UNIVERSITY OF CALIFORNIA, IRVINE

Characterization of Toxin/Immunity Complexes Involved in Bacterial Contact-Dependent Growth Inhibition

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Parker McCormick Johnson

Dissertation Committee:
Professor Celia W. Goulding, Chair
Professor Thomas Poulos
Professor Andrej Luptak

2016

DEDICATION

To everyone who helped me make it to the other side in one piece; I wouldn't have gotten this far without you:

To my family:
To my awesome Mom Ramey, who was always there to take care of me, support me, and cheer me up for as long as I can remember

To my Grandma and Grandpa Jade and Leonard, who pushed me to be better than I was

To my Dad Greg, who's always been there to help when I needed him

To my sister Skyler, for being plain awesome and never failing to make me laugh when I need it

To all of my friends, for keeping me sane, laughing, and having fun through this crazy journey

And to the best labmates a grad student could ask for, especially Heidi, Angie, and Nick, who weren't just great friends these last five years, but great mentors from the very beginning

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I am grateful to my principal investigator and mentor, Dr. Celia W. Goulding for giving me the opportunity to do research in her lab. When I entered her lab, I was inexperienced and had a weak biochemical background, but her constant support and motivation drove me to becoming the scientist I am today, a structural biologist and biochemist.

I would also like to thank the other members of my thesis committee, Dr. Thomas L. Poulos and Dr. Andrej Luptak for their support and insight.

To the Goulding lab, both past and present members: thank you all for making the days in lab as fun as they were educational. A very special thanks to Heidi Contreras, Angie Iniguez, and Nicholas Chim for helping me transition into the lab, always great mentors, co-workers, and amazing friends. Thanks for all the help, the laughs, and the stories; graduate school wouldn't have been the same without you. Thank you to Gaelle Batot for being a great friend and member of the team. Thank you to Robert Morse for training me and helping me learn X-ray crystallography.

Thank you to the two amazing undergraduate students I had the pleasure of working with: Timothy Wong and Xin 'Crystal' Liu. The two of you are brilliant and hard-working students, thanks for all of your help, and I know you'll both go on to do great things, good luck!

Thank you to all the other labs I had the pleasure of working with: the Hayes and Low labs at UC Santa Barbara, for a great collaboration and the opportunity to work on such an exciting project; and to great neighbors in the Tsai and Poulos labs here at UC Irvine.

I would also like to thank the department of Molecular Biology and Biochemistry here at UCI, and the great people who keep it running smoothly.

I would also like to thank The American Society for Biochemistry and Molecular Biology, the National Academy of Sciences, and the International Union of Crystallography for permission to reprint published material in chapters 2, 3, and the appendix, respectively.

Finally, thank you to all my family and friends who helped me get this far, I couldn't have done it without you!

CURRICULIM VITAE

Parker M. Johnson

Education

Ph.D. Molecular Biology and Biochemistry 2016
University of California, Irvine

B.S. Microbiology, Immunology and Molecular Genetics 2011
University of California, Los Angeles

Research Experience

Department of Molecular Biology and Biochemistry 2011-2016
University of California, Irvine
Doctoral thesis research conducted with Dr. Celia W. Goulding

- Conducted biochemical research on the x-ray crystal structure and catalytic function of toxic nucleases from the contact-dependent growth inhibition (CDI) pathway from pathogenic bacteria to identify potential targets for the development of novel antimicrobials.
- Trained and supervised two undergraduate students in conducting their own independently funded research projects leading to their admission into medical school.
- Presented research at numerous professional conferences and meetings and published five primary journal articles and reviews including two first-author papers.

Department of Molecular, Cellular, and Developmental Biology 2009-2011
University of California, Los Angeles
Undergraduate researcher in the lab of Dr. Chentao Lin

- Conducted independent genetic research on several plant genes relevant to modulating flowering times for agricultural purposes.
- Trained and supervised several other undergraduate researchers undergoing independent research projects.

Research Interests

Structural and functional characterization of a novel toxin/immunity gene pair from Mycobacterium tuberculosis utilizing the type VII secretion system and possibly involved in bacterial virulence (present).

Structural and functional analysis of bacterial toxin/immunity pairs involved in the contact-dependent growth inhibition (CDI) pathway from *Burkholderia pseudomallei* and Rearrangement hot-spot (Rhs) pathway in *Salmonella typhimurium* (present).

Structural and functional characterization of a unique step in contact-dependent growth inhibition (CDI) seen in uropathogenic *Escherichia coli* 536 (UPEC536) which requires toxin binding to a permissive factor, CysK (O-acetylserine sulfhydrylase), within the target cell to become catalytically and active inhibit growth (present).

Determination of the CIB1 transcription factor binding site on the *Arabidopsis thaliana* FT gene promoter (undergraduate).

Identification and characterization of ZTL (ZEITLUPE)-interacting proteins within the *Arabidopsis thaliana* proteome (undergraduate).

Publications

Johnson PM, Beck CM, Morse RP, Garza-Sanchez F, Low DA, Hayes CS, Goulding CW. Unraveling the essential role of CysK in CDI toxin activation. (Manuscript in press at Proc Natl Acad Sci.)

Johnson PM, Gucinski GC, Garza-Sanchez F, Wong T, Hung LW, Hayes CS, Goulding CW. Functional Diversity of Cytotoxic tRNase/immunity Protein Complexes from *Burkholderia pseudomallei.* (Manuscript in press at J Biol Chem.)

Johnson PM, Batot G, Goulding CW. Understanding Bacterial Toxin Diversity - Probing Nature for Novel Antimicrobial Targets. (Review article). (Manuscript in preparation).

Morse RP, Willett JLE, **Johnson PM**, Zheng J, Credali A, Iniguez A, Nowick JS, Hayes CS, Goulding CW. 2015. Diversification of β-augmentation interactions between CDI toxin/immunity proteins. J Mol Biol. 23: 3766-3784.

Tan K, **Johnson PM**, Stols L, Boubion B, Eschenfeldt W, Babnigg G, Hayes CS, Joachimiak A, Goulding CW. 2015. The structure of a contact-dependent growth-inhibition (CDI) immunity protein from *Neisseria meningitidis* MC58. Acta Cryst. F71: 702-709.

Chim N, **Johnson PM**, Goulding CW. 2013. Insights into redox sensing metalloproteins in Mycobacterium tuberculosis. J. Inorg. Biochem. 133: 118-126. (Review article).

Posters and Presentations

Johnson PM, Goulding CW. Characterization of a unique step in Contact-dependent Growth Inhibition (CDI) from uropathogenic *E. coli* 536. Mechanisms for Bacterial Toxin Delivery Symposium. Santa Barbara, CA (June 2016). Presentation.

Johnson PM, Goulding CW. Characterization of a unique step in Contact-dependent Growth Inhibition (CDI) from uropathogenic *E. coli* 536. Vertex Pharmaceuticals 'Vertex Day'. Irvine, CA (March 2016). Presentation.

Johnson PM, Morse RP, Low DA, Hayes CS, Goulding CW. Characterization of a unique step in Contact-dependent Growth Inhibition (CDI) from uropathogenic *E. coli* 536 (UPEC536). UCI Pharmaceutical Sciences 2nd Annual EWHA-UCI Joint Symposium. Irvine, CA (January 2016). Poster.

Johnson PM, Goulding CW. Characterization of a unique step in Contact-dependent Growth Inhibition (CDI) from uropathogenic *E. coli* 536. University of California, Irvine Molecular Biology and Biochemistry Departmental Retreat. Lake Arrowhead, CA (March 2015). Presentation.

Johnson PM, Morse RP, Low DA, Hayes CS, Goulding CW. Structural Analysis of Toxin/Immunity Complexes in Contact-dependent Growth Inhibition Systems. Protein Society Symposium. San Diego, CA (July 2014). Poster.

Johnson PM, Goulding CW. Structural Analysis of CDI in *Burkholderia pseudomallei* E479 and uropathogenic *E. coli* 536. Contact-dependent Growth Inhibition (CDI) collaborative meeting. Santa Barbara, CA (July 2013). Presentation.

Johnson PM, Wong T, Morse RP, Diner EJ, Hayes CS, Goulding CW. Structural Analysis of a Unique Step in Contact-dependent Growth Inhibition from Uropathogenic E. coli 536. West Coast Protein Crystallography Conference. Monterey, CA. (March 2013). Poster.

Awards and Honors:
Joseph H. Stephens Memorial Fellowship Award, granted by UCI Francisco J. Ayala School of Biological Sciences to a graduate student for excellence in biochemistry research; University of California, Irvine (2015).

Best Research Talk Award, University of California, Irvine Molecular Biology and Biochemistry Departmental Retreat; Lake Arrowhead, CA (2015).

Dean's Honors List, University of California, Los Angeles (2008-2011)

Professional Memberships
Member, The Protein Society

Teaching Appointments
Francisco J. Ayala School of Biological Sciences, *University of California,* Irvine

- M116L Experimental Molecular Biology Lab – 2012-2013, 2015-2016 (Four quarters)
- BioSci 99 Biochemistry – 2013, 2015, 2016 (Three quarters)
- BioSci 98 Molecular Biology – 2012 (One quarter)

ABSTRACT OF THE DISSERTATION

Characterizing Toxin/Immunity Complexes Involved in Bacterial Contact-Dependent Growth Inhibition

By

Parker McCormick Johnson

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2016

Professor Celia W. Goulding, Chair

Bacteria have developed complex mechanisms to survive and propagate within their environments. Contact-dependent growth inhibition (CDI) is a recently discovered mechanism of inter-bacterial competition and communication widespread amongst Gram-negative bacteria. CDI+ cells inhibit the growth of neighboring cells upon delivery of a toxic protein (CdiA-CT) into the target cell cytosol. To prevent auto-inhibition, CDI^+ bacteria express an immunity protein (CdiI) to specifically bind and inactivate cognate toxin. Here we provide further structural characterization of CDI toxin/immunity proteins to elucidate the diversity of CdiA-CT toxin activity, their potential activation, and neutralization to repurpose CDI systems for novel antimicrobials.

Burkholderia pseudomallei isolates E479 and 1026b have been shown to utilize CDI to outcompete neighboring bacteria, signifying the pathway's importance in growth and survival. The CdiA-CT toxins from these isolates are functional tRNases, yet have unique tRNA specificities and cleavage sites despite sharing no sequence similarity. We have solved the X-ray crystal structure of the *B. pseudomallei* E479 toxin/immunity protein complex and compared it to the

previously solved 1026b complex. Whereas the toxins are unique in both sequence and function, they share significant structural homology with nearly identical active sites, only differing in active site pocket size, correlating with the differing substrate cleavage sites of the two toxins. The toxin/tRNA interactions have been confirmed via molecular docking as well as small angle X-ray scattering.

The CDI system from uropathogenic *E. coli* 536 (UPEC536) is unusual, CdiA-CT exhibits toxic tRNase activity only when bound to target cell CysK, termed a "permissive factor", involved in cysteine biosynthesis. We have solved the X-ray crystal structure of the CysK/toxin/immunity ternary complex. The UPEC536 toxin inserts its C-terminus into the active site cleft of CysK, mimicking its natural interaction with CysE. A homologous toxin from the Gram-positive *Ruminococcus lactaris* has been identified to perform the same enzymatic function in a CysK-independent manner. Characterization of this toxin has shown that it's significantly more thermostable than the UPEC toxin, and is capable of binding tRNA substrate in the absence of CysK. This demonstrates a chaperone-like function for CysK, stabilizing the UPEC toxin fold and promoting its association with tRNA substrate.

CHAPTER 1

An Introduction On Bacterial Toxin Delivery Systems and Their Role in Communication and Competition

Bacterial communication and competition

In nearly every inhabitable environment on the planet, bacteria form complex networks and populations, which includes host species such as insects, plants, and animals (1). These mixed species bacterial populations are in constant flux with each other, competing for both space to grow and resources to survive and propagate. Bacteria employ multiple strategies in order to gain competitive advantage and survive within a hostile environment, or to inhibit the growth of neighboring bacterial populations (2). Bacterial competition and communication mechanisms both utilize similar pathways, whereby these signals are transmitted by secreting effector molecules into the extracellular space, or by direct delivery of molecules into neighboring cells.

Quorum sensing

Bacterial quorum sensing (QS) is a common cell-to-cell communication process and cooperative strategy by which a population of bacterial cells will ultimately produce and secrete small chemical effectors known as autoinducers (AIs) (3) or quormons (4). When these AIs are detected by protein receptors, they trigger significant changes in gene expression within a population of cells, ultimately inducing group behaviors such as virulence factor expression, bioluminescence production, genetic exchange, and biofilm production (5-9).

In response to external stimuli, bacteria begin to produce and secrete AIs into the extracellular space, whereby neighboring cells in the population will uptake and detect these small molecules. While QS systems have been identified in a wide array of both Gram-positive and Gram-negative bacteria, individual species and strains utilize different molecules, with Gram-negative bacteria using small molecules synthesized from common metabolites such as fatty-acids (10, 11), while Gram-positive bacteria tend to utilize small oligopeptides (12, 13). As a population grows, the levels of AIs in the extracellular space also increase. When their concentration reaches a certain threshold, it activates a downstream signal-transduction cascade through a QS receptor, triggering a set of gene expression changes designed for growth in a high population density. This mechanism allows bacterial populations to sample the environment and respond to the accumulation of specific AIs (8).

Bacterial Biofilms

One common downstream effect of quorum sensing is the formation of an extracellular biofilm, during which bacteria within a closely associated population will secrete multiple molecules including polysaccharides, polypeptides, lipids, and nucleic acids, forming a thick extracellular matrix and locking the cells into close proximity and promoting a dormant metabolic state (14). Biofilm formation ultimately occurs in four separate steps: adhesion, development, maturation, and dispersion (15). Adhesion occurs using surface proteins or assemblies (often motility factors such as flagella) to promote cell-to-cell and ultimately cell-to-surface adhesion, initiating the first step of biofilm formation, as biofilms cannot form without being initially adhered to a surface, either biotic or abiotic. The cells then begin to modulate from a planktonic growth state to a more dormant, immobile state, at which point the biofilm will start to mature via secretion of the extracellular matrix as well as the addition of other cells via cell-to-cell adhesion. Within a mature

biofilm, some cells will leave the extracellular matrix and return to a planktonic and motile state, free to search for nutrients or even establish another biofilm at another location. In the case of pathogenic bacteria, biofilms promote the spread of infection or even cause re-inoculation of a patient undergoing treatment.

After formation of a biofilm, the cells maintain a significant advantage compared to planktonic cells, and are capable of forming channels to distribute water, nutrients, oxygen, and even enzymes between cells within the biofilm. This growth advantage granted by biofilm formation also includes protection from environmental hazards, antibiotics, and even host immune responses when colonizing a host organism, ultimately making bacterial infections utilizing biofilms significantly more difficult to treat using conventional treatments (15-17). Biofilms also provide an excellent environment for virulence factor secretion and horizontal gene transfer to occur, that increases the likelihood of pathogenic strains causing disease, as well as acquiring and spreading antibiotic resistance genes. Due to this increase in bacterial survival, bacterial biofilms present a significant problem in food processing as well as healthcare, thereby making the development of strategies to combat biofilm formation or to bypass a biofilm to kill bacterial populations a necessity to aid in combatting life-threatening pathogens.

Inter-bacterial Competition

One widespread competitive strategy bacteria utilize is the secretion or delivery of protein toxins into neighboring cells, leading to death of non-cognate bacterial strains to reduce competition for space and resources (2, 18, 19). These protein toxins are members of several different unique toxin-delivery pathways known collectively as polymorphic toxin systems (PTS). Notably, PTS are found in many pathogenic bacteria, and have been shown to utilize multiple different secretion

pathways including type IV (T4SS) (20), type V (T5SS) (21), type VI (T6SS)- (22-24), and type VII (T7SS) (25-27) secretion systems for toxin delivery, often relying on inhibitory cell contact with the target cell. To prevent auto-inhibition, delivered toxins are neutralized by complex formation with highly specific- cognate immunity proteins. Notably, polymorphic toxin/immunity pairs are predicted to have high functional and structural diversity with dozens of unique toxin/immunity pair families identified *in silico*, and only a small subset experimentally characterized (19, 28).

Colicins

Some of the most studied and best understood PTS are colicins and S-type pyocins, toxic secreted bacteriocins produced by *Escherichia coli* and *Pseudomonas* species, respectively. Colicins were first discovered being produced by *E. coli* species within the human gastrointestinal tract, utilized by enteric bacteria to specifically kill bacteria of the same or closely related species (2, 29, 30). The secreted colicins are small, cytotoxic proteins commonly displaying a pore-forming or nuclease activity to inhibit cell growth. To prevent auto-inhibition or the targeted killing of self cells, colicin-producing bacteria also express a cognate immunity protein to inactivate the specific colicin being delivered (2, 30-33). Interestingly, colicin secretion does not use a typical secretion pathway, but instead utilizes a protein which lyses the cell-wall to release colicins into the extracellular space, resulting in the death of the colicin-producing cell. Common to PTS, colicins are modular proteins consisting of multiple domains; a conserved N-terminal translocation domain, a central receptor-binding domain, and a polymorphic C-terminal cytotoxic domain responsible for cell death. The discovery of colicins and other bacteriocins promoted the discovery and study of other PTS involved in bacterial survival and pathogenesis, such as the contact-dependent growth inhibition (CDI) pathway, and the multitude of Type VI-secreted effectors.

Contact-dependent Growth Inhibition (CDI)

Discovery of CDI

In the Low laboratory at UC Santa Barbara in 2005, *Escherichia coli* EC93 was isolated from rat intestines, it was found to be the only strain of *E. coli* present in the rat's microbiota. This *E. coli* strain demonstrated a strong growth advantage over other enteric *E. coli* strains, implying the presence of a potent inter-bacterial competition mechanism (34). Experiments showed that *E. coli* EC93 was capable of out competing *E. coli* K12 in liquid media by several thousand fold after only a few hours of growth. Interestingly, this competitive advantage was lost when the two populations of cells were separated by a membrane which prevents cell-to-cell contact, but allows small secreted molecules to pass through, signifying that this growth advantage was not reliant on the secretion of toxic molecules into the extracellular space. This competitive mechanism was then termed "contact-dependent growth inhibition" or CDI (34).

The CDI pathway is an extremely widespread inter-bacterial competition mechanism present in α, β, and γ-proteobactera, and particularly well represented within Gram-negative pathogens (19). CDI is encoded by three genes: *cdiB, cdiA, and cdiI,* located on the chromosome as a *cdiBAI* gene cluster in most Gram-negative species, and in a *cdiAIB* cluster in *Burkholderia* species (**Figure**

1.1). CdiB is a membrane transporter; CdiA, a large cell-surface protein which contains the active toxin of the CDI system; and CdiI, the immunity protein that prevents auto-inhibition (35).

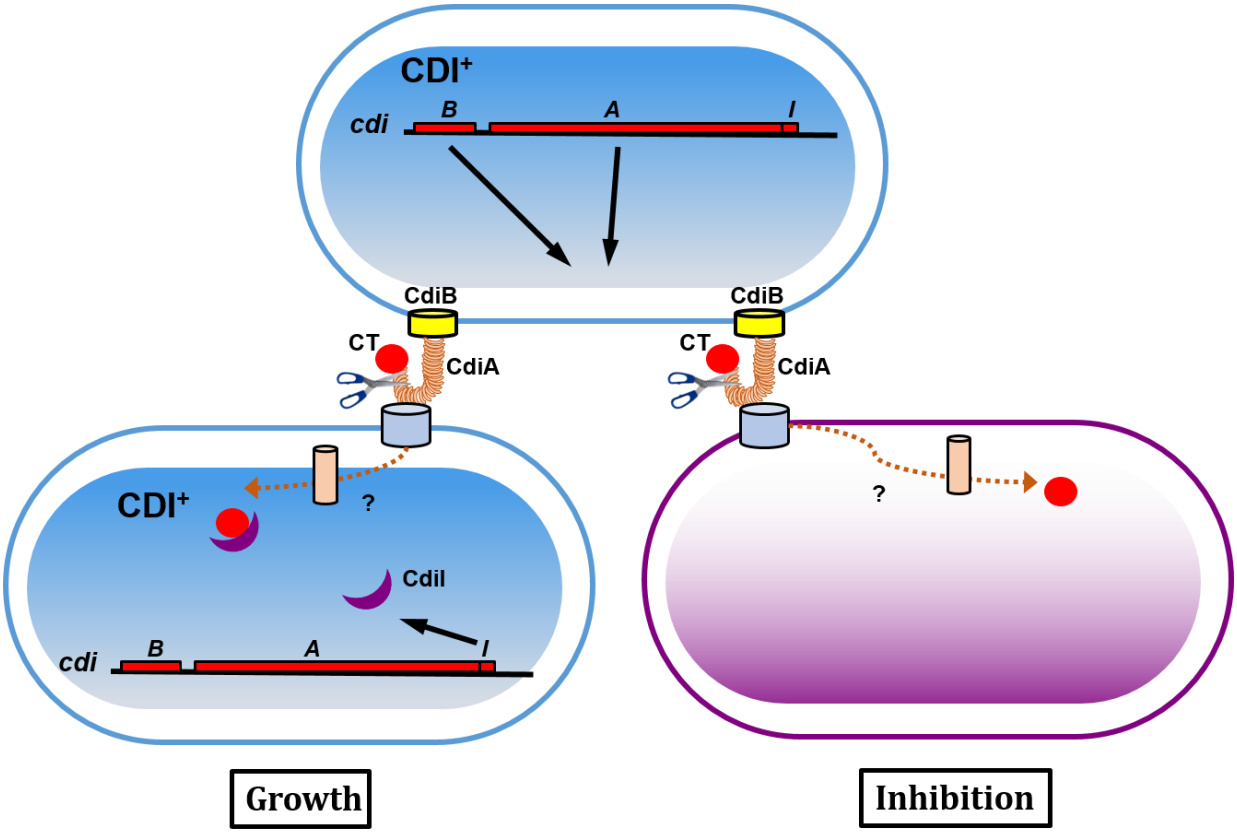


Figure 1.1: Model of Contact-dependent Growth Inhibition (CDI). CDI allows for competitive advantage in the ability to kill neighboring bacterial cells mediated by contact of the CdiA protein with a target cell, where its C-terminal domain (CdiA-CT) is imported into the target cytosol and inhibits growth of non-cognate target cells. In order to prevent susceptibility to their own toxins, CDI^+ bacteria express a cognate immunity protein (CdiI) to specifically bind and inactivate its CdiA-CT.

Mechanism of CDI

CDI functions via a CdiB/CdiA two partner T5SS by which CdiB, an Omp85-family β-barrel outer membrane protein, will export and assemble the large CdiA protein to the cell surface. CdiA proteins are large between 180 and 640 kDa in size, consisting primarily of filamentous hemagglutinin (FHA/FhaB) adhesin repeats, which are predicted to form an extended β-helical

structure, protruding several hundred angstroms from the cell surface (**Figure 1.2**) (19, 21). As predicted from the domain architecture, CdiA proteins display high conservation at the large N-terminal region, but have polymorphic C-terminal regions (CdiA-CT, 200-300 residues) demarcated by a highly conserved VENN motif in most species (**Figure 1.3**) and an (E/Q)LYN motif in *Burkholderia* species (36). These variable CdiA-CT regions contain at least two domains including the functional C-terminal toxic domain. While multiple unique catalytic activities have been identified for CdiA-CT toxins *in silico*, only a small subset have been experimentally characterized, whereby most seem to either possess pore-forming activities or nuclease activities, degrading tRNA, rRNA, or DNA (28, 35).

To prevent susceptibility to their own toxins, CDI^+ bacteria express a cognate immunity protein (CdiI) to specifically bind and inactivate CdiA-CT. Upon delivery of CdiA-CT to a target cell, if

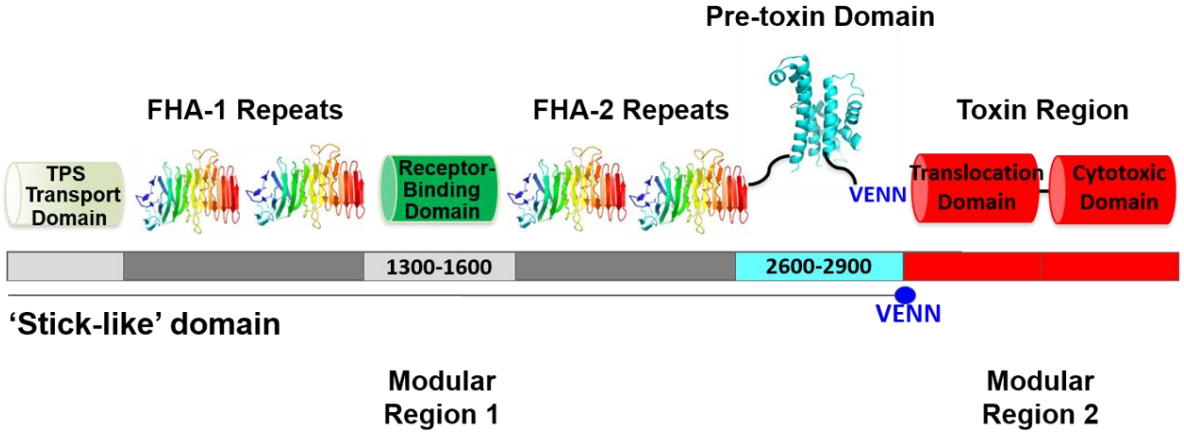


Figure 1.2: Architecture of CdiA. CdiA consists of a TPS transport domain, a series of FHA-1 repeats, a receptor-binding domain, FHA-2 repeats, a pre-toxin domain, then the CdiA-CT toxin region demarcated by the VENN motif. The toxin region consists of two domains: a translocation domain used for target cell entry, and a cytotoxic domain to inhibit target cell growth. CdiA also has two modular regions: the receptor binding domain, and the toxin region consisting of a translocation domain and the cytotoxic domain. These two regions of CdiA can be traded with those from other CDI systems, whereby replacing the receptor-binding domain can allow delivery of the C-terminal toxin to a new subset of target cells. Generating a chimeric CdiA by changing the toxin region to that of another CDI system does not affect the toxin targeting, but the chimeric CdiA protein will then deliver a non-cognate toxin to the same subset of recognizable target cells.

the target cell is CDI^+, cognate CdiI will bind to the inhibitory cell CdiA-CT, inactivating the toxin and facilitating cell growth. This interaction between cognate CdiA-CT and CdiI is highly

specific; no mechanisms of toxin inactivation by a non-cognate immunity protein have been observed, thus far . Cognate toxin/immunity pairs form highly stable complexes, binding with nanomolar affinities due to strong non-covalent interactions, with only weak interactions detected between non-cognate pairs (36-38).

When CdiA comes into contact with specific outer membrane target cell receptors, the cytotoxic CdiA-CT region is cleaved and translocated into the cytoplasm of the target bacterium by an unknown mechanism, where it will carry out its cytotoxic activity leading to growth inhibition and cell death (**Figure 1.1**) (19, 34). The target cell receptor utilized varies between CDI systems; however, it is commonly a β-barrel outer membrane protein such as BamA or OmpC/OmpF (35, 39, 40). Surprisingly, the interaction with the target cell receptor does not rely on the CdiA-CT region at the CdiA tip, but instead within a small receptor-binding domain located in the center of the large CdiA N-teminal region. As observed for colicins, CdiA proteins are modular (**Figure 1.2**); due to the target cell interaction being independent of the cytotoxic region, functional chimeric CdiA proteins can be generated (35, 39). These CdiA chimeras can be cloned at the VENN motif, trading entire C-terminal toxins between widely different CDI+ species. These chimeric CdiA proteins retain the ability to deliver non-cognate toxins into target cells utilizing the same target cell receptor and mechanism, where only cells expressing the immunity protein which can inactivate the toxic C-terminal region are resistant to growth inhibition (35, 39).

Recently, it has been shown that CdiA-CT N-terminal domains play a vital role in toxin translocation, relying on specific inner membrane proteins, commonly metabolite permeases, however their metabolic transport activities are not required for translocation of the toxins (41). It is not known whether or not these translocation domains have a direct interaction with the inner membrane proteins, however this is a highly likely mechanism of toxin entry into the target cell

cytosol, possibly by exploiting the proton motive force, which has been shown to be required for CDI (42). Surprisingly, much like the chimeric CdiA proteins discussed earlier, the translocations appear to be modular as well, where swapping two translocation domains between CdiA-CT toxins allows for translocation into non-canonical target cells (41, 43), further demonstrating the impressive modularity of the CDI system.

It should be noted that within a species of CDI+ bacteria, some strains may not express *cdiBAI* homologues, while others may contain multiple loci, however only one gene cluster tends to be expressed at any given time, capable of swapping out toxin/immunity pairs in response to environmental stimuli in an unknown mechanism (19, 28). While not classical virulence factors due to an inability to directly affect eukaryotic cells, CDI has strong implications in bacterial competition and survival within a host organism, and thus studying the structure and function of CDI toxin/immunity protein complexes may lead to the characterization of novel drug targets in a wide array of potent bacterial pathogens.

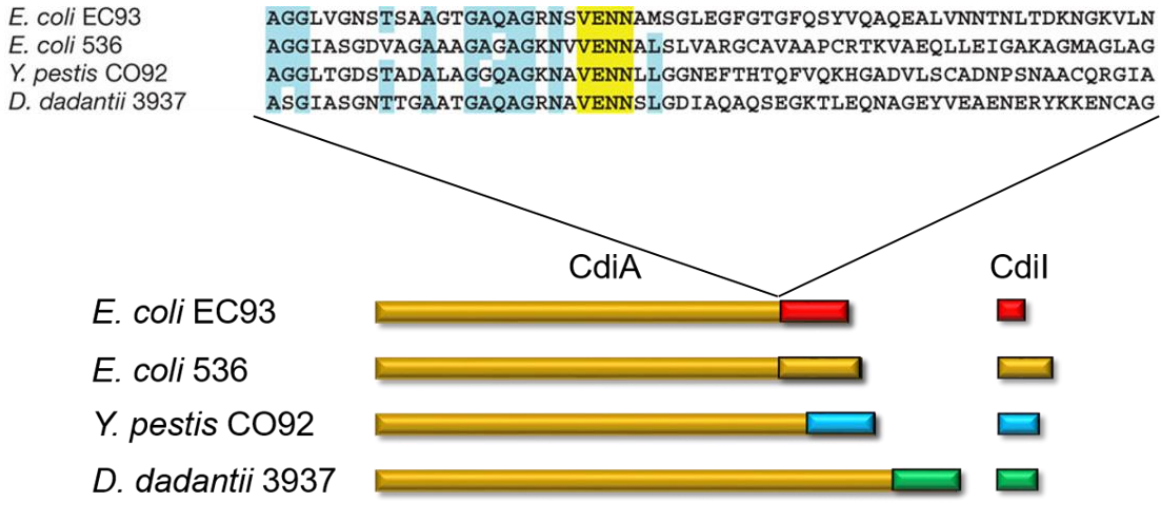


Figure 1.3: CdiA proteins have highly conserved N-termini, with polymorphic C-termini (CdiA-CT's) following a conserved VENN motif, demarcating the start of the toxin region, and demonstrating a significant loss of amino acid sequence similarity.

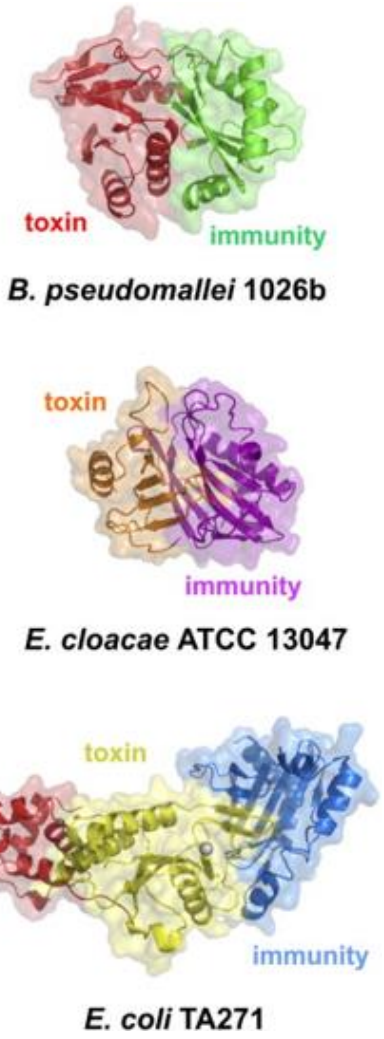


Figure 1.4: Crystal structures of CdiA-CT/CdiI toxin/immunity complexes from B. pseudomallei 1026 (PDB ID: 4G6V), E. cloacae ATCC 13047 (PDB ID: 4NTQ) and E. coli TA271 (PDB ID: 4G6U).

Characterizing CDI Toxin Structure and Function

To date, approximately 120 distinct toxin/immunity protein families have been identified in gram-negative bacteria, with homologues prevalent across multiple bacterial families. This suggests that toxin/immunity pair genes commonly undergo horizontal gene transfer, and indicates that CDI+ bacteria frequently share toxins from a large and diverse pool of effector genes (19, 28). The first CdiA-CT toxin to be experimentally characterized was that of *E. coli* EC93, which has been shown to dissipate the proton motive force by forming pores in the target cell membrane, leading to cell death (34). The most common catalytic activity amongst CDI toxins is that of a nuclease, with CDI toxins differing greatly between nucleic acid substrate, cleavage-site, or even required cofactors.

Currently, we have experimentally characterized and solved X-ray crystal structures of toxin/immunity complexes from *Burkholderia pseudomallei* E479 (2), *Burkholderia pseudomallei* 1026b (37), *E. coli* TA271 (37), uropathogenic *E. coli* 536 (Chapter 3), and *Enterobacter cloacae* (38). Each structure has a unique protein-protein interface, as well as varying modes of toxin

neutralization, by the immunity protein (**Figure 1.4**). Notably, the structures of CdiA-CT N-terminal domains are either not resolved or are truncated in the above CdiA-CT/CdiI structures.

E. coli TA271 and Burkholderia 1026b CdiA-CT/CdiI Complexes

The first crystal structures of CDI toxin/immunity complexes solved came from *E. coli* TA271 and *Burkholderia pseudomallei* 1026b (**Figure 1.4**). CdiA-CT^{TA271} has been shown to possess divalent metal ion-dependent DNase activity, demonstrating weak DNA nicking activity in the presence of Mg^{2+} ions, and extremely potent DNase activity in the presence of Zn^{2+}, capable of completely degrading plasmid DNA in a short time (37). Conversely, CdiA-CT^{Bp1026b} possesses tRNase activity, cleaving tRNA^{Ala} at the aminoacyl-acceptor stem, preventing charging of the tRNA with appropriate amino acid residues and thus halting translation and ultimately causing cell death (37).

Despite these two unique activities from widely different bacterial species, the toxin structures revealed several similarities, including an overall conserved mixed α/β-fold typical of type IIS restriction endonucleases. Both proteins consist of a central antiparallel β-sheet decorated by α-helices, forming half-β-barrel-like structures at the enzyme core. Structural alignments of the toxins reveal an rmsd of 3.9 Å, demonstrating significant structural homology, despite only ~15% sequence identity (37). Even with the significant structural homology between the two toxins, the two immunity proteins share even less sequence identity (~12%) and demonstrate no structural homology. In addition, the toxin/immunity complexes utilize widely different interaction mechanisms: the interface between CdiA-CT/CdiI^{Bp1026b} is mediated almost entirely by a large network of electrostatic interactions, with the immunity protein binding over the toxin active site and preventing access to tRNA substrate, while the CdiA-CT^{TA271}/CdiI^{TA271} complex utilizes a highly unusual mechanism of β-augmentation, whereby the toxin inserts a two-stranded β-hairpin

into a pocket within the cognate immunity protein, forming several interactions with a four-stranded antiparallel β-sheet, generating a highly stable six-stranded β-sheet with the toxin hairpin, allowing for highly specific and tight binding (37). This unique mechanism of inactivation seen in the TA271 toxin surprisingly leaves the toxin active site open to substrate binding, and thus it is currently unknown how this toxin is inactivated upon immunity protein binding, however it is known that the toxin does not undergo any conformation changes upon complex formation according to a crystal structure of the toxin in the absence of cognate immunity protein (data not shown.)

Enterobacter cloacae CdiA-CT/CdiI Complex

From the first two CdiA-CT crystal structures, it was thought that perhaps these toxins share a conserved fold despite significant differences in amino acid sequence; however, the structure of the CdiA-CT/CdiI complex from *Enterobacter cloacae* (ECL) appears distinct from other CDI toxins (**Figure 1.4**). CdiA-CT^{ECL} was predicted to possess ribonuclease activity, but did not show any cleavage of tRNA molecules *in vitro*. Surprisingly, this toxin was shown to be a potent rRNase, capable of cleaving target cell 16S rRNA and inhibiting protein translation, ultimately leading to cell death (38). Unlike most ribonucleases, CdiA-CT^{ECL} consists primarily of a twisted five-stranded antiparallel β-sheet with only a single α-helix at the N-terminus, while the cognate immunity protein forms a β-sandwich decorated by three α-helices. The toxin/immunity interface is mediated by over 20 ion-pair/H-bond interactions as well as a hydrophobic interface spanning 300 Å^2 (38). As previously mentioned, this structure differs significantly from previously solved CDI toxin structures, which commonly utilize a mixed α/β-fold, and instead shows significant structural similarity to ColE3, another well-characterized rRNase which targets 16S rRNA (44, 45). A structural alignment yields an rmsd of 2.1 Å over 76 Cα's, yet only share 18% sequence

identity. Despite these similarities in the toxins, the CdiA-CT^{ECL} and ColE3 immunity proteins show no structural homology, not only presenting different protein sizes (9.9 kDa vs. 16.9 kDa for ImE3 and CdiI^{ECL} respectively), but also present widely different folds. While both proteins are mixed α/β-folds, they show no structural superimposition and even bind to unique sites on their respective toxins, however both seem to confer immunity by blocking toxin access to the ribosomal A-site. This conservation in the toxin structures potentially demonstrates a form of convergent evolution in the toxins towards a similar fold and activity, while maintaining unique interactions with cognate immunity proteins.

Role of permissive factors in polymorphic toxin systems

While most PTS toxins possess cytotoxic domains that are active in the absence of any cofactors, some polymorphic toxins require another protein cofactor either for activation or delivery into target cells. This phenomenon was first identified in the CDI pathway of uropathogenic *E. coli* 536 (UPEC536). The CdiA-CT^{UPEC536} toxin is activated when bound to target cell CysK, involved in cysteine biosynthesis, activating the toxin's potent tRNase activity (46) (**Figure 1.5**). CysK is an extremely well conserved enzyme in Gram-negative bacteria, with ~99% sequence identity between bacterial species, however it is a non-essential enzyme (47-49). CysK-expressing cells also express an isoenzyme, CysM (49), which catalyzes the same reaction but will not bind to the UPEC toxin, allowing CysK-deficient cells to be immune to CDI by UPEC cells. If the target cell is expressing cytosolic CysK, the UPEC toxin will bind to the protein with high affinity, inserting a C-terminal tail into the CysK active site cleft, mimicking a natural interaction CysK has with another biosynthetic enzyme CysE (46, 50-53). This CysK-CdiA-CT^{UPEC536} complex then is capable of cleaving cytosolic tRNA at the anticodon loop, rendering it unusable and halting protein translation in the cell, ultimately leading to growth arrest and cell death (46). The exact mechanism

of toxin activation by CysK is not known, but extensive experimental evidence shows that CysK acts as a chaperone, stabilizing the fold of the intrinsically unstable CdiA-CT^{UPEC536} toxin and promoting its ability to bind tRNA substrate (see Chapter 3).

More recently, it was shown that Tse6 (Type VI secretion exported 6), a *Pseudomonas aeruginosa* inhibitory effector which is secreted through the Type VI secretion system, requires EF-Tu for target cell delivery (54, 55). Tse6 binds inhibitory cell housekeeping protein EF-Tu for export and secretion of Tse6 into the target cell, however EF-Tu is not required for Tse6 NAD+ glycohydrolase activity. Moreover, other toxins that specifically bind EF-Tu have been identified from strains of *Escherichia coli* (unpublished data) as well as *Salmonella typhimurium* (56) (see Chapter 4). These toxins have been shown to have some sequence similarity to known ribonucleases, potentially demonstrating a role of EF-Tu in toxin localization or substrate binding due to EF-Tu's role in tRNA binding and protein translation.

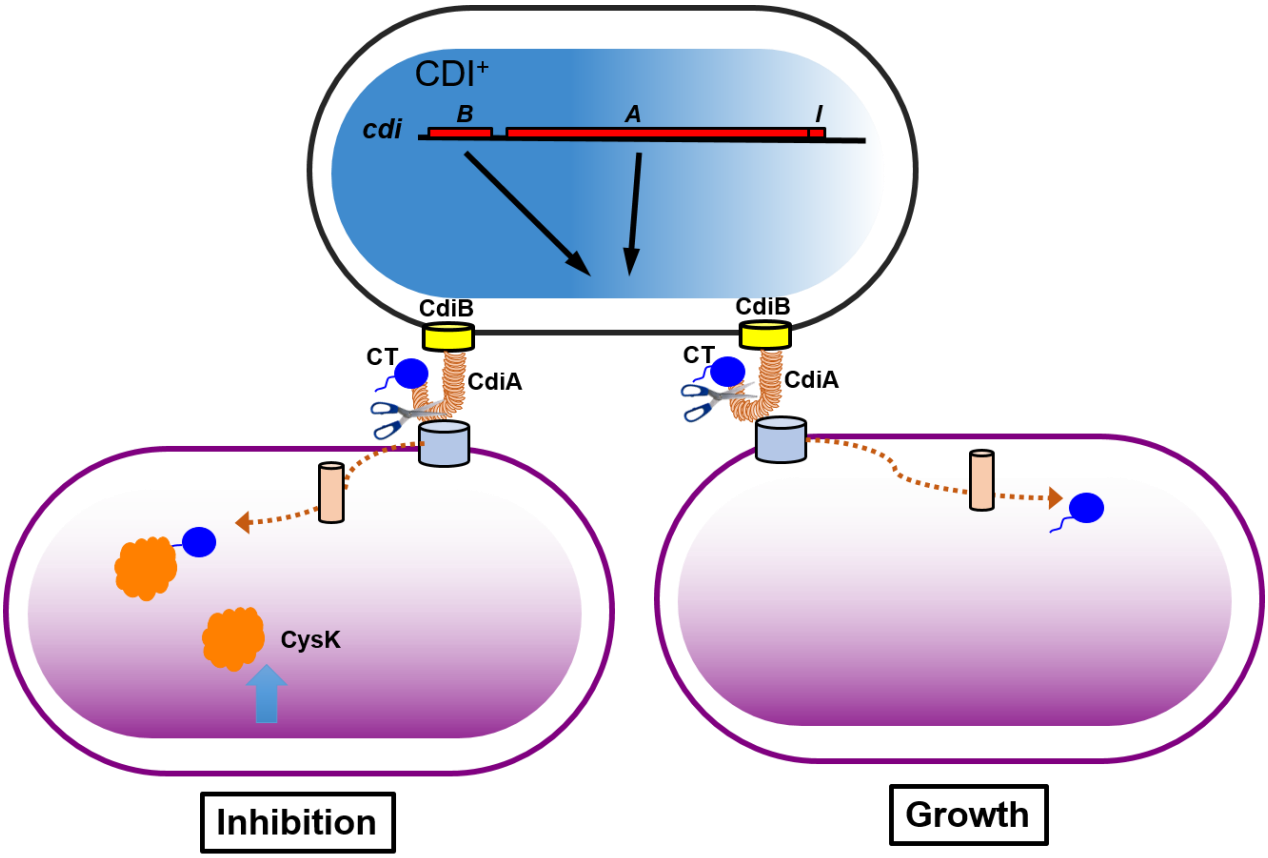


Figure 1.5: The CDI system from uropathogenic *E. coli* 536 (UPEC536) is unusual, CdiA-CT exhibits cytotoxic tRNase activity only when activated by binding target cell CysK, a permissive factor involved in cysteine biosynthesis.

