

# UC Santa Barbara

## UC Santa Barbara Electronic Theses and Dissertations

### Title

Mechanism of toxin activity and delivery in bacterial contact-dependent competition systems

### Permalink

<https://escholarship.org/uc/item/3dk7b8t4>

### Author

Beck, Christina Marie

### Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Mechanism of toxin activity and delivery in bacterial contact-dependent competition  
systems

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Molecular, Cellular, and Developmental Biology

by

Christina Marie Beck

Committee in charge:

Professor Christopher S. Hayes, Chair

Professor David A. Low

Professor Dzwokai Zach Ma

Professor Charles E. Samuel

June 2015

The dissertation of Christina Marie Beck is approved.

---

Charles E. Samuel

---

Dzwokai Zach Ma

---

David A. Low

---

Christopher S. Hayes, Committee Chair

June 2015

Mechanism of toxin activity and delivery in bacterial contact-dependent competition systems

Copyright © 2015

by

Christina Marie Beck

## ACKNOWLEDGEMENTS

I would first like to thank Elie Diner, my graduate student mentor when I joined the Hayes lab as an undergraduate student in June 2010. His work ethic, dedication, and enthusiasm for basic science inspired me to earn my PhD. I want to express my appreciation to my advisor, Chris Hayes, for supporting my ideas and letting me direct my own projects, but also for allowing me to take time out of lab to apply to medical school. I am grateful for the strong collaboration we have with Professor David Low, whose passion for science encouraged me to explore the real-life significance of my results. Every member of the Hayes and Low lab - past and present - was always willing to help me with my research. Thanks to each of you. I would especially like to thank Zach Ruhe for his technical advice, and Fernando Garza-Sanchez for his emotional support. Dr. Garza would drop everything he was doing to help me with any issue I was having, lab-related or not.

I would also like to thank my family and friends for reminding me there is such thing as “life outside of lab”. Thank you to my brother, Paul, for serving in the military for the past five years; I appreciate the sacrifices you made while I was having fun in graduate school. And finally a special thank you to my Santa Barbara family, including the boys at 6812 Pasado, the club soccer girls, and my 8 roommates at 40 Valdez Ave, for accepting me into their lives and supporting my endeavors.

I dedicate this thesis to my Uncle Tom who continues to inspire me to succeed at whatever I am doing in life, but also to allow enough time to have some fun on the side. I feel this sums up graduate school perfectly.

## Curriculum Vitae

**Christina Beck**

University of California, Santa Barbara  
Department of Molecular, Cellular, and Developmental Biology

### **Education:**

*University of California, Santa Barbara: 2011-2015*

Doctor of Philosophy in Molecular, Cellular, and Developmental Biology  
GPA 3.96

*University of California, Santa Barbara: 2007-2011*

Bachelor of Science in Cell and Developmental Biology  
Overall GPA 3.77 (College of Letters and Science High Honors)  
Major GPA 3.85

### **Awards and Honors:**

*Molecular Genetics of Bacteria and Phages Meeting*

Best Poster Award, 2014

*University of California, Santa Barbara - Department of Molecular, Cellular, and Developmental Biology*

Jean Devlin Fellowship, 2014  
Merit Fellowship, 2014  
Dean's Fellowship, 2014  
Chang Fellowship, 2014

*University of California, Santa Barbara - Recreation Sports*

Student-Athlete of the Year Award, 2012

*University of California, Santa Barbara - College of Letters and Science Undergraduate Research Colloquium*

Poster Presentation- 3<sup>rd</sup> Prize, 2011

*University of California, Santa Barbara*

Research Distinction in Major, 2011  
Jean Hodgkin's Memorial Scholarship, 2010  
Dean's List, Winter 2009, Spring 2009, Fall 2009, Fall 2010, Winter 2010

## **Research Experience:**

Doctoral Research:

*University of California, Santa Barbara*: 2011-2015

Research Advisor: Christopher Hayes, Ph.D.

- Mechanism of toxin delivery in contact-dependent growth inhibition (CDI) systems in *gammaproteobacteria*
- Characterization of toxin-antitoxin pairs in CDI systems
- Type VI secretion systems in *Enterobacter* species

Undergraduate Research:

*University of California, Santa Barbara*: 2010-2011

Research Advisor: Christopher Hayes, Ph.D.

- F pilus mediated transport of contact-dependent growth inhibition (CDI) toxins into *Escherichia coli* K12

## **Publications:**

Campanini B, Benoni R, Bettati S, **Beck CM**, Hayes CS, Mozzarelli A. "Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein." *Biochim Biophys Acta*. 2015 Feb 27. pii: S1570-9639(15)00049-7.

Ruhe ZC, Nguyen J, **Beck CM**, Low DA, Hayes CS. "The proton-motive force is required for translocation of CDI toxins across the inner membrane of target bacteria." *Molecular Microbiology*. 2014 Oct;94(2):466-81.

**Beck CM**, Diner EJ, Kim JJ, Low DA, Hayes CS. "F pili mediate import of a contact-dependent growth inhibition (CDI) toxin." *Molecular Microbiology*. 2014 Jul;93(2):276-90.

**Beck CM**, Morse RP, Cunningham DA, Low DA, Goulding CW, Hayes CS. "The CdiA protein of *Enterobacter cloacae* deploys a toxic ribosomal RNase." *Structure*. 2014 May 6;22(5):707-18.

Whitney JC, **Beck CM**, Goo YA, Russell AB, Harding B, De Leon JA, Cunningham DA, Tran BQ, Low DA, Goodlett DR, Hayes CS, Mougous JD. "Genetically distinct pathways guide effector export through the type VI secretion system." *Molecular Microbiology*. 2014 May;92(3):529-42.

Diner EJ, **Beck CM**, Webb JS, Low DA, Hayes CS. "Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI)." *Genes and Development*. 2012 Mar 1;26(5):515-25.

## **Presentations:**

### *Oral presentations:*

**Christina Beck**, Elie Diner and Christopher Hayes. “The F pilus mediates import of a contact-dependent growth inhibition (CDI) toxin.” Molecular Genetics of Bacteria and Phages Meeting. Madison, WI - August 2011

### *Poster presentations:*

**Christina Beck**, Robert Morse, Celia Goulding, Christopher Hayes. “Contact-dependent growth inhibition systems deliver toxins that inhibit bacterial translation.” Molecular Genetics of Bacteria and Phages Meeting. Madison, WI - August 2014

**Christina Beck**, David Cunningham, David Low and Christopher Hayes. “Receptor polymorphism restricts contact-dependent growth inhibition (CDI).” Molecular Genetics of Bacteria and Phages Meeting. Madison, WI - August 2013

**Christina Beck**, Elie Diner and Christopher Hayes. “The F pilus mediates import of a contact-dependent growth inhibition (CDI) toxin.” College of Letters and Science Undergraduate Research Colloquium. UCSB – May 2011

## **Teaching Experience:**

### Graduate Student Teaching Assistant:

*University of California, Santa Barbara* - Prokaryotic Genetics: Winter 2012, Biochemistry: Spring 2012, Pharmacology: Winter 2013, Animal Virology: Spring 2014

- Leading weekly discussion sections and exam review sessions

### Campus Learning Assistant Services (CLAS) Tutor:

*University of California, Santa Barbara* - Introductory Chemistry: Fall 2009-Fall 2011

- Leading biweekly group sections of undergraduate chemistry students
- Design lesson plans, worksheets, and practice exams

## **Mentoring:**

### *High School Students mentored:*

David Huang - Research Mentorship Program: June 2014-July 2014

- Sugar activation of *Enterobacter cloacae* type VI secretion system-1

Donghoon Kang - Research Mentorship Program: June 2014-July 2014

- Effector export through *Enterobacter cloacae* type VI secretion system-2

### *Undergraduate Students mentored:*



International Genetically Engineered Machine (iGEM) undergraduate students: May 2014-June 2015

- Genetically engineer *Escherichia coli* cells to express green fluorescent protein upon delivery of a contact-dependent growth inhibition (CDI) toxin

Jeff Kim - Molecular, Cellular, and Developmental Biology undergraduate student: August 2013-June 2014

- The role of contact-dependent growth inhibition (CDI) systems in kin selection in uropathogenic *E. coli*

David Cunningham - Molecular, Cellular, and Developmental Biology undergraduate student: June 2012-June 2014

- The role of outer membrane porin proteins as contact-dependent growth inhibition (CDI) receptors

*Graduate Students mentored:*

Hayley Finkelstein - Molecular, Cellular, and Developmental Biology Ph.D. student: June 2012-September 2012 (Rotation Project)

- Characterization of toxin-antitoxin pairs in contact-dependent growth inhibition (CDI) systems

Alex Fields - Molecular, Cellular, and Developmental Biology Ph.D. student: December 2011-March 2012 (Rotation Project)

- The role of contact-dependent growth inhibition (CDI) systems in biofilm formation in uropathogenic *E. coli*

## ABSTRACT

Mechanism of toxin activity and delivery in bacterial contact-dependent competition systems

by

Christina Marie Beck

Bacteria live in complex microbial communities and must face the constant challenge of limited space and resources. These harsh conditions have driven the evolution of a variety of competition and communications systems that allow bacteria to interact with the surrounding microbes and eukaryotic organisms found in their respective niches. The research presented here focuses on two such systems that Gram-negative bacteria use to deliver toxic effector molecules into neighboring bacteria. The type VI secretion system (T6SS) is a dynamic syringe-like organelle that injects toxic effectors indiscriminately into both prokaryotic and eukaryotic cells. In contrast, CDI systems are restricted to a narrow target range. CDI involves a two-partner secretion system that presents a large exoprotein, CdiA, to the surface of cells. CdiA binds to receptors on target bacteria to facilitate delivery of a toxin derived from its C-terminus.

In the introductory chapter, we will dissect what is known about each of these competition systems, including the toxin activities and delivery pathways of each system. Chapter 2 identifies a novel delivery process of a CDI toxin into *Escherichia coli*. In chapter 3 we will learn that the CDI system found in uropathogenic *E. coli* 536 is restricted to intra-strain delivery due to polymorphisms in the receptor proteins it utilizes for delivery. Chapter

4 will characterize novel CDI toxin activity while chapter 5 will explore the motive behind exploiting the metabolic enzyme, CysK, as a cofactor for a CDI toxin's activity. Finally, chapter 6 will describe the mechanism of toxin delivery in the T6SS found in the opportunistic human pathogen *Enterobacter cloacae*.

Ultimately, this thesis will explore the repertoire of toxins that are utilized by CDI and T6S systems. Chapters 3 and 5 will challenge the function of CDI as a competition system and suggests its role in kin selection. Lastly, we will examine the potential role of these competition systems in shaping microbial communities.

## TABLE OF CONTENTS

Chapter 1. Introduction .....	1
Chapter 2. F pili mediate import of a contact-dependent growth inhibition (CDI) toxin .....	29
Chapter 3. Receptor polymorphism restricts contact-dependent growth inhibition (CDI) target range.....	68
Chapter 4. Contact-dependent growth inhibition (CDI) systems deploy a toxic ribosomal RNase.....	100
Chapter 5. Exploring the role of CysK in the activation of a contact-dependent growth inhibition (CDI) toxin .....	145
Chapter 6. Characterizing Type VI secretion in <i>Enterobacter cloacae</i> .....	191
Chapter 7. Conclusion.....	235
References.....	241

## **Chapter 1. Introduction**

## General Introduction

Bacteria can be found in almost every ecological niche. Some environments, including the human intestinal tract and soil from croplands, harbor hundreds to thousands of different bacterial species<sup>1,2</sup>. In fact bacteria are rarely found in isolation, and evidence from a variety of habitats suggests the composition and diversity of bacterial species can have drastic effects on an ecosystem<sup>3-6</sup>. For example, human obesity was found to be associated with a high abundance of Firmicutes and low amount of Bacteroidetes phyla in the digestive tract while patients with inflammatory bowel disease (IBD) have a less diverse microbiota with a higher percentage of *Escherichia coli* species<sup>7-10</sup>. Given the challenge of limited space and resources in many of the environments bacteria colonize, they have developed a multitude of communication and competition systems to interact with other microbes or eukaryotic organisms. These systems can allow a simple unicellular bacterial cell to act like a multicellular organism by facilitating a sense of community and cooperation in their respective niches. Studying the mechanisms of these systems is vital if we hope to understand how they are used to shape microbial communities and thus their potential impact on an ecosystem.

In the introductory chapter, we will cover some of the systems bacteria use in intra- and inter-species communication and competition. First we will cover bacteriocins, small toxic proteins released upon cell lysis to kill neighboring bacteria. Then we will describe contact-dependent growth inhibition (CDI), a two-partner secretion system that requires physical contact with target cells to deliver a toxin domain. Lastly we will discuss the type VI secretion system (T6SS), a membrane-spanning organelle that injects toxic effector molecules into both prokaryotic and eukaryotic cells. Throughout this thesis we will try to

identify both the similarities and differences between these systems to clarify the distinct function of each system. The research presented here will explore the activities of the toxin domains encoded by these systems, the narrow target range of CDI systems, and the mechanism of T6S assembly. Collectively, this work will hopefully help elucidate the roles of these competition systems in nature.

## **Colicins**

### **Colicin Introduction**

Colicins are a well-studied example of intra-species competition in bacteria. They are soluble toxic proteins produced by certain strains of *E. coli* that get released into the environment during times of stress<sup>11,12</sup>. They specifically target *E. coli*, but producing strains do not kill their own population because they harbor an immunity protein that protects them from colicin activity<sup>13</sup>. Colicin-producing strains are abundant in nature and predominantly found in the gut of animals<sup>14</sup>. It is estimated that up to 50% of isolated strains of *E. coli* contain a colicin operon<sup>14,15</sup>. In crowded niches where resources are scarce, the stress response will induce colicin synthesis and release from a subset of the population. This will facilitate the clearing of susceptible *E. coli* strains in the area. Taken all together, colicin-harboring strains of *E. coli* are thought to have a competitive advantage in nature.

Colicins were first discovered in 1925 by Gratia, and named in 1946. They have since been discovered in many species of bacteria and are now collectively called bacteriocins. Even though they are homologous, colicins produced by other bacteria get their own label based on the name of the species producing the colicin. For example, bacteriocins produced by *Enterobacter cloacae* are called cloacins, *Yersinia pestis* produce pesticins, and so on<sup>16</sup>.

All bacteriocins behave similarly and thus the remainder of this introduction and thesis will describe and refer mostly to colicins.

Colicins are 40-80 kDa and organized into three domains; each involved with a specific process of colicin activity<sup>17-21</sup>. The middle domain is responsible for binding to a specific outer-membrane protein on the cell surface<sup>22</sup>. This leads to a restricted target range where colicins are only able to kill other *E. coli* strains that have sensitive receptors. The N-terminal domain allows for translocation through the outer membrane by targeting either the Tol or TonB machinery<sup>23</sup>. Colicins are organized into two groups based on the translocation machinery utilized by their N-terminal domain. Group A colicins (A, E1-E9, K, L, N, S4, U, and Y) are translocated through the Tol system while Group B colicins (D, H, Ia, Ib, M, 5 and 10) use the TonB system<sup>24-26</sup>. The C-terminal domain contains either pore-forming or nuclease activity which ultimately causes cell death<sup>27,28</sup>. All nucleases act in the cytoplasm targeting RNA or DNA except for colicin M which targets peptidoglycan precursors in the periplasm<sup>29-32</sup>.

### **Colicin Synthesis and Release**

Colicins are found on a colicinogenic plasmid, pCol, of which there are two types<sup>33</sup>. Type I are small multicopy plasmids that contains a mobilization element expressing only colicin genes. In contrast, type II are large single copy plasmids that carry other genes and are conjugative. Group A colicins are usually found on type I plasmids, while type II plasmids typically contain Group B colicins<sup>34</sup>. With some exceptions, colicin producing strains only harbor one specific colicinogenic plasmid. Colicins are expressed in one, two, or three gene operons under a LexA promoter<sup>35-37</sup>. During times of stress, the SOS response



will be triggered and activate the LexA promoter, thus allowing expression of the colicin operon<sup>38</sup>.

Group A colicins that contain a nuclease domain are found in three gene operons containing the colicin, followed by its immunity protein, and then a lysis protein<sup>35,39-43</sup>. The lysis protein is responsible for release of the produced colicin-immunity pair into the environment and subsequent death of the producing cell<sup>41-43</sup>. The immunity gene contains its own promoter in the open reading frame of the colicin that is constitutively on, thus protecting cells from autoinhibition or death by its siblings<sup>11,44-46</sup>. Upon SOS induction, two mRNA transcripts are produced; one that allows translation of the colicin-immunity pair only, and the other that contains mRNA for all three genes. This allows for more production of the colicin and limited amounts of the lysis protein<sup>45,47</sup>. With high enough quantities of lysis protein, the cell will lyse and release multiple copies of the colicin-immunity pair<sup>48-50</sup>. Group A pore-forming colicins are found in two gene operons because the immunity gene is found on the negative strand of the colicin<sup>51</sup>. Consequently, group A colicin release is a self-sacrificing system where some cells in the population will die to allow propagation of the remaining sibling cells.

Nuclease and pore-forming colicins in group B have the same genetic layout as group A colicins; except they do not contain a lysis gene. Therefore, the production of group B colicins does not kill the producing strain or allow release of the colicin into the environment. Interestingly, there are six secretion systems identified in Gram-negative bacteria that allow for transport of proteins into the environment without cell death. It is somewhat puzzling that colicins use this unique mechanism of export.

## Colicin Import

There are three distinct steps in colicin import. The first is binding to specific receptors on the outer membrane, the second is translocation across the cell envelope, and the third is transit through the inner membrane<sup>52-54</sup>. The first two steps are directed by the middle and N-terminal domains, respectively. Each step will be dissected in this section.

Both group A and B colicins utilize a variety of outer-membrane proteins including BtuB, OmpF, FepA, FhuA, Cir, Tsx, OmpA, and OmpW<sup>25,55,56</sup>. These receptor proteins are commonly monomeric  $\beta$ -barrels that contain extracellular loops. Crystal structures of a few colicin-bound receptors have revealed that colicins are exploiting these extracellular loops to bind target cells<sup>57,58</sup>. This binding interaction is what restricts colicins to *E. coli*, since the extracellular loops of these  $\beta$ -barrel proteins vary amongst species. While the initial binding step is similar between group A and B colicins, the subsequent translocation events differ considerably between the two colicin groups. We will first discuss group A translocation.

All group A colicins except for colicin N require another outer-membrane protein in order to initiate the translocation step. They all use OmpF with one exception. OmpF is a trimeric  $\beta$ -barrel protein that is thought to be recruited to the colicin-bound receptor protein. The interaction between the two outer-membrane proteins is predicted to form a translocon that the colicin can use to cross the membrane<sup>59</sup>. The immunity protein has been stripped off at this point so it does not enter the cell<sup>60</sup>. There are a few models of how the colicin actually crosses the cell envelope. One entails the colicin transporting through the pore of OmpF, another between the interface of OmpF and the lipid bilayer, and the last has it diffusing through the lipid bilayer itself. Regardless, the colicin translocation domain must reach into the periplasm and make contact with a component of the Tol system for transit to start<sup>61</sup>.

The Tol system consists of three inner membrane proteins, TolA, TolQ, and TolR, and a periplasmic protein, TolB<sup>62,63</sup>. Different group A colicins require distinct components of the Tol system for transport. Some colicins need every component of the Tol system, while others just need a subset<sup>64</sup>. The N-terminal domain of some group A colicins have been shown to interact directly with TolA<sup>65</sup>. Regardless of which Tol components are needed, the basic mechanism consists of the N-terminal domain reaching into the periplasm to interact with the Tol system<sup>61</sup>. Through an unknown mechanism, this allows the N-terminal and toxic domain of the colicin to enter the periplasm.

Group B colicins do not require a separate outer-membrane protein for translocation, with the exception of colicins 5 and 10. Therefore upon binding to its specific receptor protein, through an equally unknown mechanism, the N-terminal domain reaches down into the periplasm, but this time to interact with the TonB system<sup>66,67</sup>. The known components of this system are one outer-membrane protein and three inner-membrane proteins, TonB, ExbB, and ExbD<sup>68</sup>. Outer-membrane proteins involved in nutrient uptake contain a five amino acid motif called the TonB box which permits direct contact with TonB<sup>69</sup>. This interaction is thought to trigger the release of the receptor's substrate into the periplasm by harvesting energy from the proton-motive force through the TonB complex<sup>70,71</sup>. Since the N-terminal domains of group B colicins contain a TonB box, a similar mechanism of import is proposed for colicin transit.

Interestingly, the N-terminal domains of colicins can be swapped and they will still successfully enter the cell<sup>51</sup>. This suggests that even though group A and B colicins parasitize different systems, the general mechanism of crossing the outer membrane is conserved. Once inside the periplasm, the pore-forming toxins just need to be inserted into

the inner membrane to perform their toxic activity. However the nucleases must enter the cytoplasm to perform their toxic action. The mechanism of translocation across the inner membrane is even less understood than outer membrane transport. Overall, while the difference in translocation machinery is what distinguishes group A from B colicins, the general mechanism of transit into the cell is reasonably similar.

### **Colicin Toxin Domain**

The toxic action of colicins is found in the C-terminal domain which will contain either pore-forming or nuclease activity<sup>72</sup>. These domains can carry out their toxic function independent of the N-terminal and middle domains of the colicin. Immunity proteins bind to the toxic domain of its cognate colicin to prevent its lethal activity<sup>51</sup>. The structural biology of the ionophoric and nuclease colicins, as well as the mechanisms of immunity protection will be discussed in this section.

### ***Pore-formers***

There are ten pore-forming colicins. Surprisingly, they are water soluble and monomeric in producing cells which has permitted crystallization of five of them<sup>73</sup>. Each has a similar structure consisting of 10  $\alpha$ -helices packed into a bundle<sup>74-78</sup>. In the structure of colicin A, two of the  $\alpha$ -helices are hydrophobic and stay tucked inside the remaining 8 helices. It is unknown how this fold allows for insertion into the membrane upon entry into a target cell. Amazingly, it does not require the assistance of other cell proteins because *in vitro* studies have shown it can insert itself into pure lipid membranes<sup>79</sup>.

Even though the fold of each crystalized pore-forming colicin is similar, the immunity proteins are only able to protect cells from their cognate toxin. The immunity proteins

contain three or four transmembrane domains and sit in the inner membrane<sup>80-84</sup>. Swapping studies have shown that the specificity of colicin-immunity pairs comes from the ability of the immunity proteins to recognize the C-terminus of their cognate toxins<sup>23,39,85,86</sup>. The immunity protein does not prevent the toxin from inserting into the inner membrane. Rather, it is thought to prevent opening of the pore or block the pore if it does open<sup>87,88</sup>.

