*, *274*(16), 10936–10944. https://doi.org/10.1074/jbc.274.16.10936

Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M., & Hanada, K. (1982). L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *The Biochemical Journal*, *201*(1), 189–198. https://doi.org/10.1042/bj2010189

Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J., & Mekalanos, J. J. (2012). Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*, *483*, 182–186.

Beck, C. M., Morse, R. P., Cunningham, D. A., Iniguez, A., Low, D. A., Goulding, C. W., & Hayes, C. S. (2014). CdiA from Enterobacter cloacae delivers a toxic ribosomal RNase into target bacteria. *Structure*, *22*, 707–718.

Beck, C. M., Willett, J. L., Cunningham, D. A., Kim, J. J., Low, D. A., & Hayes, C. S. (2016). CdiA Effectors from Uropathogenic Escherichia coli Use Heterotrimeric Osmoporins as Receptors to Recognize Target Bacteria. *PLoS Pathog*, *12*.

Beveridge, T. J. (1981). Ultrastructure, chemistry, and function of the bacterial wall. *International Review of Cytology*, *72*, 229–317. https://doi.org/10.1016/s0074-7696(08)61198-5

Bhardwaj, K., Guarino, L., & Kao, C. C. (2004). The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J Virol*, *78*, 12218–12224.

Bhardwaj, K., Palaninathan, S., Alcantara, J. M., Yi, L. L., Guarino, L., Sacchettini, J. C., & Kao, C. C. (2008). Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J Biol Chem*, *283*, 3655–3664.

Brino, L., Urzhumtsev, A., Mousli, M., Bronner, C., Mitschler, A., Oudet, P., & Moras, D. (2000). Dimerization of Escherichia coli DNA-gyrase B provides a structural mechanism for activating the ATPase catalytic center. *The Journal of Biological Chemistry*, *275*(13), 9468–9475. https://doi.org/10.1074/jbc.275.13.9468

Burnaevskiy, N., Fox, T. G., Plymire, D. A., Ertelt, J. M., Weigele, B. A., Selyunin, A. S., Way, S. S., Patrie, S. M., & Alto, N. M. (2013). Proteolytic elimination of N-myristoyl modifications by the Shigella virulence factor IpaJ. *Nature*, *496*(7443), 106–109. https://doi.org/10.1038/nature12004

Caffarelli, E., Arese, M., Santoro, B., Fragapane, P., & Bozzoni, I. (1994). In vitro study of processing of the intron-encoded U16 small nucleolar RNA in Xenopus laevis. *Mol Cell Biol*, *14*, 2966–2974.

Caffarelli, E., Maggi, L., Fatica, A., Jiricny, J., & Bozzoni, I. (1997). A novel Mn++-dependent ribonuclease that functions in U16 snoRNA processing in X. laevis. *Biochem Biophys Res Commun*, *233*, 514–517.

Caldas, T. D., El Yaagoubi, A., & Richarme, G. (1998). Chaperone properties of bacterial elongation factor EF-Tu. *The Journal of Biological Chemistry*, *273*(19), 11478–11482. https://doi.org/10.1074/jbc.273.19.11478

Cao, Z., Casabona, M. G., Kneuper, H., Chalmers, J. D., & Palmer, T. (2016). The type VII secretion system of Staphylococcus aureus secretes a nuclease toxin that targets competitor bacteria. *Nat Microbiol*, *2*, 16183.

Cascales, E., Buchanan, S. K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S., & Cavard, D. (2007). Colicin biology. *Microbiol. Mol. Biol*, *71*, 158–229.

Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, *158*(1), 9–14. https://doi.org/10.1016/0378-1119(95)00193-a

Corbalan, N., Runti, G., Adler, C., Covaceuszach, S., Ford, R. C., Lamba, D., Beis, K., Scocchi, M., & Vincent, P. A. (2013). Functional and structural study of the dimeric inner membrane protein SbmA. *Journal of Bacteriology*, *195*(23), 5352–5361. https://doi.org/10.1128/JB.00824-13

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr*, *62*, 1002–1011.

Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. *Joint CCP*, *4*, 34–38.

Cuchillo, C. M., Nogues, M. V, & Raines, R. T. (2011). Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. *Biochemistry*, *50*, 7835–7841.