Polymorphic toxins requiring endogenous proteins such as CysK and EF-Tu are uncommon amongst currently characterized systems, although this phenomenon might be more widespread than previously thought, but difficult to detect unless verified by *in vitro* methods. The requirement for endogenous cytosolic proteins may suggest that the primary function of these polymorphic toxin systems is cell-to-cell communication with regards to biofilm formation, utilizing the toxic activity of these proteins to prevent nutrient loss from inclusion of non-self cells within a biofilm as opposed to killing off neighboring cells for the purpose of growth and dispersion; however, more experimental evidence is required to test these hypotheses.

Role of CDI in communication and biofilm formation

Initially CDI was studied as a strictly competitive mechanism, utilizing the delivery of potent toxins to inhibit the growth of neighboring cells to acquire space and nutrients to grow, and providing a possible role as a virulence factor when colonizing a human with a healthy microbiota. It was thought that the competitive advantage granted by CDI would allow for killing of established bacteria to move in and cause infection, as it has been seen that the CDI+ cells can outcompete CDI- cells in liquid media, as well as in a chicory plant host in the case of *Dickeya dadantii,* where the CDI pathway has been shown to be upregulated upon host colonization (19). Further research into the mechanisms of toxin translocation, particularly in understanding the CdiA/target receptor interaction, surprisingly showed an extremely specific targeting mechanism, whereby the set of possible target cells is limited to those of the same species as the inhibitory cell, much like with colicins and other bacteriocins (35, 39, 41, 43, 57). This interaction specificity between the CdiA receptor-binding domain and the target cell surface receptor, which typically is a large β-barrel outer membrane protein, relies on the conserved sequence of the extracellular loop regions located in the receptor protein. The target receptor of $CdiA^{EC93}$ is BamA, which, like many outer membrane proteins, contains multiple extracellular loops protruding from the β-barrel transmembrane region. While the overall structure and sequence of BamA is well conserved amongst Gram-negative bacteria, the sequences of these loop regions can vary quite considerably even between different species and strains. In particular, loops 6 and 7 of BamA have been shown to be vital for the CdiA/BamA interaction, since mutating these residues in these loop regions can completely abrogate CdiA binding and renders the target cells resistant to CDI (39).

In addition to this very limited set of susceptible target cells, recent work has also shown that the CdiA proteins may play a significant role in communication and cooperative activities. CdiB and CdiA deficient mutants of *Burkholderia thailandensis* and *Escherichia coli* EC93 not only

demonstrate a significant defect in the ability to adhere to one another as well as to surfaces, but they are completely defective in biofilm formation (58, 59), signifying that the competitive advantage granted by the toxins may not be the primary function of the pathway. Intriguingly, point mutations to inactivate the CdiA-CT toxins also show a considerable defect in biofilm formation. This demonstrates a potential role for these toxins in cell-to-cell signaling to modulate gene expression to stimulate biofilm formation within cognate CDI+ cells. While CdiA promotes intercellular adhesion, this is not dependent on interactions with BamA, as CdiA mutants that cannot bind BamA are still capable of aggregating and forming biofilm structures, however the interactions with BamA are important for formation of a fully mature biofilm (59). Overall, this work demonstrates that CDI may have a primary function in cell-to-cell communication and cooperative behaviors such as biofilm formation, promoting cellular adhesion as well as upregulation of biofilm related genes by a currently unknown mechanism, whereby the toxic activity may aid in preventing non-self cells from entering a biofilm and competing for nutrients.

Type VI-Secreted Effectors

PTS differ greatly from one another in their mechanism of secretion and toxin delivery to target cells. Colicins utilize lysis proteins to weaken the cell wall to allow for toxin secretion into the extracellular space, and CDI systems utilize the CdiB/CdiA two-partner Type V secretion system to deliver a toxin through a cell-to-cell mechanism. In recent years, another class of polymorphic toxins from *Pseudomonas* species have been shown to utilize the Type VI secretion system (T6SS) to directly deliver toxic effector molecules into the target cell (23). One key feature which sets the Type VI secreted effector toxins apart from other polymorphic toxin systems is that the Type VI secretion system can interact with both prokaryotic and eukaryotic cells (60), while CDI and colicins rely on specific bacterial outer membrane proteins for delivery and thus have a limited

subset of potential target cells. The ability of bacteria which utilize the T6SS (such as *Pseudomonas aeruginosa*) to secrete effector molecules into eukaryotic host cells has been shown to play an important role in pathogenesis and evasion of host immune response (60-62).

The T6SS has significant similarities to bacteriophage spike complexes utilized for target cell interaction and genetic material delivery. Much like the phage spike complexes, the T6SS machinery forms a proteinacious needle-like complex (**Figure 1.6**), consisting of a baseplate assembly within the inhibitory cell periplasm, and a tube-like assembly within the cysosol consisting of multiple hexameric rings of the protein Hcp (20, 23, 60, 61, 63, 64). At the tip of this Hcp tube assembly is the VgrG/PAAR spike complex, which are commonly associated with the toxic effector molecule to be delivered (24, 61, 65). It is thought that upon a conformational change in the baseplate assembly, the Hcp tube with VgrG/PAAR/Effector spike complex is pushed through the inner and outer membranes, into the extracellular space and through the target cell membrane(s), at which point the VgrG/PAAR/Effector complex will dissociate, allowing the effector molecule to perform its function either modulating host-cell gene expression or inhibiting cell growth with cytotoxic activity (**Figure 1.6**) (64). Due to the common association with virulence factor effector proteins utilized by pathogenic bacteria, the type VI secretion system and its associated factors are of utmost importance to study in order to identify novel strategies and targets for the development of increasingly necessary antimicrobials.

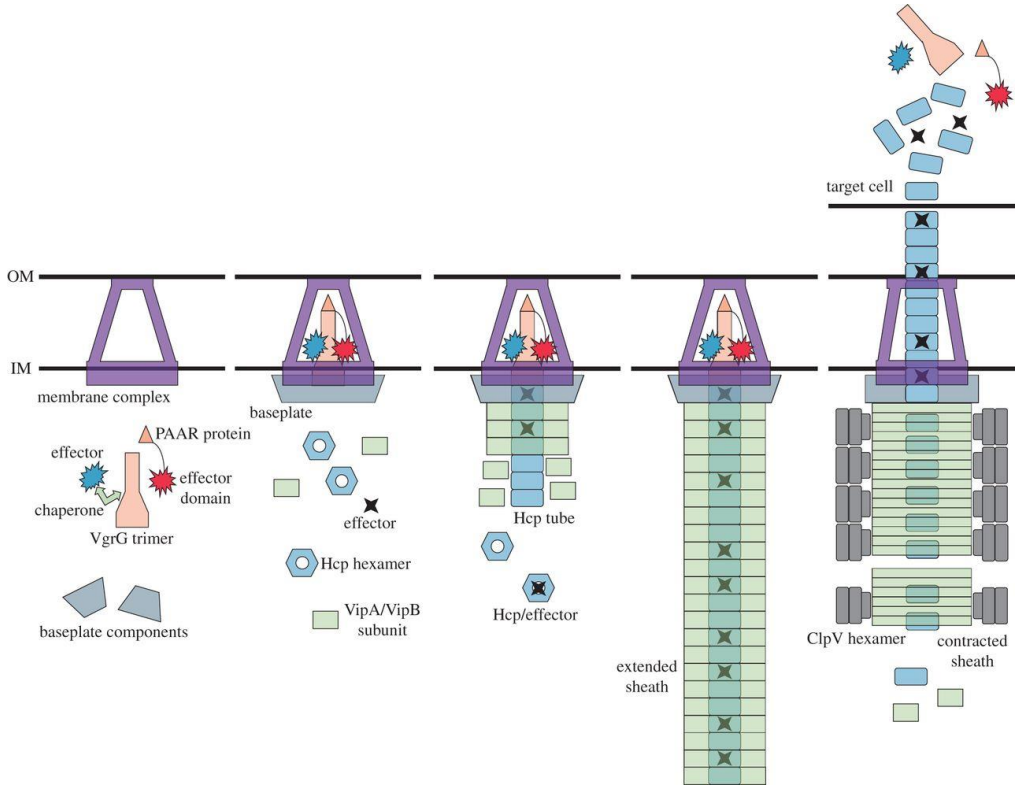


Figure 1.6 (64): Model of Type VI secretion system (T6SS). The T6SS machinery utilizes a periplasmic membrane complex with a cytosolic baseplate and tube-like complex consisting of multiple Hcp-hexamers. The tip complex consists of the VgrG/PAAR spike complex. An effector molecule can be found linked to the PAAR protein, the VgrG trimer, or an Hcp hexamer. It is thought that upon a conformational change in the baseplate/membrane complex, the needle assembly is propelled forwards through the outer membrane and into the target cell periplasm, at which point the tip assembly will dissociate and free the effector protein to act upon the target cell.

Several T6SS-secreted toxins have been identified in pathogenic *Pseudomonas* species (including *Pseudomonas aeruginosa*), referred to as Type VI secretion exported proteins (Tse's), and numbered Tse1-Tse6, with Tse5 containing N-terminal Rhs/YD- repeats. These proteins possess highly conserved N-terminal Proline-Alanine-Alanine-aRginine repeat (PAAR) domains, polymorphic C-terminal toxin domains, and highly specific toxin-neutralizing cognate immunity proteins, where the immunity gene is located downstream of that of the toxin. Much like other PTS, the effector toxins possess a wide range of cytotoxic activities such as peptidoglycan hydralases (Tse1 also called Tae1), muramidases (Tse3), NAD+ glycohydrolases (Tse6) and

DNases (54, 62, 65). Type VI amidase effectors are broadly distributed amongst Proteobacteria and can be classified in four divergent families (Tae1-4).

Rearrangement Hot-Spot (Rhs) Systems

Another polymorphic toxin system utilizing the Type VI secretion system is the Rearrangement hot-spot (Rhs) pathway. Rhs genes were first observed and reported in *E. coli* over 40 years ago, although their function was unknown (22, 66, 67). Recently, it was demonstrated that the Rhs loci encode for toxin/immunity protein pairs, where toxin delivery utilizing Type IV or Type VI secretion systems results in neighboring cell growth inhibition dependent on cell-to-cell contact (28). The genomic loci of CDI and Rhs pathways reveal several similarities: both consist of a conserved, filamentous N-terminus, with a polymorphic C-terminal catalytic toxin domain (CdiA-CT and Rhs-CT) followed by a downstream gene that encodes a highly specific cognate immunity protein (CdiI and RhsI). While CdiA proteins are marked by N-terminal hemagglutinin repeats, Rhs proteins contain N-terminal repeated Rhs motifs with the consensus sequence YDxxxGRL(I/T)b known as YD-repeats. Rhs-proteins contain a conserved PxxxxDPxGL motif to demarcate the start of the polymorphic C-terminal domain that possess nuclease activities, similar to the VENN motif in CdiA proteins (22, 68). Currently, only two Rhs-family toxin/immunity pairs have been functionally characterized: the DNase RhsA-CT from the plant pathogen *Dickeya dadantii* 3937, and the tRNase WapA-CT from the Gram-positive bacteria, *Bacillus subtilis* (69). Much like CDI, these Rhs systems provide potential targets for novel therapeutics due to their roles in bacterial competition as well as biofilm formation within a host organism during infection.

References

1. Sogin ML, *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* 103(32):12115-12120.
2. Cascales E, *et al.* (2007) Colicin biology. *Microbiology and molecular biology reviews : MMBR* 71(1):158-229.
3. Sherrer RL, *et al.* (2011) C-terminal domain of archaeal O-phosphoseryl-tRNA kinase displays large-scale motion to bind the 7-bp D-stem of archaeal tRNA(Sec). *Nucleic acids research* 39(3):1034-1041.
4. Ng WL & Bassler BL (2009) Bacterial quorum-sensing network architectures. *Annual review of genetics* 43:197-222.
5. Fuqua C & Greenberg EP (2002) Listening in on bacteria: acyl-homoserine lactone signalling. *Nature reviews. Molecular cell biology* 3(9):685-695.
6. Pappas KM, Weingart CL, & Winans SC (2004) Chemical communication in proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular signalling. *Molecular microbiology* 53(3):755-769.
7. Williams P & Camara M (2009) Quorum sensing and environmental adaptation in Pseudomonas aeruginosa: a tale of regulatory networks and multifunctional signal molecules. *Current opinion in microbiology* 12(2):182-191.
8. Hawver LA, Jung SA, & Ng WL (2016) Specificity and complexity in bacterial quorum-sensing systems. *FEMS microbiology reviews*.
9. Ng WL, *et al.* (2011) Signal production and detection specificity in Vibrio CqsA/CqsS quorum-sensing systems. *Molecular microbiology* 79(6):1407-1417.
10. Fuqua C & Greenberg EP (1998) Self perception in bacteria: quorum sensing with acylated homoserine lactones. *Current opinion in microbiology* 1(2):183-189.
11. Tiaden A, Spirig T, & Hilbi H (2010) Bacterial gene regulation by alpha-hydroxyketone signaling. *Trends in microbiology* 18(7):288-297.
12. Kleerebezem M, Quadri LE, Kuipers OP, & de Vos WM (1997) Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Molecular microbiology* 24(5):895-904.
13. Lyon GJ & Novick RP (2004) Peptide signaling in Staphylococcus aureus and other Gram-positive bacteria. *Peptides* 25(9):1389-1403.
14. Danese P, Pratt L, & Kolter R (2000) Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture. *J. Bacteriol.* 182:3593-3596.
15. Sharma G, *et al.* (2016) Escherichia coli biofilm: development and therapeutic strategies. *Journal of applied microbiology* 121(2):309-319.
16. Ito A, Taniuchi A, May T, Kawata K, & Okabe S (2009) Increased antibiotic resistance of Escherichia coli in mature biofilms. *Applied and environmental microbiology* 75(12):4093-4100.
17. Mittal S, Sharma M, & Chaudhary U (2015) Biofilm and multidrug resistance in uropathogenic Escherichia coli. *Pathogens and global health* 109(1):26-29.
18. Hayes CS & Low DA (2009) Signals of growth regulation in bacteria. *Current opinion in microbiology* 12(6):667-673.
19. Aoki SK, *et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* 468(7322):439-442.
20. Souza DP, *et al.* (2015) Bacterial killing via a type IV secretion system. *Nature communications* 6:6453.
21. Hayes CS, Aoki SK, & Low DA (2010) Bacterial contact-dependent delivery systems. *Annual review of genetics* 44:71-90.

22. Sisto A, *et al.* (2010) An Rhs-like genetic element is involved in bacteriocin production by Pseudomonas savastanoi pv. savastanoi. *Antonie van Leeuwenhoek* 98(4):505-517.
23. Hood RD, *et al.* (2010) A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. *Cell host & microbe* 7(1):25-37.
24. Hachani A, *et al.* (2011) Type VI secretion system in Pseudomonas aeruginosa: secretion and multimerization of VgrG proteins. *The Journal of biological chemistry* 286(14):12317-12327.
25. Bitter W, *et al.* (2009) Systematic genetic nomenclature for type VII secretion systems. *PLoS pathogens* 5(10):e1000507.
26. Das C, Ghosh TS, & Mande SS (2011) Computational analysis of the ESX-1 region of Mycobacterium tuberculosis: insights into the mechanism of type VII secretion system. *PloS one* 6(11):e27980.
27. Sun J, *et al.* (2015) The tuberculosis necrotizing toxin kills macrophages by hydrolyzing NAD. *Nature structural & molecular biology* 22(9):672-678.
28. Poole SJ, *et al.* (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS genetics* 7(8):e1002217.
29. Ghequire MG & De Mot R (2014) Ribosomally encoded antibacterial proteins and peptides from Pseudomonas. *FEMS microbiology reviews* 38(4):523-568.
30. Gordon DM & O'Brien CL (2006) Bacteriocin diversity and the frequency of multiple bacteriocin production in Escherichia coli. *Microbiology* 152(Pt 11):3239-3244.
31. Graille M, Mora L, Buckingham RH, 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.
32. Yajima S, *et al.* (2004) Relation between tRNase activity and the structure of colicin D according to X-ray crystallography. *Biochemical and biophysical research communications* 322(3):966-973.
33. Papadakos G, Wojdyla JA, & Kleanthous C (2012) Nuclease colicins and their immunity proteins. *Quarterly reviews of biophysics* 45(1):57-103.
34. Aoki SK, *et al.* (2005) Contact-dependent inhibition of growth in Escherichia coli. *Science* 309(5738):1245-1248.
35. Aoki SK, Poole SJ, Hayes CS, & Low DA (2011) Toxin on a stick: modular CDI toxin delivery systems play roles in bacterial competition. *Virulence* 2(4):356-359.
36. Nikolakakis K, *et al.* (2012) The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Molecular microbiology* 84(3):516-529.
37. Morse RP, *et al.* (2012) Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proceedings of the National Academy of Sciences of the United States of America* 109(52):21480-21485.
38. Beck CM, *et al.* (2014) CdiA from *Enterobacter cloacae* delivers a toxic ribosomal RNase into target bacteria. *Structure* 22:707-718.
39. Ruhe ZC, Wallace AB, Low DA, & Hayes CS (2013) Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *mBio* 4(4).
40. Webb JS, *et al.* (2013) Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. *PloS one* 8(2):e57609.
41. Willett JLE, Gucinski GC, Fatherree JP, Low DA, & Hayes CS (2015) Contact-dependent growth inhibition toxins exploit multiple independent cell-entry pathways. *Proceedings of the National Academy of Sciences* 112(36):11341-11346.
42. Ruhe ZC, Nguyen JY, Beck CM, Low DA, & Hayes CS (2014) The proton-motive force is required for translocation of CDI toxins across the inner membrane of target bacteria. *Molecular microbiology* 94(2):466-481.

43. Willett JL, Ruhe ZC, Goulding CW, Low DA, & Hayes CS (2015) Contact-Dependent Growth Inhibition (CDI) and CdiB/CdiA Two-Partner Secretion Proteins. *Journal of molecular biology* 427(23):3754-3765.
44. Lancaster LE, Savelsbergh A, Kleanthous C, Wintermeyer W, & Rodnina MV (2008) Colicin E3 cleavage of 16S rRNA impairs decoding and accelerates tRNA translocation on Escherichia coli ribosomes. *Molecular microbiology* 69(2):390-401.
45. Ng CL, *et al.* (2010) Structural basis for 16S ribosomal RNA cleavage by the cytotoxic domain of colicin E3. *Nature structural & molecular biology* 17(10):1241-1246.
46. Diner EJ, Beck CM, Webb JS, Low DA, & Hayes CS (2012) Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes & development* 26(5):515-525.
47. Burkhard P, *et al.* (1998) Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *Journal of molecular biology* 283(1):121-133.
48. Rabeh WM & Cook PF (2004) Structure and mechanism of O-acetylserine sulfhydrylase. *The Journal of biological chemistry* 279(26):26803-26806.
49. Claus MT, Zocher GE, Maier TH, & Schulz GE (2005) Structure of the O-acetylserine sulfhydrylase isoenzyme CysM from Escherichia coli. *Biochemistry* 44(24):8620-8626.
50. Zhao C, *et al.* (2006) On the interaction site of serine acetyltransferase in the cysteine synthase complex from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 341(4):911-916.
51. Francois JA, Kumaran S, & Jez JM (2006) Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the *Arabidopsis* cysteine synthase complex. *Plant Cell* 18(12):3647-3655.
52. Salsi E, *et al.* (2010) Design of O-acetylserine sulfhydrylase inhibitors by mimicking nature. *Journal of medicinal chemistry* 53(1):345-356.
53. Wang T & Leyh TS (2012) Three-stage assembly of the cysteine synthase complex from Escherichia coli. *The Journal of biological chemistry* 287(6):4360-4367.
54. Whitney JC, *et al.* (2014) Genetically distinct pathways guide effector export through the type VI secretion system. *Molecular microbiology* 92(3):529-542.
55. Whitney JC, *et al.* (2015) An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* 163(3):607-619.
56. Koskiniemi S, *et al.* (2014) Selection of orphan Rhs toxin expression in evolved Salmonella enterica serovar Typhimurium. *PLoS genetics* 10(3):e1004255.
57. Anderson MS, Garcia EC, & Cotter PA (2014) Kind discrimination and competitive exclusion mediated by contact-dependent growth inhibition systems shape biofilm community structure. *PLoS pathogens* 10(4):e1004076.
58. Garcia EC, Anderson MS, Hagar JA, & Cotter PA (2013) Burkholderia BcpA mediates biofilm formation independently of interbacterial contact-dependent growth inhibition. *Molecular microbiology* 89(6):1213-1225.
59. Ruhe ZC, *et al.* (2015) CdiA promotes receptor-independent intercellular adhesion. *Molecular microbiology* 98(1):175-192.
60. Jiang F, Waterfield NR, Yang J, Yang GW, & Jin Q (2014) A Pseudomonas aeruginosa Type VI Secretion Phospholipase D Effector Targets Both Prokaryotic and Eukaryotic Cells. *Cell host & microbe* 15(5):600-610.
61. Pukatzki S, Ma AT, Revel AT, Sturtevant D, & Mekalanos JJ (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proceedings of the National Academy of Sciences of the United States of America* 104(39):15508-15513.
62. Decoin V, *et al.* (2014) A Type VI Secretion System Is Involved in Pseudomonas fluorescens Bacterial Competition. *PloS one* 9(2).

63. Whitney JC, *et al.* (2014) Genetically distinct pathways guide effector export through the type VI secretion system. *Molecular microbiology* 92(3):529-542.
64. Basler M (2015) Type VI secretion system: secretion by a contractile nanomachine. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 370(1679).
65. Hachani A, Allsopp LP, Oduko Y, & Filloux A (2014) The VgrG Proteins Are "a la Carte" Delivery Systems for Bacterial Type VI Effectors. *Journal of Biological Chemistry* 289(25):17872-17884.
66. Lin RJ, Capage M, & Hill CW (1984) A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome. *Journal of molecular biology* 177(1):1-18.
67. Vlazny DA & Hill CW (1995) A stationary-phase-dependent viability block governed by two different polypeptides from the RhsA genetic element of Escherichia coli K-12. *Journal of bacteriology* 177(8):2209-2213.
68. Jamet A & Nassif X (2015) New players in the toxin field: polymorphic toxin systems in bacteria. *mBio* 6(3):e00285-00215.
69. Koskiniemi S, *et al.* (2013) Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America* 110(17):7032-7037.

CHAPTER 2

Functional Diversity of Cytotoxic tRNase/immunity Protein Complexes from *Burkholderia pseudomallei*

This project was performed in collaboration with members of the Hayes lab at UC Santa Barbara.

Abstract

Contact-dependent growth inhibition (CDI) is a wide-spread mechanism of inter-bacterial competition. CDI^+ bacteria deploy large CdiA effector proteins, which carry variable C-terminal toxin domains (CdiA-CT). CDI^+ cells also produce CdiI immunity proteins that specifically neutralize cognate CdiA-CT toxins to prevent auto-inhibition. Here, we present the crystal structure of the CdiA-CT/$CdiI^{E479}$ toxin/immunity protein complex from *Burkholderia pseudomallei* isolate E479. The CdiA-CT^{E479} tRNase domain contains a core α/β-fold that is characteristic of PD-(D/E)XK superfamily nucleases. Unexpectedly, the closest structural homolog of CdiA-CT^{E479} is another CDI toxin domain from *B. pseudomallei* 1026b. Though unrelated in sequence, the two *B. pseudomallei* nuclease domains share similar folds and active-site architectures. By contrast, the $CdiI^{E479}$ and $CdiI^{1026b}$ immunity proteins share no significant sequence or structural homology. CdiA-CT^{E479} and CdiA-CT^{1026b} are both tRNases, however the cleavage positions are different for each nuclease. We used a molecular docking approach to model each toxin bound to tRNA substrate. The resulting models fit into electron density envelopes generated by small-angle X-ray scattering analysis of stable complexes of tRNA bound to catalytically inactive toxin domains. CdiA-CT^{E479} is the third CDI toxin with structural homology to the PD-(D/E)XK superfamily. PD-(D/E)XK nucleases are characterized by highly variable sequences and active-site plasticity. CDI systems exploit this structural flexibility to generate toxin

diversity. These findings raise the possibility that many other uncharacterized CDI toxins may also belong to the PD-(D/E)XK superfamily.

Introduction

Bacterial contact-dependent growth inhibition (CDI) is an important mechanism of inter-cellular competition in which Gram-negative bacteria intoxicate neighboring cells upon direct contact. Genes encoding CDI systems are distributed throughout α-, β- and γ-proteobacteria and are commonly found in human pathogens such as enterohemorrhagic *Escherichia coli*, *Neisseria meningitidis, Pseudomonas aeruginosa* and *Burkholderia pseudomallei* (1-3). CDI is a function of the CdiB/CdiA family of two-partner secretion systems. CdiB is an Omp85 β-barrel protein required for the export and display of CdiA effectors on the cell surface. Based on its homology to filamentous hemagglutinin from *Bordetella* species, CdiA proteins are predicted to form filaments that project several hundred Å from the cell surface (3-5). Upon binding to specific receptors on susceptible bacteria, CdiA delivers its C-terminal toxin domain (CdiA-CT) into the target cell to inhibit growth (6-8). CDI^+ bacteria are protected from toxicity by CdiI immunity proteins, which bind to the CdiA-CT toxin domain and neutralize its activity. Though the general architecture of CdiA proteins is conserved across bacteria, the effectors vary considerably in size (180 to 640 kDa) and CdiA-CT toxin sequences are remarkably diverse (1,2,9-13). Thus, CdiA effectors collectively deploy a variety of toxin domains with distinct biochemical activities (1,3,14,15). Moreover, CdiI immunity proteins are also highly variable in sequence and only protect cells against cognate CdiA-CT toxins. The polymorphic nature of CDI toxin/immunity protein pairs and the specificity of their binding interactions suggest that the systems mediate inter-strain competition and self/nonself recognition.

We previously characterized CdiA-CT/CdiI pairs from environmental isolates of *Burkholderia pseudomallei*, using these systems as a model to explore toxin/immunity protein structure and evolution (16,17). *B. pseudomallei* is a category B pathogen and the causative agent of melioidosis, a serious human disease endemic to southeast Asia and northern Australia (18,19). *B. pseudomallei* isolates are genetically heterogeneous and different strains are thought to compete with one another for growth niches and other resources (20,21). CDI may contribute to inter-strain competition because every *B. pseudomallei* isolate carries at least one *cdi* gene cluster (16). *B. pseudomallei* CdiA proteins share 59 – 99.5% pair-wise sequence identity over the N-terminal 2,770 – 2,829 residues, but diverge abruptly after a conserved (E/Q)LYN peptide motif that demarcates the CdiA-CT region. At least 10 distinct CdiA-CT/CdiI sequence types are found in the species, but only three toxin activities have been characterized experimentally (16). The type I CdiA-CT is homologous to the nuclease domain from colicin E5, and exhibits anticodon nuclease activity against $tRNA^{His}$, $tRNA^{Tyr}$, $tRNA^{Asp}$ and $tRNA^{Asn}$ (16,22). The type II toxin cleaves all tRNAs between conserved residues T54 and Ψ55 in the T-loop; and the type V toxin preferentially cleaves near the 3´-end of $tRNA^{Ala}$ (16,17). Type VII and VIII CdiA-CT sequences are closely related in sequence and are predicted RNA deaminases (14). There are no predicted activities for the other *B. pseudomallei* CDI toxins, though the type X toxin contains a conserved RES (Arg-Glu-Ser) domain of unknown function (Pfam: PF08808).

Here, we present the crystal structure of the type II toxin/immunity protein complex from *B. pseudomallei* isolate E479 (16,20). The C-terminal tRNase domain of $CdiA\text{-}CT^{E479}$ is built upon an α/β-fold characteristic of PD-(D/E)XK nucleases. The PD-(D/E)XK superfamily includes most restriction endonucleases and other enzymes involved in DNA recombination and repair (23). A DALI server (24) search reveals that the type V toxin domain from *B. pseudomallei* 1026b is the closest structural homolog of $CdiA\text{-}CT^{E479}$. Although $CdiA\text{-}CT^{E479}$ and $CdiA\text{-}CT^{1026b}$ nuclease

domains do not share significant sequence identity, they have very similar active-site architectures with catalytic residues contributed by α1, β2, β3 and α2 elements of the PD-(D/E)XK core. By contrast, CdiI^{E479} and CdiI^{1026b} are unrelated in sequence and structure, but each immunity protein inactivates cognate toxin by binding within the nuclease active-site cleft to prevent access to tRNA. To gain insight into substrate specificity, we used molecular docking approaches to model each toxin domain bound to tRNA. Though most PD-(D/E)XK nucleases are DNA-specific phosphodiesterases, our findings indicate that these enzymes are commonly used as toxic RNases in bacterial competition. Because PD-(D/E)XK nucleases can be difficult to identify through sequence analyses (23,25), it is possible that many other uncharacterized CDI toxins also belong to the superfamily.

Results

Structure of the CdiA-CT/CdiI^{E479} toxin/immunity protein complex

We previously reported that over-expression of the CdiA-CT/CdiI^{E479}-His$_6$ complex in *E. coli* cells leads to growth arrest and concomitant tRNA degradation (16). These observations suggest that the expression construct produces insufficient immunity protein to neutralize tRNase activity. Because CdiA-CT^{E479} activity precludes protein over-production, we inactivated the toxin with the Asp285Ala mutation to allow high-level expression of the toxin/immunity protein complex (16). We purified and crystallized the SeMet-labeled complex and used SAD phasing to produce an initial partial model. The model was subsequently improved using molecular replacement with a native dataset, resulting in a final resolution of 2.0 Å (**Table 2.1**).

Table 2.1. **X-ray diffraction data and atomic refinement for the CdiA-CT/CdiI^{E479} complex.**

	Se-SAD Dataset	Native Dataset
Space Group	I4	$P22_12_1$
Unit cell dimensions (Å)	117.2 x 117.2 x 111.6	54.5 x 73.3 x 110.0
pH of crystallization condition	6.5	7.0
Protein concentration (mg/mL)	20	20
Dataset		
Wavelength, Å	0.97591	0.97591
Resolution range	50-3.3	50-2.0
Unique reflections (total)	22290 (329892)	30121 (322487)
Completeness, %*	100 (100)	99.24 (99.28)
Redundancy*	14.8 (14.9)	10.7 (10.9)
R_{merge}*,†	0.159 (0.48)	0.071 (0.494)
R_{meas}*, ‡	0.157 (0.467)	0.075 (0.519)
$R_{p.i.m.}$ *§	0.041 (0.121)	0.023 (0.156)
$CC_{1/2}$*	0.999 (0.993)	0.998 (0.966)
I/σ*	18.03 (6.95)	29.99 (5.79)
FOM	0.408	N/A
No. of Se sites	18	N/A
NCS copies	3	2
Model refinement		
Resolution range, Å		43.7-2.0
No. of reflections (working/free)		30094
No. of protein atoms		3382
No. of water molecules		155
Residues in model		CdiA-CT 201–316; CdiI 2–105
R_{work}/R_{free}, %¶		19.3/23.7
Rms deviations		
Bond lengths, Å		0.008
Bond angles		1.09
Ramachandran plot		
Most favorable region, %		98.8
Additional allowed region, %		1.2
Disallowed region		0
PDB ID code		5J4A

*Statistics for the highest-resolution shell are given in parentheses.

† $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

‡ $R_{meas} = \Sigma_{hkl} \{N(hkl) / [N(hkl) - 1]\}^{1/2} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

§ $R_{p.i.m}$ (precision-indicating R_{merge}) $= \Sigma_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

¶ $R_{work} = \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$. R_{free} was computed identically except where all reflections belong to a test set of 5% randomly selected data.

The final model includes CdiA-CT^{E479} residues Arg201 – Lys316 (numbered from Glu1 of the ELYN peptide motif), CdiI^{E479} residues Ala2 – Gly105, and 155 water molecules. The final R_{work}/R_{free} (%) was 19.3/23.7 with 98.8% of dihedral angles in favorable regions and the remaining 1.2% within allowed regions as estimated by Ramachandran plot.

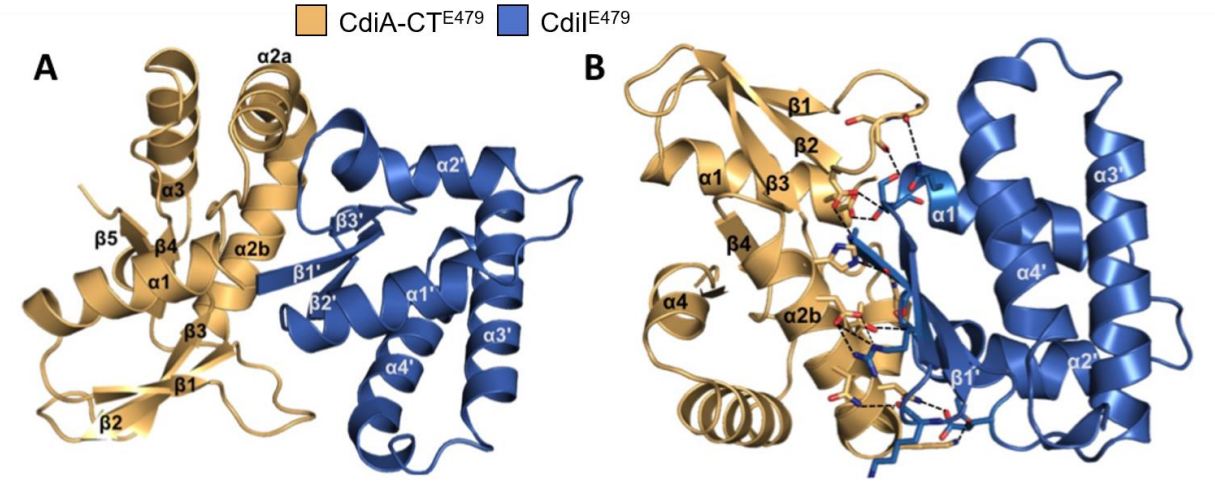


Figure 2.1: Structure of the CdiA-CT/CdiI^{E479} toxin/immunity protein complex. (**A**) The CdiA-CT/CdiI^{E479} complex is shown in cartoon representation with secondary structure elements of CdiI are indicated with prime (′) symbols. (**B**) CdiA-CT/CdiI^{E479} complex formation is mediated by electrostatic interactions. Interacting residues are in stick representation with nitrogen and oxygen atoms colored in blue and red (respectively), and bonds are shown as black dashed lines. The view is rotated 180° about the x-axis relative to panel **A**.

Like other CdiA-CT constructs (17,26,27), the N-terminal region of CdiA-CT^{E479} (residues Glu1 – Phe200) is not resolved in the final model. This unresolved region corresponds to the "translocation" domain, which is postulated to mediate CdiA-CT transport across the cytoplasmic membrane of target bacteria (28). The resolved C-terminal domain corresponds to the tRNase domain responsible for growth inhibition activity (16). The CdiA-CT^{E479} nuclease domain consists of a five-stranded mixed β-sheet decorated by four α-helices (**Figure 2.1A**). The sheet forms a half β-barrel-like structure with helix α1 running through its central cavity (**Figure 2.1A**). The C-terminal half of helix α2 (α2b) is bent 90° with respect to the N-terminal portion (α2a). The CdiI^{E479}

immunity protein consists of a slightly curved three-stranded antiparallel β-sheet decorated with four α-helices (**Figure 1A**). The CdiA-CT/CdiI^{E479} interface is largely electrostatic with 19 direct salt-bridges and hydrogen-bonds mediating the interface between the nuclease domain and immunity protein (**Figure 2.1B** & **Table 2.2**).

Table 2.2. Direct intermolecular hydrogen bonds and salt bridges in the CdiA-CT/CdiI^{E479} complex.

CdiA-CT^{E479}	CdiI^{E479}	Distance (Å)
Hydrogen bonds		
Glu204 [OE2]	Thr17 [OG1]	2.39
Gly225 [O]	Gln19 [NE2]	3.79
Ser226 [O]	Ser16 [NE2]	3.37
Asp229 [OD1]	Thr17 [N]	2.79
Gln253 [NE2]	Gly50 [O]	2.98
Asn267 [ND2]	Glu57 [OE2]	2.66
Asn270 [ND2]	Lys9 [O]	3.70
Thr271 [OG1]	Ser4 [OG]	2.39
His275 [NE2]	Gly3 [O]	2.62
Salt bridges		
Asp229 [OD2]	Ala2 [N]	3.83
Lys263 [NZ]	Glu48 [OE1]	3.64
Lys263 [NZ]	Glu48 [OE2]	2.47
Asp274 [OD1]	Arg12 [NE]	3.78
Asp274 [OD2]	Arg12 [NE]	2.82
Asp274 [OD1]	Arg12 [NH2]	2.97
Asp274 [OD2[	Arg12 [NH2]	3.47

Helix α2b of the nuclease domain sits in the curve of the CdiI^{E479} β-sheet. Additionally, helices α1′, α4′ and the β-sheet of CdiI^{E479} interact with the end of the nuclease β-sheet and the extended loop connecting α2′ and β3′ (**Figure 2.1B**). This protein-protein interface buries 1015 Å^2 of surface area, corresponding to 14% of the nuclease domain and 18% of the immunity protein total surface area. In accord with this extensive interaction surface, biolayer interferometry showed that CdiA-CT^{E479} and CdiI^{E479} form a relatively high-affinity complex with an apparent equilibrium dissociation constant of 72 ± 23 nM (**Figure 2.2**).

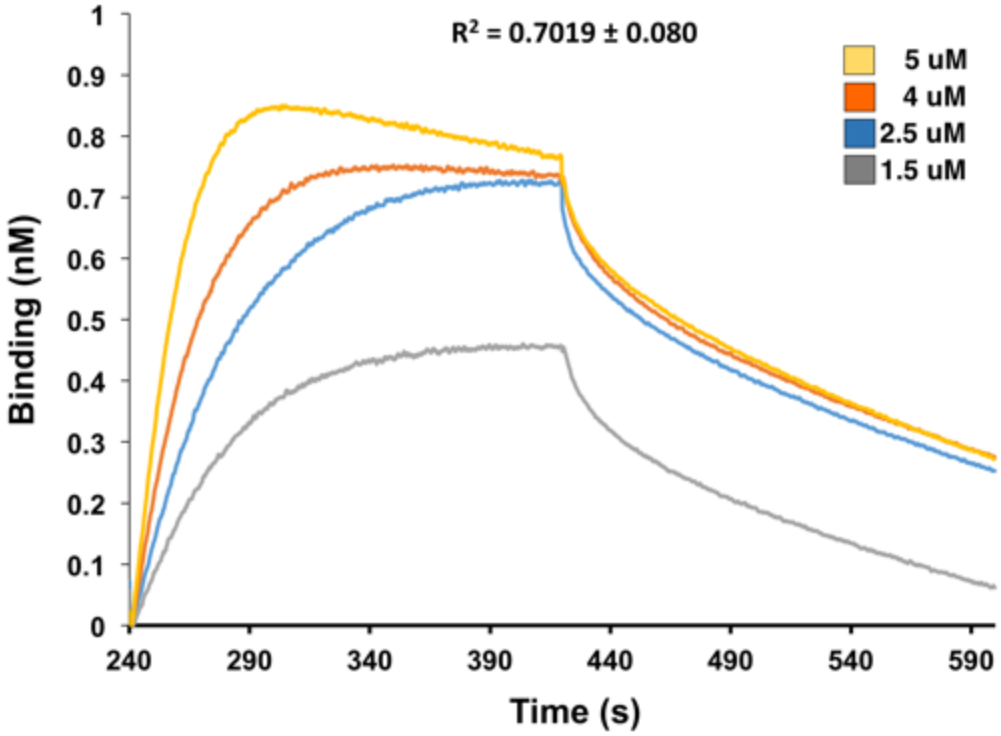


Figure 2.2: Biolayer interferometry of the CdiA-CT/CdiI^{E479} binding interaction. Immobilized CdiI^{E479}-His^6 was exposed to varying concentrations (1.5 - 5 μM) of CdiA-CT^{E479} and the binding interaction and dissociation monitored a wavelength shift (nm). Representative association and dissociation curves are presented with the overall correlation coefficient (R^2) shown for the fit.

The CdiA-CT^{E479} toxin domain is a member of the PD-(D/E)XK superfamily

We used the DALI server (24) to identify proteins with structural similarity to the CdiA-CT^{E479} nuclease domain. This search revealed that two other CdiA-CT nuclease domains from *B. pseudomallei* 1026b and *Yersinia pseudotuberculosis* YPIII exhibit the greatest similarity with CdiA-CT^{E479} (**Table 2.3**). The CdiA-CT^{1026b} and CdiA-CT^{E479} nuclease domains superimpose with root-mean-square deviation (rmsd) of 3.7 Å over 89 of 132 α-carbons, corresponding to a DALI *Z*-score of 7.0. The CdiA-CT^{YPIII} nuclease domain exhibits comparable structural similarity, though its β4-β5 hairpin element is absent from the CdiA-CT^{E479} nuclease domain (27). Other proteins identified during this search include two closely related XisH endonucleases from cyanobacteria, another CdiA-CT nuclease domain from *E. coli* EC869, and the AspBHI restriction endonuclease from *Azoarcus* sp. BH72 (**Table 2.3**).

Table 2.3. DALI server search results

Search input	**Structural homolog**	**Organism**	**PDB ID**	***Z*-score**	**rmsd (Å)[a]**
CdiA-CT^{E479}	CdiA-CT$_{II}^{1026b}$	*B. pseudomallei* 1026b	4G6V	7.0	3.7 (89/132)[b]
	CdiA-CT^{YPIII}	*Y. pseudotuberculosis* YPIII	4ZQU	7.0	3.0 (95/124)
	XisH	*N. punctiforme* PCC 73102	2INB	6.4	2.7 (87/128)
	AVA_3312	*A. variabilis*	2OKF	6.4	2.9 (90/129)
	CdiA-CT$_{o11}^{EC869}$	*E. coli* EC869	4G6U	5.8	3.3 (96/213)
	AspBHI	*Azoarcus* sp. BH72	4OC8	5.2	3.2 (91/387)
CdiI^{E479}	NS3	rice hoja blanca virus (RHBV)	3AJF	5.8	2.7 (75/90)

[a]rmsd is root-mean-square deviation.
[b]aligned over # α-carbons out of # residues.

All of these domains share the core structure of the PD-(D/E)XK nuclease superfamily, which includes most type II restriction endonucleases and various enzymes involved in DNA recombination and repair. The PD-(D/E)XK core domain is a mixed β-sheet flanked by two α-helices with αββββαβ topology. CdiA-CT^{E479} and CdiA-CT^{1026b} share this core fold, but the CdiA-CT^{E479} nuclease domain contains an insertion that forms helix α2a (**Figure 2.3A & 2.3B**). We also

note that helix α1 from CdiA-CT^{E479} is significantly shorter than the corresponding helix in the CdiA-CT^{1026b} structure (**Figure 2.3A & 2.3B**).