### ***Nucleases***

Nuclease activity is the other mechanism of killing for colicins. Colicin M targets peptidoglycan precursors in the periplasm which eventually causes cell lysis<sup>29,30</sup>. Its immunity protein has a signal sequence that anchors it to the periplasmic side of the cytoplasmic membrane to protect cells from colicin M activity<sup>51</sup>. The remaining nucleases target DNA, 16S rRNA, or tRNA in the cytoplasm<sup>51</sup>. As previously discussed, the translocation step through the inner membrane is nebulous. It is thought there is a cleavage event that allows the nuclease domain to release from the remainder of the colicin when delivered through the inner membrane<sup>89,90</sup>. In support of this model, the C-terminal domain is able to bind to and cleave its substrate independent of the translocation and receptor-binding domains. Furthermore, the immunity protein binds only to the C-terminal toxin domain of the colicin. Nevertheless, the toxic domain is able to reach the cytoplasm to either bind its substrate and cause cell death in susceptible cells, or bind its immunity protein and deactivate in protected cells.

### ***DNases***

There are four known colicins that target DNA<sup>35,91,92</sup>. The structure of two colicin DNases have been solved to show they both have a similar V-shaped fold that allows binding to the minor groove of DNA<sup>93,94</sup>. These are metal-dependent enzymes that contain an H-N-H

nuclease domain which is found in numerous nucleases in both prokaryotes and eukaryotes. This domain allows for the coordination of a divalent cation, activation of a water molecule to attack the phosphodiester backbone of DNA, and stabilization of the transition state<sup>95-97</sup>. Interestingly, the four DNase colicins are 67% identical by amino acid composition, but again the immunity proteins are specific for their cognate toxin. This is due to the sequence diversity seen at immunity binding sites. The immunities bind at an exosite to sterically hinder substrate binding, rather than directly binding to the active site cleft<sup>98</sup>.

### ***RNases***

The RNase colicins target either 16S rRNA or the anti-codon loops of specific tRNAs<sup>99-102</sup>. Both types are predicted to cleave their substrate by extracting a proton from the 2' hydroxyl of the phosphodiester backbone which promotes auto-attack and cleavage. Active-site residues help align the attack, stabilize the intermediate, and protonate the leaving group of the reaction<sup>101-105</sup>. While the mechanism of cleavage is similar between the two types of RNases, their structure and the function of their immunity proteins are distinct.

The three 16S rRNases are 80% homologous and each cleaves the 3' end of 16S rRNA in the context of 70S ribosome between nucleosides A1492 and A1493. These residues are important for the stabilization of the mRNA and anti-codon loop of the corresponding tRNA<sup>105,106</sup>. The cleavage event still allows for tRNA binding in the A-site; however the mechanism of translation inhibition and subsequent cell death is still disputed<sup>107,108</sup>. Like the DNase immunities, the 16S rRNase immunities also bind at an exosite. In this case, the negatively charged immunity protein prevents the approach to the similarly charged RNA component of the ribosome<sup>98</sup>.

Colicin D and E5 each cleave the single-stranded anticodon loop of a subset of tRNAs. Colicin D targets all four arginine tRNAs while colicin E5 targets Asn, Asp, His, and Tyr<sup>101,102</sup>. Both colicins have positively charged active sites formed by  $\beta$ -sheets and  $\alpha$ -helical bundles<sup>109,110</sup>. The tRNase immunity proteins differ from the DNase and 16S rRNase immunities in that they bind directly to the active site of their cognate toxin<sup>109,111</sup>. This is a more straight-forward mechanism of inhibition where the immunity simply blocks the site of substrate binding.

### **Colicin Conclusion**

Bacteriocins are just one of the many competition systems found in bacteria. Their synthesis and release in times of stress is an effective way to clear out sensitive strains when inhabiting a new niche, or stopping the invasion of susceptible strains in an established habitat. It has proven difficult to study the *in vivo* relevance of bacteriocins. However the mere fact that bacteriocins are extensively represented in nature suggests they provide a competitive advantage for bacteria. Furthermore, the variety of toxins found in this system provides the potential for extensive competition between strains of bacteria.

### **Contact-dependent growth inhibition (CDI)**

#### **CDI Introduction**

Contact-dependent growth inhibition (CDI) is another competition system that deploys toxins into neighboring bacteria<sup>112</sup>. It was first discovered in 2005 in *E. coli* EC93, the dominant strain of *E. coli* isolated from a rat colony. As we just learned, enteric bacteria

commonly produce soluble toxins to inhibit the growth of nearby bacteria<sup>51</sup>. However, CDI is unique in that it requires direct physical contact with target cells<sup>112</sup>.

The locus responsible for CDI activity is comprised of just three genes, *cdiB*, *cdiA* and *cdiI*. CdiB/CdiA is a two-partner secretion (TPS) system homologous to other TPS systems including FhaC/A. CdiB is a predicted  $\beta$ -barrel outer membrane protein that exports CdiA to the surface of the cell. CdiA is a large exoprotein ranging from 200-600 kDa that is predicted to form a  $\beta$ -helical structure that extends several hundred angstroms from the surface of the cell<sup>113,114</sup>. The extreme C-terminus of CdiA (CdiA-CT) encodes a toxic domain<sup>115</sup>. When CdiA comes in contact with a receptor protein, CdiA-CT gets delivered into target cells and causes growth inhibition<sup>116-119</sup>. CdiI is an immunity protein that binds to and neutralizes CdiA-CT, thus preventing CDI<sup>+</sup> cells from autoinhibition<sup>112,115</sup>.

Bioinformatic analyses have led to the discovery of CDI systems in  $\alpha$ ,  $\beta$ , and  $\gamma$ -proteobacteria. Its role in bacterial competition has since been demonstrated in multiple pathogenic species, including *Burkholderia thailandensis*, *Erwinia chrysanthemi*, and *Enterobacter cloacae*<sup>115,117,120</sup>. The prevalence of CDI systems in pathogenic bacteria, specifically its location in pathogenicity islands, suggests CDI might offer some unresolved aid to the ability of these strains to infect a host. Furthermore, given the specific target range of CDI systems, it seems plausible that these systems could also be involved in kin selection<sup>118</sup>. Insight into these speculations, as well as the mechanism of CDI will be covered in the rest of this section and throughout this thesis.

## **CDI Toxin Domain**



CdiA is conserved amongst *E. coli* until the last 200 amino acids where a conserved VENN peptide motif demarcates the toxic domain from the rest of CdiA<sup>115</sup>. The C-terminal toxin domain will be referred to as CdiA-CT for the remainder of this thesis. Consistent with this domain being polymorphic, there are numerous distinct toxic activities performed by CdiA-CTs<sup>115,117,121</sup>. Likewise, each CdiI protein is polymorphic, as only a cognate immunity protein will protect cells from CdiA-CT activity<sup>115,121</sup>. Thus far, CdiA-CTs have been identified as either pore-formers or nucleases. Interestingly, CdiA-CTs are not always found fused to CdiA. They are also found downstream of CDI systems with no ATG or GTG start site. These are called “orphans” because of their displacement from full-length CdiA and inability to be translated<sup>121</sup>. This section will discuss the various toxic activities encoded by CdiA-CTs, the mechanism of immunity protection, and provide a comparison to the toxin-immunity pairs encoded by colicins. Of note, there have been numerous publications on the predicted toxic functions of CdiA-CTs, but this section will only cover toxin activities that have been established in either *in vitro* and/or *in vivo* experiments.

### ***DNases***

The first crystal structure of a CDI DNase is from *E. coli* EC869. EC869 contains 11 orphan CdiA-CT/CdiI pairs located downstream from its CDI locus. Orphan 11 encodes for DNase activity that requires a divalent cation for activity *in vitro*. It cuts both double- and single-stranded DNA *in vitro* and seems to degrade DNA *in vivo* based on target cells lacking DAPI staining after delivery of the toxin. It has a unique interaction with its immunity protein where the two proteins bind by  $\beta$ -augmentation to produce a six-stranded antiparallel  $\beta$ -sheet. Other CDI DNases have been characterized, but not to the same extent as orphan 11.

### ***tRNases***

A cohort of CdiA-CTs cut tRNAs. However, unlike the two colicins that cut tRNA, there is a much wider variety of substrates and cut sites seen in CDI tRNases. CDI toxins from UPEC536 and *Burkholderia pseudomallei* 1026b act as general tRNases, but cleave at distinct sites in the anti-codon loop and T-loop, respectively<sup>120</sup>. In contrast, a CDI toxin from *B. pseudomallei* E479 cleaves specifically Ala1B tRNA<sup>120</sup>. The crystal structures of a handful of CDI tRNases demonstrate that immunity proteins bind in the active site of the toxins to block substrate binding (Goulding, CW & Hayes, CS unpublished data)<sup>122</sup>. However, more structures will need to be resolved to conclude this is a general mechanism of toxin inhibition for CDI tRNases.

Interestingly, CdiA-CT from uropathogenic *E. coli* 536 (CdiA-CT<sup>UPEC536</sup>) requires a cofactor protein, CysK, for both *in vitro* and *in vivo* activity<sup>123</sup>. CysK is a metabolic enzyme involved in cysteine biosynthesis. It normally binds to the enzyme involved in the previous step of cysteine synthesis, CysE, by interacting with its C-terminal tail<sup>124</sup>. Diner *et al.* found that CysK binds to the C-terminus of CdiA-CT<sup>UPEC536</sup> in a similar manner. Furthermore, the immunity protein is still able to bind to the CT even in the presence of CysK<sup>123</sup>. Chapter 5 will explore the implications of this binding event.

### ***Pore-formers***

The first CDI system discovered contains a pore-forming toxin. Extensive analyses were performed to show that CdiA-CT from the CDI system in EC93 (CdiA-CT<sup>EC93</sup>) dissipates proton-motive force and steady-state ATP levels, suggesting it forms a hole in the inner membrane<sup>125</sup>. Interestingly, AcrB is necessary for its toxic activity<sup>126</sup>. It is thought that AcrB helps insert the toxin into the membrane. This is the only pore-forming CDI toxin whose

activity has been experimentally proven. Other pore-forming CDI toxins have been identified, but only by bioinformatic analysis.

### **Delivery process of CDI**

We just learned that most CDI toxins are nucleases and must therefore enter the cytoplasm of cells to perform their toxic function. The first step of the delivery process is a binding event between CdiA and a receptor protein on the surface of target cells. This was established in CDI<sup>EC93</sup> which utilizes the monomeric  $\beta$ -barrel outer membrane protein, BamA<sup>126</sup>. Polymorphism in the extracellular loops of BamA between different species is what restricts CDI<sup>EC93</sup> to target only *E. coli* strains<sup>118</sup>. The work done in chapter 3 identifies novel receptor proteins for CdiA<sup>UPEC536</sup> which led to the identification of a 300 amino acid stretch about 2/3 the way through CdiA as the receptor-binding domain. This region of CdiA varies significantly between EC93 and other *E. coli* strains. Bioinformatic analyses have identified four different classes of CdiA in *E. coli* based on divergence specifically in the proposed receptor-binding domain region.

The receptor-binding domain has been identified, but the remainder of the delivery process through both the outer and inner membrane remains speculative. The same three models for colicin transit through the outer membrane are considered: through the pore of the  $\beta$ -barrel, at the interface of the protein and lipid and through the lipid itself. CdiA is a large 200-600 kDa protein that could itself aid in transit through the outer membrane. In contrast to colicins, CDI systems do not require TolA or TonB for delivery of CdiA-CT into the cytoplasm<sup>119</sup>. It is thought there is a cleavage event, possibly at the conserved VENN peptide motif that releases the toxic domain from the rest of CdiA. Interestingly, CDI systems are

modular. CdiA-CTs can be swapped between CDI systems and still function<sup>115</sup>. This suggests a general mechanism of delivery into the cytoplasm. Nevertheless, CdiA-CTs are able to enter the cytoplasm of susceptible cells where they can perform their toxic function or bind to their immunity protein and thus become neutralized.

### **CDI Conclusion**

CDI systems contain a repertoire of distinct toxins that allows for potentially complex bacterial competition in the environment. Some CDI toxins are homologs to colicin toxin domains. While there are certainly many similarities between the two systems, there are significant differences in active-site structure, mechanism of cleavage, substrate specificity, and function of immunity protection. The overlap between colicin and CDI toxin activities is most likely reflective of the limited number of ways to kill a cell efficiently, and these toxins have evolved to do just that. It is important to keep studying and trying to identify new ways that bacteria are finding to effectively kill one another. The topic of chapter 4 will be the identification of novel CDI toxin activity.

Colicins and CDI systems are both involved in intra-species competition. However, there are significant differences between colicin and CDI delivery that make the two systems distinct. First, colicin producing cells must sacrifice themselves whereas CDI<sup>+</sup> cells are able to persist after toxin delivery. Second, colicins get secreted into the environment in hopes of finding a susceptible target cell. This can potentially be a waste of self-sacrifice if there are no other *E. coli* strains to inhibit in the area. In contrast, CDI systems require direct physical contact with target cells, which will ensure delivery of the toxin being produced. However, the ability for colicins to diffuse into the environment gives them the potential to clear a large

area more efficiently, while CDI can only deliver into cells in the immediate area. Thus, CDI and colicins seem to offer distinct advantages and disadvantages for bacteria and therefore most likely have distinct roles in nature.

CDI systems are prevalent throughout proteobacteria and specifically pathogens. Their role in nature has yet to be determined. The narrow target range of CDI systems suggest that it plays a role in intra-species competition. Therefore, the relevance of this system in complex microbial communities may be limiting. However, the presence of an intra-species CDI system in EC93 seemed to be beneficial, as EC93 was the dominant *E. coli* strain isolated from a rat colony. Chapter 3 will illuminate the role of CDI systems in nature with the discovery of receptor proteins that restrict some CDI systems to strain-specific delivery.

## **Type VI Secretion (T6S)**

### **T6S Introduction**

Gram-negative bacteria have developed six secretion systems to translocate proteins into the extracellular space or into neighboring cells. The type VI secretion system (T6SS) is the most recent secretion system to be identified. It is a membrane-spanning organelle that injects toxins into prokaryotic or eukaryotic cells in a contact-dependent manner<sup>127</sup>. The toxins delivered are called effectors and have a wide variety of activities that lead to cell death. Contrast to colicins or CDI toxins, T6 effectors mostly target substrates in the periplasm<sup>128</sup>. Effectors contain an immunity protein located downstream to protect delivering cells from self-intoxication<sup>129,130</sup>. The first T6SS was initially found in a screen looking for a decrease in *Vibrio cholera* virulence towards *Dictyostelium discoideum*<sup>131</sup>. It

was later identified that this T6SS is able to deliver an effector into macrophages that cross-links actin molecules to prevent mobility<sup>132</sup>. Since its discovery in *V. cholera*, it seems the ability to inhibit prokaryotic cell function is rare and the main function of T6S is for bacterial competition<sup>133</sup>.

Type 6 secretion systems are found in 25 % of sequenced strains and are more heavily represented in pathogenic species<sup>131</sup>. Some strains have more than one T6 loci encoded in their genome. *Enterobacter cloacae* has two, *Pseudomonas aeruginosa* contains 3, and *Burkholderia thailandensis* holds 6 T6SSs<sup>134</sup>. The implications of having more than one locus are unknown. T6S has been demonstrated to be functional in both intra- and inter-species competition in many species including *V. cholera*, *P. aeruginosa*, *B. thailandensis*, *Serratia marcescens*, and *E. cloacae*. Additionally, functional T6S has been implicated to play a role in pathogenesis<sup>135</sup>. Clearly T6SSs are used in bacterial competition and can give strains an advantage in their respective environments.

## **T6S Apparatus**

The constituents of the type VI apparatus are encoded in large gene clusters containing all the components required to build a functional T6SS<sup>134</sup>. There are examples of constitutive, tight, and induced expression of T6 loci<sup>127</sup>. While more genes can be present in the locus, assembly of a functional type VI organelle requires 13 “core components”<sup>136-138</sup>. These components make an envelope-spanning apparatus which is composed of a tube complex, membrane complex, and baseplate<sup>139</sup>. The tube complex contains an inner tube capped with a spike that is surrounded by an outer sheath. This structure assembles onto the baseplate which serves to connect the tube to the cell envelope via the membrane complex.

The baseplate and tube complex of the type VI organelle are structurally and functionally homologous to the tail proteins of T4 phage<sup>140</sup>. These are contractile phage that use their tail to inject DNA into cells<sup>141</sup>. The type VI organelle is doing the opposite by ejecting effectors out of the cell and into neighboring targets. T4 phage use tail fibers to secure its baseplate to the outer membrane, whereas T6SSs use the membrane complex to anchor on the opposite side of the membrane. Just like T4 phage, a contractile mechanism is used to deliver effectors into target cells<sup>139</sup>. In both cases, the outer sheath contracts to propel the inner tube outwards and into their respective targets.

The nomenclature of type VI components has become complicated since its discovery. For this section and the remainder of the thesis, we will refer to the components as TssX for Type six secretion gene X. We will use the adapted names for Hcp, VgrG, ClpV, and VasK. Both adapted and conventional names of the type VI components, their homology to T4 phage proteins, and their function are listed in Figure 3.

### ***Membrane Complex***

The membrane complex is composed of TssL, TssM, and TssJ, three proteins that associate with the membrane and one another to make the complex<sup>142-146</sup>. The complex spans from the cytoplasm all the way to the outer membrane. It is unclear how it is able to transverse through the peptidoglycan layer. The role of this complex is to anchor the remaining parts of the T6S apparatus to the cell envelope<sup>147</sup>.

### ***Tube Complex***

The T6SS tube complex is composed of an inner tube engulfed by an outer sheath. The inner tube is made up of hemolysin-coregulated protein (Hcp), a structural homolog to phage tail tube proteins<sup>140,148</sup>. Hcp assembles into homohexameric rings that stack head-to-tail with

a diameter of about  $100\text{\AA}$ <sup>149,150</sup>. This Hcp tube gets engulfed by an outer sheath made up of alternating TssB and TssC units that interact with one another to stabilize the sheath<sup>151</sup>. Hcp formation is essential for sheath assembly, suggesting the inner tube acts as a scaffold for TssB/C polymerization<sup>150</sup>.

### ***Baseplate***

The baseplate serves as the site of assembly for the inner tube and sheath so they are not free-floating in the cytoplasm. The baseplate associates with the membrane complex, connecting each component of the T6 apparatus together at the cell envelope<sup>152</sup>. The known components of the baseplate are valine-glycine repeat protein G (VgrG), TssE, and TssK. VgrG and TssE share homology with phage base plate proteins gp27-gp5 and gp25 respectively<sup>140,151,152</sup>.

VgrG trimers cap the Hcp inner tube and are thought to act like a spike and help penetrate the target-cell membrane<sup>138</sup>. VgrG proteins are further capped by PAAR (proline, alanine, alanine, arginine) motif-containing proteins<sup>153</sup>. PAAR motifs are able to coordinate a zinc ion to make a cone-like structure that is thought to further sharpen the tip of the T6 apparatus<sup>153</sup>. VgrG is required for the assembly of Hcp tubules and a PAAR-containing protein is required for functional T6 firing<sup>150,153</sup>. It is therefore thought that PAAR proteins help nucleate VgrG folding which is required for inner tube formation. Chapter 6 will reveal more about the role of PAAR proteins in T6S.

Phage baseplate protein gp25 is involved with initiating the polymerization of the sheath proteins. Based on sequence homology and its ability to bind TssB/C complex, TssE likely performs the same role in T6S<sup>139,154,155</sup>.



TssK shows no similarity to phage proteins, but is required for sheath polymerization<sup>156</sup>. Furthermore, TssK binds to TssL, a component of the membrane complex<sup>156</sup>. Thus, TssK is likely the key intermediary between the baseplate/tube complex and the membrane complex. Its absence in T4 phage further validates its novel role in T6 assembly, as phage do not anchor to a membrane complex.

### ***Firing***

The triggering event for contraction and regulation of tube length during assembly remains unknown. It is known that upon contraction, which presumably happens only once everything is assembled properly, the inner tube is propelled outward<sup>139</sup>. This is supported by the presence of Hcp and VgrG proteins in the supernatant of T6<sup>+</sup> but not T6<sup>-</sup> cells<sup>132</sup>. Furthermore, the T6S apparatus has been captured in both the extended and contracted states, representing pre- and post-firing events, by electron cryo-tomography. In the extended state, the tube spans almost the entire length of the cell and has electron density inside while after contraction, it is hollow<sup>139</sup>. However, this does not elucidate how the tube spike is able to penetrate the target cell. Furthermore, how far the tubule penetrates the target cell, past the outer membrane, cell wall, or inner membrane, is unknown.

### ***Disassembly***

Once fired, the conformation of the sheath changes which allows ClpV to bind both TssB and TssC<sup>139,157</sup>. ClpV is an ATPase that uses ATP to disassemble the sheath proteins<sup>136,149,157</sup>. It is unclear whether these components can be recycled for another round of firing or if they are degraded.

### **T6S Effectors**

Just like colicin and CDI toxins, there are a wide variety of T6 effectors. Interestingly, unlike colicins and CDI, most type VI effectors act in the periplasm rather than the cytoplasm, and can be structural components of the T6SS<sup>128,131,132,153</sup>. Regardless, with the diversity of T6 effectors that continue to be identified, it is becoming more apparent that T6SSs have the potential to facilitate effective anti-bacterial activity in nature.

### ***VgrG effectors***

As previously described, VgrG is necessary for proper T6 assembly and forms the tube spike to help penetrate target cells. Interestingly, some VgrG proteins have extended C-termini which encode for a toxic domain, and are thus named “evolved” or “extended” VgrGs<sup>132,158</sup>. Extended VgrGs have been demonstrated to show toxic activity in both bacterial and prokaryotic cells<sup>132,159</sup>. The evolved VgrGs that are predicted to act in anti-bacterial activity have cognate immunity genes located immediately downstream<sup>158</sup>. The means of transport into recipient cells is clear for these effectors. Since VgrG is used to penetrate target cells, these toxins will clearly be delivered.

### ***PAAR-containing effectors***

PAAR-containing proteins are another example of a structural component of the T6 apparatus acting as an effector. However, PAAR-containing effectors can have the toxin domain at their N- or C-terminus<sup>153</sup>. There are 5 classes of PAAR-containing proteins, 4 of which contain toxic domains<sup>153</sup>. Rearrangement hot spot (Rhs) proteins are large 200-400 kDa proteins composed of a YD-repeat core and a C-terminal toxin domain demarcated by a conserved DPXGL peptide motif<sup>121</sup>. Many Rhs proteins have PAAR domains in their N-terminus and have thus been characterized as PAAR proteins<sup>153</sup>. Delivery of the toxic

domains of PAAR-proteins is similar to the proposed delivery mechanism of evolved VgrGs; they are injected into target cells as part of the cell membrane-puncturing device.

### ***Peptidoglycan-targeting effectors***

Not surprisingly, numerous T6SS effectors have enzymatic activity towards peptidoglycan, the major structural component of the bacterial cell wall. Peptidoglycan is composed of alternating units of  $\beta$ 1,4-linked N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc). Peptides of varying length attach to MurNAc and crosslink to one another, forming a mesh-like layer essential for cell structure. A type VI amidase effector (Tae) family containing 4 distinct classes has been characterized<sup>130</sup>. These enzymes are capable of cleaving the peptide crosslinks at distinct locations, thereby disrupting the meshwork and ultimately causing cell lysis<sup>129,130,160-162</sup>. Other T6 effectors target the sugar component of peptidoglycan. Muramidases and glucosaminidases can both be found in the T6SS glycoside hydrolase effector (Tge) family<sup>130,163</sup>. This family contains 3 distinct classes based on sequence-divergence and substrate specificity<sup>128</sup>. Consistent with these effectors targeting the cell wall, the cognate immunity protein of each effector contains a signal sequence that will direct it to the periplasm. Most peptidoglycan-targeting effectors are only 10-20 kDa and have been proposed to bind inside Hcp monomers as they are building the inner tubule of the T6SS apparatus and will therefore be secreted into target cells upon contraction of the sheath<sup>164</sup>.