Darbandi, A., Asadi, A., Ari, M. M., Ohadi, E., Talebi, M., Zadeh, M. H., Darb, A., Roya, E., & Kakanj, M. (2022). *Bacteriocins : Properties and potential use as antimicrobials*. *August 2021*, 1–40. https://doi.org/10.1002/jcla.24093

Davies, J. K., & Reeves, P. (1975). Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group A. *Journal of Bacteriology*, *123*(1), 102–117. https://doi.org/10.1128/jb.123.1.102-117.1975

Davis, I. W., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res*, *32*, W615-619.

de Moraes, M. H., Hsu, F., Huang, D., Bosch, D. E., Zeng, J., Radey, M. C., Simon, N., Ledvina, H. E., Frick, J. P., Wiggins, P. A., Peterson, S. B., & Mougous, J. D. (2021). An interbacterial DNA deaminase toxin directly mutagenizes surviving target populations. *ELife*, *10*. https://doi.org/10.7554/eLife.62967

Deavin, A., Mathias, A. P., & Rabin, B. R. (1966). Mechanism of action of bovine pancreatic ribonuclease. *Nature*, *211*, 252–255.

Deng, X., Hackbart, M., Mettelman, R. C., O'Brien, A., Mielech, A. M., Yi, G., Kao, C. C., & Baker, S. C. (2017). Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc Natl Acad Sci U S A*, *114*, E4251–E4260.

Dey, A., Vassallo, C. N., Conklin, A. C., Pathak, D. T., Troselj, V., & Wall, D. (2016). Sibling Rivalry in Myxococcus xanthus Is Mediated by Kin Recognition and a Polyploid Prophage. *Journal of Bacteriology*, *198*, 994–1004.

Diner, E. J., Beck, C. M., Webb, J. S., Low, D. A., & Hayes, C. S. (2012). Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes Dev*, *26*, 515–525.

Donato, S. L., Beck, C. M., Garza-Sánchez, F., Jensen, S. J., Ruhe, Z. C., Cunningham, D. A., Singleton, I., Low, D. A., & Hayes, C. S. (2020). The β-encapsulation cage of rearrangement hotspot (Rhs) effectors is required for type VI secretion. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(52), 33540–33548. https://doi.org/10.1073/pnas.1919350117

El Ghachi, M., Bouhss, A., Barreteau, H., Touzé, T., Auger, G., Blanot, D., & Mengin-Lecreulx, D. (2006). Colicin M exerts its bacteriolytic effect via enzymatic degradation of undecaprenyl phosphate-linked peptidoglycan precursors. *The Journal of Biological Chemistry*, *281*(32), 22761–22772. https://doi.org/10.1074/jbc.M602834200

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, *60*, 2126–2132.

Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A., & Mark, I. D. (2009). A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol Biol*, *498*, 105–115.

Eschenfeldt, W. H., Makowska-Grzyska, M., Stols, L., Donnelly, M. I., Jedrzejczak, R., & Joachimiak, A. (2013). New LIC vectors for production of proteins from genes containing rare codons. *J Struct Funct Genomics*, *14*, 135–144.

Eschenfeldt, W. H., Maltseva, N., Stols, L., Donnelly, M. I., Gu, M., Nocek, B., Tan, K., Kim, Y., & Joachimiak, A. (2010). Cleavable C-terminal His-tag vectors for structure determination. *J Struct Funct Genomics*, *11*, 31–39.

Ferrières, L., Hémery, G., Nham, T., Guérout, A.-M., Mazel, D., Beloin, C., & Ghigo, J.-M. (2010). Silent mischief: bacteriophage Mu insertions contaminate products of Escherichia coli random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *Journal of Bacteriology*, *192*(24), 6418–6427. https://doi.org/10.1128/JB.00621-10

Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R., & Ross, C. A. (1962). The active site and mechanism of action of bovine pancreatic ribonuclease. 7. *The Catalytic Mechanism*, *85*, 152–153.

French, S., & Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallogr A A*, *34*, 517–525.

Furano, A. V. (1975). Content of elongation factor Tu in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(12), 4780–4784. https://doi.org/10.1073/pnas.72.12.4780

Garcia-Bayona, L., Guo, M. S., & Laub, M. T. (2017). Contact-dependent killing by Caulobacter crescentus via cell surface-associated, glycine zipper proteins. *Elife*, *6*.

Garza-Sánchez, F., Janssen, B. D., & Hayes, C. S. (2006). Prolyl-tRNA(Pro) in the A-site of SecM-arrested ribosomes inhibits the recruitment of transfer-messenger RNA. *J Biol Chem*, *281*, 34258–34268.

Ghequire, M. G., & Mot, R. De. (2014). Ribosomally encoded antibacterial proteins and peptides from Pseudomonas. *FEMS Microbiol Rev*, *38*, 523–568.