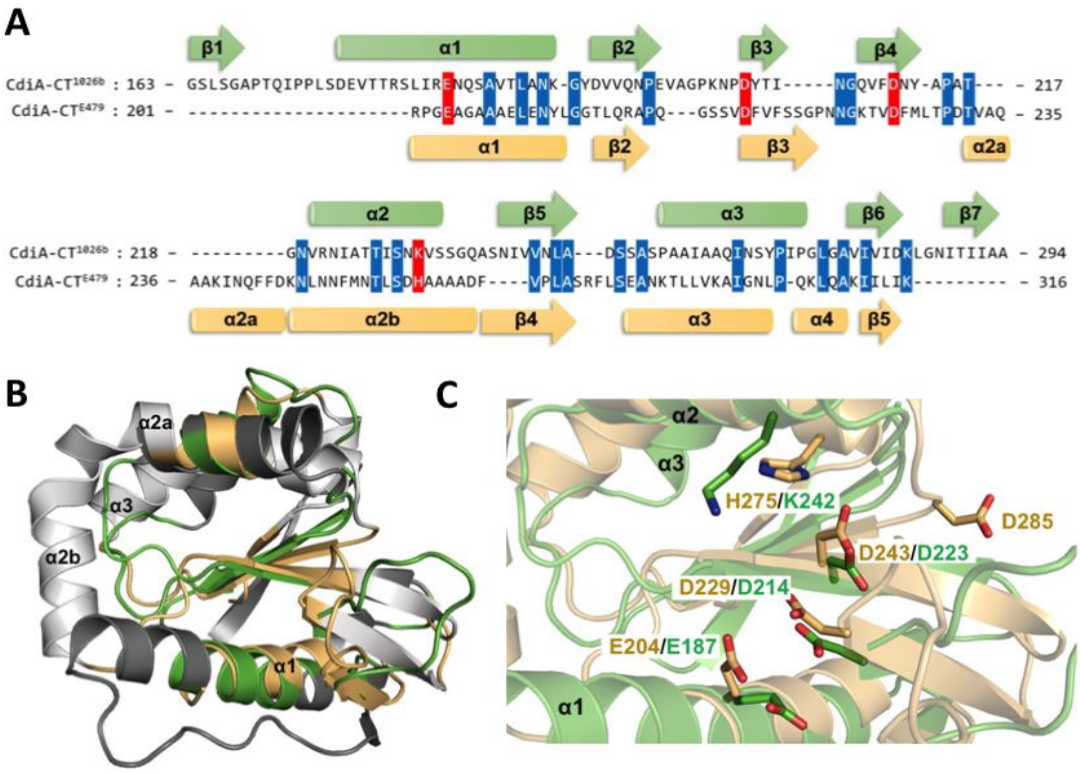


Figure 2.3: Sequence and structure alignment of CdiA-CT^{E479} and CdiA-CT^{1026b} nuclease domains. (**A**) Sequence alignment of CdiA-CT^{E479} and CdiA-CT^{1026b} toxins with proposed active-site residues highlighted in red and conserved residues in blue. Secondary structure elements are colored gold and green for CdiA-CT^{E479} and CdiA-CT^{1026b}, respectively. (**B**) Superimposition of CdiA-CT^{E479} and CdiA-CT^{1026b} nuclease domains. Secondary structure elements that superimpose are color-coded in gold (CdiA-CT^{E479}) and green (CdiA-CT^{1026b}), and those that do not align rendered in white (CdiA-CT^{E479}) and grey (CdiA-CT^{1026b}). (**C**) Active site of CdiA-CT^{E479} and CdiA-CT^{1026b} nuclease domain. Predicted active-site residues are shown in stick representation (nitrogen and oxygen atoms are colored blue and red, respectively).

Identification of the CdiA-CT^{E479} active site

We previously suggested that Asp280 and Asp285 of CdiA-CT^{E479} may function in catalysis, because mutation of these residues abrogated toxicity (16). However, structural alignment with the nuclease domain from CdiA-CT^{1026b} indicates that residues Glu204, Asp229, Asp243 and His275 of CdiA-CT^{E478} are more likely to catalyze tRNA cleavage (**Figure 2.3A**). These residues overlay well with active-site residues Glu187, Asp214, Asp223 and Lys242 of CdiA-CT^{1026b} (**Figure 2.3C**). To test this prediction, we mutated CdiA-CT^{E479} residues His275, Asp243, Asp229 and Glu204 to Ala individually and examined the growth inhibition activity of each toxin variant. Induction of wild-type CdiA-CT^{E479} expression in *E. coli* cells resulted in immediate growth inhibition (**Figure 2.4A**). By contrast, induction of domains that carry the His275Ala, Asp243Ala, Asp229Ala or Glu204Ala mutations had no effect on cell growth (**Figure 2.4A**), consistent with the loss of toxic nuclease activity. We also purified each toxin domain and tested its tRNase activity *in vitro*. Wild-type CdiA-CT^{E479} toxin cleaved a large proportion of tRNA molecules as assessed by denaturing PAGE analysis, and this activity was blocked when purified CdiI^{E479} immunity protein was included in the reaction (**Figure 2.4B**). This tRNase activity was confirmed by Northern blot analysis, which showed cleavage of tRNA^{Gly} molecules in the reactions (**Figure 2.4B**). By contrast, none of the other CdiA-CT^{E479} mutant variants exhibited detectable RNase activity (**Figure 2.4B**). To determine whether the mutations adversely affect toxin structure, we tested whether the refolded CdiA-CT^{E479} proteins still interact with cognate immunity protein using affinity co-purification. Each mutant toxin co-eluted with CdiI^{E479}-His$_6$ during Ni^{2+}-affinity chromatography (**Figure 2.4C**), indicating that the nuclease domains retain their native fold. Taken together with the crystal structure, these findings suggest that CdiA-CT^{E479} residues His275, Asp243, Asp229 and Glu204 participate in catalysis. Because CdiI^{E479} binds directly over this cluster of residues, the immunity protein presumably neutralizes toxin activity by blocking access to tRNA substrates. We previously

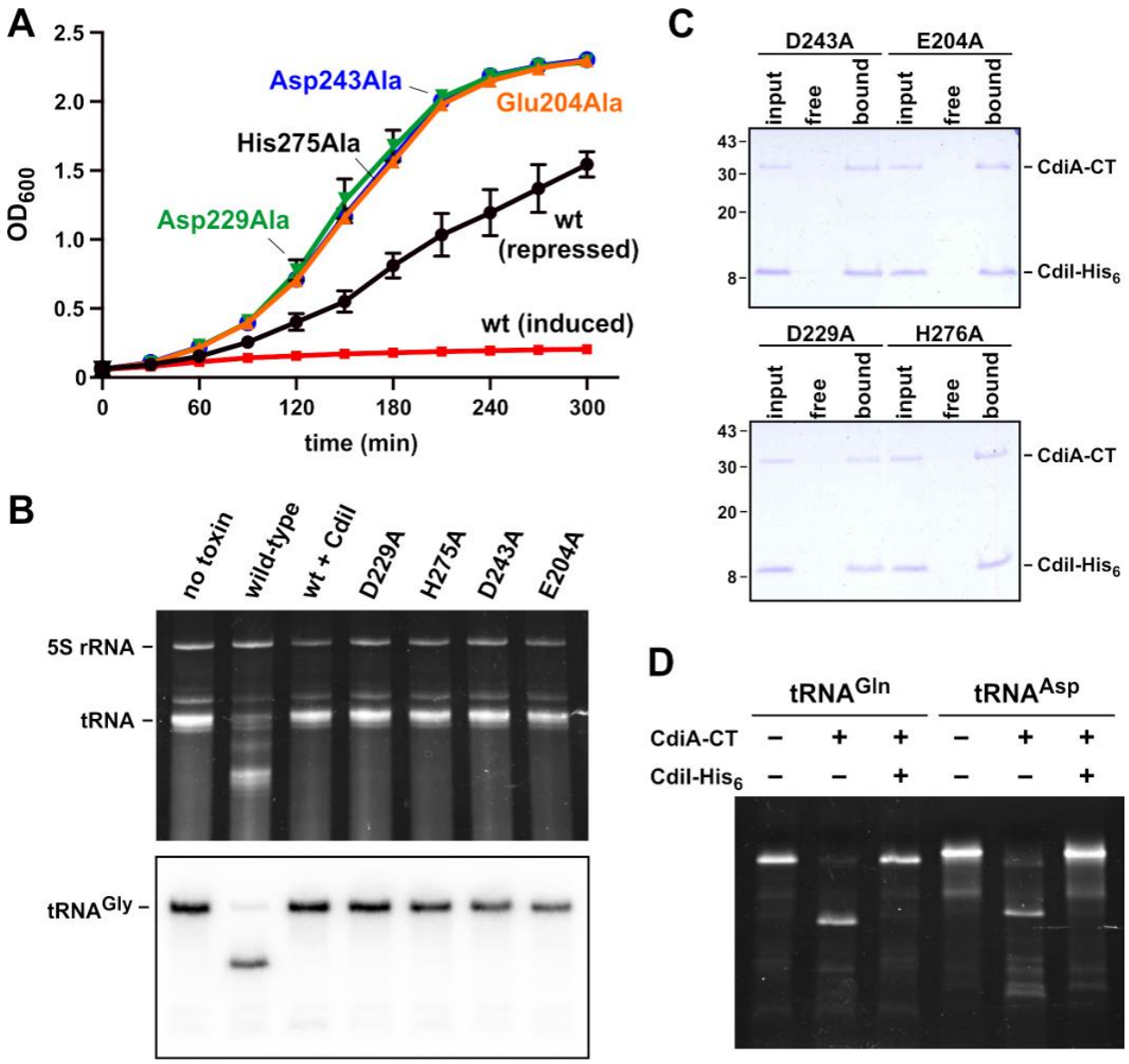


Figure 2.4 (Courtesy of Hayes lab): CdiA-CT^{E479} growth inhibition and tRNase activities. (**A**) Growth inhibition activity of CdiA-CT^{E479} variants. The indicated toxins were expressed in *E. coli* cells from a rhamnose-inducible promoter as described in Methods. Expression was induced at 0 min and cell growth monitored by measuring the optical density at 600 nm (OD_{600}). The curve labeled **repressed** corresponds to un-induced cells carrying the wild-type CdiA-CT^{E479} construct. The average ± SEM from three independent biological replicates is presented. (**B**) *In vitro* nuclease assays. The indicated CdiA-CT^{E479} variants were purified and incubated with total *E. coli* RNA. Reactions were run on denaturing 6% polyacrylamide gels and stained with ethidium bromide. (**C**) Mutant CdiA-CT^{E479} domains bind to CdiI^{E479} immunity protein. Isolated toxin domains were mixed with purified CdiI^{E479}-His_6, then subjected to Ni^{2+}-affinity chromatography. Lanes labeled **input** show the protein mixtures loaded onto the column, **free** lanes show proteins that failed to bind the column, and **bound** indicates proteins eluted from the column with imidazole. Prior work has shown that CdiA-CT^{E479} does not bind to Ni^{2+}-NTA agarose resin (16). (**D**) CdiA-CT^{E479} cleaves unmodified tRNAs produced by *in vitro* transcription. *E. coli* $tRNA^{Gln}$ and $tRNA^{Asp}$ transcripts were incubated with purified CdiA-CT^{E479} and CdiI^{E479}, and reactions analyzed on denaturing 6% polyacrylamide gels stained with ethidium bromide. Experiments in panels B, C and D were repeated twice with essentially identical results. Representative data are shown for each experiment.

showed that CdiA-CT^{E479} cleaves between residues T54 and Ψ55 in the T-loop of tRNA molecules (16). To test whether post transcriptional modifications at positions 54 and 55 are required for CdiA-CT^{E479} activity, we examined toxin activity on unmodified *E. coli* tRNA^{Asp} and tRNA^{Gln} substrates prepared by *in vitro* transcription. Each substrate was cleaved efficiently by purified nuclease and the activity completely neutralized by CdiI^{E479} immunity protein (**Figure 2.4D**). Thus, the universal T-loop modifications are not required for CdiA-CT^{E479} mediated tRNase activity.

Structural comparison of CdiI^{E479} and CdiI^{1026b} immunity proteins

Although the CdiA-CT^{E479} and CdiA-CT^{1026b} nuclease domains share a common fold and active site, the corresponding immunity proteins appear to be unrelated. Using iterative PSI-BLAST, we were unable to establish a link between CdiI^{E479} and CdiI^{1026b} sequences. Moreover, structural superimposition of the two immunity proteins reveals a poor fit between the central β-sheets and misalignments of most α-helical elements (**Figure 2.5A** & **2.5B**). CdiI^{E479} and CdiI^{1026b} align with an rmsd of 3.42 Å over only 42 of 103 α-carbons (*Z*-score of 1.8) indicating low structural similarity. We next used the DALI server to search for proteins with structural similarity

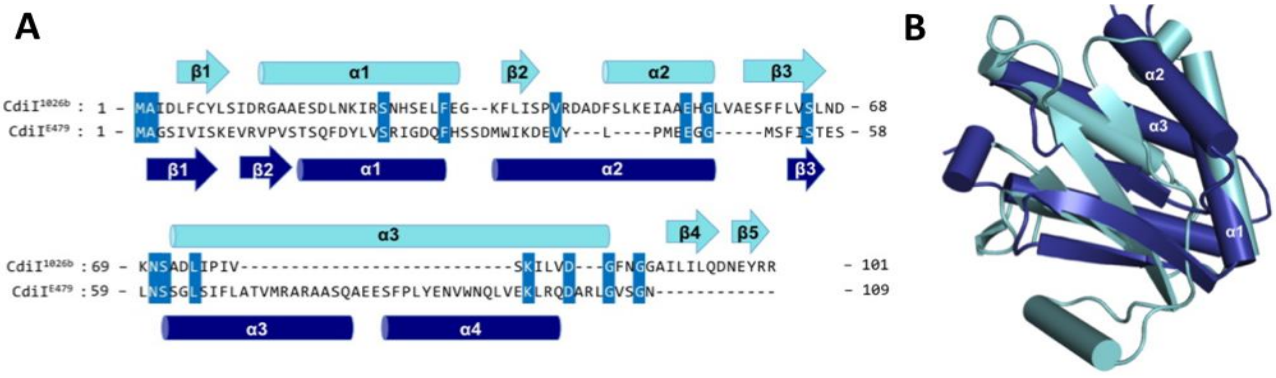


Figure 2.5: Sequence and structure comparison of CdiI^{E479} and CdiI^{1026b} immunity proteins. (**A**) Structure-based sequence alignment of CdiI^{E479} (blue) and CdiI^{1026b} (cyan) with secondary structure elements indicated above and below the sequence alignment. Conserved residues are highlighted in blue. (**B**) Superimposition of CdiI^{E479} and CdiI^{1026b} structures. Secondary elements that partially or fully superimpose are labeled.

to CdiI^{E479}. The only hit with *Z*-score > 5 was the N-terminal domain of protein NS3 from rice hoja blanca tenuivirus (**Table 2.3**). NS3 suppresses RNA interference pathways in host cells by binding to both siRNA and miRNAs (29,30). Although these proteins superimpose with rmsd of 2.7 Å over 75 of 90 α-carbons, the NS3 domain is entirely α-helical and lacks the central β-sheet found in CdiI^{E479}.

CdiA-CT/CdiI complexes have unique electrostatic interfaces

Both toxin/immunity complexes interact primarily via electrostatic interactions and shape complementation; and intriguingly, both immunity proteins use the N-terminal α-amino group to form a salt-bridge with a catalytic Asp residue in the toxin active site (**Table 2.2**) (17). However, the size, charge distribution, shape and position of these patches differ between the two immunity proteins (**Figure 2.6**). The electrostatic surface map shows that CdiI^{E479} has a positively charged protrusion (formed from α1′, α2′ and the β-sheet) that is complementary to the negatively charged active-site groove within the CdiA-CT^{E479} nuclease domain (**Figure 2.6A**). In addition, a small negatively charged patch on CdiI^{E479} interacts with a positive patch on the cognate toxin. Much like CdiI^{E479}, the CdiI^{1026b} immunity protein inserts a positively charged protrusion (formed from β3′, α3′ and the connecting loop region) into the negatively charged active-site cleft of its cognate toxin. CdiI^{1026b} also contains a negatively charged patch, which complements a positive patch adjacent to the Lys242 active-site residue within CdiA-CT^{1026b}. Collectively, these electrostatic and shape complementation interactions ensure that each nuclease domain is only neutralized by its cognate immunity protein.

Nuclease active-site architecture and tRNA specificity

CdiA-CT^{E479} and CdiA-CT^{1026b} cleave tRNA at distinct positions, and presumably this specificity is dictated by the shape of the active-site pocket. The CdiA-CT^{E479} nuclease domain

has a slightly larger, open pocket with a diameter of 12.6 Å measured from Gln253 to His275 (**Figure 2.6A**). By contrast, the active-site cleft of CdiA-CT^{1026b} is narrower with a diameter of 10.4 Å from Glu187 to Lys242 (**Figure 2.6B**). To explore how

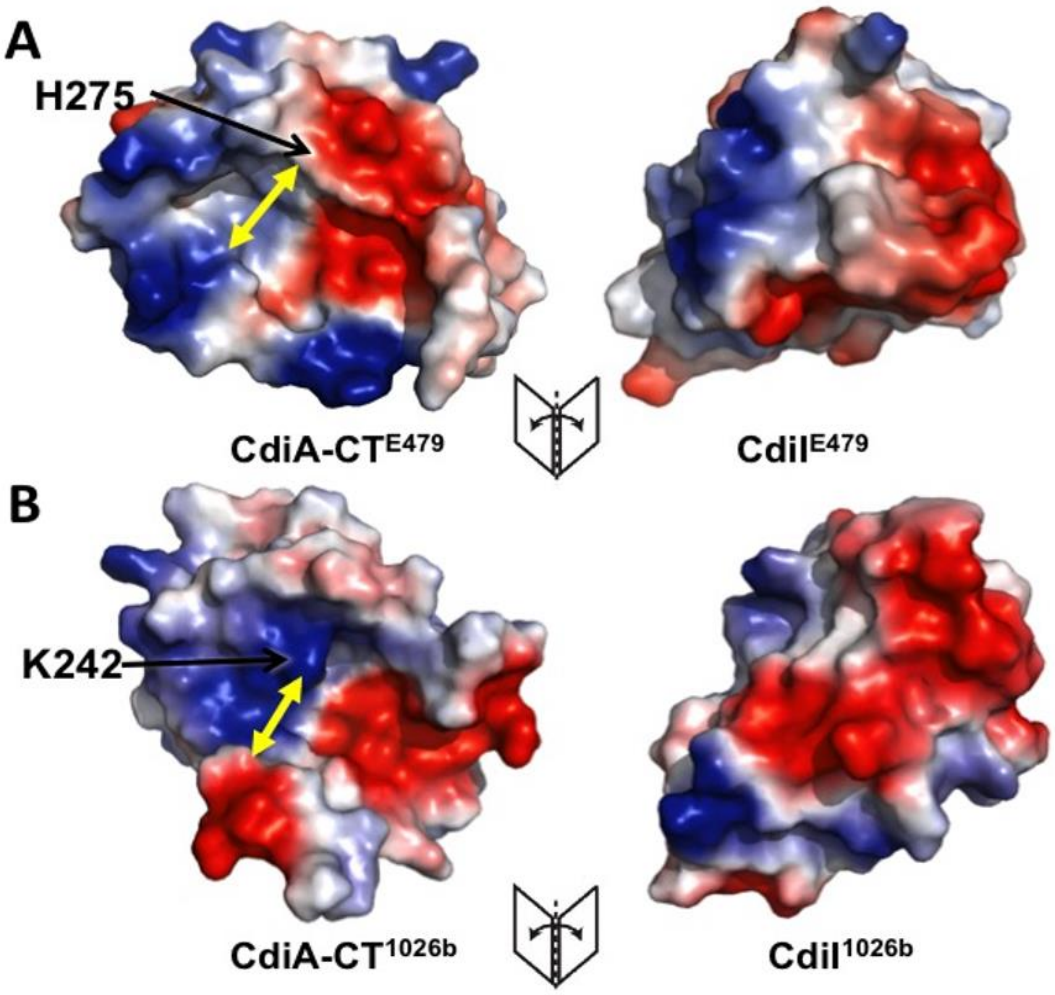


Figure 2.6: CdiA-CT/CdiI^{E479} and CdiA-CT/CdiI^{1026b} complexes interact through distinct electrostatic surfaces (**A**) Electrostatic surface map of the CdiA-CT/CdiI^{E479} complex interface. Negative, positive and neutral surfaces regions are shown in red, blue and white, respectively. (**B**) Electrostatic surface map of the CdiA-CT/CdiI^{1026b} complex interface. Yellow arrows indicate the minimal width of each active-site pocket.

the nuclease domains interact with substrate, we used Hex 8.0 to dock each toxin onto the structure of *E. coli* tRNA^{Cys}. We used experimentally determined cleavage sites for each nuclease to guide the docking searches. The models obtained from the simulations show that each nuclease active-site is positioned near the scissile bond in the tRNA substrate (**Figure 2.7A** & **2.7C**). The

calculated interaction energy for the tRNA/CdiA-CT^{E479} complex was –802 kcal/mol, and that of the tRNA/CdiA-CT^{1026b}

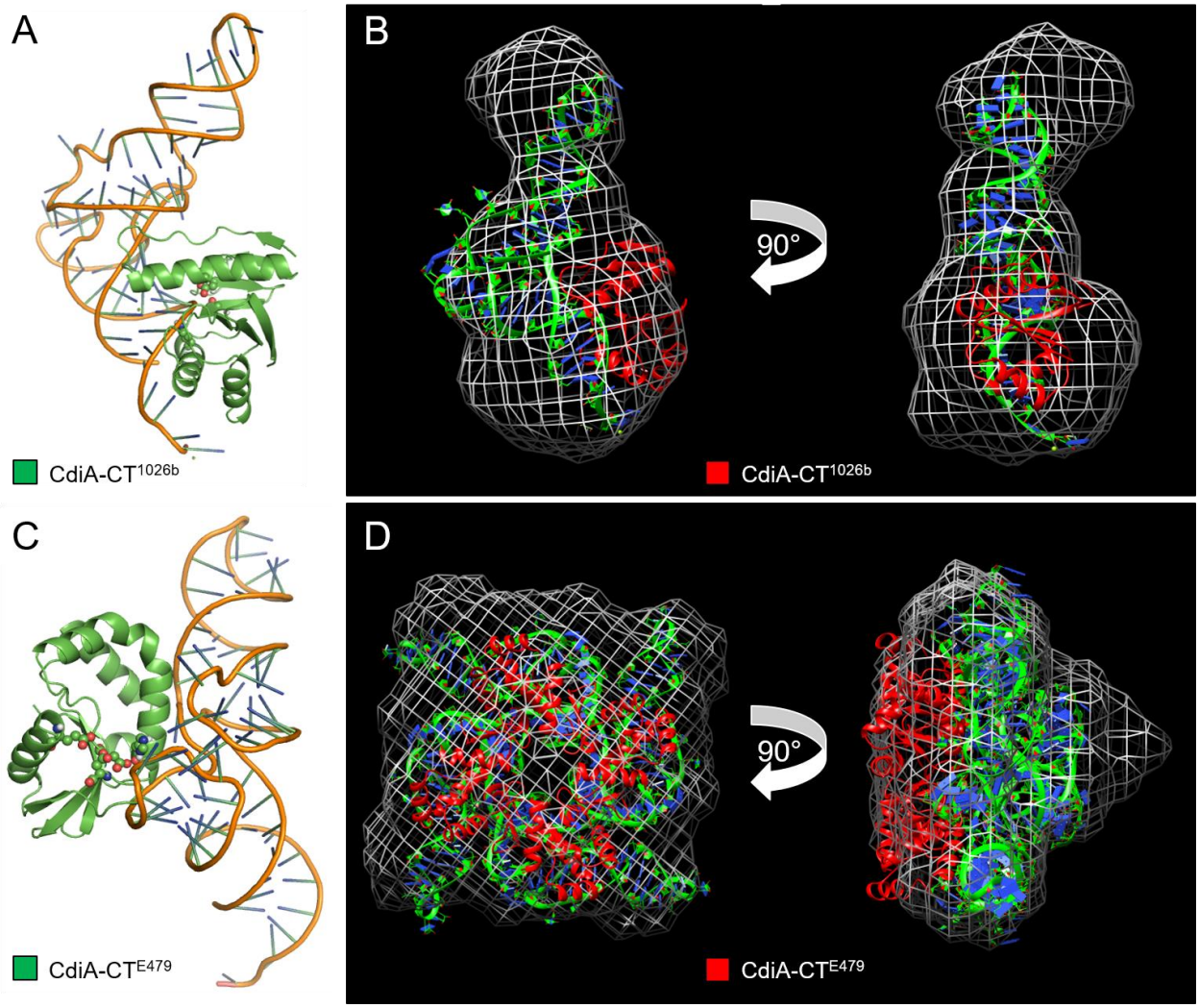


Figure 2.7: Computational modeling and SAXS analysis of tRNA/CdiA-CT complexes. (**A**) CdiA-CT^{1026b} (in green cartoon) binding with its active site residues adjacent to the backbone of the tRNA^{Cys} (PDB ID: 1B23) amino acceptor stem loop with active site residues shown as spheres (oxygen and nitrogen atoms colored red and blue, respectively). (**B**) SAXS electron density envelope (white mesh) fitted with the docking solution showing the monomeric CdiA-CT^{1026b} toxin (red cartoon) bound to tRNA^{Cys} (colored in green and blue). (**C**) CdiA-CT^{E479} (in green cartoon) binding with its active site residues adjacent to the backbone of the tRNA^{Cys} T-loop with active site residues shown as spheres oxygen and nitrogen atoms colored red and blue, respectively). (**D**) SAXS electron density envelope (white mesh) fitted with the docking solution showing the tetrameric CdiA-CT^{E479} toxin bound to four molecules of tRNA^{Cys} (colored in green and blue).

complex was –1,267 kcal/mol. These energies are comparable to those calculated for Hex-assisted docking of each toxin domain with its cognate immunity protein: –953 kcal/mol for CdiA-CT/CdiI^{E479} complex and –1,199 kcal/mol for the CdiA-CT/CdiI^{1026b} complex. A negative-control simulation using CdiI^{E479} and CysK – which do not interact with one another – yielded a considerably higher energy of –368 kcal/mol. Simulations using non-cognate toxin/immunity protein pairs produced models with lower than anticipated energies (–742 kcal/mol for CdiA-CT^{E479}/CdiI^{1026b} and –626 kcal/mol for CdiA-CT^{1026b}/CdiI^{E479}), but in each instance the solutions did not position the immunity proteins over the nuclease active sites.

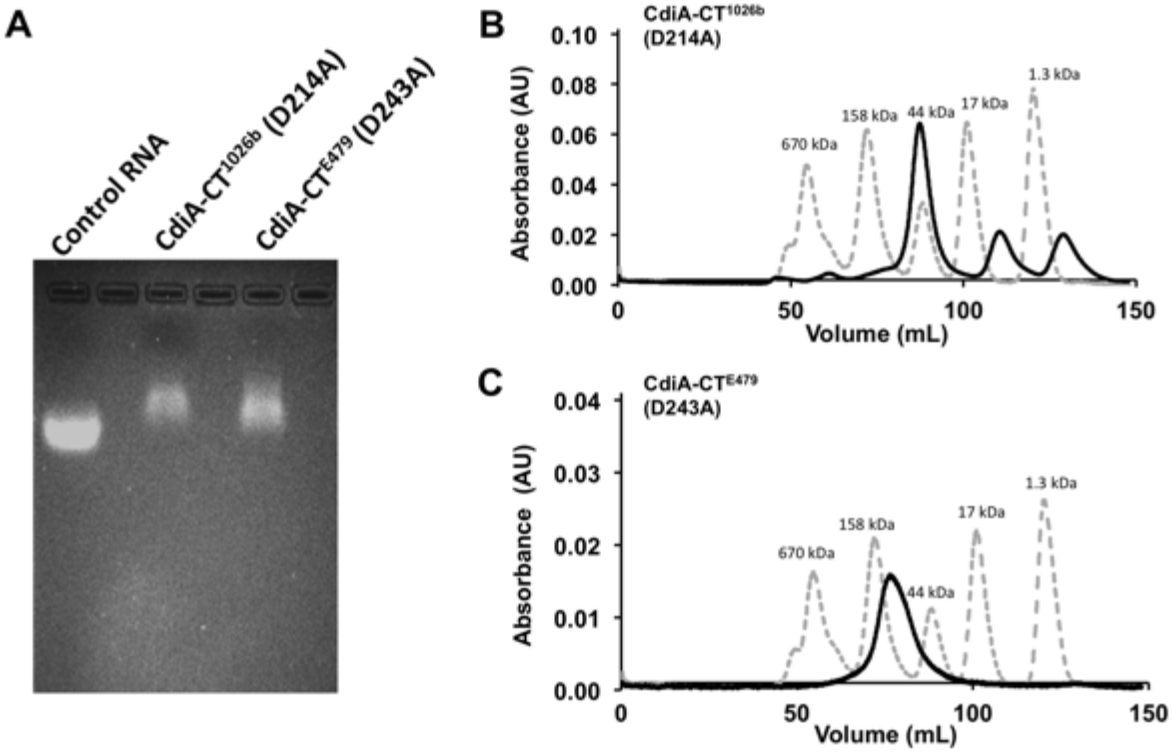


Figure 2.8: Inactive CdiA-CT^{E479} and CdiA-CT^{1026b} toxin domains bind to endogenous tRNA. (**A**) Agarose gel analysis of catalytically inactive CdiA-CT^{E479} and CdiA-CT^{1026b} toxins purified under non-denaturing conditions. Control RNA is from yeast (Sigma). (**B**) Size-exclusion chromatography of the purified tRNA/CdiA-CT(D214A)^{1026b} complex. (**C**) Size-exclusion chromatography of the purified tRNA/CdiA-CT(D243A)^{E479} complex. Chromatography migration standards are: bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa) and vitamin B_{12} (1.3 kDa). Each experiment was carried out in triplicate with similar results. Representative data are shown for each experiment.

Small-angle X-ray scattering (SAXS) analysis of toxin-tRNA complexes

To test the computational docking models, we sought to produce stable tRNA/toxin complexes for structural analysis using small-angle X-ray scattering (SAXS). We over-produced catalytically inactive toxins that carry N-terminal His_6-tags and purified the proteins by Ni^{2+}-affinity chromatography. Remarkably, large quantities of endogenous tRNA co-purified with each inactive toxin (**Figure 2.8A**), indicating that the tRNA/toxin complexes are indeed stable. The tRNA/CdiA-CT^{1026b} complex migrated at ~45 kDa on size-exclusion chromatography (SEC), indicative of a 1:1 complex with tRNA in solution (**Figure 2.8B**). By contrast, SEC showed that the tRNA/CdiA-CT^{E479} complex is ~150 kDa (**Figure 2.8C**), suggestive of a higher-order complex containing four nuclease domains bound to four tRNA molecules. We used SAXS to generate low-resolution electron density envelopes of each nucleoprotein complex. DAMAVER (31) was used to calculate normalized spatial discrepancies (NSD) of 0.979 ± 0.038 for tRNA/CdiA-CT^{1026b} and 0.959 ± 0.081 for the tRNA/CdiA-CT^{E479} complex, with no restorations rejected. These average NSD values imply reasonable stability of the solutions. The Hex-generated models for each tRNA/toxin complex were then fitted into the respective electron density envelopes using Chimera (32) (**Figure 2.7B** & **2.7D**). Supcomb (33) was used to calculate the NSD values between the averaged and filtered shape from SAXS and the structural models of 0.9371 for tRNA/CdiA-CT^{1026b} and 0.9013 for the tRNA/CdiA-CT^{E479} complex. These NSD values suggest that the average SAXS envelopes and the structural models agree quite well with each other. The four CdiA-CT^{E479} nuclease domains form a donut-like structure. Each nuclease domain fits helix α3 into the curvature of the β-sheet (strands β1-β3) on the adjacent domain. Inter-domain contacts also occur between helix α4 and β5 and the loop connecting strands β2 and β3. Within the tetramer, the four α1 helices are directed toward the center of the tetramer and the nuclease active sites project outward. The buried surface area of each monomer is consistent with a stable oligomeric state (34). The complex is further stabilized by interactions between tRNA molecules, which pack

together with their aminoacyl acceptor stems pointing into the center of the complex. Similar tight packing interactions have been observed in the crystal structure of $tRNA^{Asp}$ (35).

Discussion

The results presented here demonstrate that the CdiA-CT^{E479} nuclease domain is a member of the PD-(D/E)XK nuclease superfamily. Together with the previously characterized CdiA-CT^{1026b} and CdiA-$CT_{o11}{}^{EC869}$ nuclease domains (17,27), there are at least three CDI toxin classes that share the PD-(D/E)XK core fold. The amino acid sequences of these toxins are distinct (15 - 18% pair-wise sequence identity) and show no apparent relationship to one another through iterative PSI-BLAST analyses. However, structural superimposition of the CdiA-CT^{E479}, CdiA-CT^{1026b} and CdiA-$CT_{o11}{}^{EC869}$ toxins reveals significant similarities. The PD-(D/E)XK fold consists of a central four-stranded mixed β-sheet flanked by two α-helices with a characteristic $\alpha_1\beta_1\beta_2\beta_3\alpha_2\beta_3$ topology. The core structure serves as a scaffold to arrange catalytic residues. The canonical PD-(D/E)XK active site found in type II restriction endonucleases is built from a conserved Asp residue at the N-terminus of β2 and the (D/E)XK sub-motif within β3 of the core (23,36). However, there are several variations in the active-site configuration, with catalytic residues migrating to other secondary structure elements during evolution (23,37,38). For the CdiA-CT^{E479} nuclease domain, Asp229 and Asp243 occupy canonical positions within β2 and β3, but Glu204 and His275 are contributed by α1 and α2, respectively. This arrangement is very similar to the active site of CdiA-CT^{1026b} and the type IIS restriction endonuclease BspD6I (17,39). The DNase domains of CdiA-$CT_{o11}{}^{EC869}$ and CdiA-CT^{YPIII} have yet another configuration that was first described for EcoO109I (17,27,40,41). In these latter enzymes, Glu of the (D/E)XK sub-motif has migrated from β3 to helix α1 to produce an alternative E-PD-XXK motif. For the CdiA-$CT_{o11}{}^{EC869}$ class of toxins, the active sites use a E-(F/Y)D-SXK sequence motif (17,27). Interestingly, these latter DNases contain an additional β-hairpin inserted between α2 and β3 of the PD-(D/E)XK core. This β-hairpin

constitutes much of the binding interface with the cognate immunity protein and its sequence varies between family members (27). Analogous insertions into the PD-(D/E)XK core have been detected in other superfamily members (23), again underscoring the flexibility of the core fold.

Most PD-(D/E)XK enzymes are phosphodiesterases involved in DNA restriction, transposon excision, recombination, and repair. By contrast, there are relatively few family members with RNase activity. EndA/Sen15 tRNA splicing endonucleases were the first PD-(D/E)XK enzymes to be implicated in RNA metabolism (23,42). More recently, Rai1 has been reported to act as a phosphodiesterase to remove 5´-cap structures from eukaryotic mRNAs (43). Our findings show that the PD-(D/E)XK fold has been adopted to produce RNases with novel specificities. The CdiA-CT^{E479} nuclease domain cleaves tRNA between residues T54 and Ψ55 of the conserved TΨC-loop (16). Positions 54 and 55 are modified to thymidine and pseudouracil in eubacterial tRNAs, but these universal post-transcriptional modifications are not required for CdiA-CT^{E479} activity *in vitro*. CdiA-CT^{1026b} is a novel RNase that cleaves near the 3´-end of all tRNA (16,17). The computational docking studies reported here represent the first steps toward a detailed understanding of tRNA-binding specificity. Docking of tRNA onto the CdiA-CT^{1026b} domain provides a reasonable model for toxin binding to the aminoacyl-acceptor stem. The interaction between CdiA-CT^{E479} and substrate appears to be more complicated, and it is unclear why CdiA-CT^{E479}/tRNA complexes oligomerize in solution. It should be noted that CdiA-CT^{E479} in the absence of tRNA also forms a tetramer (data not shown). Though the models are still vague, it is tempting to speculate that the additional helix α2a within CdiA-CT^{E479} contributes to T-loop binding specificity. Helix α2a forms a ridge along the lower edge of the putative tRNA-binding surface. Residues Phe260 and Phe261, which form a prominent hydrophobic patch adjacent to the nuclease active site, may participate in substrate binding by stacking onto nucleobases. An elucidation of specific contacts must await high-resolution structural studies of toxin-substrate complexes. Given that inactive versions of each nuclease domain bind to tRNA with high affinity,

it should be possible to generate specific nucleoprotein complexes for high-resolution crystallography.

We have now reported crystal structures for four different CDI toxin classes. As described above, three of these toxins are nucleases of the PD-(D/E)XK superfamily. The other toxin, CdiA-CT^{ECL} from *Enterobacter cloacae* ATCC 13047, is an Ntox21 family member and adopts a fold common to barnase, EndoU, colicin E5/D and RelE (BECR) toxins (14,26). Sequence analyses by Aravind and colleagues indicate that CDI systems encode several other toxin families with distinct protein folds and activities (14,15). However, most CdiA-CT sequences do not have Pfam designations nor predictions for their biochemical activities (3). Given that the CdiA-CT^{E479}, CdiA-CT^{1026b} and CdiA-CT^{YPIII} toxins were not identified as PD-(D/E)XK nuclease by prior computational surveys, it remains possible that other uncharacterized CDI toxins also belong to the superfamily. Due to extreme sequence variability and catalytic residue migration, PD-(D/E)XK enzymes are notoriously difficult to identify through computational approaches (23,25,38). This problem is compounded by insertions and circular permutations of the core structure (23). Aravind and colleagues recently predicted five new restriction endonuclease-like domains (Tox-REase-2, -3, -5, -7 and -9) that are associated with prokaryotic competition systems (14). Only the Tox-REase-7 family (Pfam: PF15649) is found in CdiA effectors, and these CDI toxins appear to be limited to *Pseudomonas* and *Acinetobacter* species. These systems are under considerable positive selection to diversify, presumably due to the competitive advantage obtained with novel toxins. Similar pressures are postulated to drive the impressive diversity of restriction endonucleases, which is the result of the complex interplay between bacteria and their phages (44). Thus, it is not surprising that the versatile PD-(D/E)XK core structure has been adopted by CDI and other prokaryotic competition systems.

Materials and Methods

Plasmid constructions

Plasmids used in this study are listed in **Table 2.4**. Constructs for the over-production of CdiA-CT/CdiI^{1026b}-His$_6$ (pCH7590), CdiA-CT(D285A)/CdiI^{E479}-His$_6$ (pCH8288) and wild-type CdiA-CT/CdiI^{E479}-His$_6$ (pCH7770) complexes have been described previously (16). Active-site mutations were made in the CdiA-CT^{E479} nuclease domain using mega-primer PCR. Plasmid pCH7770 was amplified with primer E479-cdiI-Spe-rev (5´ - TTT ACT AGT ATT CCC CGA AAC TCC GAG CC) in conjunction with mutagenic forward primers: E479-E204A-for (5´ - AAA TTT AGA CCA GGT GCA GCC GGA GCA GCG GC), E479-D229A-for (5´ - GGC TCC TCG GTT GCC TTT GTA TTC AGC TCC), E479-D243A-for (5´ - AAC GGT AAG ACC GTG GCT TTT ATG CTT ACG CC), and E479-H275A-for (5´ - GAA CAC TCT TTC GGA TGC TGC GGC TGC TGC GG). The resulting products were used as mega-primers in subsequent reactions with forward primer E479-Nco-for (5´ - CGG CCA TGG CAT CGA ACG TCG AGC TTT AC). The final products were digested with NcoI and SpeI, then ligated to plasmid pET21 to generate mutant CdiA-CT/CdiI^{E479}-His$_6$ expression constructs. These plasmids were used as templates to amplify *cdiA-CT*^{E479} coding sequences with primers E479-Nco-for and E479-CT-Xho-rev (5´ - GCC ACT CGA GCC TTA CTT GAT CAG AAT AAT C). The products were digested with NcoI and Xhol, then ligated to plasmid pSCRhaB2 (45) to generate L-rhamnose inducible expression constructs to monitor growth inhibition activities. Plasmid pCH8479 was amplified with oligonucleotides 1026b-Spe-for (5´ - ATA ACT AGT GCA TCG AAC GTC GAG C) and 1026b-CT-Xho-rev (5´ - AAT CTC GAG TTA ATT CCC CTT TGG), and the resulting fragment ligated into plasmid pSH21 to generate a construct that over-produces inactive His$_6$-CdiA-CT(D214A)^{1026b}. The *cdiA-CT(D285A)*^{E479} coding sequence was amplified from pCH8427 with primers E479-CT-NdeI-H6-for (5´ - GAT CAT ATG ATG GGG GCA AGC TCA GGT AGT AAT ATC) and E479-CT-EcoRI-rev (5´ - GAT GAA TTC TCA CTT GAT CAG AAT AAT CTT

CGC CTG CAG TTT). The product was digested with NdeI/EcoRI and ligated to pET28b. The D243A mutation was introduced via site-directed mutagenesis using primers E479-CT-D243A-for (5´ - CGG TAA GAC CGT GGC GTT TAT GCT TAC GCC - 3´) and E479-CT-D243A-rev (5´ - GGC GTA AGC ATA AAC GCC ACG GTC TTA CCG - 3´) to produce an expression construct that over-produces His_6-CdiA-CT^{E479} carrying the Asp243Ala and Asp285Ala mutations.

Table 2.4. Bacterial strains and plasmids.