### ***Cell membrane-targeting effectors***

The lipid bilayer of bacterial cells is another essential component that T6S effectors target. Type VI lipase effector (Tle) proteins hydrolyze the lipids that make up the bacterial membrane, while pore-forming proteins dissipate the proton-motive force of cells<sup>128,165</sup>.

While there have only been two examples of pore-forming T6S effectors, there are 377 predicted Tle proteins in T6SS<sup>+</sup> bacteria<sup>165</sup>. The proteins have been broken up into 5 families based on bond specificity and phosphate head group preference. The families 1-4 contain a conserved GX SXG motif that is commonly found in lipases and esterases, while the Tle5 family carries HXLXXXD motifs unique to phospholipase D enzymes<sup>128</sup>. Just like peptidoglycan-targeting T6SS effectors, the cognate immunity proteins for cell membrane-targeting effectors are also predicted to function in the periplasm. The delivery process of these cell-membrane targeting effectors remains speculative. They are too large to fit inside the homohexameric ring formed by Hcp without unfolding<sup>164</sup>. Therefore, it is thought that they interact with an accessory protein to link it to the T6SS to be delivered into target cells.

### *Cytoplasmic-acting effectors*

Not all T6 effectors target substrates in the periplasm. More recently, nuclease effectors have been identified. Most of these effectors are found in PAAR-containing proteins, and especially at the C-terminus of Rhs proteins (Rhs-CT)<sup>121,153</sup>. In fact to this date all Rhs-CTs are predicted to act in the cytoplasm. Of the effectors that are predicted to bind inside Hcp rings, only 5 have predicted cytoplasmic activity, although their toxic activity is unknown<sup>128</sup>. It is unclear if the delivery of these effectors differs from delivery of periplasmic-acting effectors, as cytoplasmic-acting toxins must cross both lipid bilayers of the recipient cell.

### **T6S Conclusion**

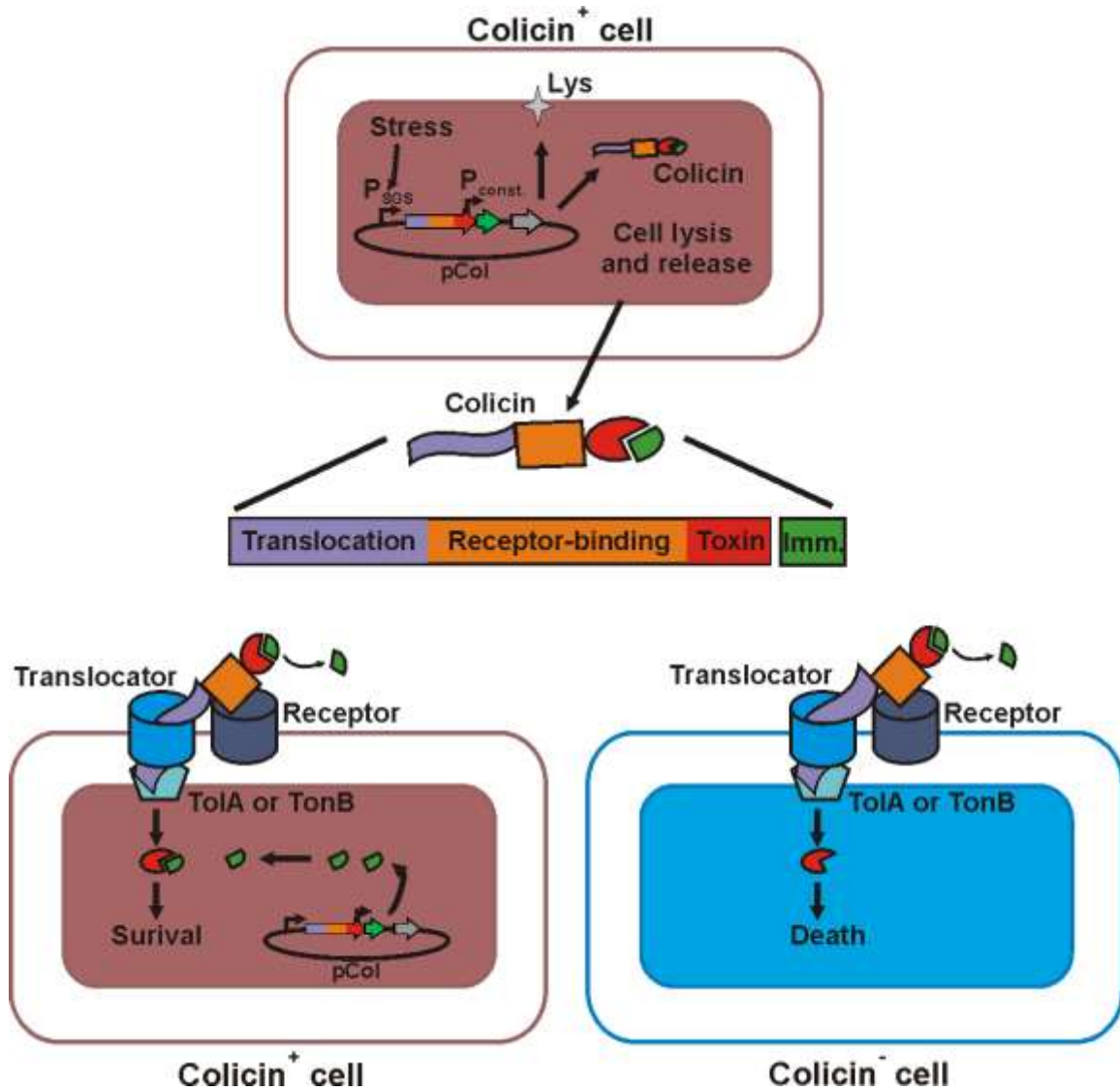
T6SSs are prevalent in proteobacteria and deploy a diverse array of effector proteins that target essential and conserved cellular structures. Taken all together, this suggests these systems play a vital role in inter-species competition. There are many parallels between T6S

and CDI systems. Both require direct contact with target cells to deliver their respective toxic effector molecule to inhibit cell growth. While CDI systems mostly target substrates in the cytoplasm and T6S effectors are more commonly periplasmic-acting, each system demonstrates the ability to severely inhibit the growth of neighboring bacteria. It seems the function of these systems is redundant. However, many strains of bacteria contain both systems, suggesting they provide their own distinct advantages. Understanding the mechanism of effector delivery, the toxins utilized and target-cell range of each of these systems will help clarify their distinct roles in nature. Chapter 6 will focus on the T6SS found in the opportunistic human pathogen, *Enterobacter cloacae*.

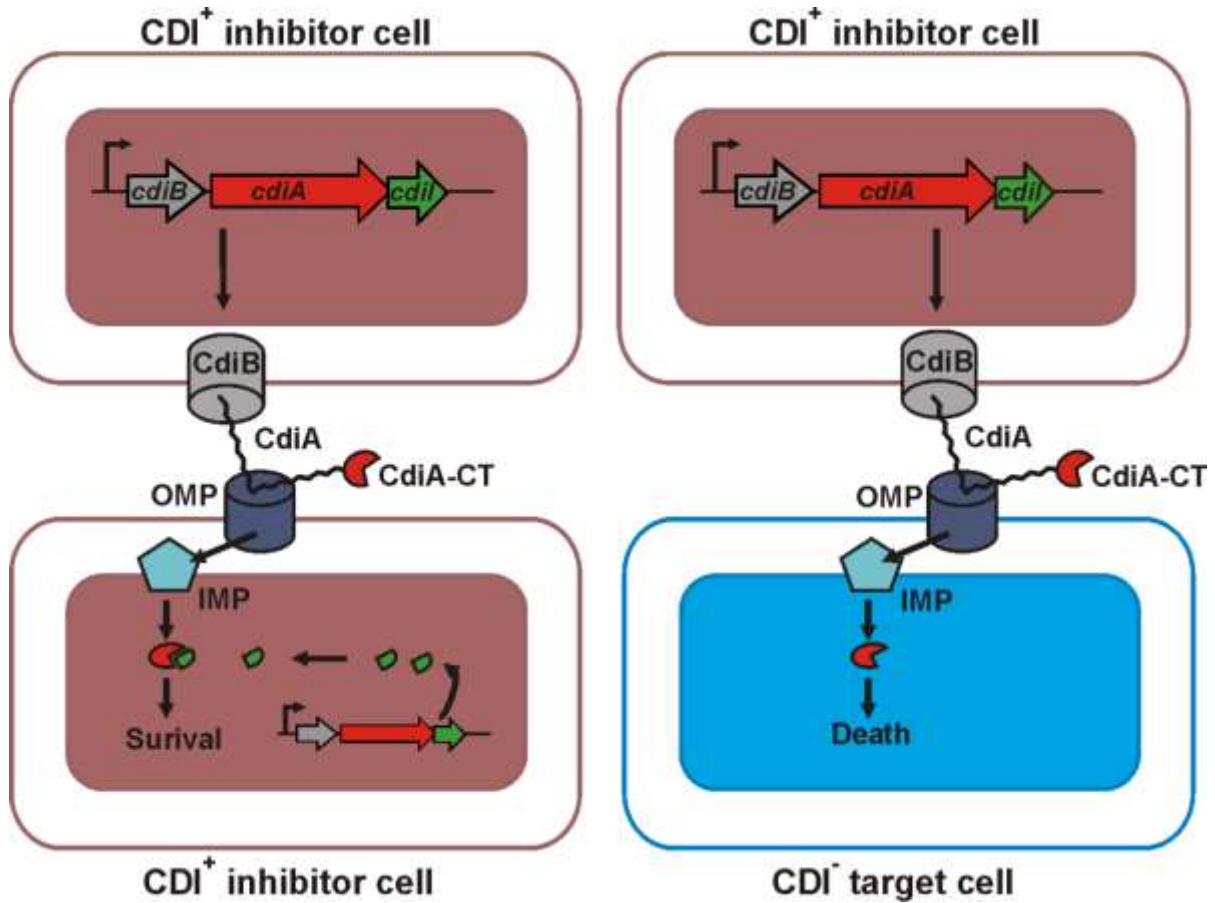
### **Conclusion of the Introduction**

Bacteria are almost exclusively found in complex microbial communities, competing with a variety of species for limited resources and space. Therefore, bacteria have developed diverse competition systems to gain a fitness advantage in such harsh environments. We have discussed 3 such competition systems that are widespread in proteobacteria and specifically pathogenic species. The diversity of toxins found in bacteriocins, CDI systems, and T6SSs demonstrate the ability of these systems to mediate antagonistic effects to a wide range of species. However, we learned bacteriocins and CDI systems are restricted to intra-species delivery, while T6SSs participate in both intra- and inter-species competition. Understanding the role of each of these systems in the environment will help elucidate how complex microbial communities are established.

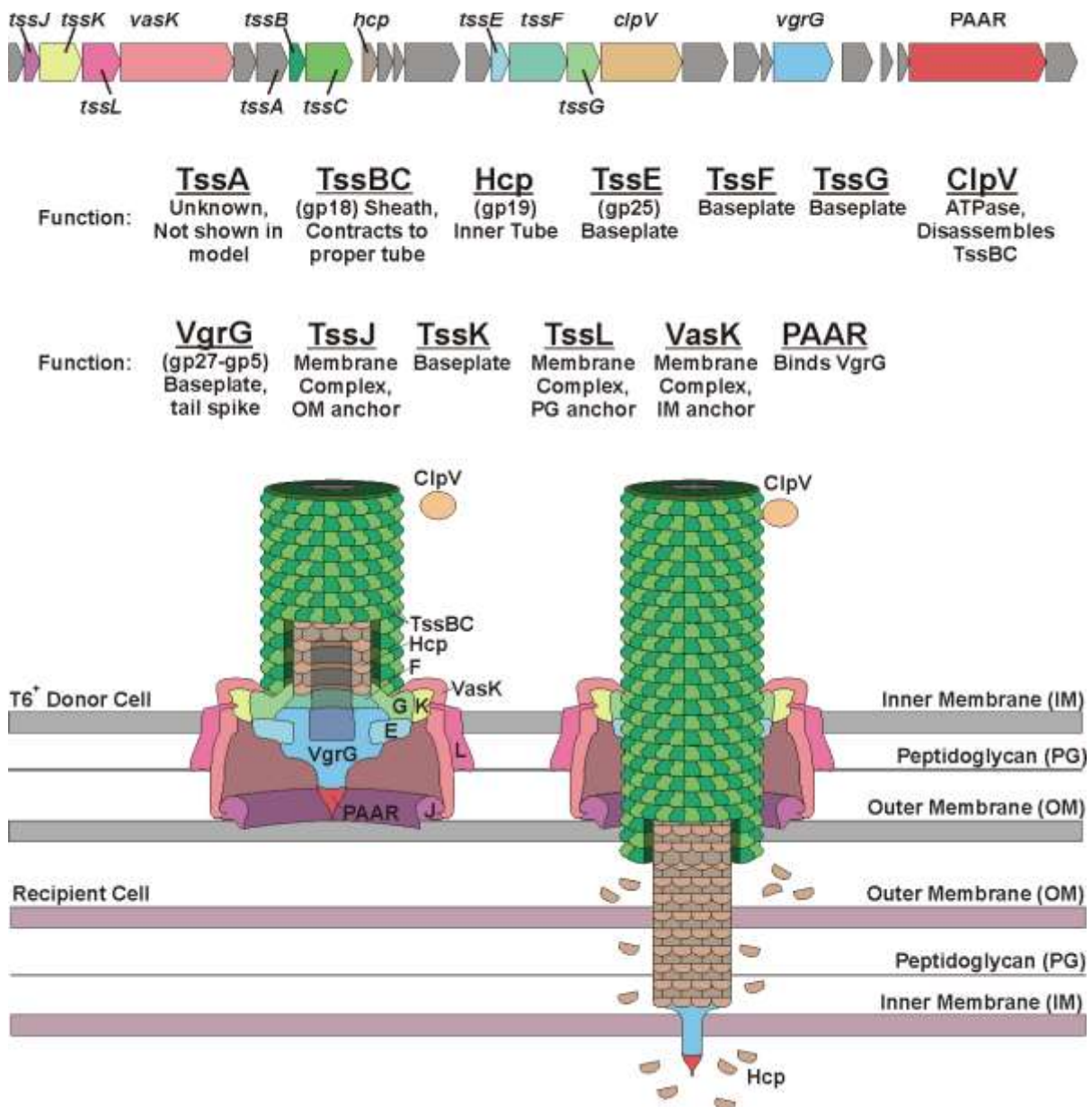
**Figure 1. Model of colicin production and import.** SOS promoter is induced upon stress leading to expression of the Colicin-Immunity pair and Lysis protein. Cell lysis allows for the release of Col-Imm. The receptor-binding domain binds its cognate outer-membrane protein which recruits a translocator to produce a translocon. The translocation domain interacts with TolA (Group A colicins) or TonB (Group B colicins) to facilitate transit through the outer membrane. Imm. is stripped away on the surface of cells. The toxin domain is delivered to the cytoplasm which causes cell death. Sibling cells will be protected by the production of cognate immunity protein.



**Figure 2. Model of contact-dependent growth inhibition (CDI).** CdiB exports CdiA onto the surface. Upon contact with a receptor protein on target cells, CdiA-CT transits into the cytoplasm where it can perform its toxic activity to cause cell death. Sibling  $CDI^+$  cells produce CdiI which binds to and inactivates CdiA-CT, thus allowing cell survival.



**Figure 3. Model of the type VI secretion system (T6SS).** Model arrangement of a T6S locus. The function of each protein in the apparatus and homology to phage proteins is listed. The inner tube comprised of Hcp homohexamers is assembled onto the baseplate and capped by PAAR-bound VgrG (spike). This is surrounded by an outer sheath made of TssBC subunits and anchored to the envelope by components of the membrane complex. Upon contraction of TssBC, the inner tube and spike are injected into recipient cells where toxic effectors are released. ClpV binds the sheath subunits for disassembly.





**Chapter 2. F pili mediate import of a contact-dependent growth inhibition (CDI) toxin**

**Note:** This research was originally published in *Molecular Microbiology*. I collaborated on this project with Elie Diner. Figure 1 is solely his work. Furthermore, Christopher Hayes and Elie Diner contributed to the writing in this chapter.

**Beck CM, Diner EJ, Kim JJ, Low DA, Hayes CS.** “F pili mediate import of a contact-dependent growth inhibition (CDI) toxin.” *Molecular Microbiology*. 2014 Jul;93(2):276-90.

## Abstract

Contact-dependent growth inhibition (CDI) is a widespread form of inter-bacterial competition that requires direct cell-to-cell contact. CDI<sup>+</sup> inhibitor cells express CdiA effector proteins on their surface. CdiA binds to specific receptors on susceptible target bacteria and delivers a toxin derived from its C-terminal region (CdiA-CT). Here, we show that purified CdiA-CT<sup>536</sup> toxin from uropathogenic *Escherichia coli* 536 translocates into bacteria, thereby by-passing the requirement for cell-to-cell contact during toxin delivery. Genetic analyses demonstrate that the N-terminal domain of CdiA-CT<sup>536</sup> is necessary and sufficient for toxin import. The CdiA receptor plays no role in this import pathway; nor do the Tol and Ton systems, which are exploited to internalize colicin toxins. Instead, CdiA-CT<sup>536</sup> import requires conjugative F pili. We provide evidence that the N-terminal domain of CdiA-CT<sup>536</sup> interacts with F pilin, and that pilus retraction is critical for toxin import. This pathway is reminiscent of the strategy used by small RNA leviviruses to infect F<sup>+</sup> cells. We propose that CdiA-CT<sup>536</sup> mimics the pilin-binding maturation proteins of leviviruses, allowing the toxin to bind F pili and become internalized during pilus retraction.

## Introduction

Contact-dependent growth inhibition (CDI) systems are distributed throughout  $\alpha$ -,  $\beta$ - and  $\gamma$ -proteobacteria, where they function in competition between closely related bacteria<sup>115,118,166,167</sup>. CDI is mediated by the CdiB/CdiA family of two-partner secretion proteins. CdiB is an outer-membrane  $\beta$ -barrel protein required for secretion of the CdiA effector. CdiA proteins are filamentous and very large, ranging from ~180 kDa in *Moraxella* to over 600 kDa in some *Pseudomonas* species<sup>115,166</sup>. CdiA extends from the surface of CDI<sup>+</sup> inhibitor cells and interacts with receptors on susceptible target bacteria. CDI has been characterized most extensively in *Escherichia coli* EC93, and its CdiA<sup>EC93</sup> effector binds the highly conserved outer-membrane protein BamA as a receptor<sup>118,168</sup>. Upon binding the receptor, CdiA<sup>EC93</sup> appears to be cleaved to release a C-terminal toxin region (CdiA-CT<sup>EC93</sup>), which is subsequently translocated into the target cell through a poorly characterized pathway<sup>115,168,169</sup>. *E. coli* EC93 inhibitor cells also deliver toxins to one another, but are protected from inhibition by a small CdiI<sup>EC93</sup> immunity protein encoded immediately downstream of *cdiA*<sup>EC93</sup><sup>112</sup>. Thus, *E. coli* EC93 cells deploy CdiA-CT<sup>EC93</sup> toxins to inhibit other non-immune strains of *E. coli*. Because CDI confers a significant growth advantage to inhibitor cells, these systems are thought to mediate inter-strain competition for environmental niches.

CDI loci encode a diverse group of CdiA-CT/CdiI toxin/immunity sequences. There are at least 18 CdiA-CT/CdiI sequence types distributed throughout *E. coli* strains, and *Burkholderia pseudomallei* strains carry 10 distinct sequence types<sup>120,167</sup>. Thus, CDI toxin/immunity pairs are highly variable, even between strains of the same species. In accord with this sequence diversity, CDI toxins exhibit a number of distinct activities. For example, CdiA-CT<sup>EC93</sup> dissipates the proton-motive force and induces the phage-shock response<sup>170</sup>,

suggesting that the toxin forms pores in the inner membrane of target cells. Many other CDI toxins have nuclease activities. CDI toxins from *Dickeya dadantii* 3937 and *E. coli* EC869 are DNases that degrade target-cell genomic DNA<sup>115,169,171</sup>. CDI<sup>+</sup> bacteria also deploy RNases that preferentially cleave tRNA molecules. CdiA-CT<sup>536</sup> from uropathogenic *E. coli* 536 (UPEC 536) is a tRNA anticodon nuclease<sup>115,123</sup>, and toxins from *B. pseudomallei* strains cleave tRNA in the anticodon loop, T-loop and aminoacyl-acceptor stem<sup>120</sup>. Each CdiA-CT toxin is specifically neutralized by its cognate CdiI protein, but not the immunity proteins from other CDI systems<sup>115,120,171</sup>. Therefore, CDI toxin/immunity complexity provides the basis for self/non-self discrimination during inter-strain competition.

The CdiA-CT toxin region is typically demarcated by a short, highly conserved peptide sequence. Most CdiA-CT regions are defined by the VENN peptide motif, but CdiA proteins from *Burkholderia* species contain an analogous (Q/E)LYN sequence<sup>115,120,172,173</sup>. The function of this motif has not been explored, but it could play two important roles in CDI. First, the universally conserved Asn residue may catalyze auto-cleavage to release the CdiA-CT for subsequent translocation into target bacteria. Asn side-chains can undergo intramolecular attack on the peptide backbone, thereby cleaving the peptide chain and producing a cyclic succinimide. Asn residue cyclization is a well-characterized reaction that mediates auto-cleavage in several secreted proteins<sup>174,175</sup>. Second, the nucleotide sequence encoding VENN may be important for recombination and subsequent expression of new *cdiA-CT/cdiI* gene pairs acquired by horizontal transfer. CDI toxins are modular and can be readily exchanged between systems. For example, *E. coli* CdiA proteins can deliver heterologous toxins from *Yersinia pestis* CO92 and *Dickeya dadantii* 3937, provided the CdiA-CTs are fused at the common VENN motif<sup>115,169</sup>. Thus, bacteria might switch toxin/immunity types

by recombining a new *cdiA-CT/cdiI* sequence onto the 3'-end of the *cdiA* gene, thereby displacing the original toxin/immunity pair. Moreover, many species carry fragmented *cdiA-CT/cdiI* gene pairs in tandem arrays downstream of the main *cdiBAI* gene cluster<sup>176</sup>. These "orphan" toxin/immunity regions usually contain interspersed sequences that are homologous to transposases and integrases, implying recent acquisition through horizontal gene transfer<sup>167,176</sup>. Taken together, these observations suggest that CdiA proteins have the intrinsic ability to auto-cleave and deliver diverse toxins across the target-cell envelope.