Ghilarov, D., Inaba-Inoue, S., Stepien, P., Qu, F., Michalczyk, E., Pakosz, Z., Nomura, N., Ogasawara, S., Walker, G. C., Rebuffat, S., Iwata, S., Heddle, J. G., & Beis, K. (2021). Molecular mechanism of SbmA, a promiscuous transporter exploited by antimicrobial peptides. *Science Advances*, *7*(37), eabj5363. https://doi.org/10.1126/sciadv.abj5363

Ghoul, M., & Mitri, S. (2016). The Ecology and Evolution of Microbial Competition. *Trends in Microbiology*, *24*(10), 833–845. https://doi.org/10.1016/j.tim.2016.06.011

Gibson, E. G., Oviatt, A. A., & Osheroff, N. (2020). Two-Dimensional Gel Electrophoresis to Resolve DNA Topoisomers. *Methods in Molecular Biology (Clifton, N.J.)*, *2119*, 15—24. https://doi.org/10.1007/978-1-0716-0323-9_2

Gilliland, G. (1997). Ribonuclease: Structures and Functions. *Academic Press, San Diego*, 306–341.

Gioia, U., Laneve, P., Dlakic, M., Arceci, M., Bozzoni, I., & Caffarelli, E. (2005). Functional characterization of XendoU, the endoribonuclease involved in small nucleolar RNA biosynthesis. *J Biol Chem*, *280*, 18996–19002.

Goldstein, S. L., & Klassen, J. L. (2020). Pseudonocardia Symbionts of Fungus-Growing Ants and the Evolution of Defensive Secondary Metabolism. *Frontiers in Microbiology*, *11*, 621041. https://doi.org/10.3389/fmicb.2020.621041

Graille, M., Mora, L., Buckingham, R. H., van Tilbeurgh, H., & de Zamaroczy, M. (2004). Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity

protein. *The EMBO Journal*, *23*(7), 1474–1482. https://doi.org/10.1038/sj.emboj.7600162

Gratia, A. (1925). Sur un remarquable exemple d'antagonisme entre deux souches de coilbacille. *CR Seances Soc. Biol. Fil.*, *93*, 1040–1041.

Guarino, L. A., Bhardwaj, K., Dong, W., Sun, J., Holzenburg, A., & Kao, C. (2005). Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J Mol Biol*, *353*, 1106–1117.

Gucinski, G. C., Michalska, K., Garza-Sánchez, F., Eschenfeldt, W. H., Stols, L., Nguyen, J. Y., Goulding, C. W., Joachimiak, A., & Hayes, C. S. (2019). Convergent Evolution of the Barnase/EndoU/Colicin/RelE (BECR) Fold in Antibacterial tRNase Toxins. *Structure (London, England : 1993)*, *27*(11), 1660-1674.e5. https://doi.org/10.1016/j.str.2019.08.010

Halvorsen, T. M., Garza-Sánchez, F., Ruhe, Z. C., Bartelli, N. L., Chan, N. A., Nguyen, J. Y., Low, D. A., & Hayes, C. S. (2021). Lipidation of Class IV CdiA Effector Proteins Promotes Target Cell Recognition during Contact-Dependent Growth Inhibition. *MBio*, *12*(5), e0253021. https://doi.org/10.1128/mBio.02530-21

Hardy, K. G., Meynell, G. G., Dowman, J. E., & Spratt, B. G. (1973). Two major groups of colicin factors: their evolutionary significance. *Molecular & General Genetics : MGG*, *125*(3), 217–230. https://doi.org/10.1007/BF00270744

Harms, A., Stanger, F. V., Scheu, P. D., de Jong, I. G., Goepfert, A., Glatter, T., Gerdes, K., Schirmer, T., & Dehio, C. (2015). Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Reports*, *12*(9), 1497–1507. https://doi.org/10.1016/j.celrep.2015.07.056

Hayes, C. S., & Sauer, R. T. (2003). Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol. Cell*, *12*, 903–911.

Hayes, C. S., Koskiniemi, S., Ruhe, Z. C., Poole, S. J., & Low, D. A. (2014). Mechanisms and biological roles of contact-dependent growth inhibition systems. *Cold Spring Harbor Perspectives in Medicine*, *4*(2). https://doi.org/10.1101/cshperspect.a010025

Hearnshaw, S. J., Chung, T. T.-H., Stevenson, C. E. M., Maxwell, A., & Lawson, D. M. (2015). The role of monovalent cations in the ATPase reaction of DNA gyrase. *Acta Crystallographica. Section D, Biological Crystallography*, *71*(Pt 4), 996–1005. https://doi.org/10.1107/S1399004715002916

Herschman, H. R., & Helinski, D. R. (1967). Comparative study of the events associated with colicin induction. *Journal of Bacteriology*, *94*(3), 691–699. https://doi.org/10.1128/jb.94.3.691-699.1967

Ho, B. T., Dong, T. G., & Mekalanos, J. J. (2014). A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe*, *15*, 9–21.

Holberger, L. E., Garza-Sanchez, F., Lamoureux, J., Low, D. A., & Hayes, C. S. (2012). A novel family of toxin/antitoxin proteins in Bacillus species. *FEBS Lett*, *586*, 132–136.