Strains or plasmids	***Description***[a]	***Reference or source***
Strain		
BL21 (DE3)	F$^-$ *ompT gal dcm lon hsdS$_B$* (r_B^- m_B^-)	Novagen
X90	F´ *lacI*^{q} *lac´ pro´/ara* Δ*(lac-pro) nal1 argE*(amb) *rif*^{r} *thi-1*, Rif^R	(61)
CH2016	X90 (DE3) Δ*rna* Δ*slyD::kan,* Rif^R Kan^R	(51)
Plasmids		
pET21b	IPTG-inducible T7 RNA polymerase expression vector, Amp^R	Novagen
pSCRhaB2	Rhamnose-inducible expression vector, Tp^R	(45)
pTrc99A	IPTG-inducible expression vector, Amp^R	GE Healthcare
pCH7590	pET21::*cdiA-CT(G123)-cdiI*$_{II}^{1026b}$, Amp^R	(16)
pCH7770	pET21::*cdiA-CT-cdiI*^{E479}, Amp^R	(16)
pCH8479	pET21::*cdiA-CT(G123/D214A)-cdiI*$_{II}^{1026b}$, Amp^R	(16)
pCH7913	pET21::*cdiI*^{E479}, Amp^R	(16)
pCH8288	pET21::*cdiA-CT(D285A)-cdiI*^{E479}, Amp^R	(16)
pCH8427	pET21::*cdiA-CT(G157/D285A)-cdiI*^{E479}, Amp^R	(16)
pPJ100	pET28::*his$_6$-cdiA-CT(G157/D243A/D285A)*^{E479}, Amp^R	This study
pCH10115	pET21::*his$_6$-cdiA-CT(G123/D214A)*^{1026b}, Amp^R	This study
pCH11617	pET21::*cdiA-CT(D229A)-cdiI*^{E479}, Amp^R	This study
pCH11618	pET21::*cdiA-CT(H275A)-cdiI*^{E479}, Amp^R	This study
pCH11619	pET21::*cdiA-CT(D243A)-cdiI*^{E479}, Amp^R	This study
pCH11620	pET21::*cdiA-CT(E204A)-cdiI*^{E479}, Amp^R	This study
pCH11648	pSCRhaB::*cdiA-CT(D229A)*^{E479}, Tp^R	This study
pCH11649	pSCRhaB::*cdiA-CT(H275A)*^{E479}, Tp^R	This study
pCH11650	pSCRhaB::*cdiA-CT(D243A)*^{E479}, Tp^R	This study
pCH11651	pSCRhaB::*cdiA-CT(E204A)*^{E479}, Tp^R	This study
pCH11669	pSCRhaB::*cdiA-CT*^{E479}, Tp^R	This study

[a]Abbreviations: Amp^R, ampicillin resistant; Kan^R, kanamycin resistant; Rif^R, rifampicin resistant; Tp^R, trimethoprim resistant

Protein over-expression and purification

CdiA-CT^{E479} (from residue Gly157, numbered from Glu1 of the ELYN motif) was co-expressed with CdiI^{E479}-His$_6$ and over-produced in *E. coli* BL21 (DE3) cells grown aerobically at 37 °C in LB medium supplemented with 50 µg/mL ampicillin. Protein expression was induced by addition of isopropyl-β-D-thiogalactoside (IPTG) to 1 mM final concentration once the culture reached an optical density at 600 nm (OD_{600}) of ~0.8. Induced cells were incubated for 4 h, then harvested by centrifugation at 5,100 x *g* for 20 min. The cell pellet was resuspended in 20 mM sodium phosphate (pH 7.0), 200 mM NaCl supplemented with 10 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride (PMSF), and the cells broken by sonication. The lysate was clarified by centrifugation at 14,000 x *g* for 30 min, and the soluble fraction passed through a 0.22 µM filter before loading onto a Ni^{2+}-charged Hi-trap column (GE Healthcare). The column was washed with 20 mM sodium phosphate (pH 7.0), 200 mM NaCl, 15 mM imidazole, and the CdiA-CT/CdiI^{E479}-His$_6$ complex eluted with a linear gradient of 15 – 250 mM imidazole. The purified complex was concentrated with 10 kDa centrifugal concentrator, then run on a Superdex 200 size-exclusion column equilibrated in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl. Selenomethionine (SeMet)-labeled proteins were over-produced in *E. coli* BL21 (DE3) cells grown in M9 minimal medium supplemented with L-leucine, L-isoleucine and L-valine at 50 mg/L; L-phenylalanine, L-lysine and L-threonine at 100 mg/L; and SeMet at 75 mg/L as described (46). The SeMet-labeled CdiA-CT/CdiI^{E479}-His$_6$ complex was purified as described above.

Crystallization and structure determination

CdiA-CT/CdiI^{E479} was crystallized by hanging drop-vapor diffusion against a 1 mL reservoir of crystallization buffer [0.1 M HEPES (pH 7.0), 20 mM $MgCl_2$, 30% (wt/vol) polyacrylic acid]. Polyacrylic acid (Sigma-Aldrich) with average molecular weight of 5,100 Da was prepared as a 50% (wt/vol) solution, filtered and used to facilitate crystallization. Hanging drops were prepared

from a 1:1 (vol:vol) mixture of protein solution (20 mg/mL) and crystallization buffer supplemented with 20 µg/mL chymotrypsin. Crystals were soaked in cryo-protectant solution containing 1:1 mixture of 40% (vol/vol) glycerol and crystallization buffer, then collected by flash freezing. A native dataset was acquired at 70 K at 0.97591 Å on beamline 7-1 at Stanford Synchrotron Radiation Lightsource (SSRL). Data were processed using HKL2000 (47), resulting in a 99.24% complete dataset to 2.0 Å resolution. The CdiA-CT/CdiI^{E479} complex crystallized in space group $P22_12_1$ with two complexes per asymmetric unit and unit-cell dimensions of 54.5 Å 5 73.3 Å 5 110.0 Å. Diffraction data were initially indexed and scaled to $P222_1$, however the best solution obtained by molecular replacement using Phaser in the PHENIX suite was $P22_12_1$. The SeMet-labeled complex was crystallized using 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (pH 6.5), 0.01 M $ZnCl_2$, 15% (wt/vol) PEG-6000 as the buffer. SeMet-labeled CdiA-CT/CdiI^{E479} crystallized in space group I4 with three complexes per asymmetric unit and unit-cell dimensions of 117.2 Å 5 117.2 Å 5 111.6 Å. A single-wavelength anomalous dataset (SAD) was collected at 70 K at 0.97591 Å on beamline 7-1 at SSRL. Data were processed using HKL2000, yielding a 99.9% complete dataset to 3.3 Å resolution. We used Autosol in the PHENIX suite (48) to detect 18 Se atoms with a figure of merit (FOM) of 0.408 and overall score of 36.2 ± 14.2. Autosol also built a partial model consisting of 416 residues with an R_{work}/R_{free} (%) of 38.4/44.9. This model showed little secondary structure except for α1 of CdiA-CT^{E479} and α2′ and α3′ of CdiI^{E479}. Molecular replacement was carried out with Phaser in the PHENIX suite using the partial model together with higher-resolution data from native crystals. The Phaser-generated model was then subjected to Autobuild and phenix.refine (48). The final model includes residues Arg201 – Lys316 of CdiA-CT^{E479} and residues Ala2 – Gly105 of CdiI^{E479} with a final R_{work}/R_{free} (%) 19.3/23.7. The Ramachandran plot shows 98.8% in the favorable allowed regions and the other 1.2% in the

allowed regions. Data collection and refinement statistics are presented in **Table 2.1**. Inter-molecular hydrogen bonds and salt-bridges were determined using PDBePISA (49).

Toxin-immunity protein binding kinetics

The apparent equilibrium dissociation constant for the CdiA-CT/CdiI^{E479} complex was determined by biolayer interferometry using a BLitz instrument (ForteBio) (50). CdiA-CT^{E479} was separated from CdiI^{E479}-His$_6$ by Ni^{2+}-affinity chromatography under denaturing conditions. The isolated toxin was refolded by dialysis and run on a Superdex S200 size-exclusion column. CdiI^{E479}-His$_6$ was purified by Ni^{2+}-affinity chromatography as described previously (16). CdiI^{E479}-His$_6$ was loaded onto a Ni^{2+}-NTA coated biosensor in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl at 25 °C. Sensor-bound immunity protein was then incubated with 1.5 – 4 μM CdiA-CT^{E479} toxin for 180 s. The sensor was then washed with buffer and toxin dissociation was monitored over 180 s. Curve fitting was run following reference subtraction using the BLitz Pro Software to calculate dissociation constants. Local fit analyses were performed for individual association-dissociation curves, followed by averaging to obtain the final apparent K_d and standard deviation.

Growth inhibition assays

E. coli X90 cells harboring rhamnose-inducible CdiA-CT^{E479} expression plasmids (**Table 2.4**) were grown to mid-log phase in LB media supplemented with 100 μg/mL trimethoprim. Cells were then diluted to OD_{600} = 0.05 in fresh LB media supplemented with 100 μg/mL trimethoprim and either 0.4% D-glucose to repress or 0.4% L-rhamnose to induce CdiA-CT^{E479} expression. Cultures were incubated with shaking at 37 °C and cell growth was monitored by measuring the OD_{600} every 30 min.

In vitro tRNase assays

Purified CdiA-CT/CdiI^{E479}-His$_6$ complexes were denatured in binding buffer supplemented with 6 M guanidine-HCl, and CdiA-CT^{E479} isolated from the void volume during Ni^{2+}-affinity chromatography (1). Toxins were refolded by dialysis against binding buffer and all purified proteins quantified by absorbance at 280 nm. Total RNA was isolated from *E. coli* X90 cells as described (51) and used as a substrate for *in vitro* nuclease assays. *E. coli* RNA (5 µg) was incubated with CdiA-CT^{E479} variants (5 µM) in reaction buffer [20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$, 0.2 mg/mL BSA] for 1 h at 37 °C. Where indicated, CdiI^{E479} was included at 17.5 µM final concentration. tRNA^{Asp} and tRNA^{Gln} substrates were generated by *in vitro* transcription using RNA polymerase from bacteriophage T7 RNA. The tRNA^{Asp} template was prepared with oligonucleotides 5´-tRNA-Asp (5´- AAT TCC TGC AGT AAT ACG ACT CAC TAT AGG AGC GGT AGT TCA GTC GGT TAG AAT ACC TG) and 3´-tRNA-Asp (5´- TGG CGG AAC GGA CGG GAC TCG AAC CCG CGA CCC CCT GCG TGA CAG GCA GGT ATT CTA AC); and the tRNA^{Gln} template with oligonucleotides 5´-tRNA-Gln (5´ - AAT TCC TGC AGT AAT ACG ACT CAC TAT AGG GGG TAT AGG GGG TAT CGC CAA GCG GTA AGG CAC CGG) and 3´-tRNA-Gln (5´ - TGG CTG GGG TAC GAG GAT TCG AAC CTC GGA ATG CCG GAA TCA GAA TCC GGT GCC TT). Annealed oligonucleotides were end-filled with Klenow fragment of DNA polymerase I. Templates were incubated with T7 RNA polymerase, 2 mM NTP, 10 mM dithiothreitol and 10 mM $MgCl_2$ for 3 h at 37 °C. Template DNA was removed with RNase-free DNase I, and the transcripts purified with the Direct-zol RNA MiniPrep Kit (Genesee Scientific). Nuclease reactions were analyzed by denaturing electrophoresis on 50% urea – 6% polyacrylamide gels in 1x Tris-borate-EDTA (TBE) buffer. Gels were stained with ethidium bromide or transferred to nylon membrane for Northern blot hybridization with 5´-radiolabeled oligonucleotide glyV probe (5´ - CTT GGC AAG GTC GTG CT) as described (16,51).

Molecular docking

Hex 8.0 (52-54) was used to dock CdiA-CT nuclease domains onto tRNA to generate models of enzyme-substrate complexes. CdiA-CT^{1026b} and CdiA-CT^{E479} nuclease domains were docked onto the structure of *E. coli* tRNA^{Cys} (PDB ID: 1B23) (55). The active site of each nuclease was positioned adjacent to the known scissile bond and the origin set to sample multiple orientations in search of the low-energy interactions. Positive control docking simulations were performed using cognate and non-cognate toxin/immunity proteins pairs from *B. pseudomallei* E479 and 1026b. As a negative control, CdiI^{E479} and CysK from *Salmonella* Typhimurium LT2 (PDB ID: 1OAS (56)) were docked onto one another.

Small-angle X-ray scattering (SAXS)

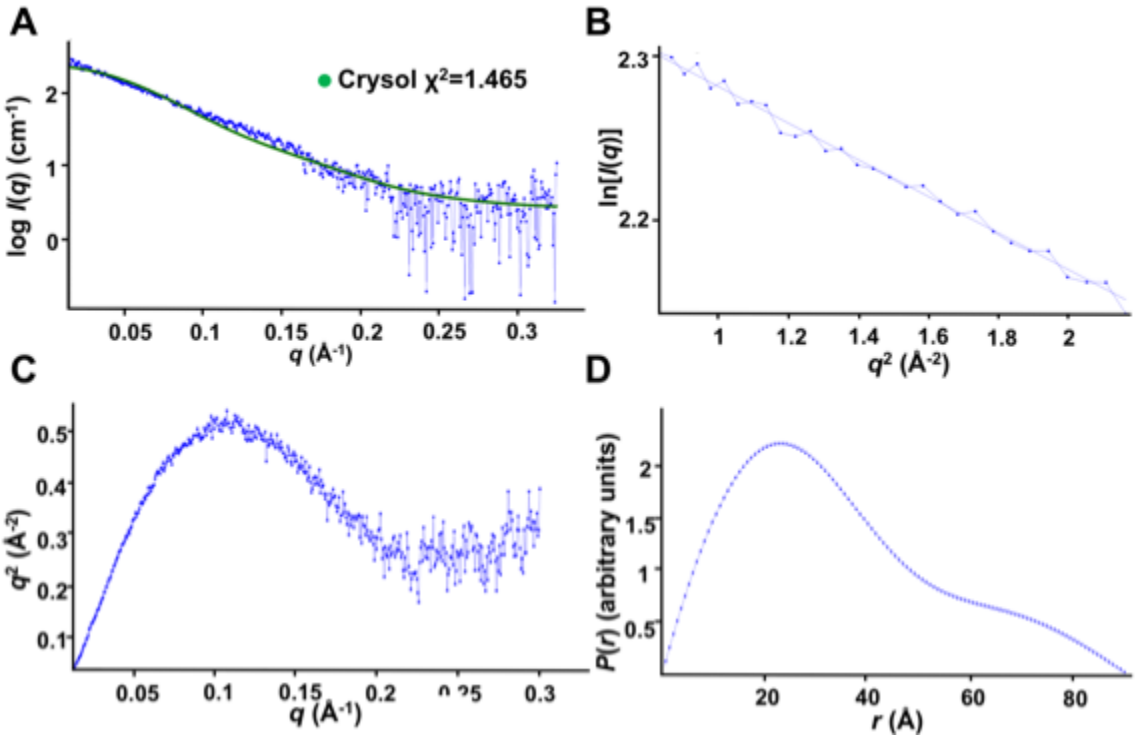
Inactive CdiA-CT(D243A/D285A)^{E479} and CdiA-CT(D214A)^{1026b} (17) toxins carrying N-terminal His_6 epitope tags were purified by Ni^{2+}-affinity chromatography under non-denaturing conditions. Under these conditions, endogenous tRNA co-purifies with the inactive nuclease domains. Toxin/tRNA complexes were exchanged into 20 mM sodium phosphate (pH 7.4), 150 mM NaCl using a Superdex S200 size-exclusion column and diluted to several concentrations ranging from 0.5 – 5 mg/mL for SAXS analysis. SAXS data were collected on SIBYLS beamline 12.3.1 at the Advanced Light Source using a Pilatus3 2M detector with exposure times of 0.5, 1, 2 and 4 s. Buffer subtracted data was analyzed using PRIMUS (31), following modification with GNOM (57), *P(r)* output files with d_{max} of 91 and 190 for tRNA/CdiA-CT^{1026b} and tRNA/CdiA-CT^{E479} complexes (respectively) were used to generate electron-density envelopes via GASBOR (58). Density envelopes (12 per tRNA/toxin complex) were averaged using DAMAVER (31), and docking solutions were fitted into the final envelopes using Chimera (32) and Crysol (59). SAXS parameters and statistics are provided in **Table 2.5** according to (60). Crysol outputs, together with Guinier, Kratky and *P(r)* plots, are presented in **Figure 2.9**.

Table 2.5. Data collection and scattering-derived parameters for tRNA/CdiA-CT complexes

Data-collection parameters	tRNA/CdiA-CT^{1026b}	tRNA/CdiA-CT^{E479}
Beam line	ALS 12.3.1	ALS 12.3.1
Wavelength (Å)	1.0	1.0
q range (Å^{-1})	0.012-0.324	0.012-0.324
Exposure time (sec)	0.5,1,2,4	0.5,1,2,4
Concentration range (mg ml^{-1})	0.5 – 5	2 – 10
Temperature (K)	283	283
Structural parameters†		
$I(0)$ (cm^{-1}) [from $P(r)$]	248.1 ± 13.2	595.9 ± 13.4
R_g (Å) [from $P(r)$]	27.3 ± 1.3	46.9 ± 3.4
$I(0)$ (cm^{-1}) [from Guinier]	246.8 ± 14.9	590.6 ± 20.8
R_g (Å) [from Guinier]	27.1 ± 1.7	46.1 ± 4.0
D_{max} (Å)#	91	190
Porod volume estimate (Å^3)	43041 ± 2353	256350 ± 2517
Dry volume calculated from Crysol[14] (Å^3) ‡	45510	43112 (176000)*
Molecular-mass determination†		
Molecular mass M_r [from Primus] (Da)	41539 ± 1675	165473 ± 6083
Calculated monomeric M_r from sequence (Da)‡	43215	41827 (167308)

† Reported for 5 mg mL^{-1} measurements.
‡ Values for tetramer shown in parentheses.
* The estimated volume of the tRNA/CdiA-CT^{E479} complex is greater than the calculated volume from sequence due to the void volume assumed by the tetramer.
# D_{max} is a model parameter in the *P(r)* calculation.

(I) tRNA/CdiA-CT^{1026b}



(II) tRNA/CdiA-CT^{E479}

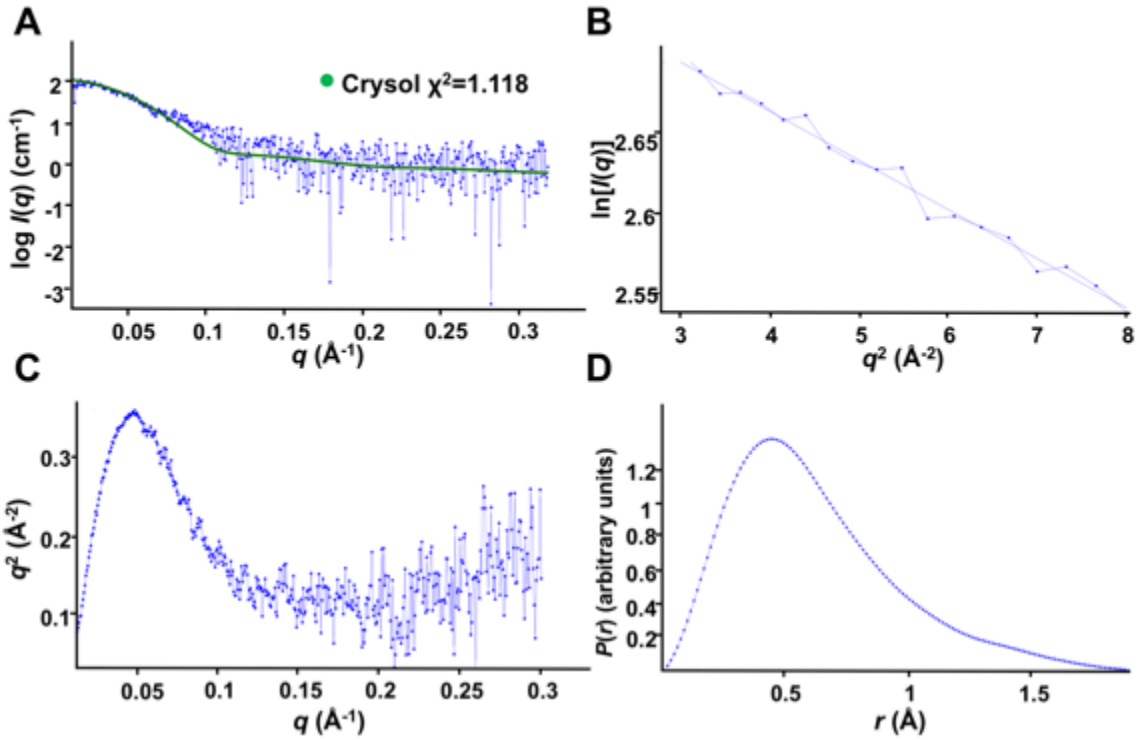


Figure 2.9: SAXS analyses of tRNA/CdiA-CT complexes. Plots for tRNA/CdiA-CT^{1026b} (I) and tRNA/CdiA-CT^{E479} (**II**) SAXS data. (**A**) log *I(q)* vs. *q* plot with experimental SAXS profile shown in blue and the corresponding structural model fitted data via Crysol (41) shown in green. (**B**) Guinier plots. (**C**) Kratky plots. (**D**) *P(r)* plots.

References

1. 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., and Low, D. A. (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**, 439-442
2. Ruhe, Z. C., Low, D. A., and Hayes, C. S. (2013) Bacterial contact-dependent growth inhibition. *Trends Microbiol* **21**, 230-237
3. Willett, J. L., Ruhe, Z. C., Goulding, C. W., Low, D. A., and Hayes, C. S. (2015) Contact-Dependent Growth Inhibition (CDI) and CdiB/CdiA Two-Partner Secretion Proteins. *J Mol Biol* **427**, 3754-3765
4. Makhov, A. M., Hannah, J. H., Brennan, M. J., Trus, B. L., Kocsis, E., Conway, J. F., Wingfield, P. T., Simon, M. N., and Steven, A. C. (1994) Filamentous hemagglutinin of *Bordetella pertussis*. A bacterial adhesin formed as a 50-nm monomeric rigid rod based on a 19-residue repeat motif rich in beta strands and turns. *J Mol Biol* **241**, 110-124
5. Kajava, A. V., Cheng, N., Cleaver, R., Kessel, M., Simon, M. N., Willery, E., Jacob-Dubuisson, F., Locht, C., and Steven, A. C. (2001) Beta-helix model for the filamentous haemagglutinin adhesin of *Bordetella pertussis* and related bacterial secretory proteins. *Mol Microbiol* **42**, 279-292
6. 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., and 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
7. Ruhe, Z. C., Townsley, L., Wallace, A. B., King, A., Van der Woude, M. W., Low, D. A., Yildiz, F. H., and Hayes, C. S. (2015) CdiA promotes receptor-independent intercellular adhesion. *Mol Microbiol* **98**, 175-192
8. Ruhe, Z. C., Wallace, A. B., Low, D. A., and Hayes, C. S. (2013) Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *mBio* **4**, e00480-00413
9. Aoki, S. K., Pamma, R., Hernday, A. D., Bickham, J. E., Braaten, B. A., and Low, D. A. (2005) Contact-dependent inhibition of growth in *Escherichia coli*. *Science* **309**, 1245-1248
10. Aoki, S. K., Poole, S. J., Hayes, C. S., and Low, D. A. (2011) Toxin on a stick: modular CDI toxin delivery systems play roles in bacterial competition. *Virulence* **2**, 356-359
11. Hayes, C. S., Aoki, S. K., and Low, D. A. (2010) Bacterial contact-dependent delivery systems. *Annu Rev Genet* **44**, 71-90
12. Poole, S. J., Diner, E. J., Aoki, S. K., Braaten, B. A., t'Kint de Roodenbeke, C., Low, D. A., and 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**, e1002217
13. Webb, J. S., Nikolakakis, K. C., Willett, J. L., Aoki, S. K., Hayes, C. S., and Low, D. A. (2013) Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. *PLoS ONE* **8**, e57609
14. Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M., and 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

15. Zhang, D., Iyer, L. M., and 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
16. 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., and Low, D. A. (2012) The toxin/immunity network of *Burkholderia pseudomallei* contact-dependent growth inhibition (CDI) systems. *Mol Microbiol* **84**, 516-529
17. Morse, R. P., Nikolakakis, K. C., Willett, J. L., Gerrick, E., Low, D. A., Hayes, C. S., and 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
18. Lazar Adler, N. R., Govan, B., Cullinane, M., Harper, M., Adler, B., and Boyce, J. D. (2009) The molecular and cellular basis of pathogenesis in melioidosis: how does Burkholderia pseudomallei cause disease? *FEMS Microbiol Rev* **33**, 1079-1099
19. Limmathurotsakul, D., and Peacock, S. J. (2011) Melioidosis: a clinical overview. *Br Med Bull* **99**, 125-139
20. Chantratita, N., Wuthiekanun, V., Limmathurotsakul, D., Vesaratchavest, M., Thanwisai, A., Amornchai, P., Tumapa, S., Feil, E. J., Day, N. P., and Peacock, S. J. (2008) Genetic diversity and microevolution of Burkholderia pseudomallei in the environment. *PLoS Negl Trop Dis* **2**, e182
21. Ngamdee, W., Tandhavanant, S., Wikraiphat, C., Reamtong, O., Wuthiekanun, V., Salje, J., Low, D. A., Peacock, S. J., and Chantratita, N. (2015) Competition between *Burkholderia pseudomallei* and *B. thailandensis*. *BMC Microbiol* **15**, 56
22. Ogawa, T., Tomita, K., Ueda, T., Watanabe, K., Uozumi, T., and Masaki, H. (1999) A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science* **283**, 2097-2100
23. Steczkiewicz, K., Muszewska, A., Knizewski, L., Rychlewski, L., and Ginalski, K. (2012) Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res* **40**, 7016-7045
24. Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549
25. Kosinski, J., Feder, M., and Bujnicki, J. M. (2005) The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. *BMC Bioinformatics* **6**, 172
26. Beck, C. M., Morse, R. P., Cunningham, D. A., Iniguez, A., Low, D. A., Goulding, C. W., and Hayes, C. S. (2014) CdiA from Enterobacter cloacae delivers a toxic ribosomal RNase into target bacteria. *Structure* **22**, 707-718
27. Morse, R. P., Willett, J. L., Johnson, P. M., Zheng, J., Credali, A., Iniguez, A., Nowick, J. S., Hayes, C. S., and Goulding, C. W. (2015) Diversification of beta-augmentation interactions between CDI toxin/immunity proteins. *J Mol Biol* **427**, 3766-3784
28. Willett, J. L., Gucinski, G. C., Fatherree, J. P., Low, D. A., and 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
29. Yang, X., Tan, S. H., Teh, Y. J., and Yuan, Y. A. (2011) Structural implications into dsRNA binding and RNA silencing suppression by NS3 protein of Rice Hoja Blanca Tenuivirus. *RNA* **17**, 903-911

30. Hemmes, H., Lakatos, L., Goldbach, R., Burgyan, J., and Prins, M. (2007) The NS3 protein of Rice hoja blanca tenuivirus suppresses RNA silencing in plant and insect hosts by efficiently binding both siRNAs and miRNAs. *RNA* **13**, 1079-1089
31. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277 - 1282
32. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612
33. Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Cryst* **34**, 33-41
34. Nooren, I. M., and Thornton, J. M. (2003) Structural characterisation and functional significance of transient protein-protein interactions. *J Mol Biol* **325**, 991-1018
35. Moras, D., Comarmond, M. B., Fischer, J., Weiss, R., Thierry, J. C., Ebel, J. P., and Giege, R. (1980) Crystal structure of yeast tRNAAsp. *Nature* **288**, 669-674
36. Pingoud, A., and Jeltsch, A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res* **29**, 3705-3727
37. Knizewski, L., Kinch, L. N., Grishin, N. V., Rychlewski, L., and Ginalski, K. (2007) Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. *BMC Struct Biol* **7**, 40
38. Gupta, R., Capalash, N., and Sharma, P. (2012) Restriction endonucleases: natural and directed evolution. *Appl Microbiol Biotechnol* **94**, 583-599
39. Kachalova, G. S., Rogulin, E. A., Yunusova, A. K., Artyukh, R. I., Perevyazova, T. A., Matvienko, N. I., Zheleznaya, L. A., and Bartunik, H. D. (2008) Structural analysis of the heterodimeric type IIS restriction endonuclease R.BspD6I acting as a complex between a monomeric site-specific nickase and a catalytic subunit. *J Mol Biol* **384**, 489-502
40. Hashimoto, H., Shimizu, T., Imasaki, T., Kato, M., Shichijo, N., Kita, K., and Sato, M. (2005) Crystal structures of type II restriction endonuclease EcoO109I and its complex with cognate DNA. *J Biol Chem* **280**, 5605-5610
41. Feder, M., and Bujnicki, J. M. (2005) Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site. *BMC Genomics* **6**, 21
42. Bujnicki, J. M., and Rychlewski, L. (2001) Unusual evolutionary history of the tRNA splicing endonuclease EndA: relationship to the LAGLIDADG and PD-(D/E)XK deoxyribonucleases. *Protein Sci* **10**, 656-660
43. Jiao, X., Xiang, S., Oh, C., Martin, C. E., Tong, L., and Kiledjian, M. (2010) Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* **467**, 608-611
44. Stern, A., and Sorek, R. (2011) The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43-51
45. Cardona, S. T., and Valvano, M. A. (2005) An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* **54**, 219-228
46. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105-124

47. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326
48. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601
49. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797
50. Concepcion, J., Witte, K., Wartchow, C., Choo, S., Yao, D., Persson, H., Wei, J., Li, P., Heidecker, B., Ma, W., Varma, R., Zhao, L. S., Perillat, D., Carricato, G., Recknor, M., Du, K., Ho, H., Ellis, T., Gamez, J., Howes, M., Phi-Wilson, J., Lockard, S., Zuk, R., and Tan, H. (2009) Label-free detection of biomolecular interactions using BioLayer interferometry for kinetic characterization. *Comb Chem High Throughput Screen* **12**, 791-800
51. Garza-Sanchez, F., Janssen, B. D., and 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
52. Ritchie, D. W. (2003) Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins* **52**, 98-106
53. Ritchie, D. W., and Kemp, G. J. (2000) Protein docking using spherical polar Fourier correlations. *Proteins* **39**, 178-194
54. Ritchie, D. W., and Venkatraman, V. (2010) Ultra-fast FFT protein docking on graphics processors. *Bioinformatics* **26**, 2398-2405
55. Nissen, P., Thirup, S., Kjeldgaard, M., and Nyborg, J. (1999) The crystal structure of Cys-tRNACys-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure* **7**, 143-156
56. Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998) Three-dimensional structure of O-acetylserine sulfhydrylase from Salmonella typhimurium. *J Mol Biol* **283**, 121-133
57. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495-503
58. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946-2953
59. Stovgaard, K., Andreetta, C., Ferkinghoff-Borg, J., and Hamelryck, T. (2010) Calculation of accurate small angle X-ray scattering curves from coarse-grained protein models. *BMC Bioinformatics* **11**, 429
60. Jacques, D. A., Guss, J. M., Svergun, D. I., and Trewhella, J. (2012) Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution. *Acta Crystallogr D Biol Crystallogr* **68**, 620-626
61. Beckwith, J. R., and Signer, E. R. (1966) Transposition of the *lac* region of *Escherichia coli*. I. Inversion of the *lac* operon and transduction of *lac* by phi80. *J Mol Biol* **19**, 254-265

CHAPTER 3

Unraveling the essential role of CysK in CDI toxin activation

This project was performed in collaboration with members of the Hayes lab at UC Santa Barbara.

Abstract

Contact-dependent growth inhibition (CDI) systems deliver CdiA-CT toxins into neighboring target bacteria and produce CdiI immunity proteins to protect against self-inhibition. The CdiA-CT^{EC536} toxin deployed by uropathogenic *Escherichia coli* 536 (EC536) is an Ntox28 family domain that only exhibits tRNase activity when bound to the cysteine biosynthetic enzyme CysK (*O*-acetylserine sulfhydrylase-A). Here, we present crystal structures of the CysK/CdiA-CT^{EC536} binary complex and the neutralized ternary complex of CysK/CdiA-CT/CdiI^{EC536}. CdiA-CT^{EC536} inserts its C-terminal GYGI peptide tail into the active-site cleft of CysK to anchor the interaction. Remarkably, *E. coli* CysE (serine *O*-acetyltransferase) uses a similar GDGI motif to form the 'cysteine synthase' complex with CysK. The cysteine synthase complex is found throughout bacteria, protozoa and plants, indicating that CdiA-CT^{EC536} exploits a highly conserved protein-protein interaction to promote its toxicity. CysK significantly increases CdiA-CT^{EC536} thermostability and is required for toxin interaction with tRNA substrates. These observations suggest that CysK stabilizes the toxin fold, thereby organizing the nuclease active site for substrate recognition and catalysis. By contrast, Ntox28 domains from Gram-positive bacteria lack C-terminal GYGI motifs, suggesting that they do not interact with CysK. We show that the Ntox28 domain from *Ruminococcus lactaris* is significantly more thermostable than CdiA-CT^{EC536}, and its intrinsic tRNA-binding properties support CysK-independent nuclease activity. The striking

differences between related Ntox28 domains suggest that CDI toxins may be under evolutionary pressure to maintain low global stability.

Introduction

Bacteria have evolved diverse mechanisms to communicate and compete with neighboring microorganisms. One such mechanism is contact-dependent growth inhibition (CDI), which mediates the transfer of protein toxins between Gram-negative bacterial cells. CDI systems are distributed throughout α, β and γ-proteobacteria and are particularly common in pathogens (1). CDI is mediated by the CdiB/CdiA family of two-partner secretion proteins. CdiB is an Omp85-family transporter that exports and assembles CdiA effectors onto the cell surface. CdiA proteins share homology with filamentous hemagglutinin adhesins and are thought to form long filaments projecting from the inhibitor cell. CdiA recognizes specific receptors on susceptible bacteria and delivers its toxic C-terminal domain (CdiA-CT) to inhibit target-cell growth. CDI^+ bacteria also produce CdiI immunity proteins, which bind CdiA-CT domains and neutralize their activity to protect against self-inhibition. The CdiA-CT region is highly polymorphic between bacteria, with sequences diverging abruptly after the VENN peptide motif within the conserved pre-toxin domain (Pfam: PF04829) (3, 4). CdiA-CT diversity reflects the variety of toxins deployed during CDI, with most experimentally characterized toxins exhibiting distinct nuclease activities (3, 5-7). CdiI immunity proteins are also variable and only neutralize their cognate CdiA-CT toxins. Thus, CDI is thought to mediate inter-strain competition, with toxin/immunity protein variability providing a mechanism to discriminate between self and non-self (1, 2).

Previous studies on the CDI toxin from uropathogenic *E. coli* 536 (EC536) revealed that it possesses latent anticodon nuclease activity against all tRNA isoacceptors (8). The $CdiA\text{-}CT^{EC536}$

region is composed of two domains that have distinct functions during CDI (9). The extreme C-terminal domain is an Ntox28 RNase family member (Pfam: PF15605), and is responsible for growth-inhibition activity (3, 8). The N-terminal domain facilitates translocation of the tethered nuclease into the cytosol of target bacteria (9). Although CdiA-CT^{EC536} rapidly cleaves tRNA *in vivo*, the purified toxin has no detectable nuclease activity *in vitro* (8). Using biochemical approaches, we discovered that CdiA-CT^{EC536} is activated when bound to the biosynthetic enzyme *O*-acetylserine sulfhydrylase-A (CysK). CysK is one of two isoenzymes (along with CysM) that catalyze the final reaction in cysteine synthesis. In bacteria and plants, CysK is found in the 'cysteine synthase' complex together with CysE – the serine *O*-acetyltransferase responsible for the penultimate step of cysteine synthesis (10). Formation of the cysteine synthase complex requires the C-terminal tail of CysE, which inserts into the CysK active site (11). The C-terminal Ile residue of CysE is particularly critical and interacts with the CysK active site using the same contacts as the enzyme substrate *O*-acetylserine (12, 13). Though CysK and CysM share 58% sequence identity, CysE does not interact stably with the CysM isoenzyme (14). The CdiA-CT^{EC536} toxin carries a C-terminal GYGI peptide motif that appears to mimic the GDGI tail of *E. coli* CysE. Moreover, *O*-acetylserine blocks the binding of both CysE and CdiA-CT^{EC536} to CysK (8, 10), indicating that CdiA-CT^{EC536} also inserts its C-terminal tail into the CysK active-site cleft. Remarkably, other proteins mimic CysE to bind CysK (15). In *Bacillus* species, the CymR transcription factor uses its C-terminal MFYI tail to bind CysK, and modulation of this interaction controls the *cys* regulon (16). Perhaps more intriguing is EGL-9, an O_2-sensing prolyl hydroxylase from *Caenorhabditis elegans* that binds a CysK homolog (CYSL-1) using a C-terminal Ile residue (17). The resulting complex senses O_2 tension indirectly through hydrogen sulfide, which accumulates during hypoxia. Sulfide binds the CYSL-1 active site and displaces the C-terminus of

EGL-9. Once liberated, EGL-9 hydroxylates HIF-1 to initiate transcriptional responses to hypoxia (17). Thus, CysK and its homologs have been coopted to regulate gene expression in bacteria and eukaryotes.

To gain mechanistic insight into toxin activation, we solved crystal structures of the CdiA-CT^{EC536} toxin in binary complex with CysK, and in ternary complex with CysK and CdiI^{EC536} immunity protein. The nuclease domain forms a small four-helix bundle with no structural similarity to other known RNase families. Two toxins bind to each CysK homodimer, and the C-terminal GYGI peptide of the nuclease inserts into the CysK active-site cleft as predicted by previous studies (8). Structure-guided mutagenesis reveals a putative catalytic triad of Asp155, His178 and Glu181 in the nuclease domain. The predicted nuclease active site is occluded by immunity protein in the CysK/CdiA-CT/CdiI^{EC536} structure, suggesting that CdiI^{EC536} blocks the binding of tRNA substrates to the toxin. Intriguingly, Ntox28 homologs from Gram-positive bacteria lack the C-terminal GYGI motif, suggesting that they do not require CysK-mediated activation. We tested this prediction using Tox28^{Rlac} from *Ruminococcus lactaris* and confirmed that the domain possesses CysK-independent tRNase activity. Moreover, Tox28^{Rlac} is significantly more stable to thermal denaturation than CdiA-CT^{EC536} and possesses intrinsic tRNA-binding activity. By contrast, CysK is required for tRNA binding to CdiA-CT^{EC536}. Collectively, these findings suggest that CysK is recruited to stabilize CdiA-CT^{EC536}, thereby organizing the active site to promote substrate-binding and catalysis.

Results

Structure of the CysK/CdiA-CT^{EC536} complex

We generated a binary complex of CysK bound to inactive CdiA-CT^{EC536} toxin that contains the His178Ala mutation (numbered from Val1 of the VENN motif). The CysK/CdiA-CT(H178A)^{EC536} complex crystallized in space group $P4_1$ and the structure was solved to 2.7 Å resolution (**Table 3.1**). Like other *O*-acetylserine sulfhydrylases (18), *E. coli* CysK is homodimeric and contains a pyridoxal 5´-phosphate (PLP) cofactor in Schiff-base linkage to Lys42. CysK in the binary complex has an "open" active-site conformation, similar to the structure of unliganded CysK from *Salmonella* Typhimurium (rmsd of 0.5 Å over all 312 α-carbons) (18). Thus, the toxin does not induce the "closed" conformation observed when CysK contains substrate covalently bound to PLP in the active site (19). As we have found in other CdiA-CT toxin structures (5, 6, 20), only the C-terminal nuclease domain (residues Lys127 – Ile227) is resolved in the final model. The CdiA-CT^{EC536} nuclease domain consists of four α-helices.

Three long helices (α1, α3 and α4) form a bundle capped by the shorter helix α2 (**Figure 3.1A**). Helices α2 and α3 are connected by the long flexible loop L2, which was modeled predominately as Ala residues. Two CdiA-CT^{EC536} nuclease domains bind to the CysK dimer, but the toxins make no contact with one another (**Figure 3.1A**), suggesting that they bind independently. The C-terminal tail of the nuclease domain inserts into the CysK active-site, with the GYGI peptide backbone forming a network of hydrogen-bonds (H-bonds) with CysK (**Figure 3.1B** & **Table 3.2**).

Table 3.1. X-ray diffraction data and refinement statistics for CysK/CdiA-CT^{EC536} complexes

	CysK/CdiA-CT(H178A)^{EC536}	CysK/^{a}CdiA-CT/CdiI^{EC536}
Space Group	$P4_1$	$C222_1$
Unit cell dimensions (Å)	64.01 x 64.01 x 365.37	81.25 x 195.54 x 175.06
pH of crystallization condition	7.9	7.1
Protein concentration (mg/mL)	20	20
Data Collection		
Wavelength, Å	1.0	0.9795
Resolution range	44.92-2.7	50-2.75
Unique reflections (total)	39795 (303136)	33877 (483433)
Completeness, %*	99.4	100.0
Redundancy*	12.8 (13.4)	14.3 (14.5)
R_{merge}*, †	0.163 (0.734)	0.279 (1.159)
R_{meas}*, ‡	0.176 (0.791)	0.289 (1.201)
$R_{p.i.m.}$*, §	0.066 (0.293)	0.077 (0.314)
$CC_{1/2}$*	0.996 (0.952)	0.995 (0.817)
I/σ*	19.96 (15.94)	10.41 (2.75)
NCS copies	2	2
Model refinement		
Resolution range, Å	44.89-2.70	48.83-2.75
No. of reflections	39640	36673
No. of protein + ligand atoms	6220	8197
No. of water molecules	110	101
Missing residues	CdiA-CT:1-126 CysK:315-323	CdiA-CT:1-126 CysK:315-323 CdiI^{b}:125-128
R_{work}/R_{free}, %¶	20.2/22.4	19.4/24.2
Rms deviations		
Bond lengths, Å	0.009	0.011
Bond angles	1.225	1.072
Ramachandran plot		
Most favorable region, %	97.08	95.83
Additional allowed region, %	2.92	3.90
Disallowed region	0	0.27
PDB ID code	5J43	5J5V

*Statistics for the highest-resolution shell are given in parentheses.

† $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

‡ $R_{meas} = \Sigma_{hkl} \{N(hkl) / [N(hkl) - 1]\}^{1/2} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

§ $R_{p.i.m}$ (precision-indicating R_{merge}) $= \Sigma_{hkl} \{1/[N(hkl) - 1]\}^{½} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

¶ $R_{work} = \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$. R_{free} was computed identically except where all reflections belong to a test set of 5% randomly selected data.

[a]CdiA-CT/CdiI^{EC536} complex is a SeMet derivatives.

[b]Missing residues for one CdiI protomer. The other CdiI protomer is only missing residue 128.

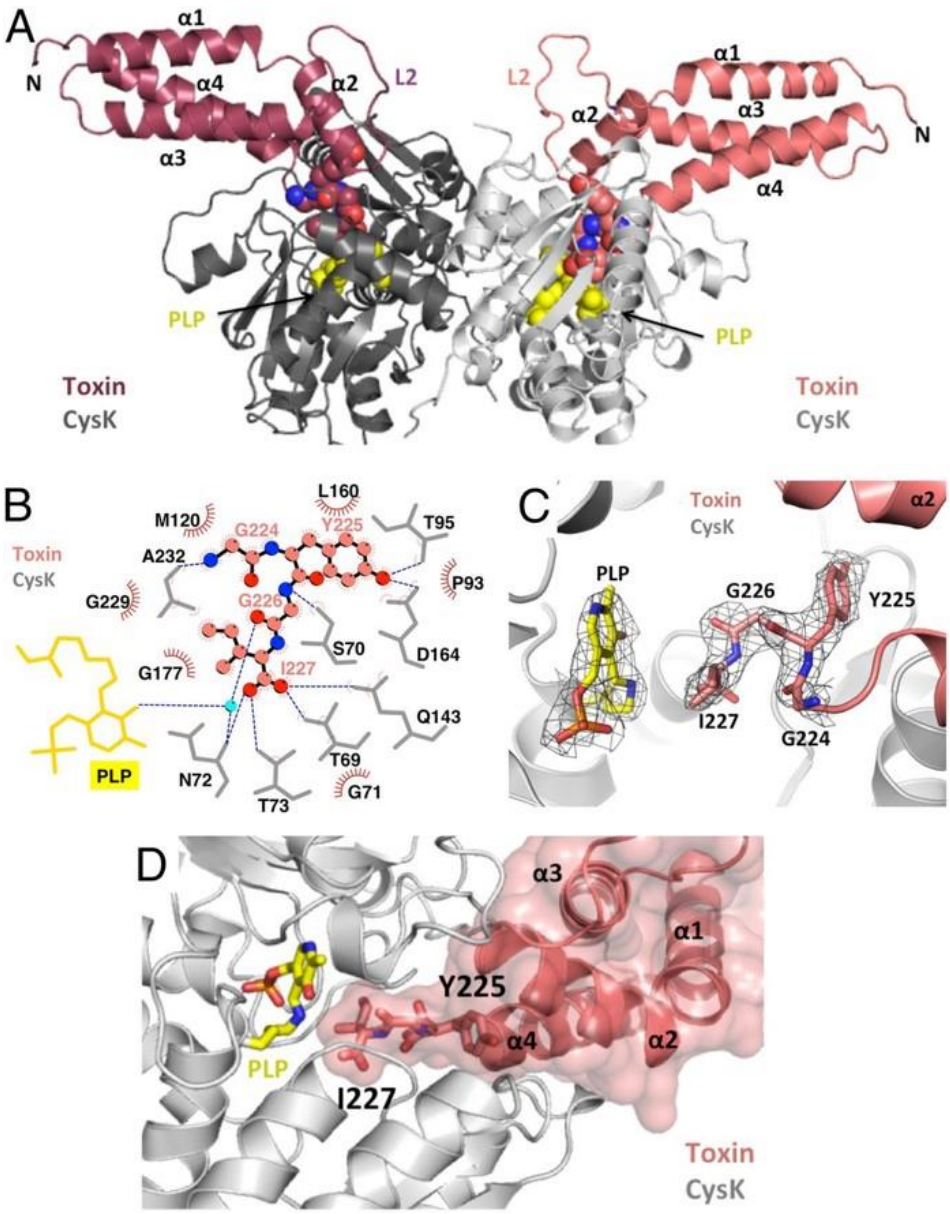


Figure 3.1: The CysK/CdiA-CT^{EC536} binary complex. (**A**) Crystal structure of the CysK/CdiA-CT^{EC536} complex. Secondary structure elements of the toxin nuclease domain are indicated together with flexible loop L2. The C-terminal GYGI peptides of CdiA-CT^{EC536} and CysK-bound pyridoxyl 5´-phosphate (PLP) are rendered as spheres. (**B**) GYGI peptide interaction network. CdiA-CT^{EC536} residues Gly224, Tyr225, Gly226 and Ile227 are shown in spheres, and CysK residues and PLP rendered as grey and yellow sticks,respectively. A water molecule is shown as a cyan sphere. Red arcs represent hydrophobic interactions and blue dashed lines indicate H-bonds. The interaction network was produced using LigPlot. (**C**) Interaction between CdiA-CT^{EC536} C-terminal GYGI peptide and the CysK active site. The CdiA-CT^{EC536} GYGI peptide and CysK PLP are shown in stick representation. The $2F_o$-F_c electron density map of the CdiA-CT^{EC536} GYGI peptide and CysK PLP is shown in grey mesh and contoured at 1.0 σ. (**D**) The CdiA-CT^{EC536} toxin domain exploits shape complementarity to bind the CysK active-site cleft. Residues Gly224 – Ile227 and PLP are shown in stick representation.