Although CDI toxins are hypothesized to translocate autonomously, CdiA<sup>EC93</sup> fragments released by *E. coli* EC93 cells do not inhibit the growth of other *E. coli* strains<sup>112</sup>. One possible explanation is that only specific CdiA-CT fragments – presumably those released in response to receptor-binding – are able to enter target bacteria. This model suggests that purified toxin may be active against bacteria if it mimics the naturally cleaved CdiA-CT fragment. Therefore, we tested whether purified CdiA-CT toxins inhibit the growth of bacterial cultures. We find that purified CdiA-CT<sup>536</sup> (from UPEC 536) inhibits *E. coli* cell growth, whereas CdiA-CT<sup>3937</sup> (from *D. dadantii* 3937) has no effect. Structure-function analyses show that the N-terminal domain of CdiA-CT<sup>536</sup> is necessary and sufficient for cell import, but surprisingly toxin translocation does not require BamA. Instead, conjugative F pili are required for the import of purified CdiA-CT<sup>536</sup> toxin. We provide evidence that CdiA-CT<sup>536</sup> binds to F pilin and hypothesize that this interaction allows the toxin to be internalized during pilus retraction. F-dependent RNA phages exploit a similar pathway to transport their genomes into cells, suggesting that CdiA-CT<sup>536</sup> mimics RNA phage to enter F<sup>+</sup> cells.

## Materials and Methods

### *Bacterial strains and plasmids used in this study*

<i>Strains or plasmids</i>	<i>Description<sup>a</sup></i>	<i>Reference</i>
<b>Strain</b>		
X90	F' <i>lacI<sup>d</sup> lac' pro' araΔ(lac-pro) nalI argE(amb) rif<sup>r</sup> thi-1, Rif<sup>R</sup></i>	177
XL-1	F':Tn10 <i>proA<sup>+</sup>B<sup>+</sup> lacI<sup>d</sup> Δ(lacZ)M15/recA1 endA1 gyrA96 thi hsdR17 (r<sub>K</sub><sup>-</sup> r<sub>K</sub><sup>+</sup>) glnV44 relA1 lac. Tet<sup>R</sup>, Nal<sup>R</sup></i>	Stratagene
MC4100	F' <i>araD139 Δ(argF-lac)U169 rpsL150 relA1 deoC1 rbsR fthD5301 fruA25 λ<sup>-</sup>. Str<sup>R</sup></i>	178
KW1070	X90 F <sup>-</sup>	Kelly P. Williams
CH2016	X90 (DE3) <i>Δrna ΔslyD::kan, Kan<sup>R</sup></i>	179
CH6479	X90 <i>ΔtolA::kan, Kan<sup>R</sup></i>	This study
CH6480	X90 <i>ΔtonB::kan, Kan<sup>R</sup></i>	This study
CH6680	X90 <i>bamA101::kan, Kan<sup>R</sup></i>	<sup>168</sup> & this study
CH6866	KW1070 <i>pOX38::gent, Gent<sup>R</sup></i>	This study
CH6939	X90 <i>ΔtraA::cat, Cm<sup>R</sup></i>	This study
CH7035	X90 <i>ΔtraT::cat, Cm<sup>R</sup></i>	This study
CH9027	KW1070 F':Tn10, Tet <sup>R</sup>	This study
CH9354	X90 <i>ΔtonB tolA::kan, Kan<sup>R</sup></i>	This study
CH11476	X90 <i>ΔtrbI::cat, Cm<sup>R</sup></i>	This study
CH11703	CH11717 <i>ΔtrbI::cat, Tet<sup>R</sup> Cm<sup>R</sup></i>	This study
CH11717	X90 Tn10, Tet <sup>R</sup>	This study
CH11718	CH11717 <i>ΔtraA::cat, Tet<sup>R</sup> Cm<sup>R</sup></i>	This study
<b>Plasmids</b>		
pBR322	Cloning vector, Amp <sup>R</sup> Tet <sup>R</sup>	180
pACYC184	Cloning vector, Cm <sup>R</sup> , Tet <sup>R</sup>	181
pTrc99A	IPTG-inducible expression vector, Amp <sup>R</sup>	GE Healthcare
pCH450	pACYC184 derivative containing <i>E. coli araC</i> and the L-arabinose-inducible P <sub>araBAD</sub> promoter, Tet <sup>R</sup>	182
pET21:: <i>cdiA-CT/cdiI</i> <sup>3937</sup>	Over-produces CdiA-CT <sup>3937</sup> and CdiI <sup>3937</sup> -His <sub>6</sub> , Amp <sup>R</sup>	115
pET21:: <i>cdiA-CT/cdiI</i> <sup>536</sup>	Over-produces CdiA-CT <sup>536</sup> and CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	115
pET21:: <i>DUF-</i>	Over-produces DUF638-CdiA-CT <sup>536</sup>	This study

<i>CT/cdiI</i> <sup>536</sup>	and CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	
pET21:: <i>VENN-less-CT/cdiI</i> <sup>536</sup>	Over-produces VENNless-CdiA-CT <sup>536</sup> and CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	This study
pET21:: <i>Cys-less-CT/cdiI</i> <sup>536</sup>	Over-produces CdiA-CT <sup>536</sup> containing Cys13Ser and Cys19Ser mutations together with CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	This study
<sub>536</sub> pET21:: <i>tRNase/cdiI</i>	Over-produces the C-terminal tRNase domain of CdiA-CT <sup>536</sup> together with CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	123
pET21:: <i>cdiA-CT(H178A)/cdiI</i> <sup>536</sup>	Over-produces catalytically inactive His178Ala variant of CdiA-CT <sup>536</sup> toxin and CdiI <sup>536</sup> -His <sub>6</sub> , Amp <sup>R</sup>	123
pET21K:: <i>cdiA-CT/cdiI</i> <sup>ECL</sup>	Over-produces CdiA-CT <sup>ECL</sup> and CdiI <sup>ECL</sup> -His <sub>6</sub> from <i>Enterobacter cloacae</i> ATCC 13047, Amp <sup>R</sup>	This study
pCH450:: <i>cdiA-CT</i> <sup>536</sup>	Expresses <i>cdiA-CT</i> <sup>536</sup> under control of P <sub>BAD</sub> promoter, Tet <sup>R</sup>	This study
pCH450:: <i>DUF-CT</i> <sup>536</sup>	Expresses <i>DUF638-cdiA-CT</i> <sup>536</sup> under control of P <sub>BAD</sub> promoter, Tet <sup>R</sup>	This study
pCH450:: <i>VENN-less-CT</i> <sup>536</sup>	Produces VENNless-CdiA-CT <sup>536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>Cys-less-CT</i> <sup>536</sup>	Produces CdiA-CT <sup>536</sup> containing Cys13Ser and Cys19Ser mutations, Tet <sup>R</sup>	This study
pCH450:: <i>tRNase</i> <sup>536</sup>	C-terminal tRNase domain of CdiA-CT <sup>536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(H178A)</i> <sup>536</sup>	catalytically inactive His178Ala variant of CdiA-CT <sup>536</sup> , Tet <sup>R</sup>	This study
pDAL776	pBR322 derivative that constitutively expresses <i>cdiI</i> <sup>536</sup> , Amp <sup>R</sup>	115
pDAL852	pBR322 derivative that constitutively expresses <i>cdiI</i> <sup>3937</sup> from <i>Dickeya dadantii</i> 3937, Amp <sup>R</sup>	115
pOX38:: <i>gent</i>	pOX38 with integrated gentamicin-resistance cassette, Gent <sup>R</sup>	183
pTrc99A:: <i>cdiI</i> <sup>536</sup>	Expresses <i>cdiI</i> <sup>536</sup> under control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pTrc99A:: <i>cdiI</i> <sup>ECL</sup>	Expresses <i>cdiI</i> <sup>ECL</sup> from <i>Enterobacter cloacae</i> ATCC 13047 under control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pTrc99A:: <i>imE5</i>	Expresses <i>imE5</i> under control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pET21:: <i>colE5-CT/imE5</i>	Over-produces ColE5-CT (residues Met429 - Gln556) and together with ImE5-His <sub>6</sub> , Amp <sup>R</sup>	This study
pET21:: <i>NT</i> <sup>536</sup> - <i>colE5-CT/imE5</i>	Over-produces a protein containing residues Val1 – Tyr82 of CdiA-CT <sup>536</sup> fused	This study



	to the nuclease domain of colicin E5 (colE5-CT) together with ImE5-His <sub>6</sub> , Amp <sup>R</sup>	
pTrc99A:: <i>traA</i>	Expresses wild-type F-pilin under the control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pTrc99A:: <i>traA(D74G)</i>	Expresses the Asp74Gly variant of F-pilin under control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pTrc99A:: <i>traA(G120C)</i>	Expresses the Gly120Cys variant of F-pilin under control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pTrc99A:: <i>trbI</i>	Expresses <i>trbI</i> under the control of the P <sub>trc</sub> promoter, Amp <sup>R</sup>	This study
pCH450:: <i>trbI</i>	Expresses <i>trbI</i> under the control of the P <sub>BAD</sub> promoter, Tet <sup>R</sup>	This study

<sup>a</sup>Abbreviations: Amp<sup>R</sup>, ampicillin resistant; Cm<sup>R</sup>, chloramphenicol resistant; Gent<sup>R</sup>, gentamicin resistant; Kan<sup>R</sup>, kanamycin resistant; Nal<sup>R</sup>, nalidixic acid resistant; Rif<sup>R</sup>, rifampicin resistant; Str<sup>R</sup>, streptomycin resistant; Tet<sup>R</sup>, tetracycline resistant

#### *Oligonucleotides used in this study*

536-Nco-for	5' - AGA CCA TGG TTG AGA ATA ATG CGC TGA G	115
536-Xho-rev	5' - GAT CTC GAG TAC AAT TAT CTG ATT GAT TTT T	115
DUF-536-Nco-for	5' - TCT CCA TGG GCG TAG ATC CGT CGA AAC TGA C	This study
VENN-less-Nco-for	5' - AAT GCC ATG GGT CTG GTT GCC AGA GG	This study
tRNase-Nco-for	5' - ATA CCA TGG GTT CCG GGG CTG CCT C	This study
536-Cys(mut)	5' - CTT TAG TCC TGC TAG GTG CTG CGA CCG CAC TGC CTC TGG CAA C	This study
pET-Sph/Pst	5' - CAA GGA ATG GTG CAT GCC TGC AGA TGG CGC CC	179
pET-Kpn/Nco	5' - AGT CCA TGG TAC CTC TCC TTC TTA AAG	This study
536-H178A-for	5' - GGA GGA TAT TGG GAT GCT ATG CAG GAA ATG C	123
536-CT-Xho-rev	5' - TTA CTC GAG GTA ATC ATA TTC CAT A	This study
536-cdiI-Eco-for	5' - AGG GAA TTC CAT ATG ATT ACC TTA CGT AAA	This study

536-cdiI-Spe-rev	5' - GTG ACT AGT TAC AAT TAT CTG ATT GAT TT	This study
colE5-M429-Nco	5' - CTG CCA TGG AAA GCA GGA AGA AG	This study
imE5-Spe	5' - AAA ACT AGT CAT CTT TAA CGT GAT AAT GAA AGC	This study
immE5-Eco-for	5' - GGA GAA TTC TCT ATG AAG TTA TCA CC	This study
pET-Pst-rev	5' - CGG CTG CAG CAG CCA ACT CAG TGG	This study
colE5-Nco-for	5' - GAG TTT CCA TGG GCG GTG GCG ATG G	This study
colE5-Bam-rev	5' - AGC GGA TCC TTT ATT ATC CTC TTT CTT CTT C	This study
colE5-Bam-for	5' - AAA GGA TCC AGA GAT GCT GAA GGC AAA C	This study
immE5-Xho-rev	5' - CAC CTC GAG CAT CTT TAA CGT GAT AAT G	This study
536-Bam-rev	5' - GCC CCG GAT CCG TAC TTA TC	This study
traA-Nco-for	5' - TAA CCC ATG GTT GCT GTT TTA AGT GTT CAG G	This study
traA-Spe-rev	5' - TTT ATA CTA GTT CAG AGG CCA ACG ACG GCC ATA CC	This study
traA-D74G	5' - TTC GGT AAG GGC TCC AGT GTT GTT AAA TGG G	This study
traA-G120C-Spe	5' - TAT ACT AGT TCA GAG GCA AAC GAC GGC C	This study
ECL-CT-Kpn-for	5' - GGG GGT ACC ATG GCT GAG AAT AAC TCG CTG GC	This study
ECL-cdiI-Nhe-rev	5' - CTC TGC TAG CGT TGT TAA GAC TAT GAT AAA AAT C	This study
ECL-cdiI-Eco-for	5' - GAT TAA GGA ATT CTG GTA TGT TTG G	This study
ECL-cdiI-Sac-rev	5' - AGT GAG CTC TGT TTC AGT TGT TAA GAC	This study
argQ probe	5' - CCT CCG ACC GCT CGG TTC G	<sup>179</sup>
tyrU probe	5' - CTT CGA AGT CTG TGA CGG CAG	<sup>120</sup>
traA-cat-for	5' - ACA TTT AAT ACA CTC TAG TTT TAT TCA TTT ATC CGA AAT TGA GGT AAC TTA TGG AGA AAA AAA TCA CTG GAT ATA CC	This study
traA-cat-rev	5' - GGA AAC GAT ATT TCT TAA GTT TAT TCT CGT CTC CCG ACA TCG TTT TAT TTC CTG TTA CGC CCC GCC CTG CCA CTC	This study
traT-cat-for	5' - AAA ACA AGA AGT TAT CAA GAG TAA AAT AAA AGA TAT TAG AGA GTA AAT ATA	This study

	TGG AGA AAA AAA TCA CTG GAT ATA CC	
traT-cat- rev	5' - CAA AGC GAG GCG TCA GTC AGG AGG CCG GTC AGA CCA GCC TCC GGA AGA TAA TTA CGC CCC GCC CTG CCA CTC	This study
trbI-cat- for1	5' - GGA AGA ACA TGA GAA ATA CAG GAG TGT GGC ATG AGT GAG AAA AAA ATC ACT GGA TAT ACC	This study
trbI-cat- for2	5' - AAG AAG TCA GGG CCG GAA ATG GCT TCG CTG GAA GCC TGG CTG GAA GAA CAT GAG AAA TAC	This study
trbI-cat- rev1	5' - GCC CCC ATA TCA GCA GGG CAA TCA GCC CCC GGC ATC TCA TGC CCC TCC CTG CCA CTC ATC	This study
trbI-cat- rev2	5' - CAC AGA TCG CCC CAG GTA CCA AGA TCG GCG GCG GCC ACA CTC TGC CCC CAT ATC AGC AGG	This study
trbI-Kpn- for	5' - TTT GGT ACC ATG AGT TCA ACG CAG AAC	This study
trbI-Xho- rev	5' - TTT CTC GAG TCT CAT GGT TCC GCC CTC	This study

### ***Bacterial strains and growth conditions***

Bacterial strains used in this study are listed in the table above. *E. coli* cells were grown in LB medium supplemented with antibiotics at the following concentrations: ampicillin (Amp), 150  $\mu\text{g mL}^{-1}$ ; chloramphenicol (Cm), 33  $\mu\text{g mL}^{-1}$ ; kanamycin (Kan), 50  $\mu\text{g mL}^{-1}$ ; rifampicin (Rif), 250  $\mu\text{g mL}^{-1}$ ; streptomycin (Str), 50  $\mu\text{g mL}^{-1}$ ; and tetracycline (Tet), 10  $\mu\text{g mL}^{-1}$ . Bacteria were grown in baffled flasks at 37 °C with shaking (215 rpm) unless otherwise indicated. The F<sup>-</sup> derivative of *E. coli* X90 (strain KW1070) was a generous gift from Dr. Kelly P. Williams. The  $\Delta\text{tolA}::\text{kan}$  and  $\Delta\text{tonB}::\text{kan}$  gene disruptions were obtained from the Keio collection<sup>184</sup> and introduced into *E. coli* X90 by bacteriophage P1-mediated transduction. The  $\Delta\text{traA}::\text{cat}$ ,  $\Delta\text{traT}::\text{cat}$  and  $\Delta\text{trbI}::\text{cat}$  disruptions were generated by phage  $\lambda$  Red-mediated recombination as described<sup>185</sup>. The chloramphenicol acetyltransferase (*cat*) open reading frame was amplified with oligonucleotides containing homology to regions

flanking *traA* (*traA*-cat-for/*traA*-cat-rev) and *traT* (*traT*-cat-for/*traT*-cat-rev). A similar procedure was used for the *trbI* gene, but the *cat* gene was amplified by two sequential reactions with *trbI*-cat-for1/*trbI*-cat-rev1 and *trbI*-cat-for2/*trbI*-cat-rev2 primer pairs. The resulting PCR products were electroporated into *E. coli* X90 cells expressing the Red proteins from plasmid pSIM6<sup>186</sup>. The F':Tn10 (from strain XL-1) and pOX38::*gent* plasmids were transferred into *E. coli* strain KW1070 by conjugation. Donor and recipient cells were mixed and spotted onto LB-agar for 4 hr at 37 °C. Exconjugants were selected on LB-agar supplemented with: Tet/Rif for F':Tn10 transfer into *E. coli* KW1070, Gent/Rif for pOX38::*gent* transfer into KW1070, and Tet/Str for F':Tn10 transfer into *E. coli* MC4100.

### ***Plasmids***

Plasmids used in this study are listed in the table above. Plasmids pET21::*cdiA-CT/cdiI*<sup>536</sup> and pET21::*cdiA-CT/cdiI*<sup>3937</sup> were used to over-produce CdiA-CT/CdiI<sup>536</sup>-His<sub>6</sub> and CdiA-CT/CdiI<sup>3937</sup>-His<sub>6</sub> complexes (respectively)<sup>115</sup>. All other CdiA-CT/CdiI<sup>536</sup>-His<sub>6</sub> over-expression constructs were generated by PCR and ligated into plasmid pET21P using NcoI/XhoI restriction sites. The DUF-CT, VENN-less and tRNase domain constructs were amplified using primers DUF-536-Nco-for, VENN-less-Nco-for and tRNase-Nco-for (respectively) in conjunction with primer 536-Xho-rev. Residues Cys13 and Cys19 of CdiA-CT<sup>536</sup> were mutated to serine by mega-primer PCR<sup>187</sup>. A fragment of plasmid pET21::*cdiA-CT/cdiI*<sup>536</sup> was amplified using primers pET-Sph/Pst and 536-Cys(mut). This product was purified and used as a mega-primer in a second PCR with primer 536-Xho-rev. The final product was digested with NcoI/XhoI and ligated to plasmid pET21P. The *colE5-CT* and *imE5* coding sequences were amplified from plasmid ColE5-099 using primers colE5-M429-

Nco and imE5-Spe, and the product digested with NcoI/SpeI and ligated to plasmid pET21P. The CdiA-CT/ColE5-CT fusion construct was generated by sequential ligation of *cdiA-CT*<sup>536</sup> and *colE5-imE5* fragments into plasmid pET21P. The *colE5-CT/imE5* sequence was amplified with primers ColE5-Bam-for and ImE5-Xho-rev and ligated to plasmid pET21P. The coding sequence for Val1 – Tyr82 of CdiA-CT<sup>536</sup> was amplified with primers 536-Nco-for and 536-Bam-rev and ligated to plasmid pET21::*colE5-CT/imE5* using NcoI and BamHI restriction sites. The CdiA-CT/CdiI<sup>ECL</sup>-His<sub>6</sub> expression construct was generated in two steps. First, the T7 promoter region of pET21P was amplified with primers pET-Sph/Pst and pET-Kpn/Nco, and the fragment ligated to PstI/NcoI-digested pET21::*cdiA-CT/cdiI*<sup>3937</sup>. The resulting pET21K::*cdiA-CT/cdiI*<sup>3937</sup> plasmid contains a unique KpnI site upstream of NcoI. The *cdiA-CT/cdiI*<sup>ECL</sup> region was then amplified with ECL-CT-Kpn-for and ECL-cdiI-Nhe-rev and ligated to KpnI/SpeI-digested plasmid pET21K::*cdiA-CT/cdiI*<sup>3937</sup>.

The *cdiI*<sup>536</sup>, *cdiI*<sup>ECL</sup> and *imE5* immunity genes were amplified from pET21 expression plasmids using 536-cdiI-Eco-for, ECL-cdiI-Eco-for and imE5-Eco-for primers (respectively) in conjunction with the pET-Pst reverse primer. All PCR products were digested with EcoRI and PstI and ligated to plasmid pTrc99A. The *traA* coding sequence from the F plasmid was amplified with primers traA-Nco-for and traA-Spe-rev and the product ligated to NcoI/SpeI-digested plasmid pTrc99A::*rhsI<sub>B</sub>*<sup>176</sup>. The Gly120Cys mutation in *traA* was generated by PCR using primers traA-Nco-for and traA-G120C-Spe. The Asp74Gly mutation in *traA* was generated by megaprimer PCR using primer traA-D74G in conjunction with the traA-Nco-for/traA-Spe-rev primer pair. The *trbI* gene was amplified with primers trbI-Kpn-for and trbI-Xho-rev and the product ligated to KpnI/XhoI-digested pTrc99KX and pCH450K to generate *trbI* expression constructs for complementation analysis.

### ***Protein purification***

All CdiA-CT/CdiI-His<sub>6</sub> complexes were over-produced in *E. coli* strain CH2016. Cultures were grown to OD<sub>600</sub> ~ 0.7 and protein production induced with 1.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for three hr. Cells were harvested over ice and collected by centrifugation at 6,000 rpm for 10 min. Cells were resuspended in 10 mL of extraction buffer [20 mM sodium phosphate (pH 7.0), 150 mM sodium chloride, 0.05% Triton X-100, 10 mM β-mercaptoethanol (β-ME), 1 mM PMSF] and broken by french press passage at 20,000 psi. Cell lysates were cleared by two consecutive centrifugations at 14,000 rpm for 10 min (SS-34 rotor) and supernatants were incubated with Ni<sup>2+</sup>-nitrilotriacetic acid (Ni<sup>2+</sup>-NTA) resin for 1.5 hr at 4 °C. The resin was washed with 20 mM sodium phosphate (pH 7.0), 150 mM sodium chloride, 0.05% Triton X-100, 10 mM β-ME, 20 mM imidazole and loaded onto a Poly-Prep column (Bio-Rad) at 4 °C. CdiA-CT/CdiI-His<sub>6</sub> complexes were eluted from the resin with native elution buffer [20 mM sodium phosphate (pH 7.0), 150 mM sodium chloride, 10 mM β-ME, 250 mM imidazole], followed by dialysis into storage buffer [20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 10 mM β-ME]. CdiA-CT toxins were isolated from CdiI-His<sub>6</sub> immunity proteins by Ni<sup>2+</sup>-affinity chromatography under denaturing conditions [6 M guanidine-HCl, 20 mM sodium phosphate (pH 7.0), 10 mM β-ME] at room temperature. Purified CdiA-CT toxins were dialyzed against storage buffer and quantified by absorbance at 280 nm.