Holm, L., & Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res*, *38*, W545-549.

Hood, R. D., Singh, P., Hsu, F., Guvener, T., Carl, M. A., Trinidad, R. R., Silverman, J. M., Ohlson, B. B., Hicks, K. G., Plemel, R. L., Li, M., Schwarz, S., Wang, W. Y., Merz, A. J., Goodlett, D. R., & Mougous, J. D. (2010). A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. *Cell Host Microbe*, *7*, 25–37.

Ivanov, K. A., Hertzig, T., Rozanov, M., Bayer, S., Thiel, V., Gorbalenya, A. E., & Ziebuhr, J. (2004). Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci U S A*, *101*, 12694–12699.

Jack, R. W., Tagg, J. R., & Ev, M. I. R. (1995). *Bacteriocins of Gram-Positive Bacteria*. *59*(2), 171–200.

Jamet, A., Jousset, A. B., Euphrasie, D., Mukorako, P., Boucharlat, A., Ducousso, A., Charbit, A., & Nassif, X. (2015). A new family of secreted toxins in pathogenic Neisseria species. *PLoS Pathog*, *11*.

Jones, A. M., Garza-Sanchez, F., So, J., Hayes, C. S., & Low, D. A. (2017). Activation of contact-dependent antibacterial tRNase toxins by translation elongation factors. *Proc Natl Acad Sci U S A*, *114*, E1951–E1957.

Jones, A. M., Low, D. A., & Hayes, C. S. (2017). Can't you hear me knocking: contact-dependent competition and cooperation in bacteria. *Emerg Top Life Sci*, *1*, 75–83.

Joseph, J. S., Saikatendu, K. S., Subramanian, V., Neuman, B. W., Buchmeier, M. J., Stevens, R. C., & Kuhn, P. (2007). Crystal structure of a monomeric form of severe acute respiratory syndrome coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J Virol*, *81*, 6700–6708.

Kang, H., Bhardwaj, K., Li, Y., Palaninathan, S., Sacchettini, J., Guarino, L., Leibowitz, J. L., & Kao, C. C. (2007). Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J Virol*, *81*, 13587–13597.

Kaper, J. B., Nataro, J. P., & Mobley, H. L. T. (2004). *Escherichia coli*. *2*(February). https://doi.org/10.1038/nrmicro818

Kindler, E., Gil-Cruz, C., Spanier, J., Li, Y., Wilhelm, J., Rabouw, H. H., Zust, R., Hwang, M., V'Kovski, P., Stalder, H., Marti, S., Habjan, M., Cervantes-Barragan, L., Elliot, R.,

Karl, N., Gaughan, C., van Kuppeveld, F. J., Silverman, R. H., Keller, M., … Thiel, V. (2017). Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLoS Pathog*, *13*.

Koch, A. L. (1998). The biophysics of the gram-negative periplasmic space. *Critical Reviews in Microbiology*, *24*(1), 23–59. https://doi.org/10.1080/10408419891294172

Koch, A. L., & Woeste, S. (1992). Elasticity of the sacculus of Escherichia coli. *Journal of Bacteriology*, *174*(14), 4811–4819. https://doi.org/10.1128/jb.174.14.4811-4819.1992

Koskiniemi, S., Garza-Sanchez, F., Sandegren, L., Webb, J. S., Braaten, B. A., Poole, S. J., Andersson, D. I., Hayes, C. S., & Low, D. A. (2014). Selection of orphan Rhs toxin expression in evolved Salmonella enterica serovar Typhimurium. *PLoS Genet*, *10*.

Koskiniemi, S., Lamoureux, J. G., & Nikolakakis, K. C. (2013). C. t'Kint de Roodenbeke, M. *D. Kaplan, D.A. Low & C.S. Hayes, Rhs Proteins from Diverse Bacteria Mediate Intercellular Competition*, *110*, 7032–7037.

Krab, I. M., te Biesebeke, R., Bernardi, A., & Parmeggiani, A. (2001). Elongation factor Ts can act as a steric chaperone by increasing the solubility of nucleotide binding-impaired elongation factor-Tu. *Biochemistry*, *40*(29), 8531–8535. https://doi.org/10.1021/bi0104930

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, *372*, 774–797.

Kudlicki, W., Coffman, A., Kramer, G., & Hardesty, B. (1997). Renaturation of rhodanese by translational elongation factor (EF) Tu. Protein refolding by EF-Tu flexing. *The Journal of Biological Chemistry*, *272*(51), 32206–32210. https://doi.org/10.1074/jbc.272.51.32206

Laneve, P., Altieri, F., Fiori, M. E., Scaloni, A., Bozzoni, I., & Caffarelli, E. (2003). Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in Xenopus laevis. *J Biol Chem*, *278*, 13026–13032.