The toxin's C-terminal Ile227 residue is positioned in close proximity to the active-site PLP and forms H-bonds with Thr69, Asn72, Thr73 and Gln143 of CysK (**Figure 3.1B** & **3.1C**). These latter contacts have been observed in structures of CysE C-terminal peptides bound to the CysK active site (13, 21, 22). The side-chain of Tyr225 H-bonds to CysK residues Thr95 and Asp164 (**Figure 3.1B** & **Table 3.2**). The GYGI tail interactions are buttressed by additional H-bonds and salt-bridges emanating from toxin helices α2, α3 and α4 (**Table 3.2**). Residues within α2 (Leu160, Ile156, Ile157, Met163), α3 (Met179, Leu186) and α4 (Leu222) also form hydrophobic interactions with CysK, and the toxin helical bundle exploits shape complementarity to fit into the CysK active-site cleft (**Figure 3.1D**). Overall, the complex interface is 1,280 $Å^2$, burying 9.2% and 19.4% of the solvent accessible surface areas of CysK and CdiA-CT^{EC536}, respectively. Dissociation constants were determined using biolayer interferometry using a Blitz instrument (ForteBio) (**Table 3.3**).

Table 3.2

A. Hydrogen bonds and salt bridges between CdiA-CT^{EC536} and CysK in the binary complex.

CdiA-CT^{EC536}	**CysK**	**Distance (Å)**
LYS 161 [NZ]	THR 95 [O]	3.09
GLY 165 [O]	ALA 310 [N]	3.02
LYS 166 [NZ]	SER308 [OG]	3.40
GLN 183 [NE2]	PRO 224 [O]	3.11
GLN 183 [NE2]	ASP 207 [OD2]	3.22
ARG 190 [NH2]	PRO 222 [O]	2.95
SER 220 [O]	LYS 121 [N]	3.17
ALA 221 [O]	MET 120 [N]	2.90
GLY 224 [N]	ALA 232 [O]	2.77
TYR 225 [OH]	THR 95 [OH]	3.00
GLY 226 [N]	SER 70 [OG]	2.64
ILE 227 [O]	THR 69 [OG1]	2.62
ILE 227 [O]	GLN 143 [NE2]	3.04

ILE 227 [OXT]	THR 73 [N]	3.14
ILE 227 [OXT]	ASN 72 [N]	3.35

B. (i) Hydrogen bonds and salt bridges between CdiA-CT^{EC536} and CysK in the ternary complex.

CdiA-CT^{EC536}	**CysK**	**Distance (Å)**
LYS 161 [NZ]	THR 95 [O]	2.71
GLY 165 [O]	ALA 310 [N]	2.76
LYS 166 [NZ]	SER308 [OG]	3.40
GLN 183 [NE2]	PRO 224 [O]	3.90
GLN 183 [NE2]	ASP 207 [OD2]	3.69
ARG 190 [NH2]	PRO 222 [O]	2.70
SER 220 [O]	LYS 121 [N]	3.12
ALA 221 [O]	MET 120 [N]	3.12
GLY 224 [N]	ALA 232 [O]	2.92
TYR 225 [OH]	THR95 [OG1]	2.86
GLY 226 [N]	SER 70 [OG]	2.65
ILE 227 [O]	THR 69 [OG1]	2.86
ILE 227 [O]	GLN 143 [NE2]	3.26
ILE 227 [OXT]	THR 73 [N]	3.57
ILE 227 [OXT]	ASN 72 [N]	3.57

(ii) Hydrogen bonds and salt bridges between CdiA-CT^{EC536} and CdiI^{EC536} in the ternary complex

CdiA-CT^{EC536}	**CdiI^{EC536}**	**Distance (Å)**
LYS 170 [NZ]	GLN 49 [O]	2.81
GLU 171 [N]	ILE 11 [O]	2.64
GLU 171 [OE2]	THR 14 [N]	3.70
ASP 177 [N]	GLN 49 [OE1]	3.00
ASP 177 [O]	GLN 49 [NE2]	2.89
HIS 178 [NE2]	SER 21 [OG]	3.40
GLU 181 [OE1]	ARG 45 [NH1]	3.03
GLU 181 [OE1]	ARG 48 [NE]	3.57
GLU 181 [OE1]	ARG 48 [NH2]	3.01
GLU 181 [OE2]	GLN49 [NE2]	2.90

GLU 181 [OE2]	ARG 48 [NE]	3.06
GLU 181 [OE2]	ARG 48 [NH2]	3.79
ASN 184 [ND2]	GLU 77 [OE2]	2.93
ASN 184 [ND2]	TYR 74 [OH]	2.77
HIS 192 [NE2]	TYR 73 [O]	2.80

E. coli CysK and CysM share related structures (rmsd of 1.8 Å over 285/292 α-carbons), but differ significantly in the loop L17 region at the entrance of the active-site cleft (**Figure 3.2A**). CysK contains 16 extra residues in loop L17 that introduce an additional α-helix (α7°) (**Figure 3.2A** & **3.3**). The differences in loop L17 have profound effects on the landscape and electrostatic surface potential of each active-site cleft (**Figure 3.2B**) (12, 23). Superimposition of CdiA-CT^{EC536} onto the CysM structure produces several clashes with the C-terminal GYGI peptide (**Figure 3.2B**). Moreover, the CysM active site contains no hydrophobic pockets capable of accommodating the side-chains of toxin residues Ile227 and Tyr225. Residues Lys223 and Gly224 at the end of toxin helix α4 and Met179 and Gln183 within helix α3 each clash with the surface of CysM loop L17 (**Figure 3.2C**). Thus, CdiA-CT^{EC536} and CysE both exploit differences in loop L17 sequence and structure to bind CysK specifically.

Table 3.3. Dissociation constants of UPEC and *R. lactaris* CDI proteins determined by biolayer interferometry

Sample	Kd
UPEC CdiI + CdiA-CT/CysK	16.1 ± 7.4 nM
UPEC CdiI/CdiA-CT + CysK	25.4 ± 17.3 nM
UPEC CdiI + CdiA-CT	27.4 ± 22.1 nM
UPEC CdiA-CT + CysK	75.1 ± 13.9 nM
Rlac Immunity + Toxin	110 ± 39 nM

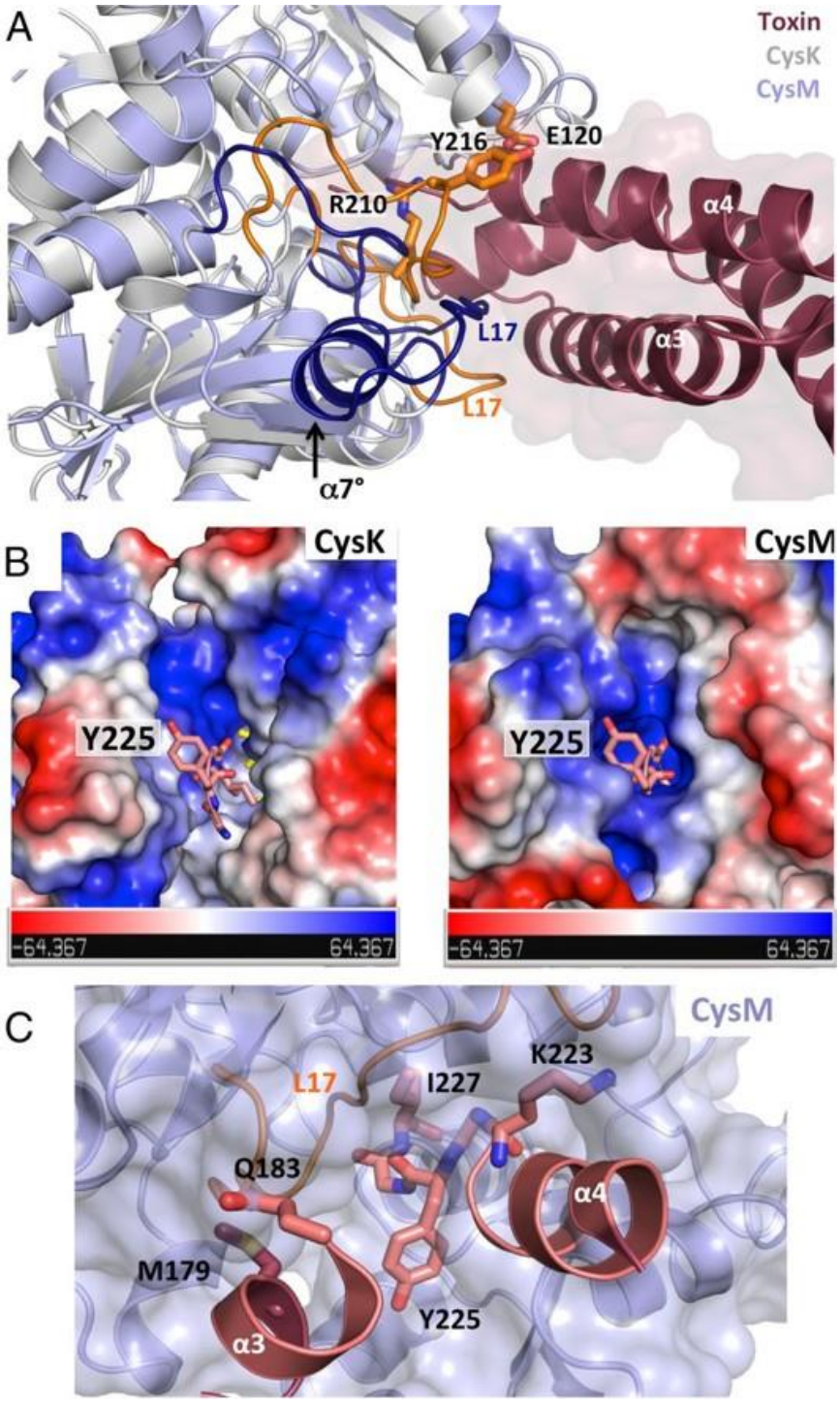


Figure 3.2: Comparison of *E. coli* CysK and CysM structures. (**A**) Superimposition of *E. coli* CysM (PDB code: 2BHT) onto the CysK/CdiA-CT^{EC536} binary complex. The divergent loop L17 regions are highlighted in orange for CysM and dark blue for CysK. The additional helix α7° in CysK is indicated, as are CysM residues that clash with the surface of the nuclease domain. (**B**) Electrostatic surface representations of CysK and CysM. Electric isopotentials of +64.4 kT/e and –64.4 kT/e are shown in blue and red, respectively. The C-terminal GYGI peptide of CdiA-CT^{EC536} is shown in stick representation. (**C**) CdiA-CT^{EC536} modeled onto CysM structure. Toxin residues that clash with loop L17 of CysM are shown in stick representation.

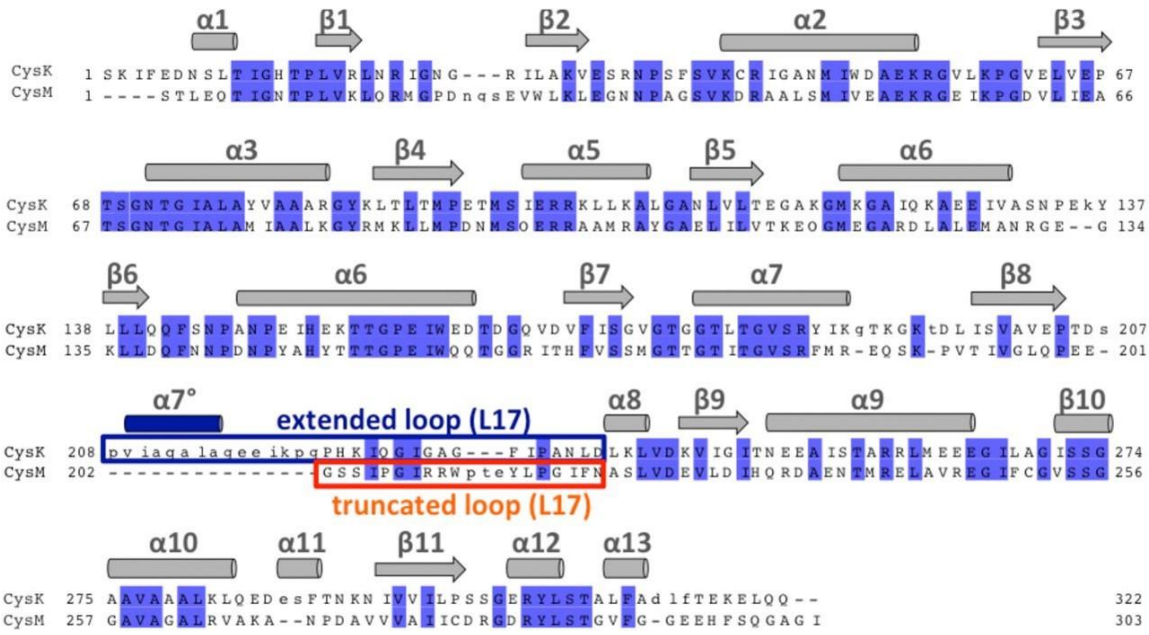


Figure 3.3: Structure and sequence alignment of *E. coli* CysK and CysM. CysK and CysM sequences were aligned using Clustal omega. Identical residues are highlighted in blue. Secondary structure elements are presented above the alignment. CysK contains an extended loop L17 region (boxed in blue) and an additional helix (α7° in blue) compared to CysM (orange).

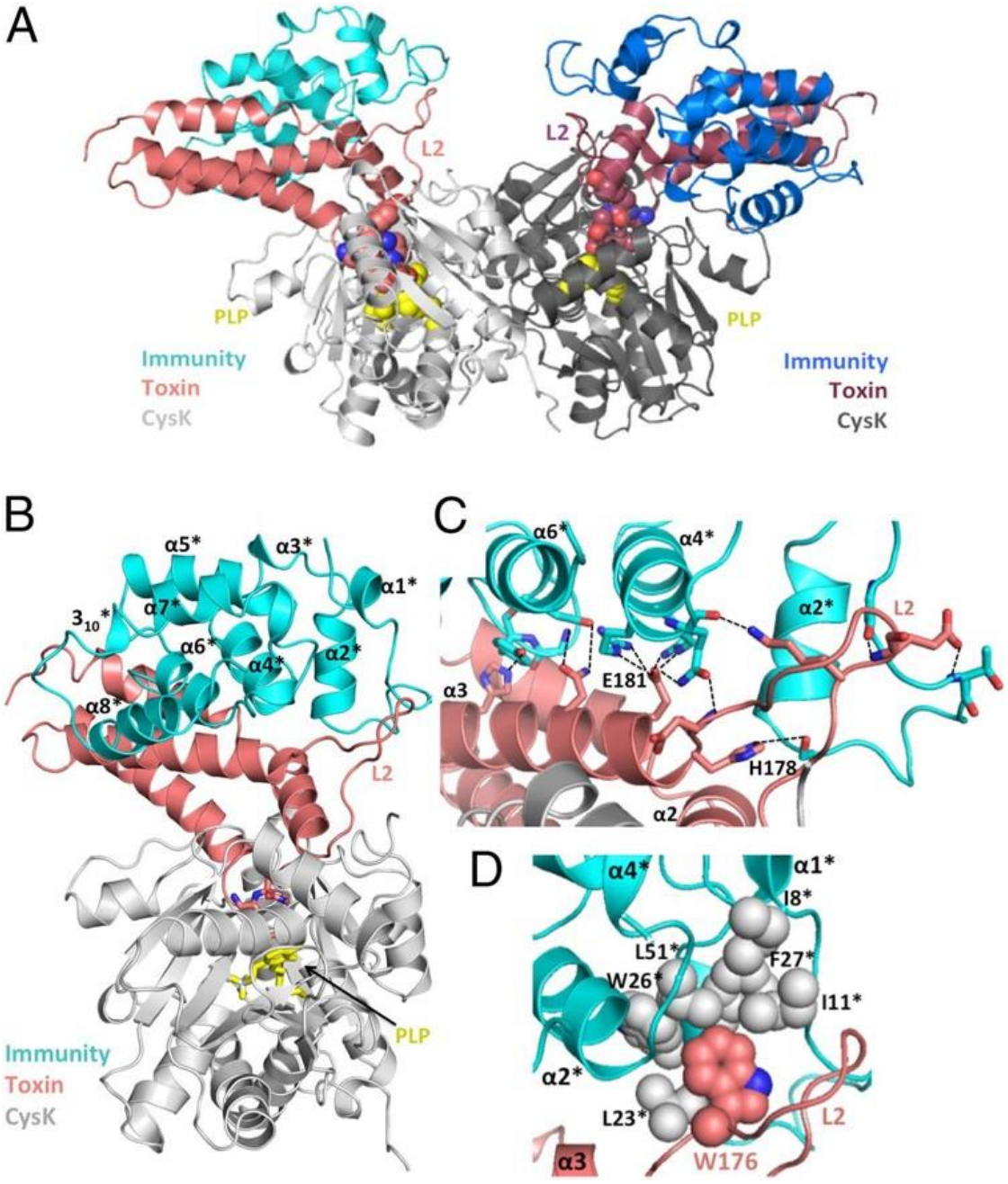


Figure 3.4: The CysK/CdiA-CT/CdiI^{EC536} ternary complex. (**A**) Crystal structure of the CysK/CdiA-CT/CdiI^{EC536} complex. The C-terminal GYGI peptide of CdiA-CT^{EC536} and CysK-bound pyridoxyl 5´-phosphate (PLP) are rendered as spheres. The ternary complex is presented in the same orientation as Fig. 1A. (**B**) Monomeric version of the ternary complex. CdiI^{EC536} secondary structure elements are outlined and indicated with superscripted asterisks (*). The C-terminal GYGI peptide of the toxin and PLP are shown as sticks. (**C**) H-bonding network between CdiA-CT^{EC536} and CdiI^{EC536}. Interacting residues are shown in stick representation and dashed lines indicated H-bonds. Active-site residues His178 and Glu181 from L2 and α3 of CdiA-CT^{EC536} interact with α2*, α4* and α6* of CdiI^{EC536}. (**D**) Hydrophobic interactions between CdiA-CT^{EC536} and CdiI^{EC536}. CdiA-CT^{EC536} Trp176 binds into a hydrophobic pocket formed by CdiI^{EC536}.

Structure of the CysK/CdiA-CT/CdiI^{EC536} ternary complex

The CdiI^{EC536} immunity protein binds to CysK/CdiA-CT^{EC536}, forming a neutralized ternary complex (8). The crystal structure of this ternary complex shows that immunity protein binds to each toxin to form a dimer of heterotrimers (**Figure 3.4A**). The CysK and CdiA-CT^{EC536} structures are very similar in the binary and ternary complexes (rmsd of 0.25 Å over all α-carbons), indicating that the immunity protein does not grossly alter toxin conformation. CdiI^{EC536} is a single domain composed of one 3_{10}- and eight α-helices arranged in four stacked layers to form an anti-parallel spiral (**Figure 3.4B**). DALI (24) searches indicate that CdiI^{EC536} does not share structural homology with other known antitoxins or immunity proteins (**Table 3.4**).

Table 3.4. DALI server search results.

Search input	**Structural homolog**	**Sequence identity** (%)	**PDB ID**	***Z*-score**	**rmsd (Å)[a]**
CdiA-CT^{EC536}	Importin subunit alpha-1	7	4ZDU	6.5	4.1 (108/416)[b]
	Unc-45	11	3NOW	6.3	2.6 (85/768)
	Gastric intrinsic factor	13	2PMV	6.3	4.4 (92/267)
	Beta-catenin	7	4EV9	6.3	4.0 (108/492)
	Plakoglobin	11	3IFG	6.3	2.5 (79/546)
	Histone binding protein N1/N2	7	1PJN	6.2	3.2 (105/427)
CdiI^{EC536}	BAG family chaperone regulator 4	8	4HWH	8.3	3.2 (79/84)
	Syntaxin 6	9	1LVF	8.3	3.0 (86/104)
	Heat shock 70 kDa protein 1	8	3A8Y	7.5	2.8 (76/100)
	Spastin	6	3EAB	7.5	3.5 (83/88)
	Diaphanous homolog 1	5	3OBV	7.2	3.5 (81/419)
	Stata protein	7	1UUR	7.2	3.2 (86/461)

[a]rmsd is root-mean-square deviation.

[b]aligned over # α-carbons out of # residues.

The immunity protein interacts exclusively with the nuclease domain of CdiA-CT^{EC536} and makes no contacts with CysK (**Figure 3.4A**). Flexible loop L2 of the nuclease domain interacts extensively with CdiI^{EC536}, stabilizing the loop and allowing its side-chains to be fully resolved in the ternary complex (**Figure 3.4C**). Fifteen H-bonds and ion-pairs connect CdiA-CT^{EC536} helix α3 to helices α2*, α4* and α6* of CdiI^{EC536} (**Figure 3.2C** & **Table 3.2**). Additionally, toxin residue Trp176 from loop L2 fits into a hydrophobic pocket formed by Ile8, Ile11, Leu23, Trp26, Phe27 and Leu51 of CdiI^{EC536} (**Figure 3.4D**). The toxin/immunity protein interface is 1,034 Å^2, burying 16.6% and 12.7% of the solvent accessible surface areas of CdiA-CT^{EC536} and CdiI^{EC536}, respectively. Comparison of different Ntox28 domains and their predicted immunity proteins reveals that many of the interacting residues are conserved throughout the family.

Identification of the CdiA-CT^{EC536} nuclease active site

DALI searches indicate that the CdiA-CT^{EC536} nuclease domain does not share structural homology with other known RNases (**Table 3.4**). The domain also lacks an obvious tRNA-binding pocket, but previous work suggests that His178 is an active-site residue (**Figure 3.5A** & 3.**5B**) (8). Therefore, we used site-directed mutagenesis to probe residues in the vicinity of His178 for roles in growth inhibition. CdiA-CT^{EC536} expression plasmids were introduced into *E. coli cysK*$^+$ and Δ*cysK* strains, and transformants were selected on media supplemented with either glucose to suppress, or arabinose to induce, toxin expression. The wild-type construct was lethal when introduced into *cysK*$^+$ cells under any condition, but had no effect on Δ*cysK* cell growth even when induced with arabinose (**Figure 3.5C**). This result illustrates *cysK*-dependent toxicity for comparison with mutated CdiA-CT^{EC536} variants. CdiA-CT^{EC536} residues Asp155 and Glu181

cluster near His178 (**Figure 3.5B**) and are completely conserved in known Ntox28 domains (**Figure 3.5A**), suggesting functional importance. Asp155Ala and Glu181Ala mutations abolished growth inhibition activity, showing the same phenotype as the His178Ala mutation (**Figure 3.5C**).

Thr185 is also positioned near His178, but mutation of this residue had no discernable effect on toxicity (**Figure 3.5C**). Asn149, Lys152 and Arg187 were tested for potential contributions to tRNA binding, but mutations at these positions also had no effect on growth inhibition (**Figure 3.5C**). Lastly, we examined loop L2 residue Trp176, which is located near the junction with helix □3 near the putative catalytic triad of Asp155, His178 and Glu181 (**Figure 3.5B**). Aromatic residues are conserved at this position in Ntox28 domains (**Figure 3.5A**), suggesting that Trp176 may contribute to tRNA substrate recognition. The Trp176Ala expression plasmid was maintained stably in *cysK*$^{+}$ cells under repressive conditions, but cell growth was inhibited upon induction with arabinose (**Figure 3.5C**). *In vitro* nuclease assays with purified toxins showed that the Asp155Ala, His178Ala and Glu181Ala mutations each blocked activity (**Figure 3.5D**). Additionally, we found that CdiA-CT^{EC536} anticodon nuclease activity requires divalent cations, suggesting that Asp155 and Glu181 may contribute to catalysis by coordinating Mg^{2+}. Further, the Thr185Ile mutation had a minor effect, but the activity of the Trp176Ala variant was significantly attenuated (**Figure 3.5D**). To exclude the possibility that the mutations prevent toxin interaction with CysK, we confirmed that each CdiA-CT^{EC536} variant interacts with

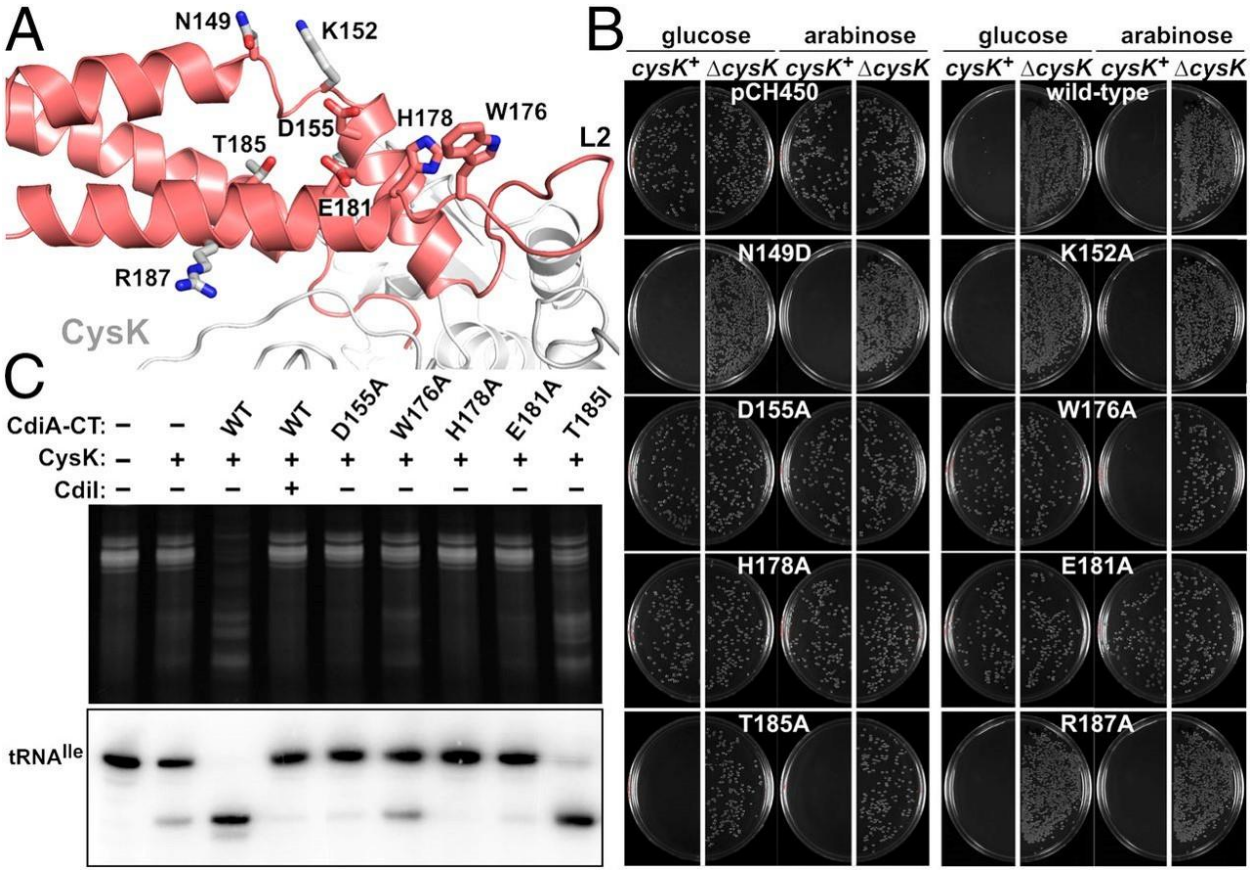


Figure 3.5 (Courtesy of Hayes lab): Identification of nuclease active-site residues. (**A**) Mutated residues of CdiA-CT^{EC536}. (**B**) Growth inhibition activity of CdiA-CT^{EC536} variants. Arabinose-inducible expression plasmids were introduced into *E. coli cysK*$^{+}$ and Δ*cysK* cells and transformants selected on media supplemented with glucose or arabinose. Plasmid pCH450 is the empty vector. (**C**) In vitro nuclease activity of CdiA-CT^{EC536} variants. Purified CdiA-CT^{EC536} proteins were incubated with total cellular RNA in the presence of CysK and CdiI^{EC536} where indicated. Reactions were run on denaturing polyacrylamide gels and visualized by ethidium bromide staining (top panel) and Northern blot hybridization (bottom panel).

CysK-His_6 using Ni^{2+}-affinity co-purification. Together, these results suggest that Asp155, His178 and Glu181 form a catalytic triad required for tRNase activity.

The Tox28^{Rlac} domain from R. lactaris is a CysK-independent nuclease

Several CdiA proteins carry Ntox28 domains, and these toxins are also found at the C-terminus of YD-repeat and WXG100 proteins from Gram-positive bacteria (4). Gram-positive Ntox28 domains have insertions in loop L3 and lack C-terminal G(Y/H)GI sequences (**Figure 3.5A**), suggesting they do not interact with CysK. We tested this prediction using the Ntox28/immunity protein pair encoded by the RUMLAC_00243/00244 loci of *R. lactaris* ATCC 29176. We first confirmed that Tox28^{Rlac} inhibits cell growth using controlled proteolysis to degrade ssrA(DAS)-tagged Imm^{Rlac} immunity protein, thereby liberating the toxin domain inside *E. coli* cells. Tox28^{Rlac} inhibited the growth of both *cysK*$^{+}$ and Δ*cysK* cells, whereas expression of CdiA-CT^{EC536} using the same approach had no inhibitory effect on Δ*cysK* cells (**Figure 3.6A**). We also found that CysK does not bind to Tox28^{Rlac} with high-affinity, though the toxin forms a stable complex with the Imm^{Rlac} immunity protein (**Figure 3.6B**). *In vitro* nuclease assays revealed that purified Tox28^{Rlac} cleaves tRNA, but in contrast to CdiA-CT^{EC536}, the addition of CysK failed to stimulate nuclease activity (**Figure 3.6C**). Together, these results demonstrate that Tox28^{Rlac} shares tRNase activity with CdiA-CT^{EC536}, but does not require activation by CysK or other proteins.

CysK stabilizes CdiA-CT^{EC536} and promotes tRNA substrate binding

The autonomy of the Tox28^{Rlac} domain raises the question of why CdiA-CT^{EC536} requires activation. Because the CdiA-CT^{EC536} nuclease domain is small and lacks an extensive

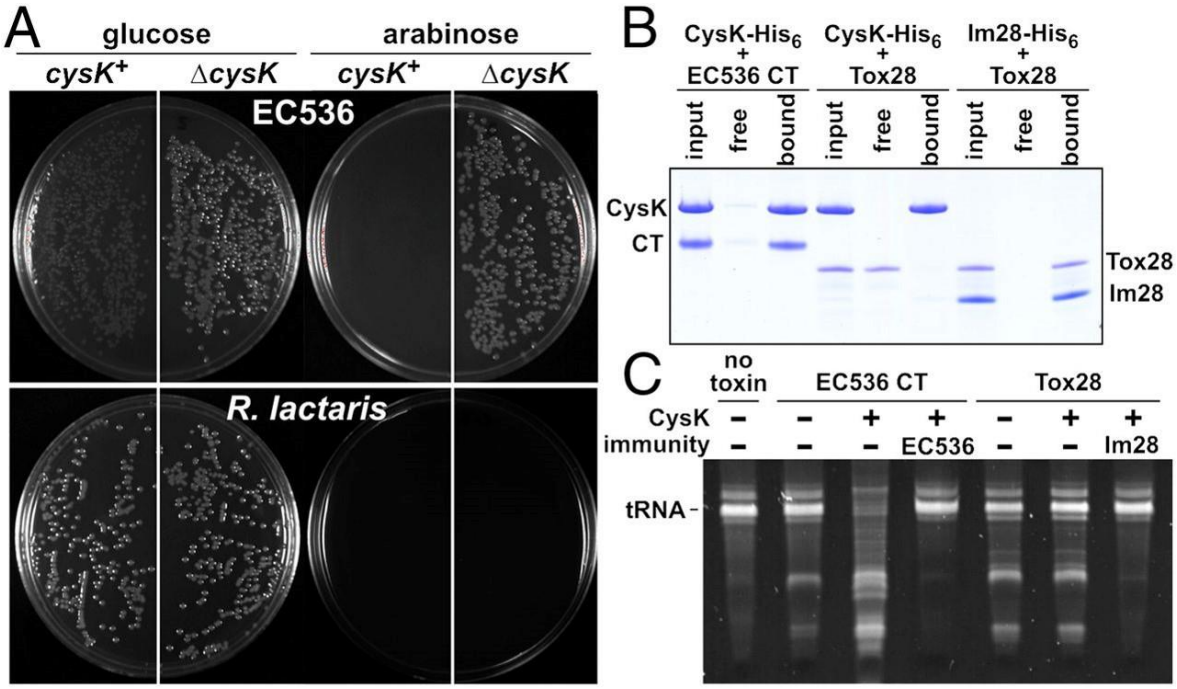


Figure 3.6 (Courtesy of Hayes lab): *R. lactaris* Tox28 is a CysK-independent tRNase. (**A**) CdiA-CT/CdiI^{EC536}-DAS and Tox/Imm^{Rlac}-DAS expression constructs were introduced into *E. coli cysK*$^{+}$ and Δ*cysK* cells and transformants selected on media supplemented with glucose or arabinose. (**B**) Protein binding assays. The indicated proteins were mixed and subjected to Ni^{2+}-affinity chromatography. The free lanes represent proteins that failed to bind the affinity matrix and the bound lanes show proteins eluted with imidazole. (**C**) In vitro nuclease assays. Purified CdiA-CT^{EC536} or Tox28^{Rlac} was incubated with total cellular RNA. Reactions were supplemented with purified CysK, CdiI^{EC536} or Imm^{Rlac} where indicated.

hydrophobic core, we explored the possibility that CysK stabilizes the toxin fold. We first monitored the stability of each toxin to thermal denaturation using circular dichroism (CD) spectroscopy. These analyses revealed that the melting temperature (T_m) for Tox28^{Rlac} is approximately 10 °C higher than that of the CdiA-CT^{EC536} toxin (**Table 3.5** & **Figure 3.7A**). We then examined the thermostability of the CysK/CdiA-CT^{EC536} complex using both CD spectroscopy and differential scanning fluorescence (DSF). Both experimental approaches showed that the CysK/CdiA-CT^{EC536} complex has a T_m of ~66 °C, which is very similar to the value obtained for Tox28^{Rlac} (**Table 3.5** & **Figure 3.7A**).

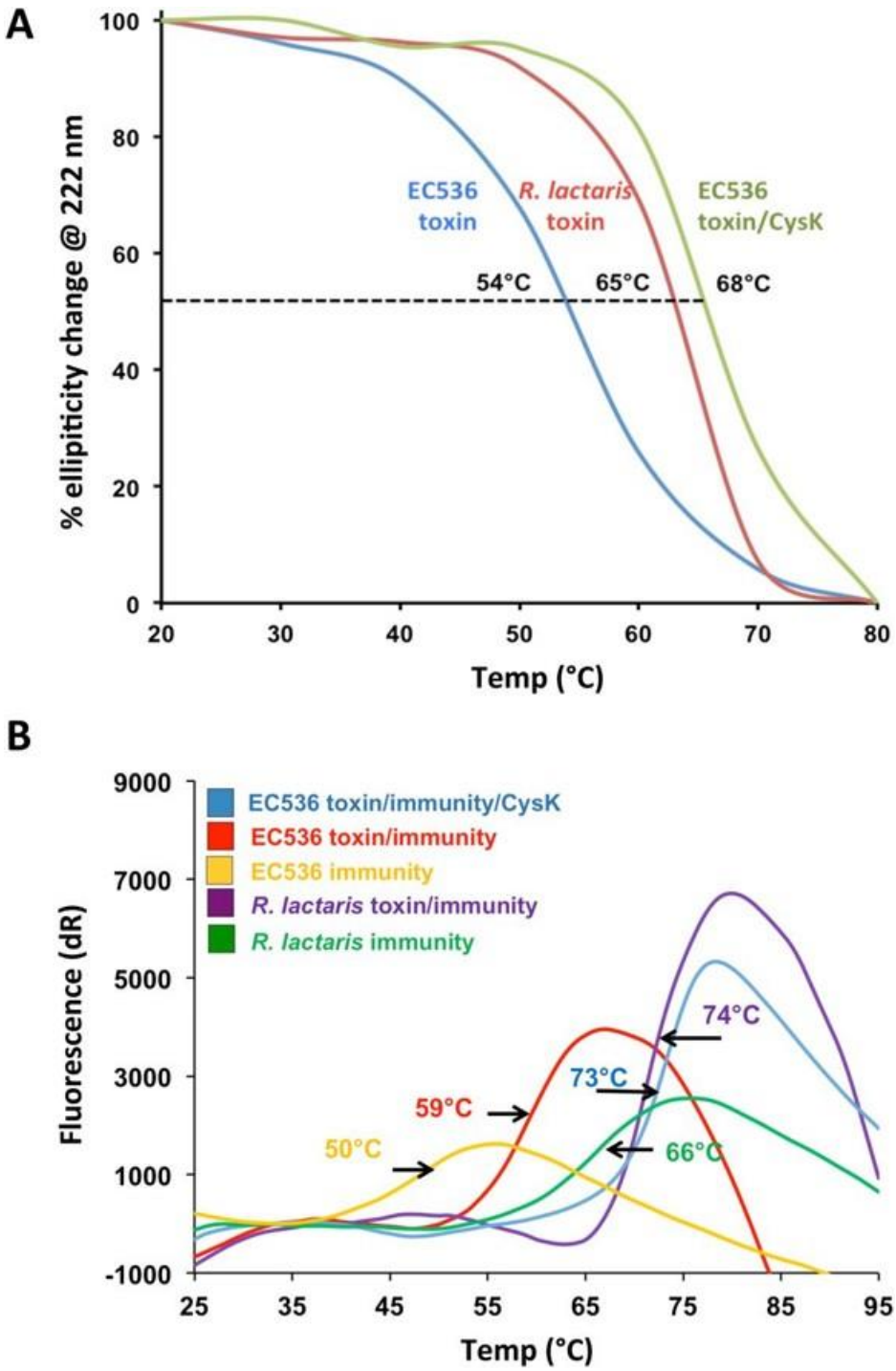


Figure 3.7: Thermal stability of Ntox28 toxins. (**A**) Thermal stability of toxin domains determined by circular dichroism (CD) spectroscopy. The fraction of folded protein was calculated and used to determine melting temperatures (T_m). (**B**) Differential scanning fluorimetry (DSF) of toxin/immunity protein complexes.

Table 3.5. Melting temperatures (T_m) of protein complexes

Protein/complex	**CD T_m (°C)**	**DSF T_m (°C)**
CdiA-CT^{EC536}	54 ± 1	ND[a]
Tox28^{Rlac}	65 ± 2	ND[a]
CdiI^{EC536}	ND[a]	50 ± 2
Imm^{Rlac}	ND[a]	66 ± 1
CysK	ND[a]	58 ± 1
CysK/CdiA-CT^{EC536}	68 ± 2	64 ± 2
CdiA-CT/CdiI^{EC536}	62 ± 2	58 ± 2
CysK/CdiA-CT/CdiI^{EC536}	ND[a]	73 ± 3
Tox28/Imm^{Rlac}	73 ± 1	74 ± 2

[a]Not determined

The T_m for isolated CysK is approximately 58 °C (**Table 3.5**), showing that the complex is more stable than the individual components. We also found that the CdiI^{EC536} and Imm^{Rlac} immunity proteins have a significant stabilizing effect, effectively increasing the T_m for each toxin (**Table 3.2** & **Figure 3.7B**). These data indicate that CdiA-CT^{EC536} is intrinsically less stable than Tox28^{Rlac}, but exhibits comparable thermostability when bound to CysK.

Finally, we tested whether CysK promotes the binding of tRNA substrates to CdiA-CT^{EC536}. We expressed His_6-epitope tagged CdiA-CT(H178A)^{EC536} and/or CysK in *E. coli*, then treated the cell lysates with formaldehyde to cross-link tRNA to the proteins. Analysis of complexes isolated from these reactions revealed that tRNA was preferentially cross-linked when both CysK and CdiA-CT^{EC536} were present in the lysates (**Figure 3.8A**). This interaction appears to reflect physiologically relevant substrate binding, because co-expression of CdiI^{EC536} reduced tRNA cross-linking to the CysK/CdiA-CT^{EC536} complex (**Figure 3.8B**). This latter result suggests that immunity protein blocks nuclease activity by occluding the tRNA-binding site. Because

Tox28^{Rlac} is an autonomous tRNase, we predicted that it should exhibit intrinsic tRNA-binding activity. We generated a His$_6$-tagged version of Tox28^{Rlac} that contains the His114Ala mutation to ablate nuclease activity and tested whether it binds substrate. Substantial amounts of tRNA co-purified with the inactive Tox28^{Rlac} domain, even without formaldehyde cross-linking (**Figure 3.8C**). The interaction between tRNA and Tox28^{Rlac} was effectively blocked by cognate Imm^{Rlac} immunity protein (**Figure 3.8C**), again suggesting that the immunity protein binds over the nuclease active site. Together, these results indicate that CysK stabilizes CdiA-CT^{EC536}, rendering the nuclease domain competent to bind substrate. By contrast, the Tox28^{Rlac} domain is intrinsically stable and binds tRNA without assistance from CysK.