### ***CdiA-CT inhibition assays***

Mid-log cells were diluted to  $OD_{600} = 0.05$  in fresh LB medium and incubated at 37 °C with shaking. After growth to  $OD_{600} = 0.1 - 0.15$ , purified CdiA-CT, ColE5-CT or CdiA-CT/CdiI complex was added to the culture at a final concentration of 100 nM. Cell growth was monitored by measuring  $OD_{600}$  every 30 min. Samples were collected periodically and total RNA isolated by guanidine isothiocyanate-phenol extraction as described<sup>179</sup>. Northern blot analysis was performed as described<sup>182</sup> using 5'-radiolabeled oligonucleotide probes to  $tRNA_{ICG}^{Arg}$  and  $tRNA_{GUA}^{Tyr}$ . Expression of plasmid-borne *traA* alleles was induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) prior to the addition of purified CdiA-CT<sup>536</sup>. The stability of CdiA-CT<sup>536</sup> in shaking broth cultures was determined by immunoblot analysis. Culture samples were removed at 0, 1, 3 and 5 hr and the media clarified by centrifugation at 14,000 5g for 10 min. The supernatants (10  $\mu$ L) were run on SDS-polyacrylamide gels and blotted onto PVDF membrane. Proteins were detected using polyclonal antibodies raised against CdiA-CT<sup>536</sup><sup>169</sup>.

### ***Bacteriophage plating and mating efficiency***

Bacteriophage resistance was determined by measuring the efficiency of plating. X90 and derivatives were incubated with 10 and 100 plaque forming units (pfu) and then plated in soft agar. The number of plaques was divided by the number of infecting particles. Phage-resistant strains were also challenged with  $10^3$  pfu, but no plaques were detected. The *traA* and *trbI* mutants were tested for mating efficiency under the same conditions as CdiA-CT inhibition assays.  $F'::Tn10$  donor cells ( $Tet^R$ ) were co-cultured with MC4100 recipient cells ( $Str^R$ ) at a 1:10 ratio in LB for 3 hr at 37 °C. Samples were collected and plated onto LB-agar

supplemented with Tet to enumerate F':Tn10 donor cells and Tet/Str quantify exconjugants. Mating efficiency was calculated by dividing the number of exconjugants by total donor cells and expressed as a percentage. The effect of purified CdiA-CT<sup>536</sup> proteins on conjugation was determined using MC4100 F':Tn10 donors and MC4100 pTrc recipients. CdiA-CT proteins were added at 1  $\mu$ M final concentration to conjugation co-cultures and mating efficiency determined as described above. Phage were inactivated with UV irradiation (64 mJ m<sup>-2</sup>) in a Stratalinker. This dose resulted in a 10<sup>6</sup>-fold reduction in R17 and M13 plaque-forming units (pfu). UV-inactivated phage ( $\sim 10^5$  particles mL<sup>-1</sup>) were added to conjugation co-cultures to determine the effect on mating efficiency.

## Results

### *Purified CdiA-CT<sup>536</sup> inhibits E. coli growth*

To test whether CdiA-CT toxins enter bacteria in the absence of cell-to-cell contact, we treated *E. coli* cultures with purified toxins from *Dickeya dadantii* 3937 (CdiA-CT<sup>3937</sup>) and uropathogenic *E. coli* 536 (CdiA-CT<sup>536</sup>). Each toxin was first purified as a complex with its cognate His<sub>6</sub>-tagged CdiI protein and then separated from the immunity protein under denaturing conditions. CdiA-CT<sup>536</sup> and CdiA-CT<sup>3937</sup> refold efficiently and regain nuclease activity when denaturant is removed by dialysis<sup>115,123</sup>. We added each toxin (100 nM final concentration) to *E. coli* X90 cultures and monitored cell growth. Purified CdiA-CT<sup>3937</sup> had no discernable effect on growth, but cells treated with CdiA-CT<sup>536</sup> were inhibited compared to the buffer-treated control (Fig. 1A). We also tested the CdiA-CT/CdiI<sup>536</sup> complex, and somewhat surprisingly found that cells treated with the toxin/immunity complex were inhibited to the same extent as cells treated with CdiA-CT<sup>536</sup> toxin alone (Fig. 1A). This



result suggests that only the toxin enters cells and immunity protein remains in the media. Alternatively, the CdiA-CT/CdiI<sup>536</sup> complex might inhibit growth by a novel mechanism that is independent of the toxin's tRNase activity. To differentiate between these possibilities, we added purified CdiA-CT<sup>536</sup> to cells that express plasmid-borne immunity genes. Cells expressing *cdiI*<sup>536</sup> were not inhibited by added CdiA-CT<sup>536</sup> toxin (Fig. 1B), indicating that immunity protein is protective when present in the cytoplasm. This protection is specific, because cells expressing non-cognate *cdiI*<sup>3937</sup> were inhibited to the same extent as cells that carry no immunity gene (Fig. 1B). To confirm that purified CdiA-CT<sup>536</sup> inhibits growth through its anticodon nuclease activity, we isolated RNA from the toxin-treated cultures and examined transfer RNA by northern blot. This analysis revealed cleaved tRNA<sub>ICG</sub><sup>Arg</sup> in all CdiA-CT<sup>536</sup>-treated cells except those expressing the cognate *cdiI*<sup>536</sup> immunity gene (Fig. 1C). Together, these results show that exogenous CdiA-CT<sup>536</sup> translocates into *E. coli* cells, where it cleaves tRNA to inhibit growth.

### ***The N-terminal domain of CdiA-CT<sup>536</sup> is required for cell import***

Like many CdiA-CTs, CdiA-CT<sup>536</sup> is composed of two domains<sup>123,171</sup>. The tRNase activity of CdiA-CT<sup>536</sup> is contained within the C-terminal 145 residues<sup>123</sup>, which corresponds to Gly3098 – Ile3242 of full-length CdiA<sup>536</sup> (Fig. 2A). However, *E. coli* cells were not inhibited when treated with the purified tRNase domain (Fig. 2B). This result suggests that the N-terminal domain of CdiA-CT<sup>536</sup> is required for cell import; or alternatively, the C-terminal tRNase domain may be insufficient to inhibit cell growth. To test inhibition activity, we constructed a plasmid that produces the tRNase domain under control of the arabinose-inducible P<sub>BAD</sub> promoter and asked whether this construct could be

introduced into *E. coli* cells in the presence of L-arabinose. As a control, we also transformed the plasmid into *E. coli*  $\Delta$ *cysK* cells, because CysK is required to activate the CdiA-CT<sup>536</sup> nuclease<sup>123</sup>. The construct expressing the tRNase domain yielded no stable transformants in the *cysK*<sup>+</sup> background, but was readily introduced into  $\Delta$ *cysK* cells (Fig. 3). Furthermore, *cysK*<sup>+</sup> cells were transformed at high-efficiency with either the empty vector plasmid (pCH450) or a construct encoding CdiA-CT<sup>536</sup> with the His178Ala mutation that abolishes tRNase activity<sup>123</sup> (Fig. 3). Together, these results demonstrate that the C-terminal tRNase domain is toxic when expressed inside cells and suggest that the N-terminal domain plays a role in cell import.

The VENN peptide motif is predicted to be part of a larger domain of unknown function, termed DUF638, which is found in many CdiA proteins<sup>115</sup>. Given the conservation of the VENN motif, we asked whether this sequence is required for CdiA-CT<sup>536</sup> import. Cells treated with CdiA-CT<sup>536</sup> lacking this sequence (VENN-less, residues Leu3023 – Ile3242) were inhibited to the same extent as cells treated with the full CdiA-CT<sup>536</sup> containing residues Val3016 – Ile3242 (Fig. 2A & 2B). Surprisingly, a larger fragment (containing residues Gly2969 – Ile3242) that includes the entire DUF638 region had no effect on cell growth (Figs. 2A & 2B), even though this protein is toxic when produced inside *cysK*<sup>+</sup> cells (Fig. 3). Although CdiA-CT sequences are diverse, their N-terminal domains often contain paired cysteine residues that presumably form disulfide linkages. We reasoned that this predicted disulfide bond could be critical for toxin import, and therefore mutated Cys3028 and Cys3034 to serine residues to generate a "Cys-less" toxin (Fig. 2A). Cells treated with the Cys-less CdiA-CT<sup>536</sup> were not inhibited (Fig. 2B), yet this toxin variant still inhibited growth when produced inside *E. coli* cells (Fig. 3). The preceding experiments results

suggest a role for N-terminal domain of CdiA-CT<sup>536</sup> in cell import. However, it is also possible that some of the toxin variants are unstable in the growth medium and do not inhibit cell growth because they are rapidly degraded. To address this possibility, we monitored CdiA-CT<sup>536</sup> antigen using immunoblot analysis, which revealed that each toxin was stable in shaking-broth cultures for up to five hours at 37 °C (data not shown). Together, these data indicate that the N-terminal domain of CdiA-CT<sup>536</sup> is required for import. However, the DUF368 region appears to block CdiA-CT<sup>536</sup> translocation, perhaps by interacting with the N-terminal region and masking cell-import epitopes.

The N-terminal domain of CdiA-CT<sup>536</sup> shares significant sequence identity with the corresponding region of CdiA<sup>ECL</sup> from *Enterobacter cloacae* ATCC 13047 (ECL), but the C-terminal nuclease domains are not related in sequence (Fig. 4A). Moreover, the C-terminal domain of CdiA-CT<sup>ECL</sup> has a distinct RNase activity that cleaves 16S ribosomal RNA<sup>188</sup>. We reasoned that if the N-terminus of CdiA-CT<sup>536</sup> mediates cell import, then purified CdiA-CT<sup>ECL</sup> (corresponding to residues Ala3087 – Asp3321) should also inhibit *E. coli* growth. We found that the purified CdiA-CT/CdiI<sup>ECL</sup> complex had no effect on *E. coli* X90 cells (data not shown), but the isolated CdiA-CT<sup>ECL</sup> toxin inhibited cell growth (Fig. 4B). This result may indicate that CdiI<sup>ECL</sup> does not dissociate from CdiA-CT<sup>ECL</sup> during translocation, or alternatively the bound immunity protein may prevent toxin import altogether. Regardless, the purified CdiA-CT<sup>ECL</sup> toxin entered cells because the *cdiI*<sup>ECL</sup> immunity gene prevented growth inhibition, whereas non-cognate *cdiI*<sup>536</sup> provided no protection (Fig. 4B). Thus, the homologous N-terminal domains of CdiA-CT<sup>536</sup> and CdiA-CT<sup>ECL</sup> appear to function similarly in cell import.

### ***The N-terminal domain of CdiA-CT<sup>536</sup> is sufficient for cell import***

The shared N-terminal domain of CdiA-CT<sup>536</sup> and CdiA-CT<sup>ECL</sup> is required for toxin import, suggesting that it carries tethered nuclease domains into the cell. We tested this hypothesis by asking whether the N-terminal domain still supports translocation when fused to a heterologous passenger domain. We fused residues Val1 – Tyr82 of CdiA-CT<sup>536</sup> to the C-terminal nuclease domain of colicin E5 (ColE5-CT, residues Ala451 – Gln556) (Fig. 5A). We chose the ColE5-CT domain as a passenger because it is similar in size and activity to the CdiA-CT<sup>536</sup> tRNase domain (Fig. 5A)<sup>101,110</sup>. Although full-length colicin E5 enters and kills *E. coli* cells<sup>101</sup>, the isolated ColE5-CT domain has no inhibition activity because it lacks receptor-binding and translocation domains (Figs. 5A & 5B). We fused the N-terminal domain of CdiA-CT<sup>536</sup> to the ColE5-CT nuclease domain and tested the purified fusion protein on *E. coli* cultures. The fusion had weak activity against *E. coli* X90 cells (data not shown), but inhibited *E. coli* XL-1 cells more profoundly (Fig. 5B). This growth inhibition was blocked when cells express the *imE5* gene (Fig. 5B), which encodes an immunity protein that specifically neutralizes ColE5-CT activity<sup>101,110</sup>. Because ColE5-CT cleaves the anticodon loops of tRNA<sup>His</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Asp</sup> and tRNA<sup>Asn</sup> molecules<sup>101</sup>, we examined tRNA<sup>Tyr</sup> by northern blot to detect nuclease activity. Cleaved tRNA<sup>Tyr</sup> was detected in treated cells that either lack immunity or express *cdiI*<sup>536</sup> (Fig. 5C). By contrast, there was less cleaved tRNA<sup>Tyr</sup> in cells that express *imE5* (Fig. 5C), consistent with the ability of ImE5 to neutralize ColE5-CT activity. Together, these results indicate that the N-terminal domain of CdiA-CT<sup>536</sup> is sufficient to translocate the colicin E5 nuclease domain into *E. coli* cells.

### ***CdiA-CT<sup>536</sup> import is independent of BamA, Tol and Ton translocation pathways***

Cell-mediated CDI requires specific receptors on target cells. Therefore, we asked whether CdiA-CT<sup>536</sup> import requires the BamA receptor using *bamA101* mutant cells, which have approximately five-fold less BamA on the cell surface and are significantly resistant to cell-mediated CDI<sup>168,169</sup>. *E. coli bamA101* cells were inhibited by purified CdiA-CT<sup>536</sup> to the same extent as *bamA*<sup>+</sup> cells (Fig. 6A), suggesting the toxin is not imported through the usual CDI pathway. This result led us to consider other mechanisms for CdiA-CT<sup>536</sup> import. Colicins exploit either the Tol or Ton systems to translocate their toxin domains into *E. coli* cells<sup>189</sup>. Therefore, we generated X90  $\Delta$ *tolA* and  $\Delta$ *tonB* strains and tested these cells for resistance to group A (colicin E5) and group B (colicin D) colicins. As expected from previous studies, we found that  $\Delta$ *tolA* cells were completely resistant to colicin E5, but inhibited by colicin D (Fig. 9). Reciprocally,  $\Delta$ *tonB* cells were resistant to colicin D and inhibited by colicin E5 (Fig. 9). Moreover, X90  $\Delta$ *tolA* cells were resistant to phage M13, but not phage R17 (Table 1), consistent with the established phenotype of this mutant. We then treated the  $\Delta$ *tolA* and  $\Delta$ *tonB* strains with purified CdiA-CT<sup>536</sup> and found that the mutants were as sensitive as *tolA*<sup>+</sup> *tonB*<sup>+</sup> cells (Figs. 6B & 6C). We also tested a  $\Delta$ *tolA*  $\Delta$ *tonB* strain to test for possible redundancy and found that the double mutant was also inhibited by CdiA-CT<sup>536</sup>, though it grew to slightly higher density than treated *tolA*<sup>+</sup> *tonB*<sup>+</sup> cells (Fig. 6D). Together, these findings suggest that purified CdiA-CT<sup>536</sup> toxin enters *E. coli* cells through a novel pathway.

### ***CdiA-CT<sup>536</sup> import requires the F pilus***

Although *E. coli* X90 and XL-1 strains are inhibited by purified CdiA-CT<sup>536</sup>, *E. coli* strain MC4100 is resistant to the toxin (data not shown). There are several genetic

differences between these strains, but both sensitive strains carry F' episomes whereas the resistant strain does not. Therefore, we tested *E. coli* X90 cells that had been cured of F' and found that this isolate is resistant to purified CdiA-CT<sup>536</sup> (Fig. 7A). Reintroduction of a different F' episome (from *E. coli* XL-1) into the cured X90 cells restored toxin sensitivity (Fig. 7A). Moreover, *E. coli* MC4100 cells that carry F' were also sensitive to purified CdiA-CT<sup>536</sup> (data not shown). Together, these results indicate that one or more genes on F are required for CdiA-CT<sup>536</sup> import. As a first step to identify F-encoded import gene(s), we tested whether cells carrying a subset of F genes are sensitive to CdiA-CT<sup>536</sup>. Plasmid pOX38::*gent* is derivative of F that contains the *tra* genes required for conjugation and a selectable gentamicin-resistance (*gent*) marker<sup>183,190</sup>. We introduced pOX38::*gent* into *E. coli* X90 F<sup>-</sup> and found that the resulting cells were sensitive to CdiA-CT<sup>536</sup> (Fig. 7A), suggesting that *tra* gene(s) could mediate toxin import. We took a candidate approach and disrupted two genes, *traA* and *traT*, that encode abundant cell-surface proteins. The *traA* gene encodes F pilin, which polymerizes to form the conjugative pilus; and *traT* encodes an outer-membrane  $\beta$ -barrel protein that functions to prevent mating between F<sup>+</sup> cells<sup>191,192</sup>. The  $\Delta traT$  mutant was still inhibited by CdiA-CT<sup>536</sup>, but  $\Delta traA$  cells were completely resistant to purified toxin (Fig. 7B). To exclude possible polar effects from the  $\Delta traA$  mutation, we complemented the mutant with plasmid-borne *traA* and restored sensitivity to the toxin (Fig. 7C). These results suggest that the F pilus is required for the import of CdiA-CT<sup>536</sup> toxin.

### ***CdiA-CT<sup>536</sup> blocks F-mediated conjugation***

We hypothesized that the N-terminal domain of CdiA-CT<sup>536</sup> binds directly to F pilin and exploits the organelle to enter cells. This import pathway could be analogous to that used

by some F-dependent bacteriophages to infect *E. coli* F<sup>+</sup> cells. Therefore, we examined *traA* mutations that confer resistance to phage R17, reasoning that these alleles may also protect cells from purified CdiA-CT<sup>536</sup>. We focused on the Asp74Gly and Gly120Cys mutations, because they produce functional conjugative pili, but interfere with phage R17 binding<sup>193,194</sup>. We introduced each mutation into plasmid-borne *traA* and used the constructs to complement X90  $\Delta traA::cat$  mutants. Each pilin variant supported conjugation, though the mating efficiency was somewhat reduced by the Asp74Gly mutation (Table 2). The mutant pilins also provided resistance to phage R17, but not phage M13 (Table 1). Cells expressing the Asp74Gly pilin were still inhibited by purified CdiA-CT<sup>536</sup>, but the Gly120Cys mutation provided partial protection against the toxin (Figs. 7D & 7E). These results suggest that CdiA-CT<sup>536</sup> binds F pili at a site that overlaps with the phage R17 binding site.

Because F-dependent bacteriophage bind directly to F pili, they can interfere with conjugation<sup>195</sup>. We reasoned that if CdiA-CT<sup>536</sup> binds directly to the F pilus, then it may also reduce mating efficiency. We first tested the effects of M13 and R17 phages on mating efficiency. The phages were inactivated with ultraviolet radiation to prevent infection and cell lysis, then added to conjugation co-cultures. The phage particles reduced mating efficiency to ~50% from about 88% for mock-treated cultures (Fig. 8). We next tested purified CdiA-CT<sup>536</sup> carrying the His178Ala mutation so that the F<sup>+</sup> donor cells were not inhibited during culture. Purified CdiA-CT<sup>536</sup> at 1  $\mu$ M reduced mating efficiency approximately 2.5-fold compared to mock-treated cells (Fig. 8). By contrast, the purified tRNase domain (which lacks the N-terminal domain of CdiA-CT<sup>536</sup>) had little effect on mating efficiency (Fig. 8). Similarly, the Cys-less and DUF638-containing toxins (see Fig. 2A) also had no substantive effect on conjugation (Fig. 8). These observations suggest that

the tRNase domain and the latter CdiA-CT<sup>536</sup> variants do not inhibit cell growth because they do not bind to F pili. Taken together with the protective effect of the *traA(G120C)* mutation, these results support a model in which the N-terminal domain of CdiA-CT<sup>536</sup> binds to the F pilus.

***trbI mutants are resistance to F-dependent bacteriophage and CdiA-CT<sup>536</sup>***

F pili are dynamic and undergo cycles of extension and retraction<sup>196</sup>. Pilus retraction is critical for F-dependent phage infection and provides a possible mechanism to internalize CdiA-CT<sup>536</sup> toxin. The *trbI* gene is thought to be required for pilus retraction<sup>197</sup>, so we generated an in-frame *trbI* deletion strain for analysis. We found that  $\Delta trbI$  cells are resistant to both R17 and M13 phage (Table 1), but still capable of conjugation, albeit at lower efficiency than wild-type cells (Table 2). These observations indicate that the  $\Delta trbI$  mutant produces functional F pili as reported previously<sup>197</sup>. We next treated  $\Delta trbI$  cells with purified CdiA-CT<sup>536</sup> and found that the mutant was fully resistant to the toxin (Fig. 7F). We attempted to complement the mutant with plasmid-borne *trbI*, but were unable to do so with both IPTG- and arabinose-inducible *trbI* constructs (data not shown). We note that Maneewannakul *et al.* could not complement *trbI* mutants with multi-copy plasmids. Moreover, they reported that *trbI* over-expression in wild-type cells phenocopies the *trbI* null mutation<sup>197</sup>. We also found that *trbI* over-expression in the *trbI*<sup>+</sup> background conferred resistance to CdiA-CT<sup>536</sup>, similar to the  $\Delta trbI$  phenotype (Fig. 7F). Therefore, pili dynamics appear to be adversely affected by under- and over-expression of *trbI*<sup>197</sup>. These results are consistent with a model in which pilus-bound toxin is internalized during retraction.



## Discussion

The experiments presented here reveal an unexpected import pathway for CDI toxins. Purified CdiA-CT<sup>536</sup> and CdiA-CT<sup>ECL</sup> enter *E. coli* F<sup>+</sup> cells and inhibit growth. These toxins inhibit growth using different C-terminal nuclease domains, but they share a common N-terminal domain. The N-terminal domain is not necessary for nuclease activity *in vitro*<sup>123</sup>, but is required for toxin translocation into F<sup>+</sup> cells. These results indicate that the N-terminal domain has autonomous import activity and carries tethered passenger domains into the cell. Indeed, the translocation domain is capable of transporting a heterologous nuclease domain from colicin E5 into F<sup>+</sup> cells. Several observations suggest that toxin import requires a binding interaction between the N-terminal domain and F pili. First, purified CdiA-CT<sup>536</sup> interferes with mating when added to conjugation co-cultures. The same phenomenon is observed with phage R17 and M13 particles, both of which bind directly to F pili. Presumably, the binding of either phage or toxin to F pili disrupts the formation of conjugation bridges. This effect is specific because the C-terminal tRNase domain of CdiA-CT<sup>536</sup> has little to no effect on mating efficiency, and mutations within the N-terminal domain that block cell-import also abrogate the toxin's effect on conjugation. Furthermore, cells expressing F pilin with the Gly120Cys mutation are significantly resistant to purified CdiA-CT<sup>536</sup>. This mutation also disrupts the interaction between F pili and phage R17<sup>194</sup>, suggesting that phage and CdiA-CT<sup>536</sup> share overlapping binding sites. Taken together, these results indicate that CdiA-CT<sup>536</sup> and CdiA-CT<sup>ECL</sup> contain a pilus-binding domain that facilitates entry into F<sup>+</sup> cells.