Laviña, M., Pugsley, A. P., & Moreno, F. (1986). Identification, mapping, cloning and characterization of a gene (sbmA) required for microcin B17 action on Escherichia coli K12. *Journal of General Microbiology*, *132*(6), 1685–1693. https://doi.org/10.1099/00221287-132-6-1685

Little, J. W., & Mount, D. W. (1982). The SOS regulatory system of Escherichia coli. *Cell*, *29*(1), 11–22. https://doi.org/10.1016/0092-8674(82)90085-x

Lopez, D., Vlamakis, H., & Kolter, R. (2010). Biofilms. *Cold Spring Harb Perspect Biol*, *2*.

MacIntyre, D. L., Miyata, S. T., Kitaoka, M., & Pukatzki, S. (2010). The Vibrio cholerae type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A*, *107*, 19520–19524.

Masaki, H., Ogawa, T., Tomita, K., Ueda, T., Watanabe, K., & Uozumi, T. (1997). Colicin E5 as a new type of cytotoxin, which cleaves a specific group of tRNAs. *Nucleic Acids Symposium Series*, *37*, 287–288.

Masaki, H., & Ogawa, T. (2002). The modes of action of colicins E5 and D, and related cytotoxic tRNases. *Biochimie*, *84*(5–6), 433–438. https://doi.org/10.1016/s0300-9084(02)01425-6

Maslowska, K. H., Makiela-Dzbenska, K., & Fijalkowska, I. J. (2019). The SOS system: A complex and tightly regulated response to DNA damage. *Environmental and Molecular Mutagenesis*, *60*(4), 368–384. https://doi.org/10.1002/em.22267

Mattiuzzo, M., Bandiera, A., Gennaro, R., Benincasa, M., Pacor, S., Antcheva, N., & Scocchi, M. (2007). Role of the Escherichia coli SbmA in the antimicrobial activity of proline-rich peptides. *Molecular Microbiology*, *66*(1), 151–163. https://doi.org/10.1111/j.1365-2958.2007.05903.x

Maxwell, A. (1993). The interaction between coumarin drugs and DNA gyrase. *Molecular Microbiology*, *9*(4), 681–686. https://doi.org/10.1111/j.1365-2958.1993.tb01728.x

McGinness, K. E., Baker, T. A., & Sauer, R. T. (2006). Engineering controllable protein degradation. *Mol. Cell*, *22*, 701–707.

Metelev, M., Osterman, I. A., Ghilarov, D., Khabibullina, N. F., Yakimov, A., Shabalin, K., Utkina, I., Travin, D. Y., Komarova, E. S., Serebryakova, M., Artamonova, T., Khodorkovskii, M., Konevega, A. L., Sergiev, P. V, Severinov, K., & Polikanov, Y. S. (2017). Klebsazolicin inhibits 70S ribosome by obstructing the peptide exit tunnel. *Nature Chemical Biology*, *13*(10), 1129–1136. https://doi.org/10.1038/nchembio.2462

Michalska, K., Gucinski, G. C., Garza-Sanchez, F., Johnson, P. M., Stols, L. M., Eschenfeldt, W. H., Babnigg, G., Low, D. A., Goulding, C. W., Joachimiak, A., & Hayes, C. S. (2017). Structure of a novel antibacterial toxin that exploits elongation factor Tu to cleave specific transfer RNAs. *Nucleic Acids Res*, *45*, 10306–10320.

Michalska, K., Quan Nhan, D., Willett, J. L. E., Stols, L. M., Eschenfeldt, W. H., Jones, A. M., Nguyen, J. Y., Koskiniemi, S., Low, D. A., Goulding, C. W., Joachimiak, A., & Hayes, C. S. (2018). Functional plasticity of antibacterial EndoU toxins. *Molecular Microbiology*, *109*(4), 509–527. https://doi.org/10.1111/mmi.14007

Minor, W., Cymborowski, M., Otwinowski, Z., & Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*, *62*, 859–866.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making Protein folding accessible to all. *Nature Methods*. https://doi.org/10.1038/s41592-022-01488-1

Mirdita, M., Steinegger, M., & S"oding, J. (2019). MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics*, *35*(16), 2856–2858. https://doi.org/10.1093/bioinformatics/bty1057

Mirdita, M., von den Driesch, L., Galiez, C., Martin, M. J., S"oding, J., & Steinegger, M. (2017). Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.*, *45*(D1), D170--D176. https://doi.org/10.1093/nar/gkw1081

Mitchell, A. L., Almeida, A., Beracochea, M., Boland, M., Burgin, J., Cochrane, G., Crusoe, M. R., Kale, V., Potter, S. C., Richardson, L. J., Sakharova, E., Scheremetjew, M., Korobeynikov, A., Shlemov, A., Kunyavskaya, O., Lapidus, A., & Finn, R. D. (2019). MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkz1035

Morse, R. P., Nikolakakis, K. C., Willett, J. L., Gerrick, E., Low, D. A., Hayes, C. S., & Goulding, C. W. (2012). Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proc Natl Acad Sci U S A*, *109*, 21480–21485.