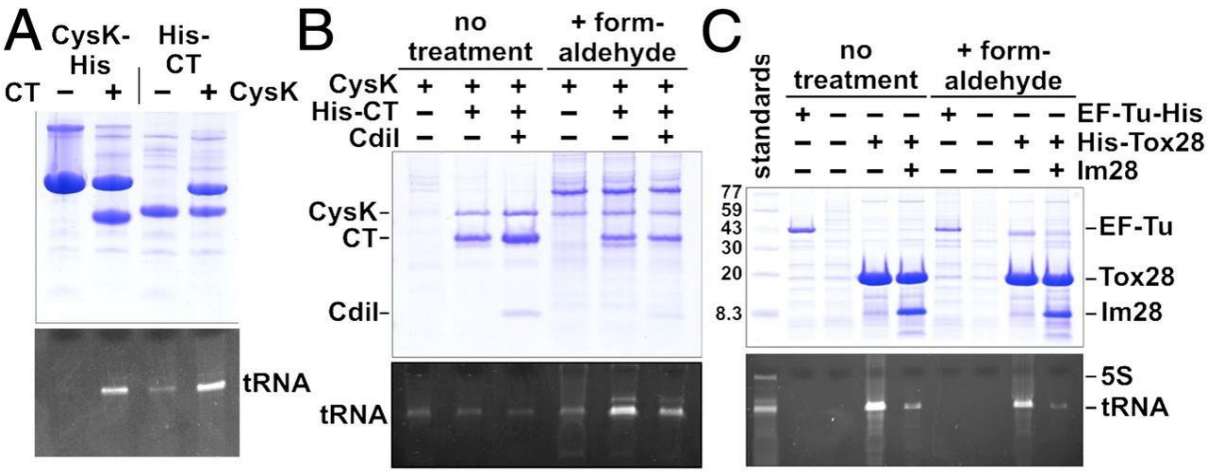


Figure 3.8 (Courtesy of Hayes lab): tRNA binding to Ntox28 nuclease domains. (**A**) Cross-linking of tRNA to CysK/CdiA-CT^{EC536} complexes. *E. coli* cell lysates containing CysK and/or CdiA-CT(H178A)^{EC536} were treated with formaldehyde and cross-linked nucleoprotein complexes purified by Ni^{2+}-affinity chromatography. Purified samples were analyzed by SDS-PAGE (top panel) and 50% urea PAGE (bottom panel) to visualize proteins and nucleic acid, respectively. (**B**) CdiI^{EC536} blocks tRNA cross-linking to CysK/CdiA-CT^{EC536}. *E. coli* cell lysates containing CysK-His$_6$, CysK-His$_6$/CdiA-CT(H178A)^{EC536} and CysK-His$_6$/CdiA-CT(H178A)/CdiI^{EC536} were treated with formaldehyde where indicated, followed by Ni^{2+}-affinity chromatography, and gel analysis as described in panel A. (**C**) Tox28^{Rlac} interacts stably with tRNA substrate. Cell lysates containing His$_6$-Tox28(H114A)^{Rlac} and Imm^{Rlac} were treated as described above. To ascertain cross-linking specificity, EF-Tu-His$_6$ was also purified and analyzed for tRNA binding.

Discussion

These results provide several insights into CdiA-CT^{EC536} toxin activation. As predicted from prior work (8), the crystal structures show that the toxin inserts its C-terminal GYGI tail into the CysK active site and anchors the complex. The toxin's C-terminal carboxylate forms important H-bond contacts with conserved CysK active-site residues in the substrate-binding loop, and these same interactions are observed in the structures of CysE peptides bound to CysK from *Haemophilus influenzae* and *Arabidopsis thaliana* (21, 22). The C-terminal Ile residues of CdiA-CT^{EC536} and CysE exploit the same hydrophobic and H-bond contacts as *O*-acetylserine to bind the CysK active site (19). Though C-terminal Ile residues are critical for binding, substitutions are tolerated at other positions within the peptide tail. Salsi *et al.* have shown that CysK binds CysE peptides altered at the penultimate and antepenultimate positions (13). This plasticity accommodates the natural variation in CysE tail sequences and accounts for the ability of *E. coli* CysK to bind both GDGI and GYGI motifs with high affinity. Structures are not available for the full cysteine synthase complex; thus it is unclear whether CdiA-CT^{EC536} mimics other features of CysE. However, we note that the CysE structure differs markedly from CdiA-CT^{EC536}. The CysE C-terminal domain forms a β-helix that terminates in a flexible tail (25). Further, the β-helical domain mediates trimerization, and two CysE trimers interact to form a larger homohexameric complex (25). If the β-helical domains of CysE interact with CysK, then the contacts are likely to be distinct from those observed with the CdiA-CT^{EC536} α-helical bundle. The same uncertainties apply to other "moonlighting" partners of CysK for which no structural information is available (15). It is intriguing that CysK has been repeatedly recruited as a binding partner by disparate proteins. Perhaps this phenomenon reflects the antiquity and immutability of cysteine synthase complexes, which remain remarkably similar in extant bacteria and plants. From the perspective

of CDI, the ubiquity and conserved active-site architecture of CysK ensures that toxins can be activated in a broad range of bacteria.

The structure of the CdiA-CT^{EC536} nuclease domain is novel, and its catalytic mechanism appears to be distinct from other anticodon nucleases. Colicins E5 and D are the only other anticodon nucleases for which structures are available (26). The nuclease domains of colicins E5 and D share an α/β-fold that characterizes the BECR (Barnase-EndoU-ColicinE5/D-RelE) family of RNases (4). Like barnase, colicins E5 and D are metal-independent nucleases that abstract a proton from the 2´-hydroxyl to initiate an intra-molecular attack on the scissile phosphodiester bond (26). By contrast, CdiA-CT^{EC536} has divalent cation-dependent nuclease activity, which is usually associated with a hydrolytic mechanism. Mutational analyses support a role for CdiA-CT^{EC536} residues Asp155, Glu181 and His178 in catalysis. Asp155 and Glu181 are candidates to coordinate Mg^{2+}, which could either activate water for hydrolysis or stabilize hydroxide ions generated by His178 acting as a general base (27). Another difference between colicins E5/D and CdiA-CT^{EC536} is the lack of a defined substrate-binding pocket in the latter nuclease. Several observations suggest that flexible loop L2 participates in tRNA binding. L2 is conserved amongst Ntox28 family members and always contains hydrophobic residues adjacent to His178. Our results show that Trp176 is important, but not strictly required, for CdiA-CT^{EC536} activity. We hypothesize that Trp176 stacks onto nucleobases within the tRNA anticodon loop. This mode of recognition is common amongst aminoacyl-tRNA synthetases, which bind cognate tRNAs using conserved hydrophobic/aromatic residues that stack onto the first and second nucleotides of the anticodon (28).

Finally, we propose that CysK promotes CdiA-CT^{EC536} nuclease activity by stabilizing the toxin's fold. The CdiA-CT^{EC536} helical bundle is relatively small and lacks an extensive hydrophobic core. Consequently, CdiA-CT^{EC536} has relatively low thermostability, raising that possibility that thermal fluctuations disrupt the active site by splaying the helices. CysK anchors toxin helices □2 and □3, thereby approximating Asp155 and Glu181 to coordinate Mg^{2+}. The CysK scaffold also anchors the ends of loop L2, which we propose is important for substrate binding. In support of this model, we find that tRNA interacts with the CysK/CdiA-CT^{EC536} complex, but not with the individual components. CdiI^{EC536} blocks the interaction with substrate, strongly suggesting that the immunity protein occludes the nuclease active site. The extensive contact between loop L2 and CdiI^{EC536}, together with the sequestration of Trp176 within the immunity protein are consistent with this model. Though CysK is critical for CdiA-CT^{EC536} nuclease activity, related toxins from Gram-positive bacteria probably do not require extrinsic activation because the Tox28 domain from *R. lactaris* has tRNase activity *in vitro*. Ntox28 domains are typically found at the C-terminus of proteins that mediate inter-bacterial competition (4, 29-31). For example, the *R. lactaris* Tox28^{Rlac} domain is part of a larger rearrangement hotspot (Rhs) repeat protein. Rhs and related YD-repeat proteins deliver toxic nuclease domains into both Gram-negative and Gram-positive bacteria (32, 33). Further, the Ntox28 homolog from *Geobacillus* sp. strain Y412MC10 (GYMC10_1092) is linked to an N-terminal ESAT-6-like domain, which is predicted to guide export through type VII secretion systems (4, 29). Given that all Ntox28 domains function in inter-cellular competition, perhaps the mechanism of toxin delivery into Gram-negative bacteria accounts for the relative instability of CdiA-CT^{EC536}. We recently discovered that CDI toxins hijack a variety of inner-membrane proteins to enter the target-cell cytoplasm (9). If CDI toxins must unfold during this translocation step, then there may be a

selective pressure for toxins with low global stability.

Materials and Methods

Plasmid constructs

Plasmids used in this study are presented in **Table 3.6**. T7 RNA polymerase based over-expression plasmids to produce CdiA-CT/CdiI^{EC536}-His$_6$ (pCH6190) and CysK-His$_6$ (pCH8215) have been described (3, 8). The *E. coli cysK* gene was amplified with primers CysK-S2G-Nco (5´ - GGC CAT GGG TAA GAT TTT TGA AGA TAA CTC) and CysK-Xho-rev (5´ - AAA CTC GAG TTA ATA ACG CTC ACC CGA TGA). The product was digested with NcoI/XhoI and ligated to pET21P and pCH450 to generate plasmids pCH11027 and pCH11014, respectively. These latter constructs produce untagged CysK. The *cdiA-CT(H178A)*^{EC536} coding sequence was amplified from pCH6938 (8) with primers 536-Nco-H6-Spe (5´ - TTC CAT GGC AAA AAG TCA TCA TCA TCA TCA CCA CAC TAG TGT TGA GAA TAA TGC GCT GAG) and 536-CT-Xho-rev (5´ - TTA CTC GAG GTA ATC ATA TTC CAT A). The product was digested with NcoI/XhoI and ligated to pET21 and pCH450 to generate His$_6$-CdiA-CT(H178A)^{EC536} over-production plasmids pCH11001 and pCH12817, respectively. Plasmid pCH6938 was also amplified with primers 536-Nco-H6-Spe and 536-cdiI-Xho-rev (5´ - GGC CTC GAG TAG TTA TAC AAT TAT CTG), and the product ligated to pET21 to generate a construct that over-produces His$_6$-CdiA-CT(H178A) and CdiI^{EC536} (pCH12515). The *cdiA-CT/cdiI*^{EC536}*(DAS)* module was excised from plasmid pCH6316 (30) with NcoI/PstI and ligated to pCH450 to generate plasmid pCH7171. His$_6$-CdiA-CT(H178)^{EC536} and His$_6$-CysK proteins with thrombin-cleavable His$_6$ epitope tags were used for biolayer interferometry studies. The *cysK* gene from plasmid pCH11014 was amplified with primers CysK-Nde-for (5´ - GAT CAT ATG ATG GGT AAG ATT TTT GAA

GAT AAC TCG CTG ACT AT) and CysK-EcoR-rev (5´ - GAT GAA TTC TCA CTC GAG ACT AGT CTG TTG CAA TTC TTT); and the *cdiA-CT(H178)*^{EC536} gene from plasmid pCH11001 was amplified with primers 536-Nde-for (5´ - GAT CAT ATG ATG ACT AGT GTT GAG AAT AAT GCG CTG AGT) and 536-EcoR-rev (5´ - GAT GAA TTC TCA TAT TCC ATA TCC TTT CAA GGC TGA TTC TAT TTT ATT AAT A). The resulting products were digested with NdeI/EcoRI and ligated to plasmid pET28a.

Table 3.6. Bacterial strains and plasmids.

Strain or plasmid	**Description**[a]	**Reference**
Strains		
BL21-Gold(DE3)	*E. coli* B, F$^-$ *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ *gal* □(DE3) *endA* Hte, Tet^R	Stratagene
X90	F´ *lacI*^{q} *lac´ pro´/ara* Δ*(lac-pro) nal1 argE*(amb) *rif*^{r} *thi-1*, Rif^R	(18)
CH1944	X90 (DE3) Δ*rna*, Rif^R	(15)
CH2016	X90 (DE3) Δ*rna* Δ*slyD::kan*, Rif^R Kan^R	(15)
CH7157	X90 Δ*clpX* Δ*clpA::kan*, Rif^R Kan^R	(19)
CH8164	X90 Δ*cysK::kan*, Rif^R Kan$^{R(}$	(2)
CH8602	X90 Δ*cysK*, Rif^R	(2)
CH8804	X90 (DE3) Δ*rna* Δ*slyD* Δ*cysK::kan*, Rif^R Kan^R	This study
CH9648	X90 (DE3) Δ*rna* Δ*slyD* Δ*tufB tufA-his*$_6$*-kan*, Rif^R Kan^R	This study
CH10013	Spontaneous rifampicin-resistant derivative of *E. coli* JCM158, Rif^R	(20)
CH10801	CH10013 Δ*cysK*, Rif^R	This study
CH11002	CH8804 pCH11014 pCH13129, Rif^R Kan^R Tet^R Amp^R	This study
CH12810	CH8804 pCH8649 pET21, Rif^R Kan^R Tet^R Amp^R	This study
CH12811	CH8804 pCH8649 pCH8501, Rif^R Kan^R Tet^R Amp^R	This study
CH12812	CH8804 pCH12817 pET21, Rif^R Kan^R Tet^R Amp^R	This study
CH12813	CH8804 pCH12817 pCH11027, Rif^R Kan^R Tet^R Amp^R	This study
CH13101	CH8804 pCH11014 pET21, Rif^R Kan^R Tet^R Amp^R	This study
CH13102	CH8804 pCH11014 pCH12515, Rif^R Kan^R Tet^R Amp^R	This study

Plasmids		
pTrc99a	IPTG-inducible expression plasmid, Amp^R	GE Healthcare
pCH450	pACYC184 derivative containing *E. coli araC* and the L-arabinose-inducible P_{ara} promoter, Tet^R	(5)
pCP20	Heat-inducible expression of FLP recombinase, Cm^R Amp^R	(21)
pET21	T7 RNA polymerase-based over-expression vector, Amp^R	(1)
pCH6190	pET21::*cdiA-CT/cdiI*^{EC536}, over-produces CdiA-CT^{EC536} and CdiI^{EC536}-His_6, Amp^R	(1)
pCH6316	pTrc99A::*cdiA-CT/cdiI*^{EC536}*(DAS)*, Amp^R	(3)
pCH6938	pET21::*cdiA-CT(H178A)/cdiI*^{EC536}, over-produces CdiA-CT(H178A)^{EC536} and CdiI^{EC536}-His_6, Amp^R	(2)
pCH7086	pCH450::*cdiA-CT(H178A)*^{EC536}, Tet^R	(22)
pCH7171	pCH450::*cdiA-CT/cdiI*^{EC536}*(DAS)*, Tet^R	This study
pCH8215	pET21::*cysK(S2G)*, over-produces CysK-His_6, Amp^R	(2)
pCH8501	pET21::*cdiA-CT(H178A)*^{EC536}, expresses inactive CdiA-CT^{EC536} without epitope tag, Amp^R	This study
pCH8649	pCH450::*cysK(S2G)-his*$_6$, Tet^R	(2)
pCH9320	pCH450::*cdiA-CT*^{EC536}, Tet^R	(22)
pCH9622	pKAN-*tufA(his*$_6$*)*, Amp^R Kan^R	This study
pCH10978	pET21::*cdiA-CT(D155A)/cdiI*^{EC536}, Amp^R	This study
pCH10980	pET21::*cdiA-CT(W176A)/cdiI*^{EC536}, Amp^R	This study
pCH10982	pET21::*cdiA-CT(E181A)/cdiI*^{EC536}, Amp^R	This study
pCH11022	pET21::*cdiA-CT(T185I)/cdiI*^{EC536}, over-produces CdiA-CT(T185I) and CdiI^{EC536}-His_6, Amp^R	This study
pCH11014	pCH450::*cysK(S2G)*, expresses untagged CysK, Tet^R	This study
pCH11016	pCH450::*cdiA-CT(D155A)*^{EC536}, Tet^R	This study
pCH11017	pCH450::*cdiA-CT(W176A)*^{EC536}, Tet^R	This study
pCH11019	pCH450::*cdiA-CT(E181A)*^{EC536}, Tet^R	This study
pCH11020	pCH450::*cdiA-CT(T185A)*^{EC536}, Tet^R	This study
pCH11022	pET21P::*cdiA-CT(T185A)/cdiI*^{EC536}, Amp^R	This study

pCH11027	pET21::*cysK(S2G)*, over-produces untagged CysK, Amp^R	This study
pCH12305	pET21::*tox/imm*^{Rlac}, over-produces Ntox28^{Rlac} and Imm^{Rlac}-His_6, Amp^R	This study
pCH12371	pCH450::*cdiA-CT(N149D)*^{EC536}, Tet^R	This study
pCH12372	pCH450::*cdiA-CT(K152A)*^{EC536}, Tet^R	This study
pCH12373	pCH450::*cdiA-CT(R187A)*^{EC536}, Tet^R	This study
pCH12334	pCH450::*tox/imm*^{Rlac}*(DAS)*, Tet^R	This study
pCH12515	pET21::*his$_6$-cdiA-CT(H178A)/cdiI*^{EC536}; over-produces His_6-CdiA-CT(H178A)^{EC536} and CdiI^{EC536}, Amp^R	This study
pCH12817	pCH450::*his$_6$-cdiA-CT(H178A)*^{EC536}, Tet^R	This study
pCH13103	pET21::*his$_6$-tox/imm*^{Rlac}, over-produces His_6-Ntox28^{Rlac} and Imm^{Rlac}, Amp^R	This study
pCH13129	pET21::*his$_6$-cdiA-CT(H178A)*^{EC536}, over-produces His_6-CdiA-CT^{EC536}, Amp^R	This study
pCH13146	pET21::*his$_6$-tox(H114A)/imm*^{Rlac}, over-produces His_6-Ntox28(H114A)^{Rlac} and Imm^{Rlac}, Amp^R	This study
pCH13153	pET21::*his$_6$-tox(H114A)*^{Rlac}, over-produces His_6-Ntox28(H114A)^{Rlac}, Amp^R	This study

[a]Abbreviations: ampicillin-resistant, Amp^R; chloramphenicol-resistant, Cm^R; kanamycin-resistant, Kan^R; rifampicin-resistance, Rif^R; tetracycline-resistant, Tet^R.

Missense mutations were introduced into the *cdiA-CT*^{EC536} coding sequence using megaprimer PCR (34). Megaprimers were amplified from template plasmid pCH6190 using 536-CT-Nco-for (5´ - AGA CCA TGG TTG AGA ATA ATG CGC TGA G) in conjunction with mutagenic reverse primers 536-CT(D155A)-rev (5´ - GAG AGT TCC AAT AAT AGC ATG ATC TTT CAG), 536-CT(W176A)-rev (5´ - CCT GCA TAT GAT CCG CAT ATC CTC CAT TC), 536-CT(E181A)-rev (5´ - GCG TAT TTT GCA TTG CCT GCA TAT GAT CC), 536-CT(T185A)-rev (5´ - CTT AAT CCT CTG AGC GCA TTT TGC ATT TCC TGC) and 536-CT(T185I)-rev (5' - TTC TTA ATC CTC TGA GGA TAT TTT GCA TTT CCT GC). Megaprimers were then used in second reactions with 536-cdiI-Xho-rev (5´ - GAT CTC GAG TAC AAT TAT CTG ATT GAT TTT T) to generate *cdiA-CT/cdiI*^{EC536} fragments for ligation to NcoI/XhoI digested pET21. The resulting constructs were amplified with primers 536-CT-Nco-for and 536-CT-Xho-rev and the products ligated to pCH450 (35) to generate arabinose-inducible

CdiA-CT^{EC536} expression plasmids. Additional *cdiA-CT* mutations were made by megaprimer PCR using 536-CT-Xho-rev in conjunction with primers 536-N149D-for (5´ - GAT AAC ACT ATA AAA GAT GCT CTG AAA GAT C), 536-K152A-for (5´ - CTA TAA AAA ATG CTC TGG CAG ATC ATG ATA TTA T), and 536-R187A-for (5´ - ATG CAA AAT ACG CTC GCA GGA TTA AGA AAT C). The resulting products were used with primer 536-CT-Nco-for to generate mutated *cdiA-CT*^{EC536} sequences for ligation into pCH450. All arabinose-inducible *cdiA-CT*^{EC536} expression constructs were initially cloned into *E. coli* X90 Δ*cysK* cells, which are resistant to CdiA-CT^{EC536} toxin activity (8).

The coding sequence for the *Ruminococcus lactaris* ATCC 29176 Ntox28 domain and immunity protein pair was custom synthesized by Integrated DNA Technologies (Coralville, IA). The synthetic DNA was digested with NcoI/SpeI and ligated to pET21P and pCH7171 to generate plasmids pCH12305 and pCH12334, respectively. pCH12305 over-produces the Ntox28/Imm^{Rlac}-His$_6$ complex, and pCH12334 expresses Ntox28^{Rlac} and ssrA(DAS)-tagged Imm^{Rlac} for *in vivo* toxicity studies. Plasmid pCH12305 was amplified with primers Rlac-his-tox-Nco-for (5´ - TTT CCA TGG CAA AAA GTC ATC ATC ATC ATC ACC ACG AAA TAG CAA GCG TTG GTT CAT CC) and Rlac-imm-Xho-rev (5´ - TTT CTC GAG ATC ACA TTA TTT TTT TGG ATA AAG TAT CTA TC), and the product ligated to pET21 to generate plasmid pCH13103, which over-produces His$_6$-Ntox28^{Rlac} and Imm^{Rlac}. The His114Ala mutation was then introduced into the His6-Ntox28^{Rlac} coding sequence using Rlac-H114A-for (5´ - GTG GAT ATT TTG ATG CTT TAG GAG AAA TGC) and Rlac-imm-Xho-rev to generate a megaprimer, which was used with primer Rlac-his-tox-Nco-for to amplify the full toxin/immunity sequence. The final product was ligated NcoI/XhoI digested pET21 to generate pCH13146. Plasmid pCH13146 was amplified with

primers Rlac-his-tox-Nco-for and Rlac-tox-Xho-rev (5´ - TTT CTC GAG TCT TGC ATC AAT TTA TAC C) and the product ligated to NcoI/XhoI digested pET21 to generate the His_6-Ntox28(H114A)^{Rlac} over-expression construct pCH13153.

E. coli strain CH9648 was generated by Red-mediated recombination to introduce the coding sequence for a C-terminal His_6 epitope to the *tufA* gene encoding EF-Tu. Homology regions were PCR amplified using primer pairs tufA-Sac (5´ - TCA GAG CTC TAC CTG TAC TGG CGT TG)/tufA(His_6)-Bam (5´ - TTT GGA TCC TTA ATG ATG GTG ATG ATG GTG GCC CAG AAC TTT AGC), and tufA-Eco (5´ - GGG CGA ATT CCA CGT TAA TTA GTT TTG)/chiA-Hind (5´ - CAT CAT AAG CTT TCG CTT TTC CCG). The resulting products were sequentially ligated into plasmid pKAN (35) using SacI/BamHI and EcoRI/HindIII restriction sites to produce plasmid pCH9622. Plasmid pCH9622 was digested with SacI and HindIII and the small fragment electroporated into Red recombinase expressing cells as described (36). Recombinants were selected on LB-agar supplemented with 50 µg/mL kanamycin.

Protein expression and purification

CysK, His_6-CdiA-CT(H178A)^{EC536} and the CdiA-CT/CdiI^{EC536}-His_6 complex were over-produced individually in *E. coli* BL21 (DE3) cells that carry pCH11027, pCH11001, pCH6190 (respectively). Cells were grown aerobically at 37 °C in LB medium supplemented with 50 µg/mL ampicillin. Once the culture reached an optical density at 600 nm (OD_{600}) of ~0.8, protein production was induced with isopropyl-β-D-thiogalactosidase (IPTG) at 1 mM final concentration. Cultures were further incubated for 4 h, then cells were collected by centrifugation and washed once with binding buffer [20 mM sodium phosphate (pH 7.0), 150 mM NaCl]. Cell pellets were

resuspended in binding buffer and combined for the preparation of CysK/His_6-CdiA-CT(H178A)^{EC536} binary and CysK/CdiA-CT/CdiI^{EC536}-His_6 ternary complexes. Mixed cell suspensions were supplemented with 10 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride (PMSF), then the cells were disrupted by sonication on ice. Unbroken cells and debris were removed by centrifugation at 18,000 5*g* for 30 min and the supernatant passed through a 0.45 μm filter. Clarified lysates were applied to a Ni^{2+}-charged HiTrap column and washed with binding buffer supplemented with 10 mM imidazole. Complexes were eluted with a linear gradient of imidazole (10 – 250 mM) in binding buffer. Fractions were collected, combined and concentrated to ~0.5 mL using a 10 kDa centrifugal concentrator. Concentrated samples were then passed over a Superdex 200 gel filtration column equilibrated with binding buffer. Protein complexes were concentrated to 20 mg/mL for crystallization trials. Selenomethionine (SeMet) labeled proteins were produced in *E. coli* BL21 (DE3) Gold cells grown in M9 minimal medium supplemented with 50 mg/L leucine, isoleucine and valine; 100 mg/L phenylalanine, lysine, and threonine; and 75 mg/L SeMet as described (37). The Ntox28/Imm^{Rlac}-His_6 complex was overproduced as described for the CdiA-CT/CdiI^{EC536}-His_6 complex, but cells were induced with IPTG for 2 h. Because Imm^{Rlac} was produced in excess of the Ntox28^{Rlac} toxin, concentrated samples were passed over a Superdex 200 gel filtration column equilibrated in binding buffer to separate Ntox28/Imm^{Rlac}-His_6 complexes from free Imm^{Rlac} immunity protein.

Crystallization and structure determination

CysK/His_6-CdiA-CT(H178A)^{EC536} crystals were grown for three weeks by hanging drop vapor-diffusion over a reservoir of 0.2 M $NaSO_4$, 0.1 M Bis-Tris propane (pH 7.9) and 20% (wt/vol) PEG 3350. Hanging drops were prepared from a 1:1 mixture of 20 mg/mL protein to

reservoir buffer. CysK/His_6-CdiA-CT(H178A)^{EC536} crystallized in space group $P4_1$ with two complexes per asymmetric unit and dimensions of 64.01 Å 5 64.01 Å 5 365.37 Å. Crystals were mounted under cryo-conditions with the addition of 40% (vol/vol) glycerol as a cryoprotectant, and a dataset was collected at 70 K with wavelength 1.0 Å on beamline 7-1 at Stanford Synchrotron Radiation Lightsource (SSRL). Images were indexed, integrated and reduced using iMOSFLM (38) resulting in a 99.4% complete dataset up to 2.70 Å resolution. Initial phases were determined by molecular replacement with autoMR in PHENIX (39) using *Salmonella* Typhimurium CysK (PDB ID: 1OAS) as the search model (18). The initial Autobuild model contained only two CysK molecules, and therefore the CdiA-CT^{EC536} molecules were built through iterative manual building in Coot, followed by subsequent Autobuild cycles (39, 40). Phenix.refine (39) was used to refine the final model, which contains residues Lys127 – Ile227 of CdiA-CT^{EC536} (numbered from Val1 of the VENN peptide motif), residues Gly2 – Ala314 for one chain and Gly2 – Ala315 for the other of CysK (where Asp314 and Leu315 are modeled as alanines) and 110 water molecules. Each CysK protomer contains pyridoxal 5´-phosphate (PLP) bound to Lys42 through Schiff-base linkage. CdiA-CT^{EC536} residues 168, 170 – 172 and 175 lack observable electron density and were modeled as alanine. The Ramachandran plot shows 97.08% and 2.92% of residues in the favorable allowed and allowed regions, respectively. Data collection and refinement statistics are presented in **Table 3.1**.

The CysK/CdiA-CT/CdiI^{EC536}-His_6 (where CdiA-CT/CdiI^{EC536}-His_6 is the SeMet derivative) ternary complex was crystallized by hanging drop-vapor diffusion over a reservoir containing 0.1 M sodium cacodylate (pH 7.1), 0.2 M ammonium sulfate and 17% (wt/vol) PEG-8000. Hanging drops were prepared as a 1:1 mixture of 20 mg/mL protein to reservoir buffer. The

complex crystallized in space group $C222_1$ with unit cell dimensions: 81.25 Å 5195.54 Å 5 175.06 Å and two complexes per asymmetric unit. Crystals were collected by flash-freezing after soaking in a cryoprotectant solution containing a 1:1 mix of 40% glycerol (vol/vol) in crystallization buffer. A native dataset was collected at 70K at 0.9795 Å on beamline 7-1 at SSRL. Data processing was conducted using the HKL2000 suite (41), resulting in a 100% complete dataset up to 2.75 Å resolution. Molecular replacement was performed with Phaser in PHENIX (39) using the CysK/CdiA-CT(H178A)^{EC536} complex structure as the search model. The initial Phaser-generated model was subjected to Autobuild and phenix.refine (39). The CdiI^{EC536} structure and final model were built through several cycles of manual building in Coot and structure refinement in phenix.refine. The final model includes residues Leu132 – Ile227 of CdiA-CT^{EC536}, residues Gly2 – Ala314 of CysK (where Asp314 is modeled as Ala), and CdiI includes residues Ile2 – Ile127 of CdiI^{EC536} for one protomer and Ile2 – Asn124 for the other. Each CysK protomer contains PLP bound to Lys42, as observed for the binary complex, and the final model includes 101 water molecules. The Ramachandran plots shows 95.83% of dihedral angles in favorable regions, 3.9% in allowed regions and 0.27% (Asp111 of one CdiI protomer and Glu108 of both CdiI protomers) in the disallowed regions. Residues in the disallowed regions fit the electron density well. Data collection and refinement statistics are presented in **Table 3.1**. All molecular graphics were prepared using PyMol (42).

In vitro toxin-binding and RNase assays

Purified complexes of toxins bound to His_6-tagged immunity proteins were denatured in binding buffer supplemented with either 8 M urea or 6 M guanidine-HCl, and the toxins were isolated from the void volume during Ni^{2+}-affinity chromatography (3). Toxins were refolded by dialysis against binding buffer. CysK-His_6 and $CdiI^{EC536}$-His_6 were purified by Ni^{2+}-affinity chromatography under non-denaturing conditions as described (7). Imm^{Rlac}-His_6 was refolded by dialysis after Ni^{2+}-affinity chromatography under denaturing conditions. All purified proteins were quantified by absorbance at 280 nm. Purified CdiA-CT^{EC536}, CysK-His_6 and CdiI-His_6 were mixed at 10 μM final concentration in binding buffer and protein-protein interactions assessed by co-purification during Ni^{2+}-affinity chromatography as described (3, 7). Total RNA was isolated from *E. coli* X90 cells as described (43) and used as a substrate for *in vitro* nuclease assays. Total RNA (5 μg) was incubated with nuclease domains (1 μM) in reaction buffer [20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM $MgCl_2$] for 1 h at 37 °C. Where indicated, tRNase reactions contained CysK and/or immunity proteins at 1 μM final concentration. Reactions were analyzed by denaturing electrophoresis on 50% urea – 6% polyacrylamide gels in 15 Tris-borate-EDTA buffer. Gels were stained with ethidium bromide or transferred to nylon membrane for northern blot hybridization to 5´-radiolabeled oligonucleotide Ile1 probe, (5´ - ACC GAC CTC ACC CTT ATC AG) as described (43).

In vivo toxicity assays

E. coli X90 and X90 Δ*cysK* cells were transformed with 100 ng of arabinose-inducible CdiA-CT^{EC536} expression plasmid expression constructs followed by recovery for 1 h at 37 °C in

LB media supplemented with 0.4% glucose. Stable transformants were selected on LB-agar supplemented with 12.5 µg/mL tetracycline and 0.4% D-glucose or L-arabinose. Ntox28^{Rlac} toxicity was assessed using an arabinose-inducible construct (pCH12334) that produces Ntox28^{Rlac} and ssrA(DAS)-tagged Imm^{Rlac} protein. Ntox28^{Rlac} inhibition activity from this latter construct was compared to CysK-dependent toxicity of an analogous construct (pCH7171) that produces CdiA-CT^{EC536} and ssrA(DAS)-tagged CdiI^{EC536}.

tRNA cross-linking

To assess tRNA binding to CysK/CdiA-CT^{EC536} complexes, *E. coli* strains CH12810, CH12811, CH12812 and CH12813 were grown individually in 250 mL of LB media supplemented with 150 µg/mL ampicillin and 12.5 µg/mL tetracycline. Cultures were grown to mid-log phase (OD_{600} ~ 0.7), then adjusted to 0.4% L-arabinose and 1 mM IPTG to induce CysK and CdiA-CT(H178A)^{EC536} production for 2.5 h. Cells were collected by centrifugation and resuspended in lysis buffer [20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM $MgCl_2$ and 0.01% Triton X-100]. Cells were broken by two passages through a French press at 18,000 psi and unbroken cells removed by centrifugation at 16,000 5*g* for 15 min. Clarified lysates were incubated with Ni^{2+}-NTA agarose resin at 4 °C for 1 h, then treated with 0.1% formaldehyde for 10 min at ambient temperature. Reactions were quenched with L-glycine at 125 mM. The resins were washed with lysis buffer supplemented with 30 mM imidazole and proteins eluted in lysis buffer supplemented with 300 mM imidazole. Elutions were analyzed by SDS-PAGE to detect proteins and 50% urea-polyacrylamide gel electrophoresis and ethidium bromide staining to detect nucleic acids.

The effects of CdiI^{EC536} and Imm^{Rlac} immunity proteins on tRNA cross-linking were determined in a similar manner. *E. coli* strains CH11002 and CH13102 were grown individually in 150 mL of LB media supplemented with 150 µg/mL ampicillin, 12.5 µg/mL tetracycline and 0.4% L-arabinose to induce CysK production. At mid-log phase, the cultures were adjusted to 1 mM IPTG to induce the production of His$_6$-CdiA-CT(H178A)^{EC536} or the His$_6$-CdiA-CT(H178A)/CdiI^{EC536} complex for 30 min. For Ntox28/Imm^{Rlac} reactions, cultures of *E. coli* CH2016 carrying pET21, pCH13146 or pCH13153 were induced with 1.5 mM IPTG for 1 h 30 min. *E. coli* strain CH9648, which expresses His$_6$-tagged translation factor EF-Tu, was also tested as a specificity control. Cells were broken by French press as described above and each lysate pre-incubated with Ni^{2+}-NTA agarose in 20 mM sodium phosphate (pH 8.0), 150 mM NaCl for 1 h at 4 °C. Each sample was split in two, and one half adjusted to 0.1% formaldehyde. After 10 min at ambient temperature, the reactions were quenched with 125 mM L-glycine. Resins were washed and nucleoprotein complexes eluted as described above.

Protein binding affinities

Dissociation constants for protein complexes were obtained using a BLitz biolayer interferometer (ForteBio) as described previously (6). Purified His$_6$-tagged proteins were diluted to 150 µg/mL in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl and loaded onto Ni^{2+}-NTA coated biosensors at 25 °C. Sensor-bound proteins were then exposed to binding partners at 0.5 - 10 µm for 180 s, and then dissociation into buffer was monitored over 180 s. Dissociation constants were calculated following reference subtraction using BLitz Pro Software (ForteBio).

Protein thermal stability measurements

Protein thermal stability was determined using differential scanning fluorimetry (DSF) and circular dichroism (CD) spectroscopy. For DSF measurements, proteins were incubated with 25 or 40 µM SYPRO orange dye in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl. Samples were heated from 25 – 96°C at 1 °C/min using an Mx3005P QPCR machine (Agilent Technologies). The dye was excited at 492 nm and fluorescence emission monitored at 610 nm. Melting curves were obtained in duplicate and each experiment was conducted independently three times. Melting temperatures (T_m) were determined using nonlinear regression to determine melting-curve inflection points. Due to the high baseline fluorescence of SYPRO orange bound to CdiA-CT^{EC536} toxin at 25 °C, we were unable to determine T_m values for this toxin using DSF. Therefore, we monitored toxin thermal stability using a Jasco J-810 spectropolarimeter. CdiA-CT^{EC536} and Ntox28^{Rlac} toxins concentrated to 0.5 mg/mL in 20 mM sodium phosphate (pH 7.4), and initial CD spectra (190 – 260 nm) were collected at 20 °C using a 1 mm cell to determine the secondary structure content. Proteins were heated from 20 – 80 °C and ellipticity at 222 nm measured every 2 °C. The percentage of folded protein was calculated by measuring the change in ellipticity at 222 nm (44). Thermal melting curves were plotted and T_m values were acquired from the inflection point of the curve. T_m values for CdiA-CT/CdiI^{EC536} and Ntox28/Imm^{Rlac} complexes were also determined by CD spectrometry to correlate with the values determined by DSF.

References

1. Willett JL, Ruhe ZC, Goulding CW, Low DA, & Hayes CS (2015) Contact-Dependent Growth Inhibition (CDI) and CdiB/CdiA Two-Partner Secretion Proteins. *Journal of molecular biology* 427(23):3754-3765.
2. Ruhe ZC, Wallace AB, Low DA, & Hayes CS (2013) Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *mBio* 4(4).
3. Aoki SK, *et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* 468(7322):439-442.

4. Zhang D, de Souza RF, Anantharaman V, Iyer LM, & 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.
5. Beck CM, *et al.* (2014) CdiA from *Enterobacter cloacae* delivers a toxic ribosomal RNase into target bacteria. *Structure* 22:707-718.
6. Morse RP, *et al.* (2012) Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proc Natl Acad Sci U S A* 109:21480-21485.
7. Nikolakakis K, *et al.* (2012) The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Molecular microbiology* 84(3):516-529.
8. Diner EJ, Beck CM, Webb JS, Low DA, & Hayes CS (2012) Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes & development* 26(5):515-525.
9. Willett JL, Gucinski GC, Fatherree JP, Low DA, & Hayes CS (2015) Contact-dependent growth inhibition toxins exploit multiple independent cell-entry pathways. *Proceedings of the National Academy of Sciences of the United States of America* 112(36):11341-11346.
10. Kredich NM, Becker MA, & Tomkins GM (1969) Purification and characterization of cysteine synthetase, a bifunctional protein complex, from *Salmonella typhimurium*. *J. Biol. Chem.* 244(9):2428-2439.
11. Zhao C, *et al.* (2006) On the interaction site of serine acetyltransferase in the cysteine synthase complex from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 341(4):911-916.
12. Chattopadhyay A, *et al.* (2007) Structure, mechanism, and conformational dynamics of O-acetylserine sulfhydrylase from Salmonella typhimurium: comparison of A and B isozymes. *Biochemistry* 46(28):8315-8330.
13. Salsi E, *et al.* (2010) Design of O-acetylserine sulfhydrylase inhibitors by mimicking nature. *J. Med. Chem.* 53(1):345-356.
14. Campanini B, *et al.* (2005) Interaction of serine acetyltransferase with O-acetylserine sulfhydrylase active site: evidence from fluorescence spectroscopy. *Protein Sci.* 14(8):2115-2124.
15. Campanini B, *et al.* (2015) Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein. *Biochimica et biophysica acta* 1854(9):1184-1193.
16. Tanous C, *et al.* (2008) The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in *Bacillus subtilis*. *J. Biol. Chem.* 283(51):35551-35560.
17. Ma DK, Vozdek R, Bhatla N, & Horvitz HR (2012) CYSL-1 interacts with the O2-sensing hydroxylase EGL-9 to promote H2S-modulated hypoxia-induced behavioral plasticity in *C. elegans*. *Neuron* 73(5):925-940.
18. Burkhard P, *et al.* (1998) Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *Journal of molecular biology* 283(1):121-133.
19. Burkhard P, Tai CH, Ristroph CM, Cook PF, & Jansonius JN (1999) Ligand binding induces a large conformational change in O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *Journal of molecular biology* 291(4):941-953.
20. Morse RP, *et al.* (2015) Diversification of beta-Augmentation Interactions between CDI Toxin/Immunity Proteins. *Journal of molecular biology* 427(23):3766-3784.

21. Huang B, Vetting MW, & Roderick SL (2005) The active site of O-acetylserine sulfhydrylase is the anchor point for bienzyme complex formation with serine acetyltransferase. *J. Bacteriol.* 187(9):3201-3205.
22. Francois JA, Kumaran S, & Jez JM (2006) Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the *Arabidopsis* cysteine synthase complex. *Plant Cell* 18(12):3647-3655.
23. Claus MT, Zocher GE, Maier TH, & Schulz GE (2005) Structure of the O-acetylserine sulfhydrylase isoenzyme CysM from Escherichia coli. *Biochemistry* 44(24):8620-8626.
24. Holm L & Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic acids research* 38(Web Server issue):W545-549.
25. Hindson VJ, Moody PC, Rowe AJ, & Shaw WV (2000) Serine acetyltransferase from *Escherichia coli* is a dimer of trimers. *The Journal of biological chemistry* 275(1):461-466.
26. Papadakos G, Wojdyla JA, & Kleanthous C (2012) Nuclease colicins and their immunity proteins. *Quarterly reviews of biophysics* 45(1):57-103.
27. Dupureur CM (2008) Roles of metal ions in nucleases. *Curr Opin Chem Biol* 12(2):250-255.
28. Beuning PJ & Musier-Forsyth K (1999) Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers* 52(1):1-28.
29. Zhang D, Iyer LM, & 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 research* 39(11):4532-4552.
30. Poole SJ, *et al.* (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet.* 7(8):e1002217.
31. Holberger LE, Garza-Sanchez F, Lamoureux J, Low DA, & Hayes CS (2012) A novel family of toxin/antitoxin proteins in *Bacillus* species. *FEBS letters* 586(2):132-136.
32. Koskiniemi S, *et al.* (2013) Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America* 110(17):7032-7037.
33. Whitney JC, *et al.* (2014) Genetically distinct pathways guide effector export through the type VI secretion system. *Molecular microbiology* 92(3):529-542.
34. Aiyar A & Leis J (1993) Modification of the megaprimer method of PCR mutagenesis: improved amplification of the final product. *Biotechniques* 14(3):366-369.
35. Hayes CS & Sauer RT (2003) Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol. Cell* 12(4):903-911.
36. Hayes CS, Bose B, & Sauer RT (2002) Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *The Journal of biological chemistry* 277(37):33825-33832.
37. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, & Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *Journal of molecular biology* 229(1):105-124.
38. Battye TG, Kontogiannis L, Johnson O, Powell HR, & Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* 67(Pt 4):271-281.

39. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
40. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.
41. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
42. Delano WL (2002) The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
43. Garza-Sánchez F, Janssen BD, & Hayes CS (2006) Prolyl-tRNA(Pro) in the A-site of SecM-arrested ribosomes inhibits the recruitment of transfer-messenger RNA. *The Journal of biological chemistry* 281(45):34258-34268.
44. Chim N, Harmston CA, Guzman DJ, & Goulding CW (2013) Structural and biochemical characterization of the essential DsbA-like disulfide bond forming protein from Mycobacterium tuberculosis. *BMC Struct Biol* 13:23.