F pili are dynamic structures that extend in search of potential mating partners and retract to bring donor and recipient cells together to establish conjugation bridges<sup>196,198</sup>. This

cycle of extension and retraction also provides a possible mechanism to import CdiA-CT<sup>536</sup> toxin across the cell envelope. Retraction is driven by pilus disassembly into pilin monomers within the inner membrane<sup>199</sup>. Therefore, pilus-bound CdiA-CT<sup>536</sup> could be carried directly to the inner membrane during retraction. If this model is correct, then the toxin must remain stably associated with the pilus as it retracts through the lumen of the type IV secretion assembly. Type IV secretion assemblies are large, barrel-shaped structures that span the entire cell envelope<sup>200,201</sup>. The F-pilus assembly is composed of a hetero-oligomeric complex of TraV, TraK and TraB, with a tetradecameric ring of TraV forming an aperture in the outer membrane through which the pilus emerges<sup>202</sup>. Presumably, TraV forms a tight seal around the pilus to prevent the loss of periplasmic contents to the media. Therefore, the associated CdiA-CT<sup>536</sup> must pass through the TraV aperture and not be stripped from the pilus surface. Pilus retraction is also critical for F-dependent phage infection, and a number of phage-resistance mutations appear to disrupt pilus dynamics<sup>203,204</sup>. These resistance mutations are unmapped, and to our knowledge, *trbI* is the only F gene to be specifically linked to defects in pilus retraction<sup>197</sup>. As predicted by the retraction-import model, in-frame deletion of *trbI* protects F<sup>+</sup> cells from CdiA-CT<sup>536</sup>, just as it provides resistance to F-dependent phages<sup>197</sup>. However, we note that pili dynamics have not been directly examined in  $\Delta trbI$  mutants. The retraction-import model would be strengthened if  $\Delta trbI$  retraction defects could be confirmed through real-time video microscopy as described by Silverman and colleagues<sup>196</sup>.

As outlined above, there are parallels between CdiA-CT<sup>536</sup> import and infection by F-dependent phages. Two bacteriophage families exploit conjugative pili as host-cell receptors. The *Inoviridae* are filamentous, single-strand DNA viruses (e.g. phages fd and M13) that bind to the tip of the pilus using the g3p capsid protein<sup>193,194</sup>. These phages also require TolA

as a co-receptor<sup>205</sup>. However, CdiA-CT<sup>536</sup> import does not require TolA, suggesting that the toxin is internalized through another pathway. Moreover, CdiA-CT<sup>536</sup> appears to bind F pili in a manner similar to phage R17 and other leviviruses. The *Leviviridae* are icosahedral, single-strand RNA viruses that attach to the side of conjugative pili using a single copy of the maturation (or assembly) protein present in the phage capsid<sup>206-208</sup>. Upon binding the pilus, the maturation protein is released from the capsid and proteolytically processed into two fragments<sup>209-211</sup>. These peptide fragments remain associated with the phage genome and are internalized with viral RNA during infection<sup>209,212,213</sup>. The internalization process also requires TraD motor function and other plasmid-transfer initiation proteins<sup>214-217</sup>. Together, these observations suggest that the maturation protein guides the tethered genome into the cytoplasm. A similar mechanism may underlie CdiA-CT<sup>536</sup> import into F<sup>+</sup> cells. Although the N-terminal domain of CdiA-CT<sup>536</sup> lacks homology with known maturation proteins, we note that these proteins are quite diverse between different phage<sup>218,219</sup>. For example, the maturation proteins from MS2, GA and Q $\beta$  phages are only 26 – 49% identical to one another, yet are all thought to bind along the shafts of F pili. Given this sequence diversity, it seems reasonable to posit that the N-terminal domain of CdiA-CT<sup>536</sup> mimics a leviviral maturation protein.

In principle, pilus-mediated import provides an additional mechanism to deliver CDI toxins into target bacteria. However, it is not clear that this mode of delivery occurs naturally. We have previously reported that *E. coli* EC93 cells release CdiA fragments into the extracellular milieu, but these fragments are not inhibitory<sup>112</sup>. Similarly, CdiA<sup>536</sup> fragments from UPEC 536 culture supernatants do not inhibit the growth of F<sup>+</sup> cells (C.M.B & C.S.H., unpublished data). Another possibility is that the binding of CdiA<sup>536</sup> to F pili augments cell-

cell interactions to potentiate CDI. We have found that F<sup>+</sup> cells are generally better CDI targets than F<sup>-</sup> cells, but this is also true for CdiA effector proteins that lack the pilus-interaction domain (C.M.B & C.S.H, unpublished data). Thus, there is no evidence that full-length CdiA<sup>536</sup> on the surface of inhibitor cells binds to F pili on target cells. If the interaction with F pili plays no role in cell-mediated CDI, then what is the biological significance of these findings? We hypothesize that these phenomena reflect a critical function for the N-terminal domain in transport across the target-cell inner membrane during CDI. In this model, the unusual pilus-binding activity of CdiA-CT<sup>536</sup> allows the toxin to by-pass the CDI receptor and gain entry into the periplasm. Once in the periplasm, the toxin is predicted to resume its normal translocation pathway, using the N-terminal domain to cross the inner membrane and enter the cytoplasm. Though this model has not been tested directly, there is evidence that the N-terminal domain of CdiA-CT<sup>536</sup> is critical for CDI. Mutation of residues Cys3028/Cys3034 within full-length CdiA<sup>536</sup> abrogates CDI, but still allows transfer of CdiA-CT<sup>536</sup> toxin antigen to the surface of target bacteria (Julia S. Webb & D.A.L., unpublished data). These observations indicate that CDI is disrupted at a later stage in the pathway, as expected for a defect in toxin translocation. If this model is correct, then our results suggest that CDI toxins and leviviral genomes may use the same basic mechanism to cross the inner membrane of bacteria.

**Table 1. Bacteriophage plating efficiency<sup>a</sup>**

<b>Bacterial strain</b>	<b>Bacteriophage</b>	
	<b>R17</b>	<b>M13</b>
X90 $\Delta tolA$	0.93 $\pm$ 0.12	0.0 <sup>b</sup>
X90 $\Delta tonB$	1.2 $\pm$ 0.26	1.0 $\pm$ 0.20
X90 $\Delta tolA \Delta tonB$	1.1 $\pm$ 0.15	0.0 <sup>b</sup>
X90 $\Delta traA::cat$ pTrc	0.0 <sup>b</sup>	0.0 <sup>b</sup>
X90 $\Delta traA::cat$ pTrc:: <i>traA</i>	1.2 $\pm$ 0.18	1.2 $\pm$ 0.15
X90 $\Delta traA::cat$ pTrc:: <i>traA(D74G)</i>	0.0 <sup>b</sup>	1.1 $\pm$ 0.09
X90 $\Delta traA::cat$ pTrc:: <i>traA(G120C)</i>	0.09 $\pm$ 0.005 <sup>c</sup>	0.77 $\pm$ 0.06 <sup>c</sup>
X90 $\Delta trbI::cat$	0.1 $\pm$ 0.04 <sup>c</sup>	0.1 $\pm$ 0.03 <sup>c</sup>

<sup>a</sup>The indicated bacterial strains were infected with 10 and 100 plaque forming units (pfu).

The number of plaques for each strain was divided by the plaques obtained with wild-type X90 cells. The mean plating efficiency  $\pm$  SEM is reported ( $n = 4$ ).

<sup>b</sup>No plaques were detected when bacteria were incubated with  $10^3$  pfu of phage.

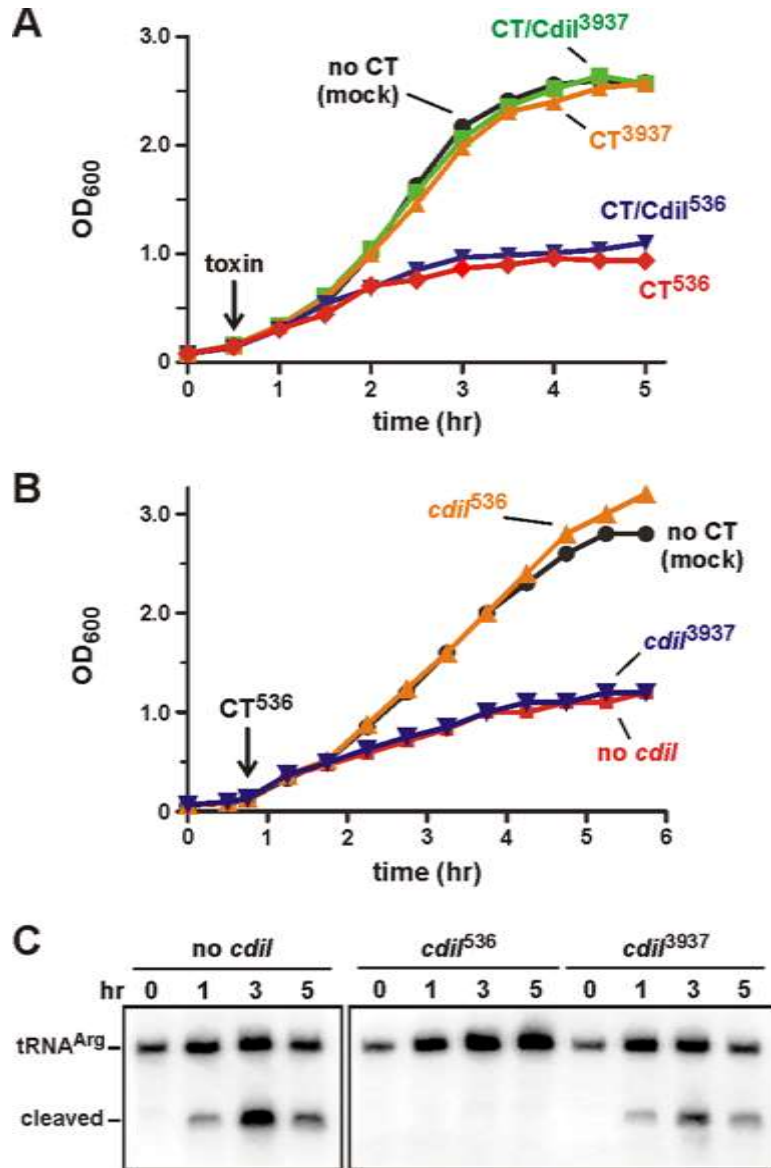
<sup>c</sup>Plaques were turbid.

**Table 2. Mating efficiency<sup>a</sup>**

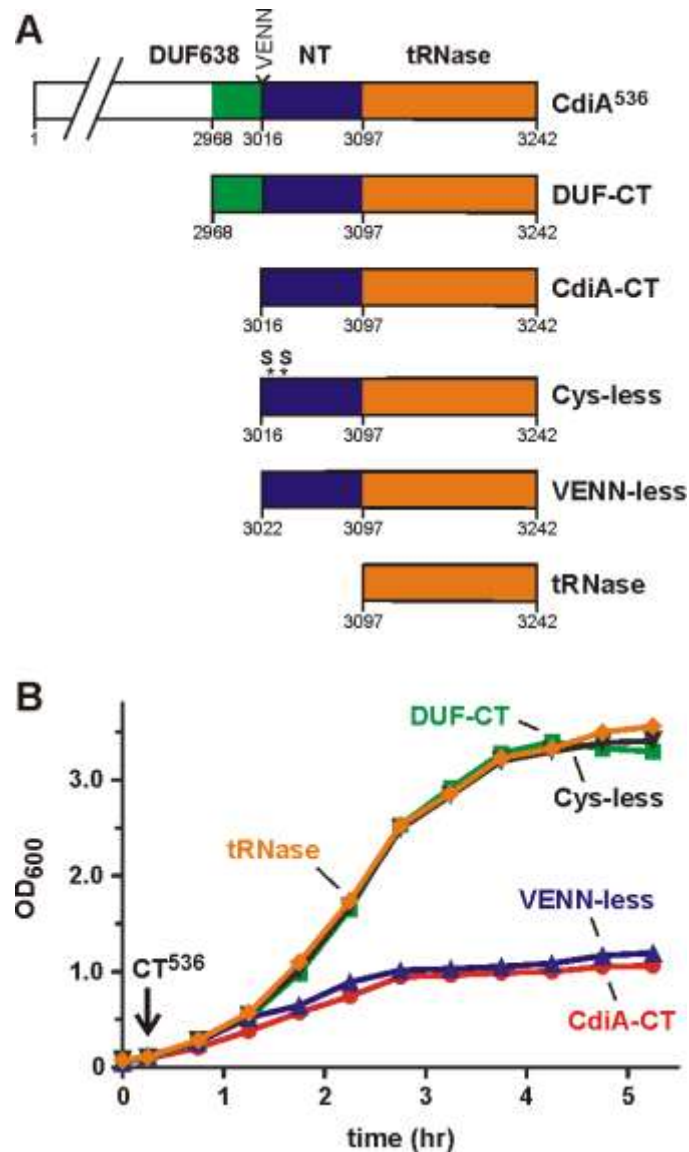
<b>F<sup>+</sup> donor strain</b>	<b>Percent exconjugants (%)</b>
CH11717	79 ± 2.8
CH11717 $\Delta traA::cat$	0.0
CH11717 $\Delta traA::cat$ pTrc	0.0
CH11717 $\Delta traA::cat$ pTrc:: <i>traA</i>	75 ± 5.4
CH11717 $\Delta traA::cat$ pTrc:: <i>traA(D74G)</i>	58 ± 5.7
CH11717 $\Delta traA::cat$ pTrc:: <i>traA(G120C)</i>	72 ± 0.5
CH11717 $\Delta trbI::cat$	13 ± 3.6

<sup>a</sup>Donor strains were derived from strain CH11717, which is X90 transduced with the Tn10 marker from *E. coli* XL-1. Donors were cultured at a 1:10 ratio with Str<sup>R</sup> recipient cells as described in Methods. Mating efficiency was determined by dividing the number of Tet<sup>R</sup> Str<sup>R</sup> colonies by the number of total Tet<sup>R</sup> colonies and expressed as a percent. The mean plating efficiency ± SEM is reported for three independent experiments.

**Figure 1. Purified CdiA-CT<sup>536</sup> inhibits *E. coli* cell growth.** **A)** Purified CdiA-CT (CT) or CdiA-CT/CdiI (CT/CdiI) complex was added to *E. coli* X90 cultures at 30 min (indicated by the arrow) and cell growth monitored by measuring the optical density at 600 nm (OD<sub>600</sub>). **B)** *E. coli* X90 cells carrying plasmid-borne *cdiI*<sup>536</sup> or *cdiI*<sup>3937</sup> immunity genes were treated with purified CdiA-CT<sup>536</sup> where indicated the arrow. **C)** Northern blot analysis of CdiA-CT<sup>536</sup> treated cells. Total RNA was isolated from the cells in panel B and tRNA<sub>ICG</sub><sup>Arg</sup> analyzed by northern blot. The migration positions of full-length and cleaved tRNA<sub>ICG</sub><sup>Arg</sup> are indicated.

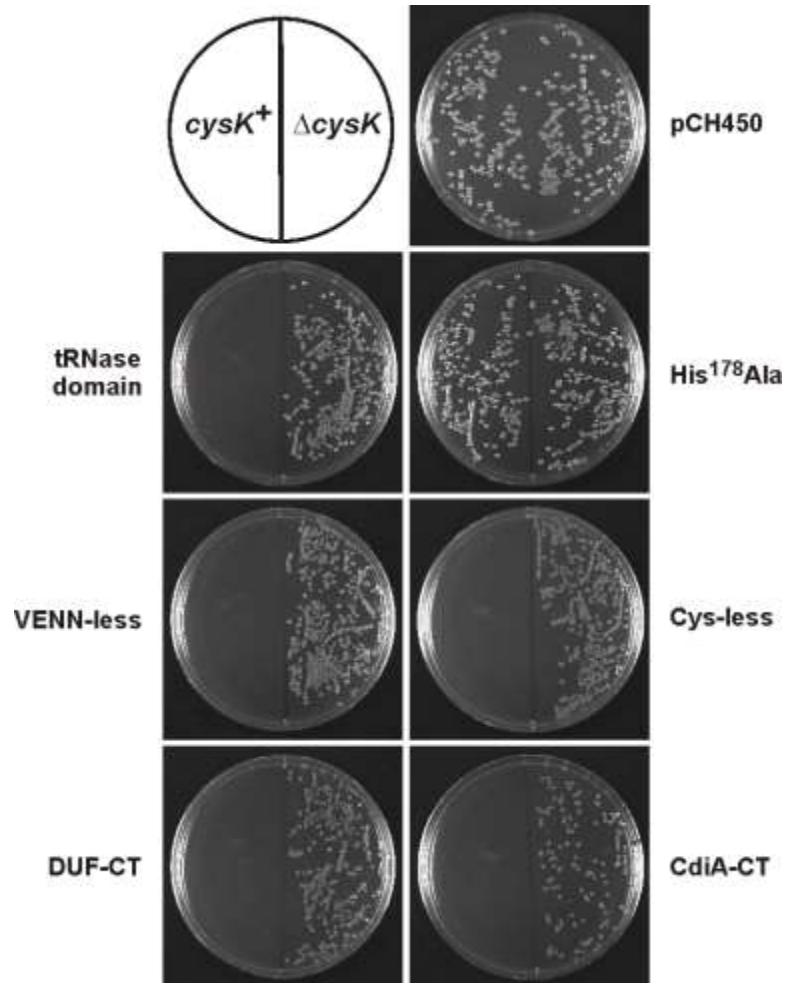


**Figure 2. The N-terminal domain of CdiA-CT<sup>536</sup> is required for import.** A) CdiA-CT<sup>536</sup> constructs. Predicted domain organization for the C-terminal region of CdiA<sup>536</sup> is depicted with the corresponding residue numbers. Residues Val3016 – Asn3019 comprise the VENN peptide motif. DUF638 is a domain of unknown function corresponding to Pfam PF04829. B) *E. coli* X90 cells were treated with the indicated CdiA-CT<sup>536</sup> fragments and cell growth monitored by OD<sub>600</sub> measurements.

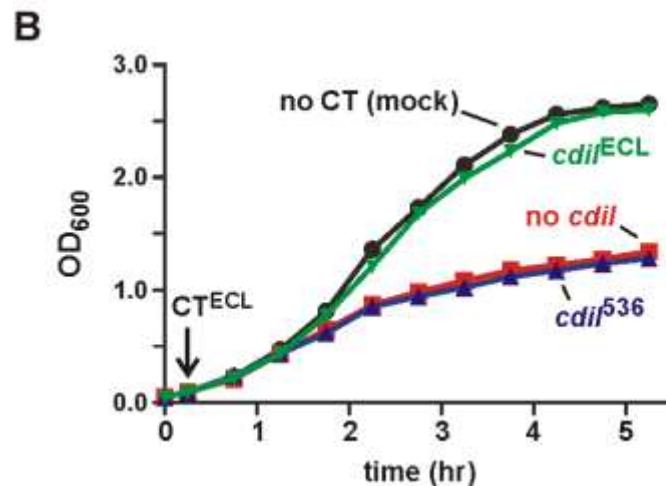
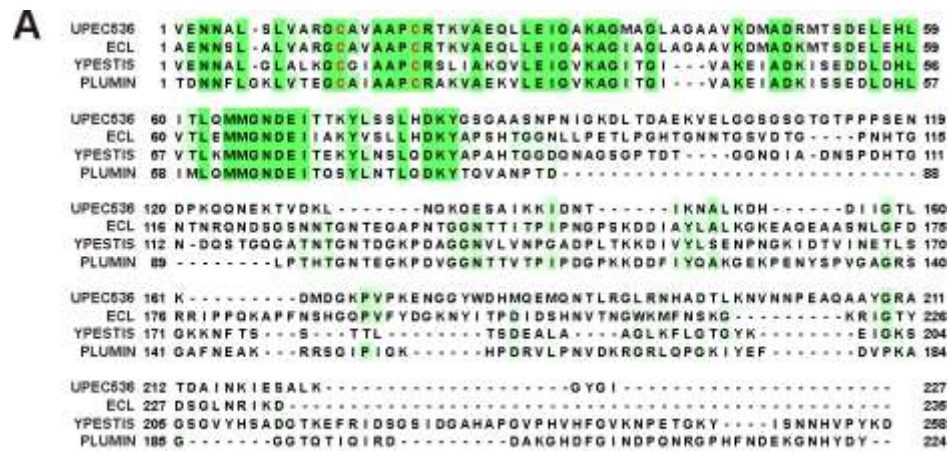




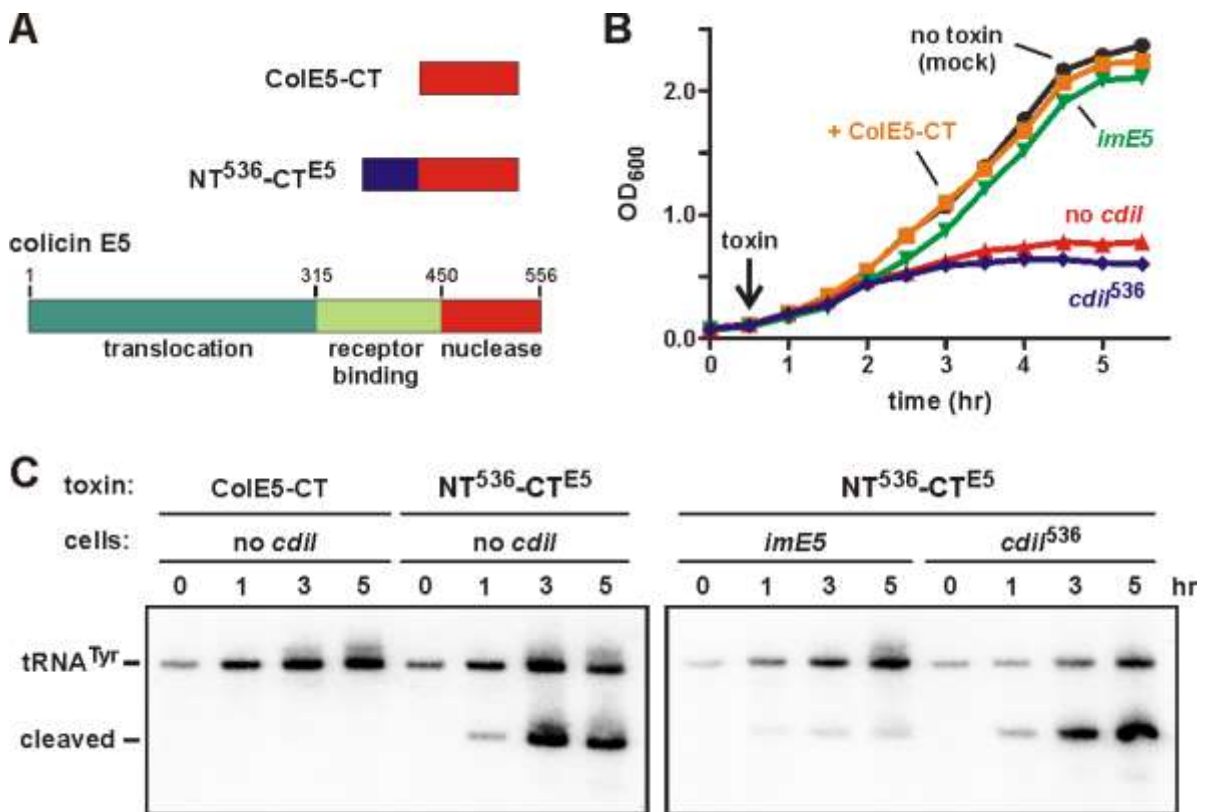
**Figure 3. CdiA-CT<sup>536</sup> toxicity.** Arabinose-inducible *cdiA-CT<sup>536</sup>* expression plasmids were introduced into *E. coli* X90 and transformants selected on LB-agar supplemented with tetracycline and L-arabinose. CdiA-CT<sup>536</sup> requires activation by CysK, therefore *E. coli*  $\Delta$ *cysK* cells were used to control for transformation efficiency. Construct nomenclature corresponds to that introduced in Fig. 2A. The His178Ala mutation ablates tRNase activity and plasmid pCH450 is the empty vector.