Morse, R. P., Willett, J. L. E., Johnson, P. M., Zheng, J., Credali, A., Iniguez, A., Nowick, J. S., Hayes, C. S., & Goulding, C. W. (2015). Diversification of β-Augmentation Interactions between CDI Toxin/Immunity Proteins. *Journal of Molecular Biology*, *427*(23), 3766–3784. https://doi.org/10.1016/j.jmb.2015.09.020

Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, *53*, 240–255.

Myers-Morales, T., Sim, M. M. S., DuCote, T. J., & Garcia, E. C. (2021). Burkholderia multivorans requires species-specific GltJK for entry of a contact-dependent growth inhibition system protein. *Molecular Microbiology*, *116*(3), 957–973. https://doi.org/10.1111/mmi.14783

Nagel de Zwaig, R., & Luria, S. E. (1967). Genetics and physiology of colicin-tolerant mutants of Escherichia coli. *Journal of Bacteriology*, *94*(4), 1112–1123. https://doi.org/10.1128/jb.94.4.1112-1123.1967

Nedialkova, D. D., Ulferts, R., van den Born, E., Lauber, C., Gorbalenya, A. E., Ziebuhr, J., & Snijder, E. J. (2009). Biochemical characterization of arterivirus nonstructural protein 11 reveals the nidovirus-wide conservation of a replicative endoribonuclease. *J Virol*, *83*, 5671–5682.

Ng, W. L., & Bassler, B. L. (2009). Bacterial quorum-sensing network architectures. *Annu Rev Genet*, *43*, 197–222.

Nikolakakis, K., Amber, S., Wilbur, J. S., Diner, E. J., Aoki, S. K., Poole, S. J., Tuanyok, A., Keim, P. S., Peacock, S., Hayes, C. S., & Low, D. A. (2012). The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Mol Microbiol*, *84*, 516–529.

Nomura, M., & Witten, C. (1967). Interaction of colicins with bacterial cells. 3. Colicin-tolerant mutations in Escherichia coli. *Journal of Bacteriology*, *94*(4), 1093–1111. https://doi.org/10.1128/jb.94.4.1093-1111.1967

Nyholm, S. V, & McFall-Ngai, M. J. (2021). A lasting symbiosis: how the Hawaiian bobtail squid finds and keeps its bioluminescent bacterial partner. *Nature Reviews. Microbiology*, *19*(10), 666–679. https://doi.org/10.1038/s41579-021-00567-y

Ogle, J. M., Carter, A. P., & Ramakrishnan, V. (2003). Insights into the decoding mechanism from recent ribosome structures. *Trends in Biochemical Sciences*, *28*(5), 259–266. https://doi.org/10.1016/S0968-0004(03)00066-5

Ogunrinola, G. A., Oyewale, J. O., Oshamika, O. O., & Olasehinde, G. I. (2020). The Human Microbiome and Its Impacts on Health. *International Journal of Microbiology*, *2020*, 8045646. https://doi.org/10.1155/2020/8045646

Ohr, R. J., Anderson, M., Shi, M., Schneewind, O., & Missiakas, D. (2017). EssD, a Nuclease Effector of the Staphylococcus aureus ESS Pathway. *Journal of Bacteriology*, *199*.

Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. *In: Proceedings of the CCP*, *26*(1991), 80–85.

Padilla, J. E., & Yeates, T. O. (2003). A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr D Biol Crystallogr*, *59*, 1124–1130.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, *25*, 1605–1612.

Pham, T. D. M., Ziora, Z. M., & Blaskovich, M. A. T. (2019). Quinolone antibiotics. *MedChemComm*, *10*(10), 1719–1739. https://doi.org/10.1039/c9md00120d

Pickard, J. M., Zeng, M. Y., Caruso, R., & Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunological Reviews*, *279*(1), 70–89. https://doi.org/10.1111/imr.12567

Poole, S. J., Diner, E. J., Aoki, S. K., Braaten, B. A., de Roodenbeke, C. T., Low, D. A., & Hayes, C. S. (2011). Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet*, *7*.