CHAPTER 4

Preliminary Structural Analysis of an EF-Tu-binding Rhs-family Toxin/Immunity Complex in *Salmonella* Typhimurium

Abstract

Salmonella Typhimurium is one of the most common foodborne illnesses worldwide. With the uprising of *S.* Typhimurium drug-resistant strains (1, 2), discovering novel proteins or understanding pathways that may aid in virulence is of vital importance. The rearrangement hot-spot (Rhs)-family proteins are of particular interest, as they are commonly found in pathogenic bacteria and grant a competitive advantage over other neighboring cells. The Rhs pathway utilizes toxic proteins which can be delivered to neighboring cells in a contact-dependent manner utilizing the type IV or type VI secretion systems (T4SS or T6SS respectively) to inhibit their growth (3). Similarly to the contact-dependent growth inhibition (CDI) pathway, Rhs-family proteins are large with highly conserved N-termini and polymorphic C-termini which possess cytotoxic activity (Rhs-CT's). Downstream of these toxin regions are small genes encoding immunity proteins, termed RhsI's, which tightly bind and inactive cognate Rhs-CT toxins to form stable protein complexes, preventing self-inhibition by Rhs-CT cytotoxicity. Rhs genes have been identified in *S.* Typhimurium, encoding a functional Rhs-CT/RhsI toxin/immunity protein complex. Surprisingly, the Rhs-CT^{Styph}/$RhsI^{Styph}$ complex interacts with the abundant and vital bacterial host cell protein elongation factor Tu (EF-Tu) to form a ternary complex, implicating that EF-Tu may be required to activate the Rhs-CT toxin. We aim to structurally and functionally characterize the *S.* Typhimurium Rhs toxin/immunity protein complex to shed light on the function of the Rhs-

CT^{Styph} and its neutralization by $RhsI^{Styph}$ and to understand the role of EF-Tu in toxin function; Rhs proteins may also prove to be potent targets for novel antimicrobial development. We have purified and crystallized the protein complex and produced diffracting crystals and collecting a native dataset to 3.0 Å resolution. We have been unable to solve the crystal structure using molecular replacement, and are currently attempting to use experimental phasing methods including crystallization of SeMet-derivative protein complex as well as soaking heavy metals into native crystals.

Introduction

Rearrangement hot spot (Rhs)-type genes were first discovered in 1979, where regions of high frequency genetic rearrangement were observed in genomic regions containing YD-peptide repeats (4, 5). Rhs genes were reported in several bacterial genomes, but had no known function. Vlazny *et al* were able to link RhsA from *E. coli* K12 to growth-inhibitory activity, lost upon expression of a small gene downstream of RhsA, implicating the first known function of an Rhs (YD-repeat) locus (6). In 2010, it was shown that an Rhs-family gene from *Pseudomonas savastanoi* is involved in bacteriocin production (7). Analyzing the functions and behaviors of the Rhs proteins and loci show several parallels with those of the CDI pathway. Recently, it has been demonstrated that the Rhs loci from the plant pathogen *Dickeya dadantii* encode a toxin/immunity protein pair that is capable of inhibiting neighboring cell growth by delivering a cytotoxic nuclease in a type VI secretion system (T6SS)-dependent manner (3, 8). Similar to CDI, this growth advantage phenotype was nullified upon loss of cell-to-cell contact, loss of the *D. dadantii* VgrG gene (required for T6SS), or expression of the cognate $RhsI_A$ in the target cells. These results suggest that indeed Rhs proteins function in a similar manner to CDI systems.

The CDI and Rhs pathways show many functional similarities, both utilizing large toxin proteins with polymorphic C-terminal catalytic regions (CdiA-CTs and Rhs-CTs). These toxins are both inhibited by cognate immunity proteins expressed by a small gene just downstream of the toxin (CdiI and RhsI). The primary difference between these two inhibition pathways lies in the mechanism of toxin secretion and delivery to target cells. CDI utilizes a type V secretion system (T5SS) using CdiB to export and present CdiA on the cell surface, where CdiA may contact target cell receptors, resulting in cleavage and translocation of the toxic CdiA-CT protein into the target cell cytosol. Rhs primarily utilizes the T6SS, utilizing a proteinacous bacterial needle-like complex, with a VgrG/PAAR spike complex as the needle "tip" required for target cell recognition and membrane puncturing. The effector molecule (in this case Rhs-CT) is then often found linked to either VgrG, the PAAR protein, or even an Hcp hexamer. Upon membrane puncturing, the effector-containing spike complex dissociates, releasing the Rhs-CT toxin delivery into the periplasm and allowing toxin translocation into the target cell cytosol (9, 10).

Salmonella enterica serovar Typhimurium is a common enteric bacterial pathogen and one of the world's most prevalent foodborne diseases. Combined with the recent uprising of multi-drug resistant *S.* Typhimurium (1, 2), identification and characterization of proteins and pathways involved in virulence and cell survival is important in combating pathogenesis. *Salmonella* has been shown to utilize the T6SS along with several pathogenic bacteria to deliver effector molecules into the cytosol of neighboring cells to inhibit growth or modulate gene expression (9-11). Downstream of the T6SS gene locus reside protein-coding genes with the consensus sequence YDxxxGRL(I/T) known as YD-repeats, marking them as Rhs-family proteins. Rhs-proteins contain a conserved PxxxxDPxGL motif to demarcate the start of a polymorphic C-terminal

domain that displays nuclease activity, similar to the VENN motif in CdiA proteins, demonstrating the presence of an Rhs-family toxin/immunity pair in *S.* Typhimurium.

Initial expression and purification of the *S.* Typhimurium Rhs-CT/RhsI proteins showed that the toxin/immunity complex co-purifies with endogenous elongation factor Tu (EF-Tu) protein to form a stable ternary complex. EF-Tu is a three-domain GTPase responsible for binding and recruiting aminoacylated tRNA (aa-tRNA) to the mRNA-bound ribosome during protein translation (12-15). This function, and thus this enzyme, is essential to protein translation and cellular growth. As a GTPase, EF-Tu required hydrolysis of bound GTP in order to function, forming an "active" ternary complex consisting of EF-Tu, GTP, and the aa-tRNA. This ternary complex is known to bind to the ribosome within the A-site, promoting GTP hydrolysis, transfer of the aa-tRNA to the ribosome, and release of the "inactive" GDP-bound EF-Tu (16). In order to reactivate the EF-Tu, EF-Ts will then exchange the EF-Tu-bound GDP for a molecule of GTP, allowing the GTP-bound EF-Tu to bind another aa-tRNA molecule and continue its role in translation (14, 17-20). This difference between the "active" vs. "inactive" EF-Tu forms is thought to be due to a significant conformational change when EF-Tu is bound to GTP or GDP, whereby the three domains adopt a "compact" or "extended" conformation, respectively (21-25).

Despite extensive research carried out on EF-Tu and its role in translation, it is currently unknown what role EF-Tu binding might have on a cytotoxic protein such as the Rhs-CT from *S.* Typhimuirum. However, we hypothesize that it may function in a similar manner to that of CysK in the UPEC536 CDI pathway (26). The UPEC536 CdiA-CT toxin has been shown to require association with the cysteine biosynthetic enzyme, CysK, to aid in recruitment of the tRNA substrate. We predict that EF-Tu's-binding is used to localize the Rhs-CT^{STyph} toxin to the EF-Tu-bound tRNA substrate. The catalytic function of Rhs-CT^{STyph} is still unknown, however the

toxin shows sequence homology to several ribonucleases from *E. coli*, and considering its association with EF-Tu likely acts upon a tRNA substrate. This is not the first example of a cytotoxic protein interacting with EF-Tu; a predicted CDI-family CdiA-CT ribonuclease from *E. coli* NC101 also co-purifies with EF-Tu, forming a ternary complex with cognate immunity protein. The function of EF-Tu binding CdiA-CT^{NC101} is still unknown, but this demonstrates a potentially common mechanism utilized by bacteria to partake in cell-to-cell communication or potentially even as a novel form of virulence factor. In addition, recent work has shown that a type VI-secreted toxin from *Pseudomonas* aeruginosa, Tse6, binds to EF-Tu with high affinity (27). Tse6 is an toxic NAD+ glycohydrolase, capable of cleaving NAD+ and NADP+ resulting in cell growth inhibition. Tse6 is delivered via the T6SS proteinacous needle complex, associating with the VgrG tip assembly to be directly delivered to the target cell periplasm, at which point target cell cytosolic EF-Tu is required for translocation of Tse6 into the target cell cytosol to inhibit growth. While EF-Tu is not required for Tse6 catalytic activity as is thought of canonical permissive factors, Tse6 mutants which lack the ability to bind EF-Tu are incapable of bypassing the inner membrane to enter a target cell, showing that EF-Tu is required for toxin delivery. This further demonstrates EF-Tu binding as being a widespread mechanism utilized by polymorphic toxin systems to aid in inhibition of neighboring cells, marking it as highly important to study in order to better understand bacterial communication and competition mechanisms such as CDI and Rhs. Due to the competitive advantage granted by the Rhs pathway and the possible role in bacterial survival during host colonization, the Rhs proteins may also prove to be potent drug targets, which may lead to the development of novel antimicrobials or more efficient strategies to combat potent human pathogens.

Preliminary Results

Salmonella Rhs Complex Purification and Crystallization

Native Rhs-CT^{Styph} toxin fused to an N-terminal thioredoxin A (TrxA) was co-expressed with cognate His$_6$-tagged RhsI^{Styph} immunity protein. Ni^{2+}-affinity chromatography resulted in the co-purification of the complex along with endogenous *E. coli* EF-Tu. The *S.* Typhimurium Rhs-CT/RhsI/EF-Tu ternary complex displays a 1:1:1 ratio of each protein as confirmed by analytical size exclusion chromatography (data not shown) and SDS-PAGE (**Figure 4.1A**). Sparse matrix

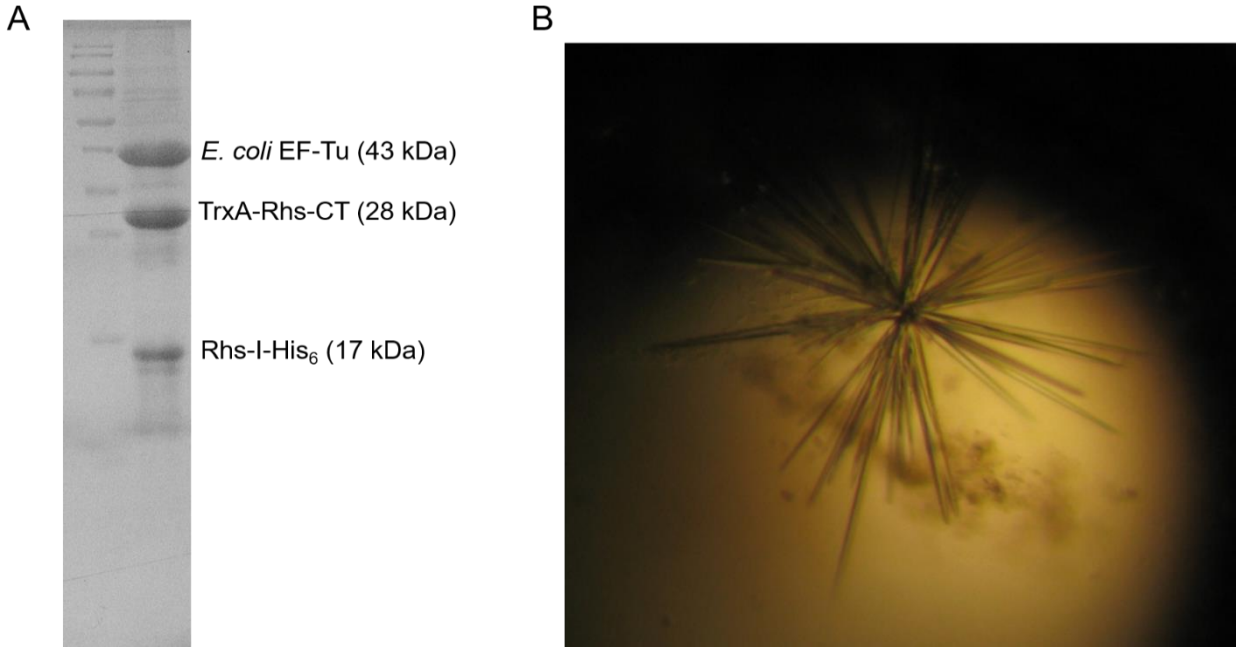


Figure 4.1: Purification and crystallization of *Salmonella* Typhimurium Rhs proteins. (**A**) SDS-PAGE showing purified *S.* Typhimurium TrxA-Rhs-CT/RhsI in a ternary complex with *E. coli* EF-Tu. (**B**) Protein crystals of *S.* Typhimurium ternary Rhs complex at 4.85 mg/mL.

crystal screens were setup with the ternary complex at 4.85 mg/mL and trace amounts of chymotrypsin at a 1:1000 protease:protein ratio (28). After 5 months, crystals grew in several different conditions. A single crystal grown in 0.2 M sodium thiocyanate pH 6.9, 20% PEG 3350 diffracted to 2.6 Å, although optimization resulted in non-diffracting crystals. Another condition (0.1 M bis-tris propane pH 6.5, 0.2 M sodium bromide, and 20% PEG 3350) yielded crystals which

diffracted consistently to 7 Å. Optimization of this condition resulted in crystals that diffract to 3-3.5 Å (**Figure 4.1B**). A native dataset was collected at a 1.0 Å wavelength and indexed to space group C2 with unit cell dimensions 100.2 Å x 47.9 Å x 81.6 Å x 90˚ x 108.2˚ x 90˚ with a highest resolution of 3.0 Å (**Figure 4.2A**). Attempting to calculate the Matthews coefficient revealed that the unit cell determined is far too small to accommodate the entire protein complex of 88 kDa. In the full-length complex, EF-Tu is the largest protein at 43 kDa, with TrxA-Rhs-CT^{Styph} at 28 kDa (10 kDa for TrxA and 18 kDa for the toxin) and RhsI^{Styph} at 17 kDa. Analysis of the crystals by

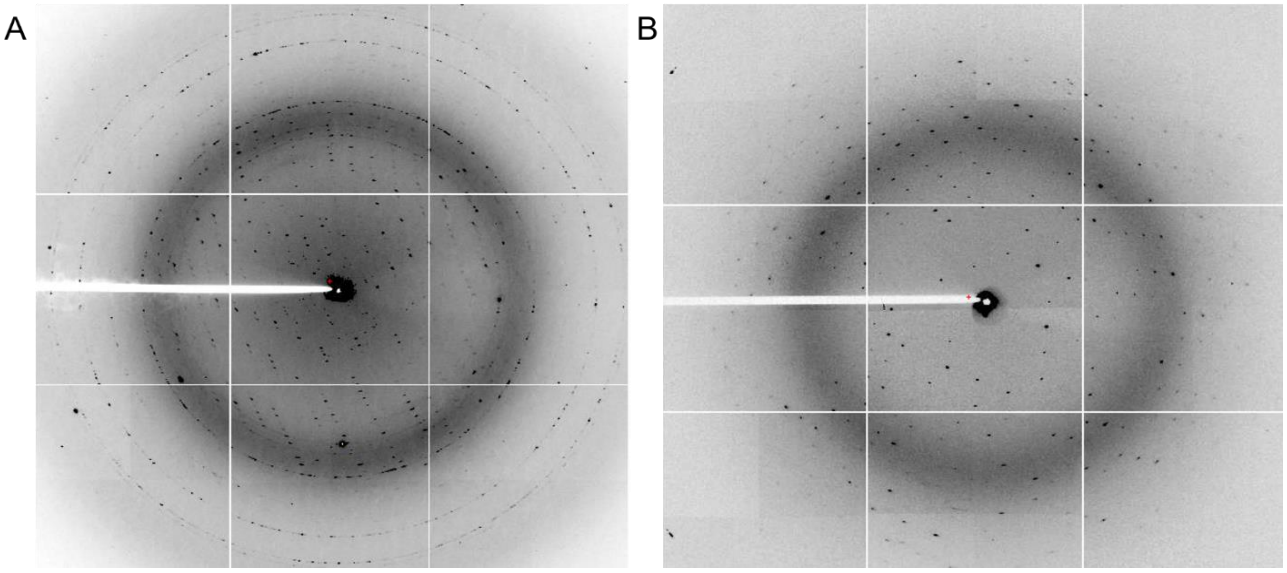


Figure 4.2: Crystal diffraction and data collection from *S.* Typhimurium ternary Rhs complex crystals. (**A**) Diffraction images of *S.* Typhimurium ternary Rhs complex crystals diffracting poorly to ~3.5 Å. (**B**) Diffraction of optimized *S.* Typhimurium ternary Rhs complex crystals. A high quality 3.0 Å dataset was collected, indexing to the space group C2 with unit cell: 99.6 Å x 47.2 Å x 80.6 Å x 90˚ x 108.3˚ x

SDS-PAGE and MALDI-TOF mass spectrometry exhibited three separate proteins at 23 kDa, 18 kDa, and 10 kDa for a total complex size of 51 kDa (**Figure 4.3**), indicating that two or more of the proteins within the crystal have been truncated. Using this complex size, the Matthews coefficient correlates with one complex per asymmetric unit.

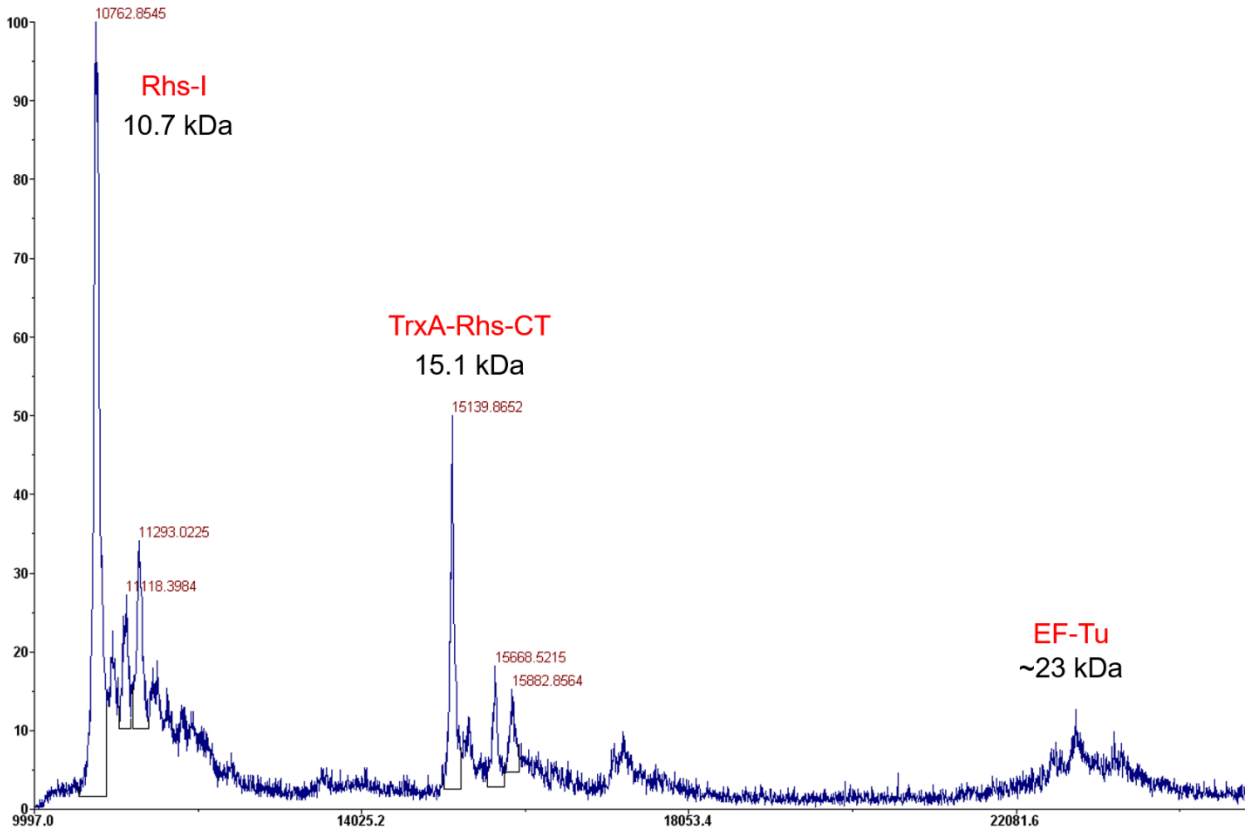


Figure 4.3: MALDI-TOF spectra for *Salmonella* Typhimurium ternary Rhs complex crystals shows three peaks at 10.7 kDa, 15.1 kDa, and 23 kDa indicative of truncated RhsI, Rhs-CT, and EF-Tu, respectively.

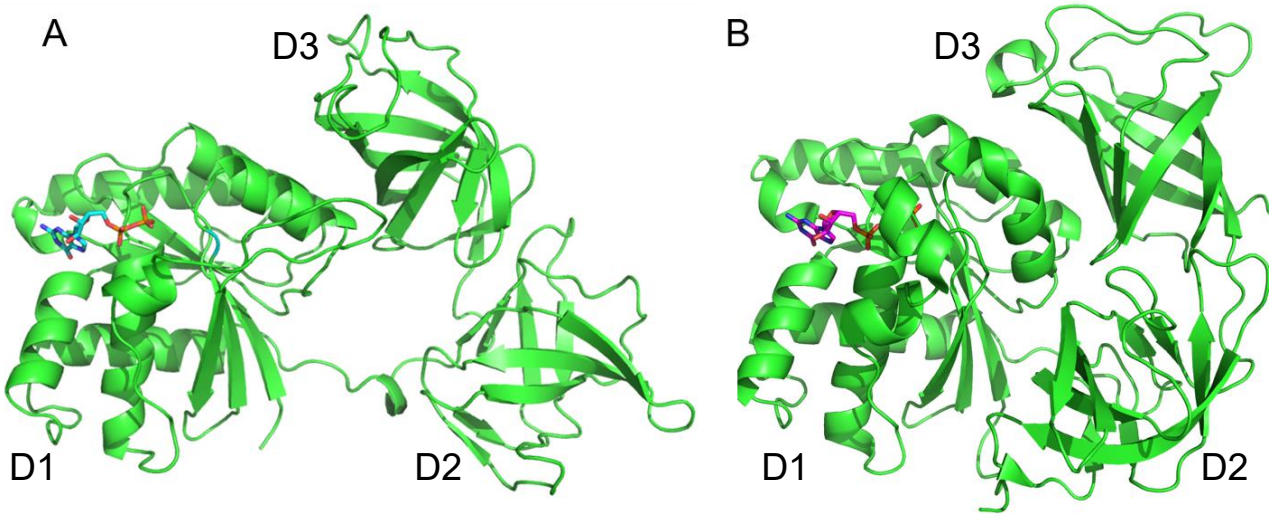


Figure 4.4: EF-Tu can be found in two structural conformations depending on its association with either GDP or GTP with domains 1-3 labeled as D1, D2, and D3 respectively. (**A**) Structure of *Thermus aquaticus* EF-Tu bound to GDP (PDB ID: 1TUI) with GDP shown in cyan. (**B**) Structure of *Thermus thermophilus* EF-Tu in the "active" form bound to GTP (PDB ID: 1EFT) with GTP shown in purple.

Structure determination of Salmonella Rhs complex

EF-Tu consists of three globular domains and can be found in two different conformations of its three domains depending on its guanosine-based ligand. When bound to GDP, EF-Tu is in its "inactive form," showing a reduced affinity for tRNA and triggering release form the ribosome. This GDP-bound conformation shows a weak association between the three domains (D1, D2, and D3), with no visible contacts being made between D1 and D2. EF-Tu binds GTP, the protein undergoes a conformation shift to its "active conformation" (**Figure 4.4**). This shift involves multiple contacts being formed between D1 and both D2 and D3, winding the extended linker region into the C-terminal α-helix on D1 to form a more compact "active" fold. Unlike the "extended" GDP-bound structure, the overall structure adopts a "compact" fold, with extensive contacts being formed between D1 and D2. Despite the significant change in protein conformation, the overall structures of D2 and D3 remain largely unchanged, overlaying between the two structures with an rmsd of 0.8 Å. We predicted that the endogenous EF-Tu may be a mix of these two conformations, preventing proper crystal packing until the proteins had degraded to a just domains D2 and D3. To lock all of the EF-Tu into the "inactive" conformation, we prepared crystal screens with the addition of 10-fold excess GDP to the ternary EF-Tu/Rhs-CT/RhsI complex and in the absence of protease. Shortly after preparation of these trays, we observed growth of protein crystals (5 mg/mL) with a similar morphology to the previous needles, within only 2 weeks, and in a different condition (0.1 M HEPES pH 7.0, 0.2 M NaCl, and 25% PEG 3350). These crystallization conditions were optimized and resulted in crystals that diffracted to 2.5-3.0 Å (**Figure 4.2B**). One dataset was collected to 3.0 Å resolution at 1.0 Å wavelength and indexed also to space group C2 with unit cell dimensions 99.6 Å x 47.2 Å x 80.6 Å x 90° x 108.3° x 90°. Despite the absence of protease in these drops, MALDI-TOF showed a similar pattern of

truncated proteins, implying that the protein degradation may be a spontaneous process due to unstable or overly flexible regions of the proteins, or by protease contamination.

Table 4.1. **X-ray diffraction data for the *Salmonella* EF-Tu/Rhs-CT/RhsI complex**

	First Dataset	**Second Dataset**
Space Group	C2	C2
Unit cell dimensions (Å)	100.2, 47.9, 81.6, 90, 108.2, 90	99.6, 47.2, 80.6, 90, 108.3, 90
pH of crystallization condition	6.5	7.0
Protein concentration (mg/mL)	4.83	5.0
Dataset		
Wavelength, Å	1.0	1.0
Resolution range	50-3.0	50-3.0
Unique reflections (total)	7107 (22513)	7312 (50443)
Completeness, %	94.0	99.3
Redundancy*	3.2 (3.2)	6.9 (5.5)
R_{merge}*,†	0.203 (0.510)	0.133 (0.464)
R_{meas}*, ‡	0.243 (0.610)	0.144 (0.509)
$R_{p.i.m.}$ *,§	0.132 (0.329)	0.054 (0.205)
$CC_{1/2}$*	0.730 (0.848)	0.999 (0.906)
I/σ*	7.0 (2.9)	19.9 (3.2)
NCS copies	1	1

*Statistics for the highest-resolution shell are given in parentheses.

† $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

‡ $R_{meas} = \Sigma_{hkl} \{N(hkl) / [N(hkl) - 1]\}^{1/2} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

§ $R_{p.i.m}$ (precision-indicating R_{merge}) $= \Sigma_{hkl} \{1/[N(hkl) - 1]\}^{½} \Sigma_i |I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

Comparing the statistics of these two datasets (**Table 4.1**) it is evident that this crystal resulted in a much higher quality native dataset, providing an excellent starting point to attempt solving the crystal structure. Molecular replacement was performed with both datasets using the crystal structure of *E. coli* EF-Tu. Several homology models for the Rhs-CT toxin and RhsI immunity protein were generated using iTASSER (29-31) and Phyre2 (32). Models were also generated from the recently solved *Escherichia coli* NC101 EF-Tu/CdiA-CT/CdiI complex crystal structure, which we believe may have a similar structure and binding mechanism between toxin and EF-Tu as the *Salmonella* Rhs complex. The structures of the NC101 EF-Tu/CDI complex were initially used as direct search models for molecular replacement. However, despite a predicted similar

toxin structure and interface with EF-Tu did not result in any viable solutions. The search models were then stripped of all side-chains, resulting in protein models consisting of only alanine residues, allowing the molecular replacement search to look strictly at the α-carbon backbone; this strategy can aid in determination of a solution between proteins with variable amino acid compositions but with conserved secondary structures and tertiary folds. In addition to the poly-alanine models, we also utilized the program *Sculptor* (33) to generate new models of the NC101 toxin/EF-Tu complex, but replacing the NC101 toxin amino acid sequence with that of the *Salmonella* Rhs-CT toxin, likely determining a solution if the Rhs-CT^{Styph} toxin and CdiA-CT^{NC101} toxin share a conserved overall fold. Using the sequence-threaded CdiA-CT^{NC101} generated a partial model of the complex, building primarily random strings of alanine and glycine residues, however this model also demonstrates a clear four-stranded antiparallel β-sheet at the center of the model (**Figure 4.5**). This β-sheet matches the sequence of the Rhs-CT toxin, with appropriate side chains fitting into the electron density (contoured at 1.5 σ) quite well; however, due to the high R_{work}/R_{free} (41.38/48.42) this may be due to model bias from the molecular replacement search. We are currently attempting to optimize this model further by iterative cycles of building followed by refinement in an attempt to lower the R-values and acquire a high-confidence model. Despite the relatively high confidence of these several models and the good quality of the diffraction data, we have been unable to successfully generate a complete model of the complex by molecular replacement. We hypothesize that the presence of multiple proteins in the crystals, (two of them with little structural information and all of them truncated), makes resolving the crystal structure by non-experimental phasing methods highly unlikely, forcing us to utilize alternative methods of overcoming the phase-problem of X-ray crystallography. One commonly used and highly effect

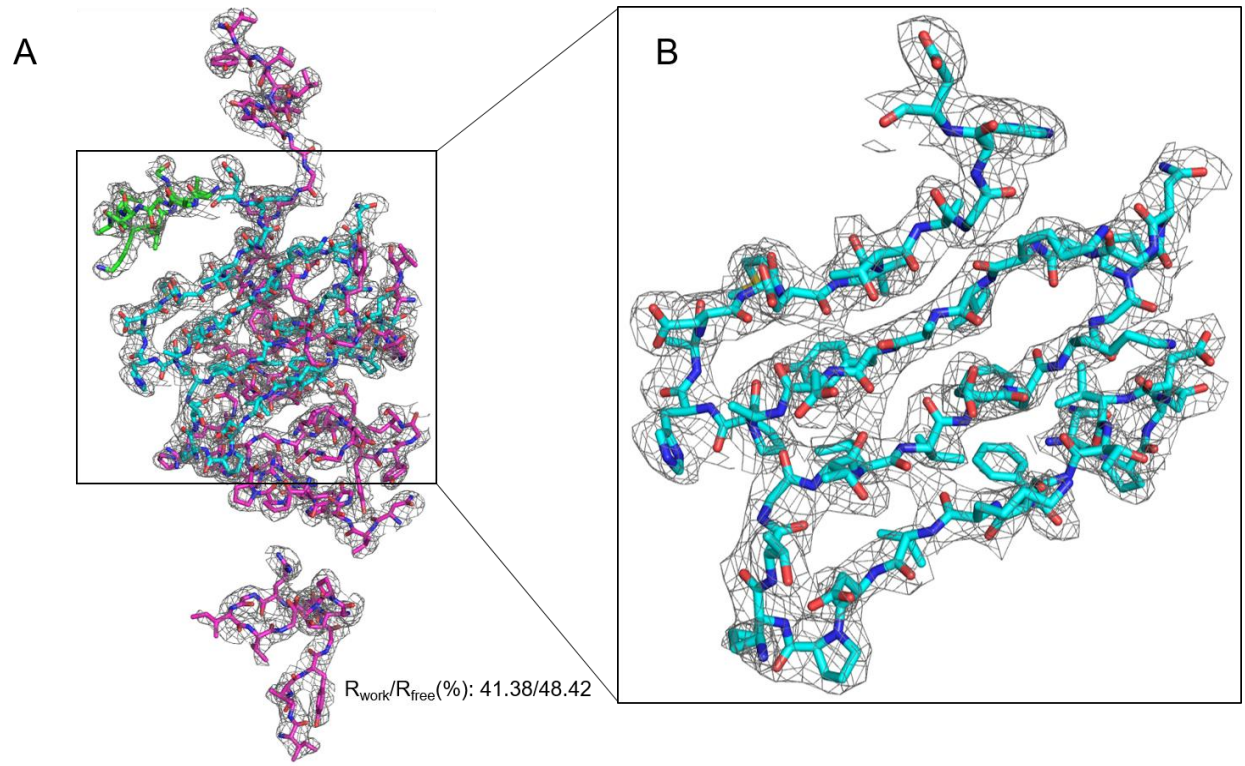


Figure 4.5: Partial solution of *Salmonella* Typhimurium Rhs-CT toxin with electron density map contoured at 1.5 σ. (**A**) Partial solution from molecular replacement using NC101 CdiA-CT as a search model, with the toxin shown in cyan and the identity of the purple and green chains unknown. (**B**) The partial structure of the Rhs-CT toxin shows a four-stranded antiparallel β-sheet, with residue side chains matching the known toxin sequence as well as fitting the density quite well.

mechanism of experimental phasing is the production of selenomethionine (SeMet)-substituted proteins. By expressing a protein of interest in minimal media lacking native methionine and instead adding methionine residues with a selenium atom as opposed to a sulphur atom, we can generate a protein with an inherent anomalous signal which can be used to solve the phase problem of crystallography without relying on molecular replacement via imperfect search models (34). In our efforts to solve the structure of the *Salmonella* Rhs complex by experimental phasing, we first prepared selenomethionine (SeMet)-substituted *S.* Typhimurium Rhs-CT/RhsI (co-purifying with endogenous SeMet-substituted EF-Tu) and setup screens of the SeMet-ternary complex. After two months of growth, several new crystal hits appeared in a range of conditions including the two conditions that previously generated diffracting native crystals. Both crystallization conditions for the SeMet-ternary complex have been extensively optimized yet have not produced crystals

capable of diffracting to a resolution higher than 5.5 Å. While the SeMet protein crystals tend to diffract to a lower resolution than the native crystals, we can instead soak a large quantity of native ternary complex crystals using heavy metal soaks, introducing bromide or iodide ions into the crystals to allow for determination of the phase angle utilizing heavy metal ion anomalous dispersion (i.e. SAD or MAD) (35, 36). We hope that acquisition of a reasonable-quality anomalous dataset by either SeMet substitution or Br/I-soaking will allow us to overcome the phase problem and determine an initial low resolution model of the complex which can then be utilized as a search model for molecular replacement using our high quality 3.0 Å native dataset, allowing for a complete structure of the ternary complex and elucidation of the protein interactions and mechanisms.

References

1. Hoffmann M, *et al.* (2013) Complete Genome Sequence of a Multidrug-Resistant Salmonella enterica Serovar Typhimurium var. 5- Strain Isolated from Chicken Breast. *Genome announcements* 1(6).
2. Tamamura Y, *et al.* (2013) Complete nucleotide sequences of virulence-resistance plasmids carried by emerging multidrug-resistant Salmonella enterica Serovar Typhimurium isolated from cattle in Hokkaido, Japan. *PloS one* 8(10):e77644.
3. Koskiniemi S, *et al.* (2013) Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America* 110(17):7032-7037.
4. Capage M & Hill CW (1979) Preferential unequal recombination in the glyS region of the Escherichia coli chromosome. *Journal of molecular biology* 127(1):73-87.
5. Lin RJ, Capage M, & Hill CW (1984) A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome. *Journal of molecular biology* 177(1):1-18.
6. Vlazny DA & Hill CW (1995) A stationary-phase-dependent viability block governed by two different polypeptides from the RhsA genetic element of Escherichia coli K-12. *Journal of bacteriology* 177(8):2209-2213.
7. Sisto A, *et al.* (2010) An Rhs-like genetic element is involved in bacteriocin production by Pseudomonas savastanoi pv. savastanoi. *Antonie van Leeuwenhoek* 98(4):505-517.
8. Poole SJ, *et al.* (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS genetics* 7(8):e1002217.

9. Hachani A, *et al.* (2011) Type VI secretion system in Pseudomonas aeruginosa: secretion and multimerization of VgrG proteins. *The Journal of biological chemistry* 286(14):12317-12327.
10. Pukatzki S, Ma AT, Revel AT, Sturtevant D, & Mekalanos JJ (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proceedings of the National Academy of Sciences of the United States of America* 104(39):15508-15513.
11. Hood RD, *et al.* (2010) A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. *Cell host & microbe* 7(1):25-37.
12. Miller DL & Weissbach H (1977) Factors involved in the transfer of aminoacyl-tRNA to the ribosome. *Molecular Mechanism of Protein Biosynthesis*:323-373.
13. Thomposon RC & Dix DB (1982) Accuracy of protein biosynthesis. A kinetic study of the reaction of poly(U)-programmed ribosomes with a leucyl-tRNA2-elongation factor Tu-GTP complex. *The Journal of biological chemistry* 257(12):6677-6682.
14. Rodnina MV, Pape T, Fricke R, & Wintermeyer W (1995) Elongation factor Tu, a GTPase triggered by codon recognition on the ribosome: mechanism and GTP consumption. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 73(11-12):1221-1227.
15. Rodnina MV, Savelsbergh A, & Wintermeyer W (1999) Dynamics of translation on the ribosome: molecular mechanics of translocation. *FEMS microbiology reviews* 23(3):317-333.
16. Knudsen C, Wieden HJ, & Rodnina MV (2001) The importance of structural transitions of the switch II region for the functions of elongation factor Tu on the ribosome. *The Journal of biological chemistry* 276(25):22183-22190.
17. Burnett BJ, *et al.* (2013) Elongation factor Ts directly facilitates the formation and disassembly of the Escherichia coli elongation factor Tu.GTP.aminoacyl-tRNA ternary complex. *The Journal of biological chemistry* 288(19):13917-13928.
18. Liu W, *et al.* (2015) EF-Tu dynamics during pre-translocation complex formation: EF-Tu.GDP exits the ribosome via two different pathways. *Nucleic acids research* 43(19):9519-9528.
19. Louie A & Jurnak F (1985) Kinetic studies of Escherichia coli elongation factor Tu-guanosine 5'-triphosphate-aminoacyl-tRNA complexes. *Biochemistry* 24(23):6433-6439.
20. Romero G, Chau V, & Biltonen RL (1985) Kinetics and thermodynamics of the interaction of elongation factor Tu with elongation factor Ts, guanine nucleotides, and aminoacyl-tRNA. *The Journal of biological chemistry* 260(10):6167-6174.
21. la Cour TF, Nyborg J, Thirup S, & Clark BF (1985) Structural details of the binding of guanosine diphosphate to elongation factor Tu from E. coli as studied by X-ray crystallography. *The EMBO journal* 4(9):2385-2388.
22. Nissen P, *et al.* (1995) Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science* 270(5241):1464-1472.
23. Nissen P, Kjeldgaard M, Thirup S, Clark BF, & Nyborg J (1996) The ternary complex of aminoacylated tRNA and EF-Tu-GTP. Recognition of a bond and a fold. *Biochimie* 78(11-12):921-933.
24. Abel K, Yoder MD, Hilgenfeld R, & Jurnak F (1996) An alpha to beta conformational switch in EF-Tu. *Structure* 4(10):1153-1159.
25. Polekhina G, *et al.* (1996) Helix unwinding in the effector region of elongation factor EF-Tu-GDP. *Structure* 4(10):1141-1151.

26. Diner EJ, Beck CM, Webb JS, Low DA, & Hayes CS (2012) Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes & development* 26(5):515-525.
27. Whitney JC, *et al.* (2015) An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* 163(3):607-619.
28. Dong A, *et al.* (2007) In situ proteolysis for protein crystallization and structure determination. *Nature methods* 4(12):1019-1021.
29. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* 9:40.
30. Roy A, Kucukural A, & Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols* 5(4):725-738.
31. Yang J, *et al.* (2015) The I-TASSER Suite: protein structure and function prediction. *Nature methods* 12(1):7-8.
32. Kelley LA, Mezulis S, Yates CM, Wass MN, & Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* 10(6):845-858.
33. Birmanns S, Rusu M, & Wriggers W (2011) Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes. *J Struct Biol* 173(3):428-435.
34. Hendrickson WA, Horton JR, & LeMaster DM (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *The EMBO journal* 9(5):1665-1672.
35. Sun PD, Radaev S, & Kattah M (2002) Generating isomorphous heavy-atom derivatives by a quick-soak method. Part I: test cases. *Acta crystallographica. Section D, Biological crystallography* 58(Pt 7):1092-1098.
36. Kim Y & Haren AM (1995) The application of crystal soaking technique to study the effect of zinc and cresol on insulinotropin crystals grown from a saline solution. *Pharmaceutical research* 12(11):1664-1670.

CHAPTER 5

Concluding Remarks and Future Directions: Understanding Polymorphic Toxin Systems and Their Role in Competitive and Cooperative Behaviors

Summary

Contact-dependent growth inhibition (CDI) is a unique inter-bacterial competition mechanism by which CDI+ cells can deliver polymorphic CdiA-CT toxins to inhibit cell growth and promote cooperative behaviors such as cell-to-cell adhesion and biofilm formation (1-5). In order to prevent auto-inhibition and protect against inhibition from self-cells, CDI+ bacteria express CdiI immunity proteins that tightly and specifically bind cognate CdiA-CT's and inhibit cytotoxic activity. Here we presented further structural and functional characterization of novel CDI toxin/immunity complexes as well as the preliminary characterization of an Rhs-family protein complex. These structures yielded additional insight into the significant structural and functional diversity amongst CDI toxins, whereby even two *B. pseudomallei* CdiA-CT toxins with no sequence similarity and unique tRNase functions share a conserved protein fold, yet still show unique interfaces with cognate immunity proteins. To add to toxin diversity, some even require binding to a target cell protein (known as a "permissive factor"), for growth inhibitory activity; we have shown that the CdiA-CT toxin from uropathogenic *E. coli* 536 binds to target cell CysK to become an active tRNase. CysK activates the CdiA-CT toxin by stabilizing its fold and promoting its association with tRNA substrate to allow for cleavage. While CysK is the currently the only experimentally characterized permissive factor, many other toxin systems, including an Rhs-family toxin/immunity complex from *Salmonella* Typhimurium, have been

shown to tightly interact with the endogenous protein EF-Tu, implying that permissive factors maybe a widespread strategy utilized by polymorphic toxin systems. For future research, we are interested in further characterizing the cell-to-cell contact and translocation mechanisms involved in toxin delivery into target cells mediated by the CdiA "stick" region, as well as further understanding the role of permissive factors in polymorphic toxin function or delivery. While it has been shown that CDI plays an important role in communication and cooperative behaviors, it is currently unknown exactly how CdiA-CT toxin delivery effects gene expression regulation to induce significant changes in cellular function such as adhesion and biofilm formation.

Elaborating on CDI Toxin/Immunity Complex Diversity

As of now, over 100 CdiA-CT/CdiI toxin/immunity protein families have been identified, possessing a wide variety of activities, structures, and unique complex interfaces (6-11). We sought to investigate the variability of toxin/immunity complex structures and interfaces between two environmental isolates of *Burkholderia pseudomallei*, E479 and 1026b. These two CdiA-CT toxins are both tRNases, yet show no significant sequence homology with each other. In addition, the two toxins cleave different subsets of tRNA at unique cut-sites: CdiA-CT^{1026b} cleaves tRNA^{Ala} at the aminoacyl acceptor stem, while CdiA-CT^{E479} cleaves most cellular tRNA's at the T-loop (8). Comparing the two crystal structures revealed that the two toxins share significant structural homology, both consisting of a central β-sheet decorated by α-helices, typical of a canonical ribonuclease fold, even sharing nearly identical active sites. Further inspection of the active site pockets shows that CdiA-CT^{E479} has a larger active site pocket, potentially better suited for fitting a wider variety of tRNA flexible T-loops, while the more

specific CdiA-CT^{1026b} toxin has a narrower active site pocket capable of accommodating the backbone of the aminoacyl acceptor stem of tRNA substrate, as has been confirmed with both molecular docking simulations and SAXS. Despite the significant structural similarity seen between these two toxins from the same bacterial species, we do not see any cross-protection against non-cognate toxins by these CdiI immunity proteins. We hope to continue solving novel CdiA-CT/CdiI complex structures from uncharacterized protein families to learn more about protein diversity as well as identify more potential targets for novel antimicrobials.