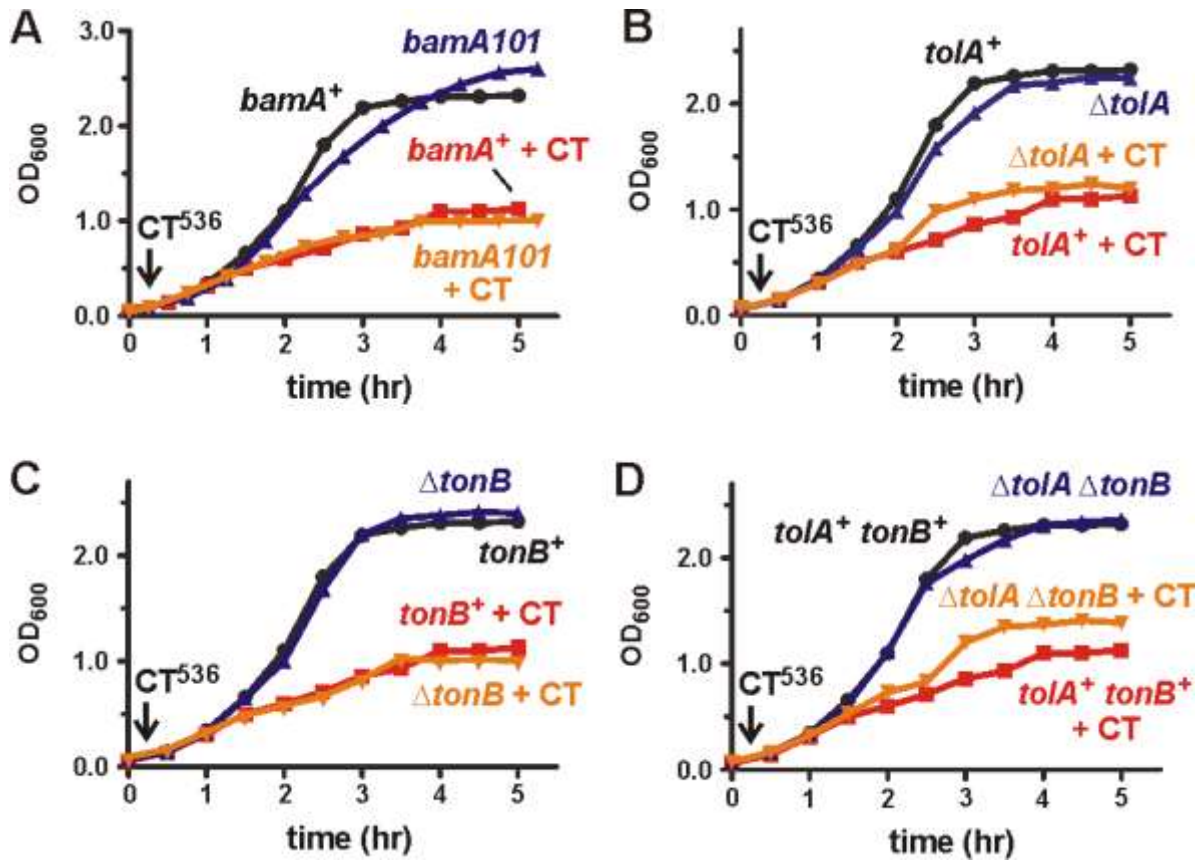
**Figure 4. CdiA-CT<sup>ECL</sup> from *Enterobacter cloacae* inhibits *E. coli* cell growth. A)** Alignment of the CdiA-CT regions from UPEC 536 (Uniprot: Q0T963), *Enterobacter cloacae* ATCC 13047 (D5CBA0), *Yersinia pestis* 91001 (Q74T84) and *Photobacterium luminescens* TT01 (Q7MB60). Proteins were aligned using Clustal-W and rendered with Jalview 2.8 at 30% sequence identity. The conserved cysteine residues within the N-terminal domain are depicted in red. **B)** Purified CdiA-CT<sup>ECL</sup> was added to *E. coli* X90 cultures when indicated (downward arrow) and cell growth monitored by OD<sub>600</sub> measurements. Where indicated, cells carried plasmid-borne *cdiI*<sup>536</sup> (*cdiI*<sup>536</sup>) or *cdiI*<sup>ECL</sup> immunity genes.



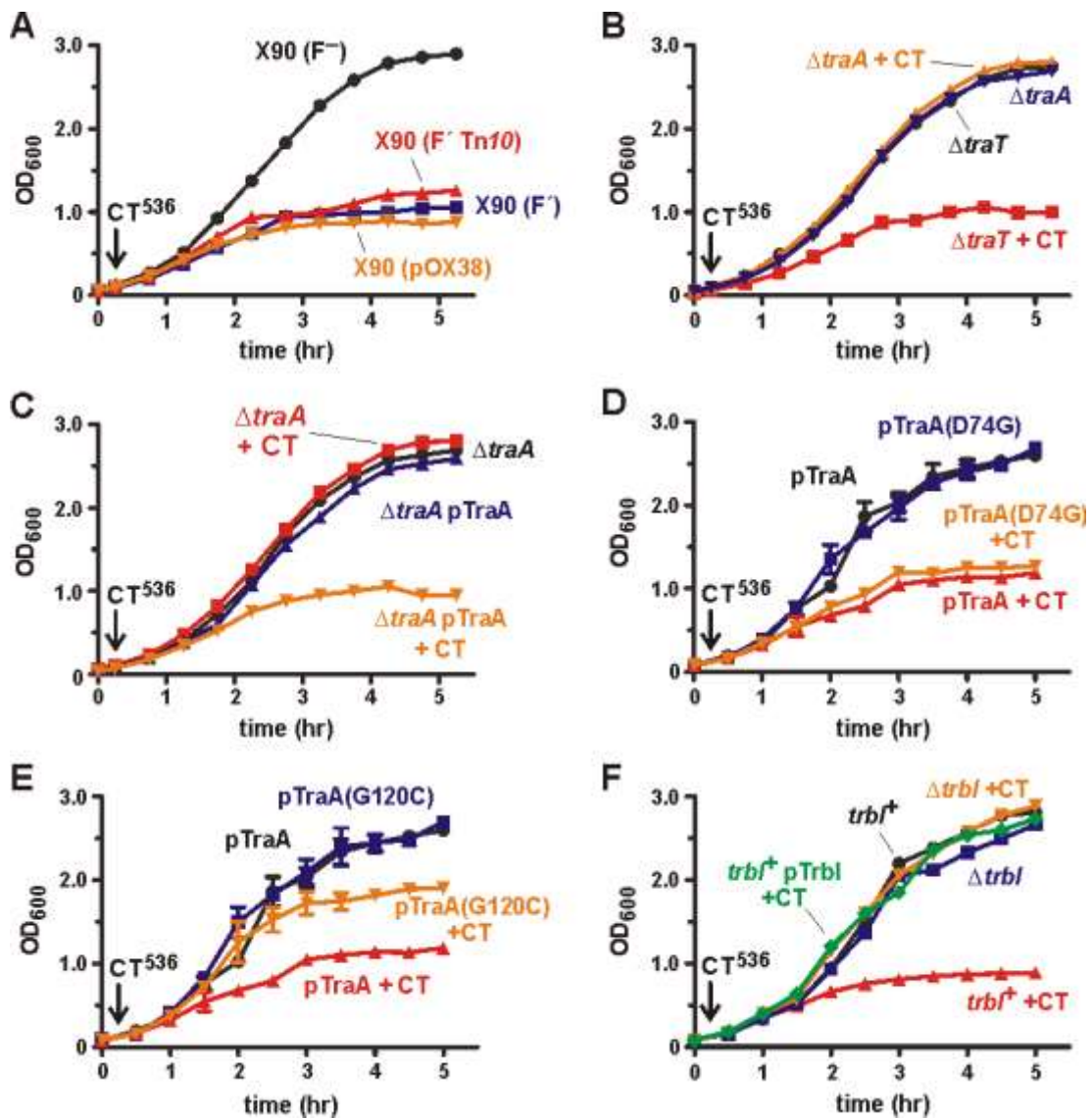
**Figure 5. The N-terminal domain of CdiA-CT<sup>536</sup> is sufficient for import.** **A)** Colicin E5 is composed of translocation, receptor-binding and nuclease domains. The C-terminal domain (ColE5-CT) has anticodon nuclease activity. The NT<sup>536</sup>-CT<sup>E5</sup> protein contains residues Val3016 – Gly3097 of CdiA<sup>536</sup> fused to the ColE5-CT nuclease domain. **B)** *E. coli* XL-1 cells were treated with purified ColE5-CT or NT<sup>536</sup>-CT<sup>E5</sup> fusion at 30 min (indicated by the arrow) and growth monitored by OD<sub>600</sub> measurements. Where indicated, cells carried plasmid-borne *cdiI*<sup>536</sup> (*cdiI*<sup>536</sup>) or *imE5* immunity genes. **C)** Northern blot analysis of toxin-treated cells. Total RNA was isolated from the cells in panel B and tRNA<sup>Tyr</sup> analyzed by northern blot. The migration positions of full-length and cleaved tRNA<sup>Tyr</sup> are indicated.



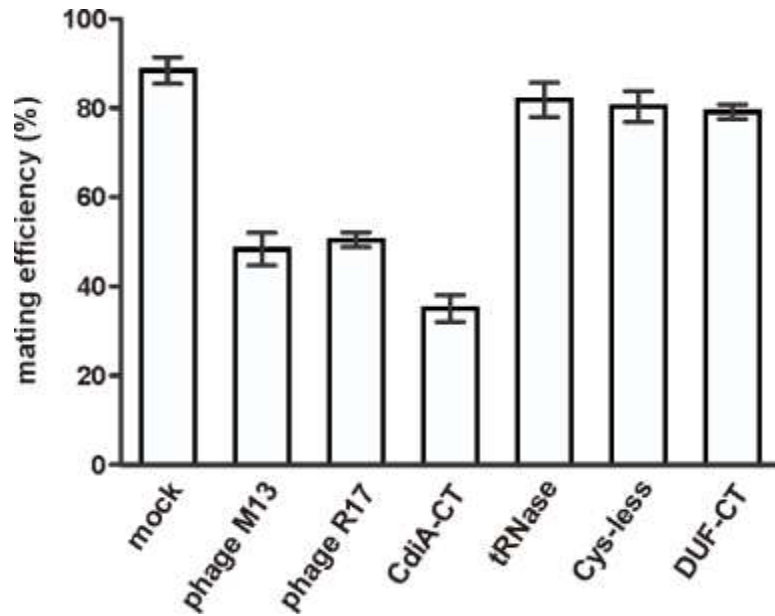
**Figure 6. CdiA-CT<sup>536</sup> is not translocated via known toxin-import pathways.** Wild-type *E. coli* X90 and the indicated mutants were treated with purified CdiA-CT<sup>536</sup> where indicated (+CT), and cell growth was monitored by OD<sub>600</sub> measurements. **A)** *E. coli* X90 (*bamA*<sup>+</sup>) and *bamA101* mutants. **B)** *E. coli* X90  $\Delta$ *tolA* mutants. **C)** *E. coli* X90  $\Delta$ *tonB* mutants. **D)** *E. coli* X90  $\Delta$ *tolA*  $\Delta$ *tonB* mutants.



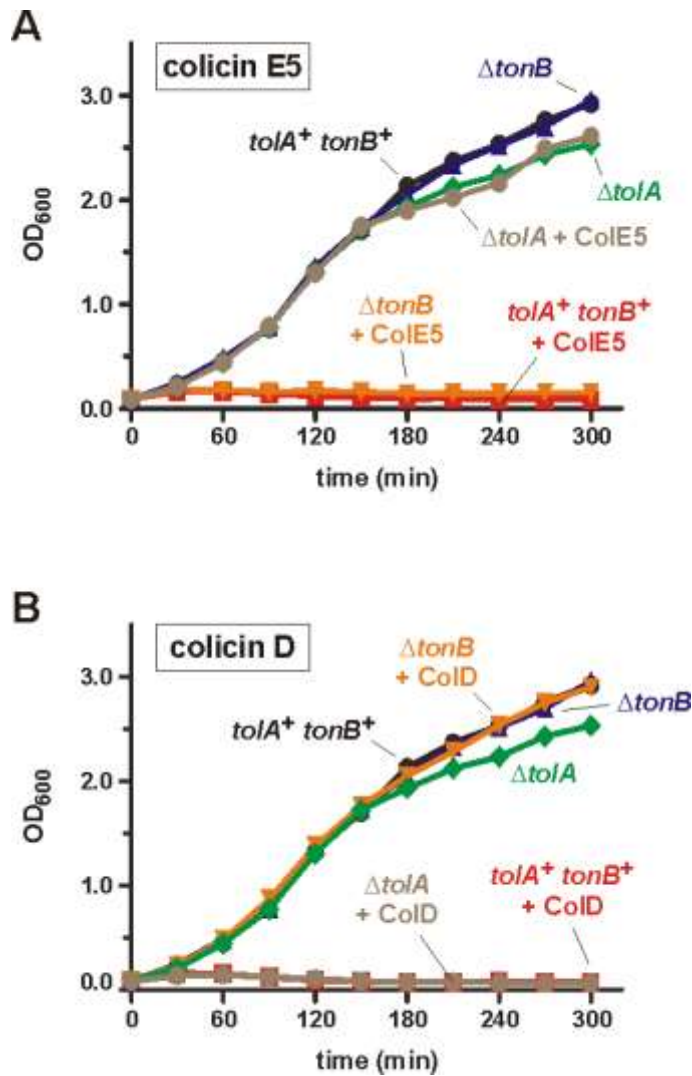
**Figure 7. The F pilus is required for CdiA-CT<sup>536</sup> import.** *E. coli* X90 and its derivatives were treated with purified CdiA-CT<sup>536</sup> where indicated (+CT), and cell growth was monitored by OD<sub>600</sub> measurements. **A)** Wild-type X90 is indicated as F<sup>-</sup> and cured cells as F<sup>+</sup>. The F<sup>+</sup>::Tn10 episome and pOX38::*gent* (pOX38) were introduced into cured X90 and the resulting cells treated with CdiA-CT<sup>536</sup>. **B)** *E. coli* X90  $\Delta traA::cat$  and  $\Delta traT::cat$  mutants. **C)** *E. coli* X90  $\Delta traA::cat$  cells complemented with plasmid-borne *traA*. **D)** and **E)** *E. coli* X90  $\Delta traA::cat$  cells complemented with plasmid-borne *traA(D74G)* and *traA(G120C)*. Error bars correspond to the standard error of the mean for three independent experiments. **F)** *E. coli* X90  $\Delta trbI::cat$  mutant. Wild-type (*trbI*<sup>+</sup>) cells expressing *trbI* from a plasmid (pTrbI) were also tested for toxin-resistance.



**Figure 8. Purified CdiA-CT<sup>536</sup> interferes with F-mediated conjugation.** The efficiency of conjugation was determined using the F':Tn10 episome as described in methods. Cell suspensions were treated with UV-inactivated phage particles or the indicated CdiA-CT<sup>536</sup> constructs (see Fig. 2A). The CdiA-CT<sup>536</sup> construct contains the His178Ala mutation to ablate tRNase activity.



**Figure 9. *E. coli*  $\Delta tolA$  and  $\Delta tonB$  mutants are resistant to colicin E5 and D, respectively. A) *E. coli* X90 was treated with colicin E5 (ColE5, group A) at 0 min and cell growth monitored by optical density at 600 nm (OD<sub>600</sub>). B) *E. coli* X90 was treated at 0 min with colicin D (ColD, group B) and cell growth monitored by OD<sub>600</sub>. Strains carried  $\Delta tolA$  or  $\Delta tonB$  deletions where indicated.**



**Chapter 3. Receptor polymorphism restricts contact-dependent growth inhibition (CDI) target range**



## Abstract

Contact-dependent growth inhibition (CDI) systems are distributed throughout  $\alpha$ -,  $\beta$ - and  $\gamma$ -proteobacteria and are thought to function in inter-cellular competition. CDI<sup>+</sup> inhibitor cells express CdiA effector proteins on their surface. CdiA binds to receptors on target bacteria and delivers a toxin derived from its C-terminus. CDI<sup>+</sup> cells also express an immunity protein that binds the CDI toxin and protects the cell from autoinhibition. Recent work has shown that CdiA<sup>EC93</sup> from *E. coli* EC93 exploits the highly conserved Bama protein as a receptor, restricting its target range to *E. coli* strains only. Here, we identify OmpC and OmpF as novel receptor proteins for CdiA<sup>UPEC536</sup> from uropathogenic *E. coli* strain 536 (UPEC536). OmpC is subject to very strong positive selection and exhibits significant variability in its extracellular loops between strains of *E. coli*. OmpC polymorphism has a dramatic effect on target cell susceptibility to CDI<sup>UPEC536</sup> with our data indicating that loops 4 and 5 form the binding-epitope of CdiA<sup>UPEC536</sup>. Remarkably, many OmpC alleles are completely resistant to CDI<sup>UPEC536</sup>, while OmpC<sup>UPEC536</sup> serves as an excellent receptor. These results demonstrate that CDI<sup>UPEC536</sup> activity is restricted to a subset of target strains, raising questions about the utility of these systems in intercellular competition. Given that the CDI<sup>UPEC536</sup> appears to be optimized for toxin delivery into sibling UPEC536 cells, we propose that these systems function in kin selection.

## Introduction

Contact-dependent growth inhibition (CDI) is a bacterial competition system discovered in the *Escherichia coli* isolate, EC93<sup>112</sup>. Aoki *et al.* were curious as to how this strain was able to establish itself as the dominant strain of *E. coli* in the flora of a rat colony. Enteric bacteria commonly produce soluble toxins to hinder the growth of neighboring bacteria, but EC93 demonstrated the unique ability to inhibit the growth of other *E. coli* cells upon direct physical contact<sup>51,112</sup>. This competitive phenotype was termed CDI and the locus responsible was found to be comprised of just three genes, *cdiB*, *cdiA*, and *cdiI*<sup>112</sup>. CDI works as a two-partner secretion system where CdiB exports the toxic effector, CdiA, onto the surface of cells<sup>112</sup>. CdiA is a large 200-600 kDa hemagglutinin repeat protein that is predicted to form a long  $\beta$ -helical filament with a toxic domain at its C-terminus<sup>220</sup>. Upon contact with neighboring cells, the C-terminal toxic domain (CdiA-CT) gets delivered into target cells and causes growth inhibition<sup>115,116</sup>. CdiA-CTs encode for a variety of toxins, including DNases, RNases, and pore-forming proteins<sup>115</sup>. CDI<sup>+</sup> cells are protected from their own toxin domain by an immunity protein that binds to and inactivates its cognate CdiA-CT<sup>115,122</sup>.

The delivery process of CDI toxins remains speculative. Most CdiA-CTs act as nucleases that target substrates in the cytoplasm; thus they must cross the cell envelope of target cells<sup>120,121</sup>. Extensive studies have been performed on CDI<sup>EC93</sup> to show that inhibition of target cells is dependent on the expression of the outer-membrane protein BamA (YaeT)<sup>126</sup>. BamA is a monomeric  $\beta$ -barrel protein with five loops that are predicted to extend into the extracellular space above the outer leaflet<sup>221</sup>. It functions in the BAM assembly complex to assist other  $\beta$ -barrel proteins to insert properly into the outer membrane<sup>222</sup>. A transposon insertion resulting in a 10-fold decrease in BamA expression (BamA101) results in target-cell

resistance to CDI<sup>EC93</sup><sup>126</sup>. Furthermore, BamA101 target cells have decreased binding to CDI<sup>+</sup> cells, and addition of anti-BamA antibodies blocks CDI<sup>EC93</sup> activity<sup>126</sup>. Taken all together, these results establish BamA as the receptor protein for CDI<sup>EC93</sup>, though the receptor-binding domain of CdiA<sup>EC93</sup> is not known.

BamA is highly conserved amongst *E. coli* strains, but varies among even closely related species<sup>118</sup>. The regions of variability map to the extracellular loops of BamA<sup>223</sup>. Recently, Ruhe *et al.* established that these polymorphisms is what restricts the target range of CDI<sup>EC93</sup><sup>118</sup>. They demonstrated that *E. coli* targets expressing BamA alleles from *Enterobacter cloacae* and *Salmonella enterica* are not susceptible to CDI<sup>EC93</sup><sup>118</sup>. Likewise, these targets are unable to bind CDI<sup>EC93</sup> inhibitors<sup>118</sup>. Further analysis revealed that BamA loops 6 and 7 form the binding-epitope for CDI<sup>EC93</sup><sup>118</sup>. This work raises questions about the use of CDI systems in nature. Bacteria are constantly encountering a multitude of species, so they would presumably want a competition system with a broad target range. However, CDI did give EC93 a competitive advantage in the flora of a rat colony. At this point it is unclear if all *E. coli* CdiA proteins utilize BamA as a receptor protein. Nevertheless, the fact that CDI<sup>EC93</sup> is restricted to targeting *E. coli* suggests CDI systems are optimized for intra-species delivery and may play a role in kin selection.

Here we demonstrate that the CDI system found in uropathogenic *E. coli* 536 (CDI<sup>UPEC536</sup>) does not utilize BamA as a receptor protein. Genetic selection and subsequent analyses reveal both OmpC and OmpF, the two main porins in *E. coli*, as the receptor proteins for CDI<sup>UPEC536</sup>. Aligning CdiA<sup>EC93</sup> and CdiA<sup>UPEC536</sup> gives some insight into identifying the receptor-binding domain as the two proteins have limited regions of variability. Furthermore, we demonstrate polymorphisms in the extracellular loops of OmpC

restrict CDI<sup>UPEC536</sup> to an even narrower target range than CDI<sup>EC93</sup>. Moreover, CDI<sup>UPEC536</sup> seems to be optimized for self-delivery. This work further supports the theory that CDI systems are involved in kin selection and reveal a potential receptor-binding domain in *E. coli* CdiA.