Raines, R. T. (1998). Ribonuclease A. *Chemical Reviews*, *98*(3), 1045–1066. https://doi.org/10.1021/cr960427h

Raivio, T. L., Leblanc, S. K. D., & Price, N. L. (2013). The Escherichia coli Cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. *Journal of Bacteriology*, *195*(12), 2755–2767. https://doi.org/10.1128/JB.00105-13

Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, *42*(Database issue), D503-9. https://doi.org/10.1093/nar/gkt953

Renzi, F., Caffarelli, E., Laneve, P., Bozzoni, I., Brunori, M., & Vallone, B. (2006). The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *Proc Natl Acad Sci U S A*, *103*, 12365–12370.

Ricagno, S., Egloff, M. P., Ulferts, R., Coutard, B., Nurizzo, D., Campanacci, V., Cambillau, C., Ziebuhr, J., & Canard, B. (2006). Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc Natl Acad Sci U S A*, *103*, 11892–11897.

Rice, A. J., Park, A., & Pinkett, H. W. (2014). Diversity in ABC transporters: type I, II and III importers. *Critical Reviews in Biochemistry and Molecular Biology*, *49*(5), 426–437. https://doi.org/10.3109/10409238.2014.953626

Riley, M. A. (1993). Molecular mechanisms of colicin evolution. *Molecular Biology and Evolution*, *10*(6), 1380–1395. https://doi.org/10.1093/oxfordjournals.molbev.a040081

Riley, M. A. (1993). Positive selection for colicin diversity in bacteria. *Molecular Biology and Evolution*, *10*(5), 1048–1059. https://doi.org/10.1093/oxfordjournals.molbev.a040054

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, *42*, W320-324.

Ruhe, Z. C., Nguyen, J. Y., Chen, A. J., Leung, N. Y., Hayes, C. S., & Low, D. A. (2016). CDI Systems Are Stably Maintained by a Cell-Contact Mediated Surveillance Mechanism. *PLoS Genet*, *12*.

Ruhe, Z. C., Nguyen, J. Y., Xiong, J., Koskiniemi, S., Beck, C. M., Perkins, B. R., Low, D. A., & Hayes, C. S. (2017). CdiA Effectors Use Modular Receptor-Binding Domains To Recognize Target Bacteria. *MBio*, *8*.

Ruhe, Z. C., Low, D. A., & Hayes, C. S. (2013). Bacterial contact-dependent growth inhibition. *Trends in Microbiology*, *21*(5), 230–237. https://doi.org/10.1016/j.tim.2013.02.003

Ruhe, Z. C., Subramanian, P., Song, K., Nguyen, J. Y., Stevens, T. A., Low, D. A., Jensen, G. J., & Hayes, C. S. (2018). Programmed Secretion Arrest and Receptor-Triggered Toxin Export during Antibacterial Contact-Dependent Growth Inhibition. *Cell*, *175*(4), 921-933.e14. https://doi.org/10.1016/j.cell.2018.10.033

Ruhe, Z. C., Wallace, A. B., Low, D. A., & Hayes, C. S. (2013). Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *MBio*, *4*(4). https://doi.org/10.1128/mBio.00480-13

Russell, A. B., Peterson, S. B., & Mougous, J. D. (2014). Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol*, *12*, 137–148.

Salomón, R. A., & Farías, R. N. (1995). The peptide antibiotic microcin 25 is imported through the TonB pathway and the SbmA protein. *Journal of Bacteriology*, *177*(11), 3323–3325. https://doi.org/10.1128/jb.177.11.3323-3325.1995

Schwarz, D. S., & Blower, M. D. (2014). The calcium-dependent ribonuclease XendoU promotes ER network formation through local RNA degradation. *J Cell Biol*, *207*, 41–57.

Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallogr A*, *64*, 112–122.

Shi, Y., Li, Y., Lei, Y., Ye, G., Shen, Z., Sun, L., Luo, R., Wang, D., Fu, Z. F., Xiao, S., & Peng, G. (2016). A Dimerization-Dependent Mechanism Drives the Endoribonuclease Function of Porcine Reproductive and Respiratory Syndrome Virus nsp11. *J Virol*, *90*, 4579–4592.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J., & Gorbalenya, A. E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol*, *331*, 991–1004.

Souza, D. P., Oka, G. U., Alvarez-Martinez, C. E., Bisson-Filho, A. W., Dunger, G., Hobeika, L., Cavalcante, N. S., Alegria, M. C., Barbosa, L. R., Salinas, R. K., Guzzo, C. R., & Farah, C. S. (2015). Bacterial killing via a type IV secretion system. *Nat Commun*, *6*, 6453.

Stroukova, D., & Lakey, J. H. (2015). Pore-Forming Colicins: Unusual Ion Channels -- Unusually Regulated. In A. H. Delcour (Ed.), *Electrophysiology of Unconventional Channels and Pores* (pp. 185–208). Springer International Publishing. https://doi.org/10.1007/978-3-319-20149-8_8

Tan, K., Johnson, P. M., Stols, L., Boubion, B., Eschenfeldt, W., Babnigg, G., Hayes, C. S., Joachimiak, A., & Goulding, C. W. (2015). The structure of a contact-dependent growth-inhibition (CDI) immunity protein from Neisseria meningitidis MC58. *Acta Crystallogr F Struct Biol Commun*, *71*, 702–709.

Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., & Oppenheim, A. (2007). Recombineering: genetic engineering in bacteria using homologous recombination. In *Current protocols in molecular biology / edited by Frederick M*. Unit 1 16.

Unterweger, D., Miyata, S. T., Bachmann, V., Brooks, T. M., Mullins, T., Kostiuk, B., Provenzano, D., & Pukatzki, S. (2014). The Vibrio cholerae type VI secretion system employs diverse effector modules for intraspecific competition. *Nat Commun*, *5*, 3549.

Vassallo, C. N., Cao, P., Conklin, A., Finkelstein, H., Hayes, C. S., & Wall, D. (2017). Infectious polymorphic toxins delivered by outer membrane exchange discriminate kin in myxobacteria. *Elife*, *6*.

Walker, D., Lancaster, L., James, R., & Kleanthous, C. (2004). Identification of the catalytic motif of the microbial ribosome inactivating cytotoxin colicin E3. *Protein Science*, *13*(6), 1603–1611.

Wang, Y., Jiang, Y., Meyering-Voss, M., Sprinzl, M., & Sigler, P. B. (1997). Crystal structure of the EF-Tu.EF-Ts complex from Thermus thermophilus. *Nature Structural Biology*, *4*(8), 650–656. https://doi.org/10.1038/nsb0897-650

Weiland-bräuer, N. (2021). *Friends or Foes — Microbial Interactions in Nature*.

Whitney, J. C., Beck, C. M., Goo, Y. A., Russell, A. B., Harding, B. N., Leon, J. A. De, Cunningham, D. A., Tran, B. Q., Low, D. A., Goodlett, D. R., Hayes, C. S., & Mougous, J. D. (2014). Genetically distinct pathways guide effector export through the type VI secretion system. *Mol Microbiol*, *92*, 529–542.

Whitney, J. C., Peterson, S. B., Kim, J., Pazos, M., Verster, A. J., Radey, M. C., Kulasekara, H. D., Ching, M. Q., Bullen, N. P., Bryant, D., Goo, Y. A., Surette, M. G., Borenstein, E., Vollmer, W., & Mougous, J. D. (2017). A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. *Elife*, *6*.

Willett, J. L., Gucinski, G. C., Fatherree, J. P., Low, D. A., & Hayes, C. S. (2015). Contact-dependent growth inhibition toxins exploit multiple independent cell-entry pathways. *Proc Natl Acad Sci U S A*, *112*, 11341–11346.

Willett, J. L., Ruhe, Z. C., Goulding, C. W., Low, D. A., & Hayes, C. S. (2015). Contact-Dependent Growth Inhibition (CDI) and CdiB/CdiA Two-Partner Secretion Proteins. *J Mol Biol*, *427*, 3754–3765.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*, *67*, 235–242.

Xu, X., Zhai, Y., Sun, F., Lou, Z., Su, D., Xu, Y., Zhang, R., Joachimiak, A., Zhang, X. C., Bartlam, M., & Rao, Z. (2006). New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J Virol*, *80*, 7909–7917.

Yamazaki, Y., Niki, H., & Kato, J. (2008). Profiling of Escherichia coli Chromosome database. *Methods in Molecular Biology (Clifton, N.J.)*, *416*, 385–389. https://doi.org/10.1007/978-1-59745-321-9_26

Yorgey, P., Lee, J., Kördel, J., Vivas, E., Warner, P., Jebaratnam, D., & Kolter, R. (1994). Posttranslational modifications in microcin B17 define an additional class of DNA gyrase inhibitor. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(10), 4519–4523. https://doi.org/10.1073/pnas.91.10.4519

Zahran, H. H. (1999). Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiology and Molecular Biology Reviews : MMBR*, *63*(4), 968–989, table of contents. https://doi.org/10.1128/MMBR.63.4.968-989.1999

Zechiedrich, E. L., Khodursky, A. B., & Cozzarelli, N. R. (1997). Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in Escherichia coli. *Genes & Development*, *11*(19), 2580–2592. https://doi.org/10.1101/gad.11.19.2580

Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M., & Aravind, L. (2012). Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct*, *7*, 18.

Zhang, D., Iyer, L. M., & Aravind, L. (2011). A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res*, *39*, 4532–4552.

Zhang, M., Li, X., Deng, Z., Chen, Z., Liu, Y., Gao, Y., Wu, W., & Chen, Z. (2017). Structural Biology of the Arterivirus nsp11 Endoribonucleases. *J Virol*, *91*.