Elucidating the Role of Host-Cell Permissive Factors in Polymorphic Toxin Systems

Upon first characterizing the cytotoxic activity of the CdiA-CT toxin from uropathogenic *E. coli* 536, it was surprising to find that it is actually inactive *in vitro*. A series of tRNase assays demonstrated that the biosynthetic enzyme CysK is both necessary and sufficient for the activity of CdiA-CT^{UPEC536} *in vitro* (12). In addition, CysK-deficient cells were completely resistant to the growth inhibitory activity of CdiA-CT^{UPEC536}, implicating CysK as being required for CDI *in vivo* as well. Through comparison to a CysK-independent CdiA-CT^{UPEC536} homolog from the Gram-positive organism *Ruminococcus lactaris*, we have determined that CysK binding has a chaperone-like effect to stabilize the UPEC536 toxin fold and promote its association with tRNA substrate. We are currently attempting to solve the X-ray crystal structure of the *R. lactaris* toxin/immunity complex to further study what structural properties this toxin evolved to be CysK-independent in comparison to the UPEC536 toxin. While we have determined how CysK activates UPEC536 toxin activity, we are still not aware of *why* CysK would be necessary where a large majority of experimentally characterized toxins do not require a permissive factor: is this

simply an inefficient evolutionary branch that persisted across multiple bacterial species, or is it part of a more complex mechanism that we have not yet identified?

Initially the permissive factor phenomenon was thought to be unique to UPEC536; however, recently numerous polymorphic toxins from *Pseudomonas aeruginosa*, *Escherichia coli* NC101, and *Salmonella* Typhimurium have been shown to associate with EF-Tu. Tse6 is a toxic NAD+ glycohydrolase from *P. aeruginosa* which is secreted by the type VI secretion system (T6SS) to neighboring cells to inhibit growth (13). Recent work has shown that target cell EF-Tu is required for translocation of Tse6 across the inner membrane, however once within the cytosol, interaction of Tse6 with EF-Tu is not required for toxin activity. The CDI-family CdiA-CT toxin from *E. coli* NC101 and the Rhs-family Rhs-CT toxin from *S.* Typhimurium are currently uncharacterized, though predicted to be ribonucleases based on sequence homology; both of these toxins have been shown to form stable ternary complexes between the toxins, cognate immunity proteins, and EF-Tu. While no information is available yet on the function of EF-Tu binding for either system, the structure of the NC101 EF-Tu/CdiA-CT/CdiI ternary complex has recently been solved (unpublished data), showing that EF-Tu makes significant contacts with both the toxin and immunity protein, as opposed to only the toxin as seen with the UPEC CysK/CdiA-CT/CdiI ternary complex. This raises several questions on the potential function of EF-Tu binding; we hypothesize that EF-Tu remains bound to both toxin and immunity protein until the complex comes into contact with the appropriate RNA substrate, at which point the complex will dissociate and the toxin will become active and inhibit cell growth. We are currently attempting to solve the crystal structure of the *S.* Typhimurium EF-Tu/Rhs-CT/RhsI ternary complex and hope that this will further elucidate the role of both EF-Tu and permissive factors in general in polymorphic toxin systems.

Role of CDI in Bacterial Cooperative Behaviors

Polymorphic toxin systems such as CDI and Rhs play an important role in bacterial competition, granting a significant competitive advantage over neighboring bacterial populations (6, 14-16). Initially this growth advantage was thought to be primarily through the growth inhibitory activity granted by toxin delivery. Recent work has shown that this inhibitory activity has a much more specific set of potential target cells than expected for a competitive mechanism. CdiA proteins interact with receptors on the outer membrane of target cells, typically targeting β-barrel proteins such as BamA and OmpC/OmpF (17). These proteins are highly conserved amongst Gram-negative bacteria, but display sequence diversity amongst the small extracellular loops between β-sheet strands. It has been shown that the sequences of these loops are important for interaction with CdiA, as mutating these loop residues can prevent CdiA binding, or even change the specificity for a CdiA protein from a different strain or species. If CDI was a strictly competitive mechanism, it seems counter-productive to have such a small set of viable target cells. Recent work has demonstrated that CdiA promotes cell-to-cell adhesion, and cells lacking CdiA or even with active site mutations in CdiA-CT toxins are deficient in biofilm formation (16, 18-20). This implies that the primary role of CDI is communication and cell-to-cell signaling, functioning like a contact-dependent mechanism of quorum sensing to promote cooperative behaviors such as biofilm formation, utilizing the toxin/immunity pairs to keep non-self cells from leeching nutrients from a biofilm.

Concluding Remarks

This dissertation presents a structural and functional analysis of multiple novel protein complexes within the CDI system. We determined specific CdiA-CT cytotoxic activities as well as their neutralization by cognate CdiI immunity proteins, including why immunity proteins do not cross-protect against non-cognate toxins. We also highlighted the mechanism of activation of a CdiA-CT toxin by a target cell permissive factor, a unique step in CDI that appears to be more widespread amongst polymorphic toxin systems than initially thought. Polymorphic toxin systems are very widespread amongst α, β, and γ-proteobacteria, and are particularly well represented amongst human pathogens (6, 7); due to the rise in antibiotic resistant bacterial pathogens, mechanisms that bacteria utilize to increase virulence or survival within a host organism may provide targets for novel antibiotics. In addition to providing potent drug targets, the CDI pathway has been shown to play an important role in biofilm formation, making a bacterial population more resistant to both conventional treatment methods as well as the host immune system. Designing antimicrobials against CDI proteins may be multi-faceted, preventing the bacteria from forming biofilms as well as inhibiting their growth or decreasing their viability within a human host.

References

1. Hayes CS, Aoki SK, & Low DA (2010) Bacterial contact-dependent delivery systems. *Annual review of genetics* 44:71-90.
2. Aoki SK, Poole SJ, Hayes CS, & Low DA (2011) Toxin on a stick: modular CDI toxin delivery systems play roles in bacterial competition. *Virulence* 2(4):356-359.
3. Ruhe ZC, Low DA, & Hayes CS (2013) Bacterial contact-dependent growth inhibition. *Trends in microbiology* 21(5):230-237.
4. Webb JS*, et al.* (2013) Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. *PloS one* 8(2):e57609.

5. Willett JL, Ruhe ZC, Goulding CW, Low DA, & Hayes CS (2015) Contact-Dependent Growth Inhibition (CDI) and CdiB/CdiA Two-Partner Secretion Proteins. *Journal of molecular biology* 427(23):3754-3765.
6. Aoki SK*, et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* 468(7322):439-442.
7. Poole SJ*, et al.* (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS genetics* 7(8):e1002217.
8. Nikolakakis K*, et al.* (2012) The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Molecular microbiology* 84(3):516-529.
9. Morse RP*, et al.* (2012) Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proceedings of the National Academy of Sciences of the United States of America* 109(52):21480-21485.
10. Beck CM*, et al.* (2014) CdiA from *Enterobacter cloacae* delivers a toxic ribosomal RNase into target bacteria. *Structure* 22:707-718.
11. Morse RP*, et al.* (2015) Diversification of beta-Augmentation Interactions between CDI Toxin/Immunity Proteins. *Journal of molecular biology* 427(23):3766-3784.
12. Diner EJ, Beck CM, Webb JS, Low DA, & Hayes CS (2012) Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes & development* 26(5):515-525.
13. Whitney JC*, et al.* (2015) An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* 163(3):607-619.
14. Aoki SK*, et al.* (2005) Contact-dependent inhibition of growth in Escherichia coli. *Science* 309(5738):1245-1248.
15. Koskiniemi S*, et al.* (2013) Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America* 110(17):7032-7037.
16. Jamet A & Nassif X (2015) New players in the toxin field: polymorphic toxin systems in bacteria. *mBio* 6(3):e00285-00215.
17. Ruhe ZC, Wallace AB, Low DA, & Hayes CS (2013) Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *mBio* 4(4).
18. Garcia EC, Anderson MS, Hagar JA, & Cotter PA (2013) Burkholderia BcpA mediates biofilm formation independently of interbacterial contact-dependent growth inhibition. *Molecular microbiology* 89(6):1213-1225.
19. Anderson MS, Garcia EC, & Cotter PA (2014) Kind discrimination and competitive exclusion mediated by contact-dependent growth inhibition systems shape biofilm community structure. *PLoS pathogens* 10(4):e1004076.
20. Ruhe ZC*, et al.* (2015) CdiA promotes receptor-independent intercellular adhesion. *Molecular microbiology* 98(1):175-192.

APPENDIX:

The structure of a contact-dependent growth-inhibition (CDI) immunity protein from *Neisseria meningitidis* MC58

This project was performed in collaboration with members of the Joachimiak lab at Argonne National Laboratory

Abstract

Contact-dependent growth inhibition (CDI) is an important mechanism of intercellular competition between neighboring Gram-negative bacteria. CDI systems encode large surface-exposed CdiA effector proteins that carry a variety of C-terminal toxin domains (CdiA-CTs). All CDI^+ bacteria also produce CdiI immunity proteins that specifically bind to the cognate CdiA-CT and neutralize its toxin activity to prevent auto-inhibition. Here, the X-ray crystal structure of a CdiI immunity protein from *Neisseria meningitidis* MC58 is presented at 1.45 Å resolution. The CdiI protein has structural homology to the Whirly family of RNA-binding proteins, but appears to lack the characteristic nucleic acid-binding motif of this family. Sequence homology suggests that the cognate CdiA-CT is related to the eukaryotic EndoU family of RNA-processing enzymes. A homology model is presented of the CdiA-CT based on the structure of the XendoU nuclease from *Xenopus laevis*. Molecular-docking simulations predict that the CdiA-CT toxin active site is occluded upon binding to the CdiI immunity protein. Together, these observations suggest that the immunity protein neutralizes toxin activity by preventing access to RNA substrates.

Introduction

Bacteria have developed several complex mechanisms to interact and communicate with neighboring microbes in the environment. One such mechanism is contact-dependent growth inhibition (CDI), a form of interbacterial competition found in several important human pathogens including uropathogenic *Escherichia coli*, *Burkholderia pseudomallei* and *Neisseria meningitidis* (1, 2). CDI is mediated by the CdiB/CdiA family of two-partner secretion proteins. CdiB is an outer membrane β-barrel protein that exports and displays the CdiA effector protein on the surface of CDI^+ inhibitor cells (3). CdiA proteins are very large, ranging from 180 kDa to over 600 kDa depending on the bacterial species, and are characterized by hemagglutinin-peptide repeats that suggest a filamentous structure (4). CdiA proteins are predicted to extend several hundred angstroms from the surface of inhibitor cells to interact with specific receptors on the surface of susceptible target bacteria. Upon contact with its receptor, CdiA delivers a toxin domain derived from its extreme C-terminus (CdiA-CT) into the target (1, 5) . CdiA-CT toxins vary considerably between bacteria and even between different strains of the same species (1). This sequence diversity corresponds to a variety of toxin activities ranging from the formation of membrane pores to the degradation of ribosomal RNA (6). CDI^+ bacteria protect themselves from auto-inhibition by expressing CdiI immunity proteins, which bind to the CdiA-CT domain and neutralize its toxin activity. CdiI immunity proteins are specific for their cognate CdiA-CT and do not protect cells from the toxins of other CDI^+ bacteria. Thus, CDI systems encode a complex network of toxin–immunity protein pairs that are deployed for intercellular competition.

N. meningitidis is a parasitic, aerobic, Gram-negative bacterium responsible for pyogenic meningitis and meningococcal septicemia. It is a major cause of disease worldwide, resulting in

hearing loss, brain damage and death in 4–10% of sufferers (7, 8). Every year, approximately 3000–4000 cases of *N. meningitidis*-linked meningitis are reported in the United States (7). Because this pathogen poses a serious threat to global health, a greater understanding of its growth control could be leveraged to develop novel therapeutics targeted specifically to *Neisseria*. All *N. meningitidis* isolates carry at least one CDI system, and some strains have multiple complex loci that contain two *cdiA* genes and tandem arrays of `orphan' *cdiA-CT*/*cdiI* gene pairs (9, 10). Orphan *cdiA-CT* gene fragments often share significant regions of homology with the upstream *cdiA* gene and therefore can undergo homologous recombination to fuse the orphan *cdiA-CT*/*cdiI* module onto *cdiA*. This process can abruptly change the toxin deployed by the cell (11). The large number of CDI-associated toxin/immunity genes carried by *N. meningitidis* suggests that these systems mediate interstrain competition. This hypothesis is supported by a recent study by Tommassen and coworkers (12). Here, we report the crystal structure of $CdiI_{o2}^{MC58\text{-}1}$, an orphan CDI immunity protein from *N. meningitidis* MC58. In addition, we have generated structural models for the cognate CdiA-$CT_{o2}^{MC58\text{-}1}$ toxin and its corresponding toxin–immunity protein complex.

Results

Overall structure of $CdiI_{o2}^{MC58\text{-}1}$

$CdiI_{o2}^{MC58\text{-}1}$ crystallized in space group $P2_1$ with two molecules in the asymmetric unit. The structure was determined using a bromide derivative and SAD phasing, and the final model was refined to a resolution of 1.45 Å (**Table A.1**).

The $CdiI_{o2}^{MC58-1}$ structure represents an α/β-fold comprising a four-stranded antiparallel β-sheet, against which a five-helix bundle is packed (**Figure A.1A**). The helical bundle includes three helices (α1–α3) from the N-terminus and two (α5 and α6) from the C-terminus. Additionally, there are two consecutive helices with one short 3_{10}-helix (colored blue in **Figure A.1A**) followed by another short α-helix (α4) within a large loop that connects strands β2 and β3. Helix α4 is located on the edge of the β-sheet on the face opposite to the helical bundle and close to the C-terminus. With the exception of the surface interacting with the large loop, the remainder of one face of the β-sheet is completely exposed to solvent.

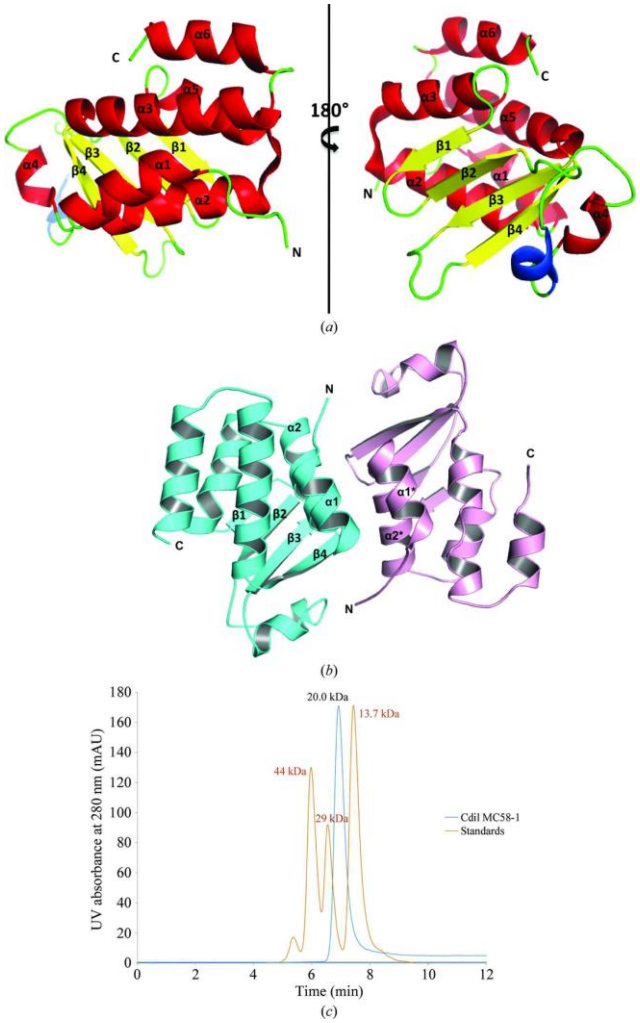


Figure A.1 (Courtesy of Joachimiak lab): (**A**) Ribbon cartoon of the $CdiI_{o2}^{MC58-1}$ structure with α-helices, β-strands and the 3_{10}-helix colored red, yellow and blue, respectively. (**B**) The interface produced by helices α1 and α2 from each monomer in the asymmetric unit. (**C**) Analytical size-exclusion chromatography using an SRT-SEC-150 column suggests that $CdiI_{o2}^{MC58-1}$ (blue trace; standards, orange trace) is monomeric in solution

Within the asymmetric unit, $CdiI_{o2}^{MC58-1}$ appears to form a nearly perfect noncrystallographic twofold-symmetric dimer, with helices α1 and α2 from each monomer packed against each other in an antiparallel mode (**Figure A.1B**). This type of helical bundle is a common structural motif at protein–protein interfaces (30). The buried surface area owing to dimerization is about 1065 $Å^2$ per monomer as determined by *PDBePISA* (31). To test whether

$CdiI_{o2}^{MC58\text{-}1}$ is dimeric in solution, we analyzed the immunity protein by analyticalsize-exclusion chromatography and found it to be predominately monomeric (**Figure A.1C**). Thus, the dimeric assembly observed in the crystal structure is perhaps an artifact of crystallization.

Structural comparison

$CdiI_{o2}^{MC58\text{-}1}$ has moderate structural similarity to two eukaryotic nucleic acid-binding proteins. The closest structural homolog, as determined using the *DALI* server (32, 33), is a mitochondrial Whirly protein (Why2; PDB entry 4kop) from *Arabidopsis thaliana* (34), and $CdiI_{o2}^{MC58\text{-}1}$ superimposes upon Why2 with an r.m.s.d. of 2.5 Å over 80 of 145 C^{α} atoms (**Figure A.2A, A.2B, and A.2D**). Whirly family members are single-stranded DNA-binding proteins that modulate DNA repair in plant chloroplasts (35, 36). The next closest structural homolog is mitochondrial RNA-binding protein 1 (MRP1; PDB entry 2gia) from *Trypanosoma brucei* (37), and MRP1 superimposes onto $CdiI_{o2}^{MC58\text{-}1}$ with an r.m.s.d. of 3.0 Å over 85 of 132 C^{α} atoms (**Figure A.2C and A.2E**). MRP1 forms a heterotetramer with MRP2, and together the two proteins function in RNA editing by promoting the hybridization to guide RNAs to their target mRNAs (38). Although the top structural homologs are nucleic acid-binding proteins, $CdiI_{o2}^{MC58\text{-}1}$ lacks the structural elements used by these proteins to bind DNA or RNA. These observations suggest that $CdiI_{o2}^{MC58\text{-}1}$ is unlikely to bind nucleic acids.

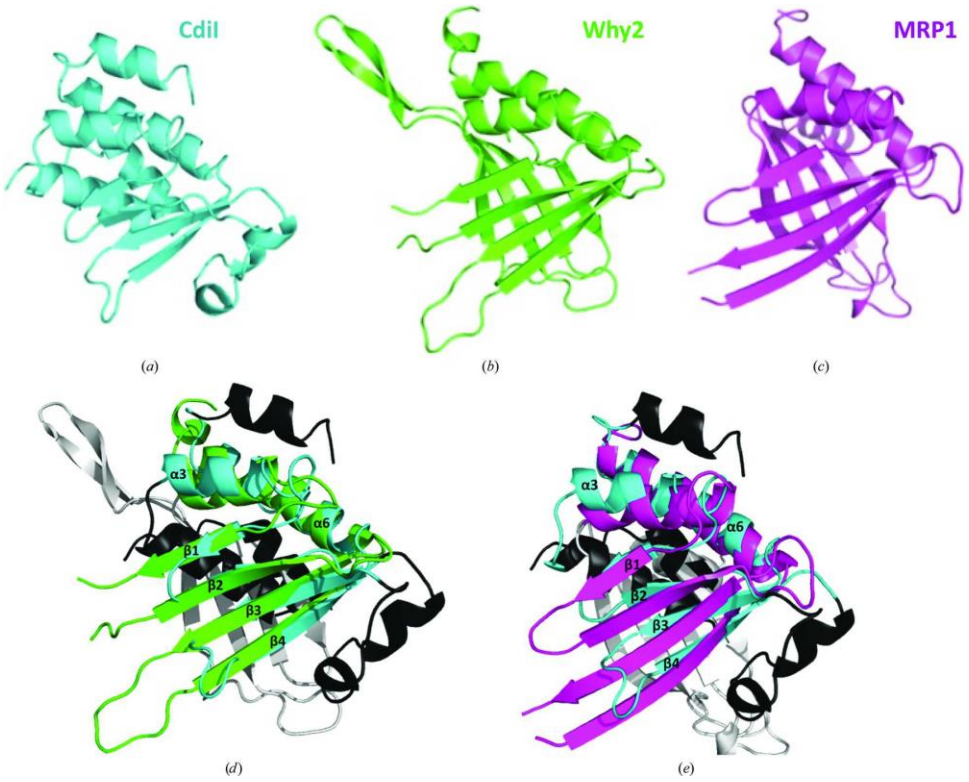


Figure A.2 (Courtesy of Joachimiak lab): $CdiI_{o2}^{MC58-1}$ structural homologs in ribbon representation. (**A**) $CdiI_{o2}^{MC58-1}$ (PDB entry 4q7o), (**B**) Why2 (PDB entry 4kop) and (**C**) MRP1 (PDB entry 2gia). (**D**) Superimposition of $CdiI_{o2}^{MC58-1}$ and Why2. Secondary-structure elements of $CdiI_{o2}^{MC58-1}$ and Why2 that superimpose with good agreement are colored cyan and green, respectively. Structural elements that do not superimpose are colored black and white for $CdiI_{o2}^{MC58-1}$ and Why2, respectively. (**E**) Superimposition of $CdiI_{o2}^{MC58-1}$ and MRP1. Secondary-structure elements of $CdiI_{o2}^{MC58-1}$ and MRP1 that superimpose with good agreement are colored cyan and pink, respectively. Structural elements that do not superimpose are colored black and white for $CdiI_{o2}^{MC58-1}$ and MRP1, respectively.

Predicted function of the CdiA-CT_{o2}^{MC58-1} toxin

No experimental or structural information is available for CdiA-CT_{o2}^{MC58-1}, which is the predicted toxin encoded by the adjacent NMB0502 gene. Aravind and coworkers have predicted that the C-terminal domain of CdiA-CT_{o2}^{MC58-1} is related to the EndoU nucleases (39), which comprise a superfamily of Mn^{2+}-dependent RNA-processing enzymes found mostly in eukaryotes, although family members are also found in the cyanobacterium *Nostoc punctiforme* (24) and at the C-terminus of a MafB RNase toxin from *N. meningitidis (40)*. CdiA-CT_{o2}^{MC58-1} has diverged substantially from the eukaryotic enzymes and shares only 16% sequence identity with the C-terminal nuclease domain (residues Ile142–Tyr292) of XendoU, a poly(U)-specific endoribonuclease from *Xenopus laevis* (**Figure A.3A**) (41). The crystal structure of XendoU has been solved (24), and therefore we used it as a guide to generate a model of CdiA-CT_{o2}^{MC58-1} (**Figure A.3C**). The resulting model shows that the XendoU active-site residues His162, His178 and Lys224 correspond to His504, His521 and Lys563 in CdiA-CT_{o2}^{MC58-1} (24, 42). Together, these observations suggest that CdiA-CT_{o2}^{MC58-1} may possess a similar Mn^{2+}-dependent RNA-processing/degrading activity to other members of the EndoU family.

Modeling of the CdiA-CT_{o2}^{MC58-1}–$CdiI_{o2}^{MC58-1}$ complex

To gain insight into the toxin–immunity protein binding interactions, we first tested whether CdiA-CT_{o2}^{MC58-1} forms a complex with $CdiI_{o2}^{MC58-1}$. We co-expressed the toxin with His_6-tagged $CdiI_{o2}^{MC58-1}$ and then purified the immunity protein by Ni^{2+}-affinity chromatography. The untagged toxin co-eluted with His_6-tagged $CdiI_{o2}^{MC58-1}$ (**Figure A.3B**), indicating that CdiA-

CT_{o2}^{MC58-1} and $CdiI_{o2}^{MC58-1}$ do indeed form a complex. We then conducted docking simulations of the monomeric $CdiI_{o2}^{MC58-1}$ immunity protein structure onto the CdiA-CT_{o2}^{MC58-1} model. The

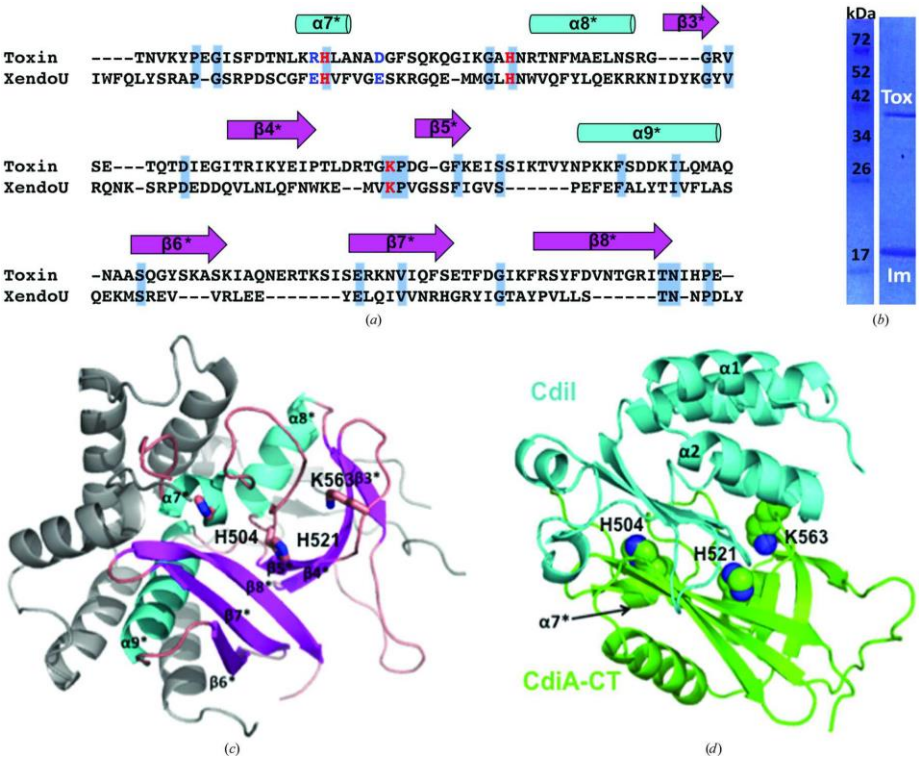


Figure A.3: Modeled structures of CdiA-CT_{o2}^{MC58-1} and its complex with $CdiI_{o2}^{MC58-1}$. (**A**) Sequence alignment by*ClustalW* of the C-terminal region of CdiA-CT_{o2}^{MC58-1} with the C-terminal domain of XendoU (PDB entry2c1w). Residues on a gray/blue background are conserved, catalytic residues are colored red and RNA-binding residues are colored blue. (**B**) SDS–PAGE of the Ni^{2+}-affinity purified complex of untagged CdiA-CT_{o2}^{MC58-1} (Tox) with His_6-tagged $CdiI_{o2}^{MC58-1}$ (Im). (**C**) Modeled structure of the C-terminal domain of CdiA-CT_{o2}^{MC58-1} with α-helices colored cyan and β-strands colored salmon. Elements in light gray are those of the N-terminal domain of the XendoU structure that have no sequence homology to CdiA-CT_{o2}^{MC58-1}. The predicted active-site residues of CdiA-CT_{o2}^{MC58-1} are shown in stick representation with C, O and N atoms colored pink, red and blue, respectively. (**D**) The docked complex structure in ribbon representation of the toxin (colored green, secondary-structure elements indicated with asterisks) with the immunity protein (colored cyan). The predicted CdiA-CT_{o2}^{MC58-1} active-site residues, shown in sphere representation with C and N atoms colored green and blue, respectively, are occluded from solvent by the docked immunity protein $CdiI_{o2}^{MC58-1}$.

lowest interaction energy obtained from these simulations was -776.7 kcal mol^{-1}, which is considerably lower than the energy calculated for non-interacting proteins. The *Hex*-generated

model predicts that $CdiI_{o2}^{MC58-1}$ binds directly over the active site of CdiA-CT_{o2}^{MC58-1} (**Figure A.3D**), likely neutralizing the toxin by preventing access to RNA substrates.

Discussion

We have elucidated the structure of $CdiI_{o2}^{MC58-1}$, a predicted immunity protein encoded within the CDI-1 locus of *N. meningitidis* MC58. $CdiI_{o2}^{MC58-1}$ has moderate structural homology to Whirly-like proteins found in plastids, but appears to lack the characteristic Whirly RNA-binding site. In addition, we modeled the structure of the associated CdiA-CT_{o2}^{MC58-1} toxin domain, which is proposed to have a similar active-site motif and RNA-processing activity as eukaryotic EndoU nucleases. Molecular-docking simulations predict that $CdiI_{o2}^{MC58-1}$ occludes the active site of CdiA-CT_{o2}^{MC58-1} in the toxin–immunity protein complex. Experiments to test the biochemical activity of CdiA-CT_{o2}^{MC58-1} and its proposed active-site residues are under way.

Materials and methods

Cloning of the N. meningitidis CdiA-CT_{o2}^{MC58-1}/$CdiI_{o2}^{MC58-1}$ genes

A fragment containing NMB0502 and NBM0503 (encoding CdiA-CT_{o2}^{MC58-1} and $CdiI_{o2}^{MC58-1}$, respectively) was amplified from *N. meningitidis* MC58 genomic DNA using 5′ -GTC TCT CCC ATG GTG AAA AAT AAT CAG CTT AGC GAC AAA GAG as the forward primer and 5′ -TGG TGG TGC CCA GCG GTT TCA TGC AGG CTA CAG TTT GTT TGA as the reverse primer. The gel-purified PCR product was treated with phage T4 DNA polymerase and dTTP as described previously (13) and ligated to plasmid pMCSG58, which appends a noncleavable His_6 tag to the

C-terminus of $CdiI_{o2}^{MC58-1}$ (14). The identity of the cloned insert was confirmed by DNA sequencing.

Expression and purification of N. meningitidis $CdiI_{o2}^{MC58-1}$

The construct was introduced into *E. coli* BL21 (DE3) cells for overexpression and protein purification. The cells were grown at 37°C in LB medium supplemented with 100 µg ml^{-1} ampicillin. After the cells had grown to an optical density at 600 nm of ∼0.6, the culture was cooled to 18°C and protein expression was induced with 0.5 m*M* isopropyl β-D-1-thio-galactopyranoside (IPTG) overnight. Under these growth conditions, only the $CdiI_{o2}^{MC58-1}$ immunity protein was overproduced. The cells were harvested by centrifugation, resuspended in 50 m*M* Tris pH 8.0, 500 m*M* NaCl, 10 m*M* β-mercaptoethanol (BME), 10% glycerol and lysed with Fast Break reagent (Promega) containing 10 µg ml^{-1} lysozyme and protease-inhibitor cocktail (Roche). The cell lysate was centrifuged at 10 000 rev min^{-1} for 1 h and the supernatant was passed through a 0.22 µm filter. The clarified lysate was then loaded onto an Ni^{2+}-Sepharose HisTrap column (GE Healthcare) and proteins were eluted with a 20–250 m*M* linear gradient of imidazole in resuspension buffer. Fractions were pooled and loaded onto a HiLoad 26/60 Superdex 75 size-exclusion column equilibrated with 20 m*M* Tris pH 7.5, 150 m*M* NaCl, 2 m*M* dithiothreitol. Fractions containing purified $CdiI_{o2}^{MC58-1}$ immunity protein were pooled and concentrated for crystallization using an Amicon Ultra centrifugal filter device with a 3000 Da cutoff (Millipore).

Size-exclusion chromatography

Analysis of the purified $CdiI_{o2}^{MC58\text{-}1}$ was performed using a Dionex HPLC system with an analytical size-exclusion column from Sepax (SRT-SEC-150, Sepax Technologies). $CdiI_{o2}^{MC58\text{-}1}$ was diluted to 5 mg ml^{-1} in standard running buffer (20 m*M* Tris pH 7.8, 150 m*M* NaCl, 2 m*M* dithiothreitol). The sample-injection volume was 20 µl and the flow rate of the analysis was 1.0 ml min^{-1}. $CdiI_{o2}^{MC58\text{-}1}$ was run in duplicate. Each run took approximately 15 min. The molecular-weight determination of $CdiI_{o2}^{MC58\text{-}1}$ was calculated using linear regression data analysis with ovalbumin (44 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa) as migration standards.

Crystallization of the $CdiI_{o2}^{MC58\text{-}1}$ immunity protein

Native $CdiI_{o2}^{MC58\text{-}1}$ crystals were grown at 4°C using sitting drops that consisted of 10 mg ml^{-1} protein in 0.2 *M* $MgCl_2$, 0.1 *M* bis-tris pH 5.5, 20% PEG 3350. Bromide derivatives were prepared by dipping crystals into a solution of 1.0 *M* KBr, 0.2 *M* $MgCl_2$, 0.1 *M* bis-tris pH 5.5, 20% PEG 3350, 15% glycerol for approximately 10 s. Bromide-derivatized crystals were subsequently cryocooled in liquid nitrogen and used to collect X-ray diffraction data for phase determination (15).

X-ray data collection, structure determination and refinement

A set of single-wavelength anomalous diffraction (SAD) data was collected near the bromine absorption peak (12.40 keV) at 100 K from one $CdiI_{o2}^{MC58\text{-}1}$ crystal. Data were obtained on the 19-ID beamline of the Structural Biology Center at the Advanced Photon Source at Argonne National Laboratory using the *SBCcollect* program (16). Data-set intensities were integrated, scaled and merged using the *HKL*-3000 program suite (17) (**Table A.1**). From the Matthews correlation coefficient, two $CdiI_{o2}^{MC58\text{-}1}$ molecules were predicted in one asymmetric unit. 17 Br sites were located using *SHELXD* (18) and were used for phasing with *MLPHARE* from *CCP*4 (19). After density modification, a partial model of 138 residues (46% of a dimer) without side chains was built in three cycles of *ARP*/*wARP* model building (20). All of the abovementioned programs are integrated within the *HKL*-3000 suite (17). The final $CdiI_{o2}^{MC58\text{-}1}$ model was completed manually using *Coot* (21) and was refined with *phenix.refine* (22) (**Table A.1**).

Table A.1 Data-collection and crystallographic statistics for $CdiI_{o2}^{MC58\text{-}1}$

Data collection	
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 45.38, b = 53.53, c = 59.70, β = 98.04
Molecular weight† (Da)	16727
No. of residues†	143
Molecules in asymmetric unit	2
Wavelength (Å)	0.9193 [Br peak]
Resolution (Å)	30.0–1.45 (1.48–1.45)
No. of unique reflections	49973‡
Multiplicity	3.4 (2.2)
Completeness (%)	99.5 (94.9)
R_{merge} (%)	10.7 (38.4)
$\langle I/\sigma(I)\rangle$	24.4 (2.3)
Solvent content (%)	41.7
Phasing	
R_{Cullis} (anomalous) (%)	89
Figure of merit (%)	19.4§
Refinement	
Resolution (Å)	30.0–1.45

No. of reflections (work/test)	47390/2544
R_{cryst}/R_{free} (%)	14.9/18.9
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.005
Bond angles (°)	0.967
No. of atoms	
Protein	2229
Heteroatoms	300
Mean *B* value (Å^2)	
Main chain	11.25
Side chain	14.24
Ramachandran plot statistics, residues in (%)	
Most favored regions	93.0
Additional allowed region	7.0
Generously allowed regions	0
Disallowed region	0
PDB code	4q7o

†Not including a three-residue N-terminal tag, SNA.
‡Including Bijvoet pairs.
§Before density modification.

Expression and purification of the N. meningitidis CdiA-CT$_{o2}^{MC58-1}$–CdiI$_{o2}^{MC58-1}$ complex

The construct was introduced into *E. coli* BL21 (DE3) cells and grown at 37°C in LB medium supplemented with 100 µg ml^{-1} ampicillin. After the cells had grown to an optical density at 600 nm of ∼0.8, protein expression was induced with 1.5 m*M* IPTG for 2.5 h at 37°C. The cells were harvested by centrifugation and resuspended in 20 m*M* Tris pH 8.0, 150 m*M*NaCl, 10 m*M* BME, 1 m*M* phenylmethylsulfonyl fluoride, 10 µg ml^{-1} lysozyme. The cells were lysed using a microfluidizer and centrifuged at 10 000 rev min^{-1} for 1 h and the supernatant was passed through a 0.22 µm filter. The clarified lysate was then loaded onto Ni^{2+}–NTA resin (GE Healthcare) and nonspecifically bound proteins were eluted with resuspension buffer with no BME and 20 m*M* imidazole under gravity. The imidazole concentration was increased to 250 m*M* to elute the toxin–immunity protein complex.

Docking of predicted toxin and immunity proteins

A model of the CdiA-$CT_{o2}^{MC58\text{-}1}$–$CdiI_{o2}^{MC58\text{-}1}$ binding interaction was generated through docking simulations. A computational model of the CdiA-$CT_{o2}^{MC58\text{-}1}$ structure was generated with *Sculptor* (23) using the three-dimensional structure of the homologous XendoU nuclease (PDB entry 2c1w) (24) as a guide. The CdiA-$CT_{o2}^{MC58\text{-}1}$sequence was fitted into the XendoU structure while maintaining the overall fold and alignment of the predicted active-site histidine residues.

Hex 8.0 (25, 26) was used to dock the $CdiI_{o2}^{MC58\text{-}1}$ immunity protein onto the *Sculptor*-modeled CdiA-$CT_{o2}^{MC58\text{-}1}$ structure. The proteins were oriented and the origins were set to allow free rotation of the two molecules during the search for low-energy binding interactions based on complementary shape and electrostatics. Energies for each model were calculated by adding all intermolecular interactions after a round of molecular-mechanics energy minimization. Typical Hex simulations produce binding energies of −600 to −1000 kcal mol^{-1}(27). A control docking simulation using CdiA-CT^{EC536} toxin and CysK, which are known to interact (28), generated a low energy of interaction of −988 kcal mol^{-1}. Simulations for proteins that do not interact (*E. coli* CysK and *B. pseudomallei* E479 CdiI) (29) yielded a much higher energy of −368 kcal mol^{-1}.

References

1. Aoki SK, *et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* 468(7322):439-442.

2. Aoki SK, Poole SJ, Hayes CS, & Low DA (2011) Toxin on a stick: modular CDI toxin delivery systems play roles in bacterial competition. *Virulence* 2(4):356-359.
3. Aoki SK, *et al.* (2005) Contact-dependent inhibition of growth in Escherichia coli. *Science* 309(5738):1245-1248.
4. Kajava AV, *et al.* (2001) Beta-helix model for the filamentous haemagglutinin adhesin of Bordetella pertussis and related bacterial secretory proteins. *Molecular microbiology* 42(2):279-292.
5. Ruhe ZC, Nguyen JY, Beck CM, Low DA, & Hayes CS (2014) The proton-motive force is required for translocation of CDI toxins across the inner membrane of target bacteria. *Molecular microbiology* 94(2):466-481.
6. Ruhe ZC, Low DA, & Hayes CS (2013) Bacterial contact-dependent growth inhibition. *Trends in microbiology* 21(5):230-237.
7. Thigpen MC, *et al.* (2011) Bacterial meningitis in the United States, 1998-2007. *The New England journal of medicine* 364(21):2016-2025.
8. Nikulin J, Panzner U, Frosch M, & Schubert-Unkmeir A (2006) Intracellular survival and replication of Neisseria meningitidis in human brain microvascular endothelial cells. *International journal of medical microbiology : IJMM* 296(8):553-558.
9. Bentley SD, *et al.* (2007) Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS genetics* 3(2):230-240.
10. Poole SJ, *et al.* (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS genetics* 7(8):e1002217.
11. Koskiniemi S, *et al.* (2014) Selection of orphan Rhs toxin expression in evolved Salmonella enterica serovar Typhimurium. *PLoS genetics* 10(3):e1004255.
12. 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.
13. Eschenfeldt WH, *et al.* (2010) Cleavable C-terminal His-tag vectors for structure determination. *Journal of structural and functional genomics* 11(1):31-39.
14. Eschenfeldt WH, *et al.* (2013) New LIC vectors for production of proteins from genes containing rare codons. *Journal of structural and functional genomics* 14(4):135-144.
15. Dauter Z, Dauter M, & Rajashankar KR (2000) Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta crystallographica. Section D, Biological crystallography* 56(Pt 2):232-237.
16. Rosenbaum G, *et al.* (2006) The Structural Biology Center 19ID undulator beamline: facility specifications and protein crystallographic results. *Journal of synchrotron radiation* 13(Pt 1):30-45.
17. 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* 62:859-866.
18. Sheldrick GM (2008) A short history of SHELX. *Acta crystallographica. Section A, Foundations of crystallography* 64(Pt 1):112-122.
19. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* 67(Pt 4):235-242.

20. Cohen SX, *et al.* (2004) Towards complete validated models in the next generation of ARP/wARP. *Acta crystallographica. Section D, Biological crystallography* 60(Pt 12 Pt 1):2222-2229.
21. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* 60(Pt 12 Pt 1):2126-2132.
22. Afonine PV, *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* 68(Pt 4):352-367.
23. Birmanns S, Rusu M, & Wriggers W (2011) Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes. *J Struct Biol* 173(3):428-435.
24. Renzi F, *et al.* (2006) The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *Proceedings of the National Academy of Sciences of the United States of America* 103(33):12365-12370.
25. Ritchie DW (2003) Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins* 52(1):98-106.
26. Ritchie DW & Venkatraman V (2010) Ultra-fast FFT protein docking on graphics processors. *Bioinformatics* 26(19):2398-2405.
27. Gupta UK, Mahanta S, & Paul S (2013) In silico design of small peptide-based Hsp90 inhibitor: a novel anticancer agent. *Medical hypotheses* 81(5):853-861.
28. Diner EJ, Beck CM, Webb JS, Low DA, & Hayes CS (2012) Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes & development* 26(5):515-525.
29. Nikolakakis K, *et al.* (2012) The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Molecular microbiology* 84(3):516-529.
30. Norel R, Lin SL, Wolfson HJ, & Nussinov R (1995) Molecular-Surface Complementarity at Protein-Protein Interfaces - the Critical Role Played by Surface Normals at Well Placed, Sparse, Points in Docking. *Journal of molecular biology* 252(2):263-273.
31. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* 372(3):774-797.
32. Holm L, Kaariainen S, Rosenstrom P, & Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24(23):2780-2781.
33. Holm L & Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic acids research* 38(Web Server issue):W545-549.
34. Cappadocia L, Parent JS, Sygusch J, & Brisson N (2013) A family portrait: structural comparison of the Whirly proteins from Arabidopsis thaliana and Solanum tuberosum. *Acta crystallographica. Section F, Structural biology and crystallization communications* 69(Pt 11):1207-1211.
35. Cappadocia L, *et al.* (2010) Crystal Structures of DNA-Whirly Complexes and Their Role in Arabidopsis Organelle Genome Repair. *Plant Cell* 22(6):1849-1867.
36. Desveaux D, Marechal A, & Brisson N (2005) Whirly transcription factors: defense gene regulation and beyond. *Trends in plant science* 10(2):95-102.
37. Schumacher MA, Karamooz E, Zikova A, Trantirek L, & Lukes J (2006) Crystal structures of T. brucei MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* 126(4):701-711.

38. Aphasizhev R, Aphasizheva I, Nelson RE, & Simpson L (2003) A 100-kD complex of two RNA-binding proteins from mitochondria of Leishmania tarentolae catalyzes RNA annealing and interacts with several RNA editing components. *Rna-a Publication of the Rna Society* 9(1):62-76.
39. Zhang D, de Souza RF, Anantharaman V, Iyer LM, & 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.
40. Jamet A & Nassif X (2015) New players in the toxin field: polymorphic toxin systems in bacteria. *mBio* 6(3):e00285-00215.
41. Laneve P, *et al.* (2003) Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in Xenopus laevis. *The Journal of biological chemistry* 278(15):13026-13032.
42. Gioia U, *et al.* (2005) Functional characterization of XendoU, the endoribonuclease involved in small nucleolar RNA biosynthesis. *The Journal of biological chemistry* 280(19):18996-19002.