## Materials and Methods

### *Bacterial strains and plasmids used in this study*

<i>Strains or plasmids</i>	<i>Description<sup>a</sup></i>	<i>Reference</i>
<b>Strain</b>		
EPI100	F <sup>-</sup> <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> ) φ80 <i>dlacZ</i> Δ <i>M15</i> Δ <i>lacXcZ</i> Δ <i>M15</i> Δ <i>lacX</i> <i>recA1</i> <i>endA1</i> <i>araD139</i> Δ( <i>ara, leu</i> )7697 <i>galU galK</i> λ <sup>-</sup> <i>rpsL nupG</i>	Epicentre
MFDpir	RP4-2 Tc::[Δ <i>Mu1</i> :: <i>aac(3)IV</i> -Δ <i>aphA</i> -Δ <i>nic35</i> - Δ <i>Mu2</i> :: <i>zeo</i> ] Δ <i>dapA</i> ::( <i>erm-pir</i> ) Δ <i>recA</i>	?
DL4905	MC4100 l640-13 P <sub><i>papBA</i></sub> - <i>gfp-mut3</i> , Kan <sup>R</sup>	118
DL6536	UPEC536 Δ <i>kps15</i> :: <i>cat</i> Δ <i>araBAD</i> <i>spec_araC_pBAD_cdiBAI</i> , Cm <sup>R</sup> Strpt <sup>R</sup> Spec <sup>R</sup>	115
DL63881	UPEC536 Δ <i>kps15</i> :: <i>cat</i> Δ <i>araBAD</i> <i>spec_araC_pBAD_cdiBA</i> Δ <i>CT-I</i> , Cm <sup>R</sup> Strpt <sup>R</sup> Spec <sup>R</sup>	115
CH2016	X90 (DE3) Δ <i>rna</i> Δ <i>slyD</i> :: <i>kan</i> , Kan <sup>R</sup>	179
CH10226	MC4100 Δ <i>bam</i> :: <i>cat</i> pZS21:: <i>bamA</i> <sup><i>E.coli</i></sup> , AmpR	This study
CH10227	MC4100 Δ <i>bam</i> :: <i>cat</i> pZS21:: <i>bamA</i> <sup><i>E.cloacae</i></sup> , AmpR	This study
CH10013	JCM158 spontaneous rifampicin resistance, Rif <sup>R</sup>	117
CH10099	JCM158Rif <sup>R</sup> Δ <i>ompC</i>	This study
CH10100	JCM158Rif <sup>R</sup> Δ <i>ompF</i>	This study
CH10136	JCM158Rif <sup>R</sup> Δ <i>ompC</i> Δ <i>ompF</i>	This study
CH12134	JCM158Rif <sup>R</sup> Δ <i>ompF ompC</i> <sup>UPEC536</sup> :: <i>kan</i> , Kan <sup>R</sup>	This study
CH12345	JCM158 Δ <i>wzb</i> Δ <i>ompF</i> Δ <i>ompC</i> Δ <i>recA</i> :: <i>kan</i> , Kan <sup>R</sup>	This study
CH12180	JCM158 <i>ompC</i> <sup>UPEC536</sup> Δ <i>ompF</i> Δ <i>wzb</i> Δ <i>recA</i> :: <i>kan</i> , Kan <sup>R</sup>	This study

<i>Plasmids</i>		
pSC189kan	Encodes transposase and mariner transposon, requires Pir for replication.	?
pBR322::cysK	Constitutive expression of CysK, Amp <sup>R</sup>	This study
pCH450	pACYC184 derivative containing <i>E. coli</i> <i>araC</i> and the L-arabinose-inducible P <sub>araBAD</sub> promoter, Tet <sup>R</sup>	182
pCH450Kpn::DsRed	Arabinose-inducible expression of DsRed, Tet <sup>R</sup>	This study
pZS21amp::bamA <sup>E.coli</sup>	Constitutive expression of bamA <sup>E.coli</sup> , Amp <sup>R</sup>	118
pZS21amp::bamA <sup>E.cloacae</sup>	Constitutive expression of bamA <sup>E.cloacae</sup> , Amp <sup>R</sup>	118
pZS21kan	pZS31 derivative carrying an ampicillin-resistance cassette, Kan <sup>R</sup>	This study
pZS21kan::ompF <sup>K12</sup>	Constitutive expression of ompF <sup>E.coli</sup> , Kan <sup>R</sup>	This study
pET21::colE5-immE5	Over-produces ColE5-ImmE5 -His <sub>6</sub> , Amp <sup>R</sup>	224
pTrc99aKX	IPTG-inducible expression vector with KpnI/XhoI cloning sites, Amp <sup>R</sup>	117
pTrc99aKX::ompF <sup>K12</sup>	Expresses OmpF <sup>K12</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompF <sup>S.enterica</sup>	Expresses OmpF <sup>S.enterica</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompF <sup>E.cloacae</sup>	Expresses OmpF <sup>E.cloacae</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>K12</sup>	Expresses OmpC <sup>K12</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>S.enterica</sup>	Expresses OmpC <sup>S.enterica</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>E.cloacae</sup>	Expresses OmpC <sup>E.cloacae</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UPEC536</sup>	Expresses OmpC <sup>UPEC536</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>F11</sup>	Expresses OmpC <sup>F11</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>CFT073</sup>	Expresses OmpC <sup>CFT073</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UT189</sup>	Expresses OmpC <sup>UT189</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>EC93</sup>	Expresses OmpC <sup>EC93</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>EC869</sup>	Expresses OmpC <sup>EC869</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC(E188G) <sup>EC869</sup>	Expresses OmpC(E188G) <sup>EC869</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UT12</sup>	Expresses OmpC <sup>UT12</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UT15</sup>	Expresses OmpC <sup>UT15</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UT18</sup>	Expresses OmpC <sup>UT18</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC <sup>UT110</sup>	Expresses OmpC <sup>UT110</sup> , Amp <sup>R</sup>	This study
pTrc99aKX::ompC(536L5) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study

pTrc99aKX:: <i>ompC</i> (536 L7) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pTrc99aKX:: <i>ompC</i> (536 L8) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pTrc99aKX:: <i>ompC</i> (536 L7,8) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pTrc99aKX:: <i>ompC</i> (536 L5,7) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pTrc99aKX:: <i>ompC</i> (536 L5,8) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pTrc99aKX:: <i>ompC</i> (F1 IL4) <sup>CFT073</sup>	Expresses the indicated OmpC, Amp <sup>R</sup>	This study
pDAL660Δ1-39	Expresses CDI <sup>EC93</sup> , Amp <sup>R</sup>	112
pDAL866	Expresses CDI <sup>UPEC536</sup> , Amp <sup>R</sup> Cm <sup>R</sup>	116
pWEB:: <i>TNC</i>	pWEB:: <i>TNC</i> (CDI $\Gamma$ ), Amp <sup>R</sup> , Cm <sup>R</sup>	112
pCP20	Heat-inducible expression of FLP recombinase, Cm <sup>R</sup> Amp <sup>R</sup>	225

<sup>a</sup>Abbreviations: Amp<sup>R</sup>, ampicillin resistant; Cm<sup>R</sup>, chloramphenicol resistant; Kan<sup>R</sup>, kanamycin resistant; Rif<sup>R</sup>, rifampicin resistant; Str<sup>R</sup>, streptomycin resistant; Tet<sup>R</sup>, tetracycline resistant

### ***Bacterial strains and growth conditions***

The bacterial strains and plasmids used in this study are listed in the table above. Bacteria were grown in LB broth or LB agar unless otherwise noted. Media were supplemented with antibiotics at the following concentrations: ampicillin, 150  $\mu\text{g ml}^{-1}$ ; kanamycin, 50  $\mu\text{g ml}^{-1}$ ; chloramphenicol, 66  $\mu\text{g ml}^{-1}$ ; spectinomycin, 50  $\mu\text{g ml}^{-1}$ ; rifampin, 200  $\mu\text{g ml}^{-1}$ ; streptomycin, 100  $\mu\text{g ml}^{-1}$ . The *bamA*::*cat* allele was transduced from EPI100  $\Delta\text{bamA}$ ::*cat* pZS21-*bamA*<sup>+</sup> cells into MC4100 pZS21-*bamA*<sup>+</sup> by bacteriophage P1-mediated transduction<sup>118</sup>. The  $\Delta\text{ompC}$ , *ompF*, *wzb*, and *recA*::*kan* gene disruptions were obtained from the Keio collection and introduced into *E. coli* JCM158Rif<sup>R</sup> by P1 transduction<sup>226</sup>. The Kan<sup>R</sup> cassettes were removed by transforming pCP20, selecting on LB-Amp plates at 30°C, then isolating colonies at 37°C.

pZS21kan::*ompF*<sup>K12</sup> was made by PCR amplifying *E. coli* K12 genomic DNA (gDNA) with oligos 5' - TTT GGA TCC ATG AGG GTA ATA AAT AAT GAT G - 3' and 5' - TTT TCT AGA GGT GTG CTA TTA GAA CTG - 3', digesting with BamHI/XbaI and ligated to digested pZS21kan. pTrc99aKX::*ompF* alleles were made by PCR amplifying *E. coli*, *S. enterica*, or *E. cloacae* gDNA with 5' - TTT GGT ACC ATG ATG AAG CGC AAT ATT C - 3' (K12-ompF-for) and 5' - TTT CTC GAG TTA GAA CTG GTA AAC GAT ACC C - 3' for K12, 5' - TTT GGT ACC ATG AAA CTT AAG TTA GTG GCA G - 3' and 5' - TTT CTC GAG ATT AGA ACT GGT AGT TCA GAC C - 3' for *S. enterica*, and K12-ompF-for and 5' - TTT CTC GAG ATT AGA ACT GGT AAA CCA GAC - 3' for *E. cloacae*, digested with KpnI/XhoI and ligated into digested pTrc99aKX vector.

All pTrc99aKX::*ompC* *E. coli* alleles were amplified with 5' - GTT GGT ACC ATG AAA GTT AAA GTA CTG TCC - 3' and 5' - TCA CTC GAG ATT AGA ACT GGT AAA CCA GAC - 3', digested with KpnI/XhoI and ligated into digested pTrc99aKX vector. OmpC(E188G)<sup>EC869</sup>, OmpC(536L5)<sup>CFT073</sup>, OmpC(536L7)<sup>CFT073</sup>, OmpC(536L8)<sup>CFT073</sup>, OmpC(536L7,8)<sup>CFT073</sup>, OmpC(536L5,7)<sup>CFT073</sup>, OmpC(536L5,8)<sup>CFT073</sup>, OmpC(F11L4)<sup>CFT073</sup> were made via megaprimer PCR using oligos 5' - CGT TCT GAC GCA GTG CTC CAC GAC CGT TGT TAG T - 3' for E188G, 5' - GCT CCA AAC GTA CCG ATG CTC AGA ACA CCG CTG CTT ACA TAG GCA ACG GCG ACC GTG CTG - 3' for 536L5, 5' - AGG TAA AAA CCT GGG TAC TAT CGG TAC TCG TAA CTA CGA CGA C - 3' for 536L7, 5' - CCA GTT CAC CCG CGA CGC TGG CAT CAA CAC TGA TAA CAT CGT AGC TC - 3' for 536L8, and 5' - CTT TAC GAC CAT TGT TGG TCA TGC CTT CAC CGC TTA CGC TGC C - 3' for F11L4.

### ***Growth competitions***

Inhibitor and the described *E. coli* JCM158Rif<sup>R</sup> target cells were grown to mid-log and mixed at a 10:1 ratio in liquid LB without antibiotics for 5 hours at 37°C with vigorous shaking. Samples of the co-cultures were removed, serially diluted, and plated onto selective LB agar to enumerate target cells as CFU ml<sup>-1</sup>. CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, CDI inhibitors are *E. coli* EPI100 cells containing pDAL660Δ1-39 (CDI<sup>EC93</sup>), pDAL866 (CDI<sup>UPEC536</sup>) or pWEB::TNC (CDI<sup>-</sup>). UPEC536 CDI<sup>+</sup> or CDI inhibitors are DL6536 and DL6381, respectively. These latter inhibitor cells were grown in 0.2% arabinose for induction of its CDI locus, and 0.2% arabinose was added to the competition media. In figure 2B, the OmpC is being expressed in pTrc99aKX and OmpF is in pZS21kan. These are both *E. coli* K12 alleles.

### ***Cell-cell binding***

Overnight cultures of *E. coli* strain DL4905 carrying cosmid pDAL660Δ1-39 (CDI<sup>EC93</sup>), pDAL866 (CDI<sup>UPEC536</sup>) or pWEB::TNC (CDI<sup>-</sup>) were diluted into fresh tryptone broth (TB) and grown to mid-log phase at 30°C. These cells were then mixed at a 5:1 ratio with the described *E. coli* JCM158Rif<sup>R</sup> target cells carrying pCH450::DsRed and pTrc99aKX::OmpC or OmpF. In figure 2B, OmpC is being expressed in pTrc99aKX and OmpF is pZS21kan. The OmpC used here is from UPEC536 for more efficient binding. OmpF is from *E. coli* K12. The cell suspension was incubated with aeration for 15 min at 30°C, diluted 1:50 into filtered 1× phosphate-buffered saline (PBS), and then analyzed on an Accuri C6 flow cytometer using FL1 (533/30 nm, GFP) and FL2 (585/40 nm, DS-Red) fluorophore filters (Becton Dickinson).

### ***Western blot analysis***



Cells were grown to mid-log in 2mL liquid LB with aeration and harvested for protein extraction with urea lysis buffer. Equal amounts of protein were subjected to 6M urea 10% acrylamide SDS-PAGE for 4 hours at 100V. The gel was transferred onto PVDF membrane, and incubated with 1:15,000 of anti-OmpC/F/A antibody, a kind gift from Thomas Silhavy.

### ***Colicin E5 purification and growth curve***

Colicin E5 was purified as previously described<sup>224</sup>. Cells were grown to mid-log in LB-Amp and back diluted to OD<sub>600</sub> 0.05. Growth was analyzed by measuring the optical density every 60 minutes for 5 hours. 1 nM of colicin was added after 30 minutes of growth.

## **Results**

### ***CDI<sup>UPEC536</sup> is not restricted by BamA<sup>E. coli</sup>***

An alignment of CdiA<sup>EC93</sup> and CdiA<sup>UPEC536</sup> reveals two regions of variability (Fig. 1A). To test whether CDI<sup>UPEC536</sup> utilizes BamA as its receptor protein, we asked if *E. cloacae* or *S. enterica* BamA alleles could support CDI<sup>UPEC536</sup>. EPI100 cells expressing CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or CDI cosmids were co-cultured with MC4100 *bamA*<sup>-</sup> target cells complemented with *E. coli* or *E. cloacae* BamA on plasmid pZS21amp. Cells were mixed at a 10:1 ratio in LB broth for 5 hours and selected on LB-antibiotic plates to count cell viability (Fig. 1B). Firstly, each target cell grew ~1.5 logs when incubated with the control CDI inhibitors, indicating there is no growth defect from expressing heterologous BamA alleles. Next, targets expressing BamA<sup>E. coli</sup> lost 3.5 and 2 logs of cell viability when competed against CDI<sup>EC93</sup> and CDI<sup>UPEC536</sup> respectively. As previously reported, target cells expressing BamA<sup>E. cloacae</sup> are resistant to CDI<sup>EC93</sup>, however, these targets are still inhibited 2 logs when competed

against CDI<sup>UPEC536</sup>. These data indicate that CDI<sup>UPEC536</sup> either does not require BamA as a receptor protein, or the epitope targeted on BamA is conserved between the two alleles tested.

### ***OmpC and OmpF are required for CDI<sup>UPEC536</sup>***

To look for novel receptor proteins targeted by CDI<sup>UPEC536</sup>, we set up a genetic selection to identify mutants that conferred resistance to CDI<sup>UPEC536</sup>. CdiA-CT<sup>UPEC536</sup> requires the presence of CysK in target cells for activation<sup>123</sup>. Thus, plasmid-borne CysK was included in target cells to prevent *cysK*<sup>-</sup> targets from coming through the screen. JCM158 pBR322::*cysK* cells were subjected to transposon mutagenesis using a mariner-based system.

Approximately 50,000 cells containing transposon insertions were pooled and incubated with UPEC536 CDI<sup>+</sup> inhibitor cells at a 10:1 ratio for three rounds of enrichment. After each round, target cell viability increased about one log and the population was completely resistant after round three, indicating the presence of mutants resistant to CDI<sup>UPEC536</sup>. After proving resistance was linked to the transposon, individual mutants were subjected to arbitrary PCR to locate the insertion site. Ten unique transposon insertions were found in the open reading frames of *ompC* and *ompF*, five in each gene (Fig. 2A).

To rule out polar effects of the transposon insertion, clean deletions of *ompC* and *ompF* were made in the original target strain using alleles from the Keio collection<sup>226</sup>. These targets were competed against UPEC536 CDI<sup>+</sup> and CDI<sup>-</sup> mock inhibitors at a 10:1 ratio for 5 hours in liquid broth. Both *ompC*<sup>-</sup> and *ompF*<sup>-</sup> targets were completely resistant to CDI<sup>UPEC536</sup> while WT targets lost 2 logs of viability (data not shown). Each target strain grew to full capacity against CDI<sup>-</sup> inhibitors (data not shown). Second, plasmid-borne copies of OmpC and OmpF

were used to complement an *ompC ompF* strain. When complemented with empty vectors, or just OmpC or OmpF, target cells were resistant to CDI<sup>UPEC536</sup>, but when both OmpC and OmpF were present, targets became susceptible (Fig. 2B). Furthermore, each target cell tested grew to full capacity against UPEC536 CDI inhibitors (Fig, 2B). Taken all together, we have established both *ompC* and *ompF* are required for CDI<sup>UPEC536</sup>.

### ***OmpC and OmpF are required for cell-cell binding***

OmpC and OmpF are the two major porin proteins present in the outer membrane of *E. coli*<sup>227</sup>. They share 60% homology and have similar  $\beta$ -barrel structures, though they have different pore sizes which facilitates selective uptake of molecules<sup>228,229</sup>. Both form homotrimers in the membrane with 7 extracellular loops per monomer. With this information it is reasonable to hypothesize OmpC and OmpF act as receptor proteins for CdiA<sup>UPEC536</sup>. As mentioned previously, CDI<sup>EC93</sup> requires the presence of BamA to bind target cells<sup>126</sup>. Thus, we asked if both OmpC and OmpF are required for cell-cell binding.

To do this, MC4100 GFP<sup>+</sup> cells containing cosmids expressing CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or no CDI genes were mixed with target cells expressing DsRed protein and assayed for their ability to bind one another by flow-cytometry. The presence of green/red aggregates represents inhibitor cells bound to targets. This binding event is quantified as the percent of targets bound to inhibitors. CDI<sup>EC93</sup> inhibitors are able to bind ~ 70% of each target cell population due to the presence of BamA (Fig. 2C). CDI inhibitors bind ~0.5% of each target cell population, representing “background” levels of binding. CDI<sup>UPEC536</sup> inhibitors bind targets expressing either OmpC or OmpF at background levels, but bind ~45% of OmpC<sup>+</sup> OmpF<sup>+</sup> cells (Fig. 2C). Albeit this is lower than CDI<sup>EC93</sup> binding, it is still significant and

specific to targets expressing both OmpC and OmpF. Therefore, both OmpC and OmpF are required for cell-cell binding.

### ***CDI<sup>UPEC536</sup> targets the extracellular loops of OmpC and OmpF***

We hypothesized CDI<sup>UPEC536</sup> is targeting the extracellular loops of OmpC and OmpF, just as CDI<sup>EC93</sup> exploits loops 6 and 7 in BamA<sup>118</sup>. An alignment of OmpC or OmpF to a variety of species demonstrates the greatest regions of variability map to the extracellular loops (Fig 3A & 4A). We used this natural variation to test our hypothesis. To do this, we complemented JCM158 *ompC*<sup>-</sup> or *ompF*<sup>-</sup> target cells with *ompC* or *ompF* alleles from various species and asked if they were susceptible to CDI<sup>UPEC536</sup>. Each target was incubated with UPEC536 CDI<sup>+</sup> or CDI inhibitors at a 10:1 ratio for 5 hours. All target cells expressing an empty vector and all target cells incubated with CDI inhibitors grew ~1 log during co-culture, demonstrating the foreign alleles do not affect growth (Fig. 3B & 4B). JCM158 *ompC*<sup>-</sup> targets expressing OmpC<sup>*E.coli*</sup> or OmpC<sup>*E.cloacae*</sup> lost 1-2 logs of viability while those expressing OmpC<sup>*S.enterica*</sup> did not lose any viability (Fig. 3B). Similar to the OmpC results, JCM158 *ompF*<sup>-</sup> targets expressing OmpF<sup>*E.coli*</sup> or OmpF<sup>*E.cloacae*</sup> lost about 2.5 logs of viability, while OmpF<sup>*S.enterica*</sup> targets barely lost any viability (Fig. 4B). Thus, different alleles of OmpC and OmpF lead to varying levels of target cell susceptibility.

Cell-cell binding was performed to see if the levels of inhibition reflected the ability of inhibitor cells to bind targets. To examine this, the same cell-cell binding assay was performed as previously described. The ability of CDI<sup>UPEC536</sup> inhibitor cells to bind targets roughly reflects the competition results (Fig. 3C & 4C). Western blot analysis was attempted to show equal levels of OmpC and OmpF in each target cell. However, the antibody used was

raised to *E. coli* OmpC so this is not a reasonable method to analyze protein levels. The antibody did not detect OmpF<sup>*E. cloacae*</sup> or OmpF<sup>*S. enterica*</sup>, but it did show equal levels of the OmpC alleles tested (data not shown). Taken all together, target cells expressing different alleles of OmpC and OmpF lead to varying levels of inhibition, which can be reflected in the ability of inhibitor cells to bind targets. Considering the main regions of variability between sensitive and resistant alleles are in the extracellular loops of both OmpC and OmpF, these data suggest CDI<sup>UPEC536</sup> targets the extracellular loops of these porins.

### ***OmpC<sup>E.coli</sup> polymorphism restricts CDI<sup>UPEC536</sup>***

Not only do OmpC and OmpF vary amongst species, but they are two of the most positively selected proteins in *E. coli*<sup>230</sup>. This is most likely due to their function as receptor proteins for various bacteriophage and colicins along with being substrates for the immune system<sup>230</sup>. An alignment of OmpF from various strains of *E. coli* reveals two areas of deviation, mapping to extracellular loops 4 and 5 (Fig. 5A). Even more striking, a similar alignment of OmpC shows incredibly variation, including up to 8 residue insertions and 4 residue deletions in loops 4, 5 and 7 (Fig. 5B & 5C). Given the polymorphisms seen in OmpF is restricted to just a few residues, we reasoned it would not vastly affect CDI<sup>UPEC536</sup> and focused on OmpC. To test whether different *E. coli* OmpC alleles exhibit different levels of sensitivity to CDI<sup>UPEC536</sup>, 12 of the most divergent alleles were used to complement JCM158 *ompC* target cells on plasmid pTrc99a and competed against UPEC536 CDI<sup>+</sup> or CDI<sup>-</sup> cells. Each target cell grew ~ 1 log during co-culture with CDI inhibitors, indicating the variant alleles do not affect *E. coli* K12 growth (Fig. 6A). Shockingly, 5 of the alleles tested were completely resistant to CDI<sup>UPEC536</sup> while the other 7 showed various levels of sensitivity (Fig.

6A).  $OmpC^{UPEC536}$  was one of the most sensitive alleles, indicating UPEC536  $CDI^+$  cells are capable of self-delivery. Western blot analysis demonstrates that each  $OmpC$  allele tested was expressed at similar levels in target cells (Fig. 6B). These results further demonstrate that  $CDI^{UPEC536}$  targets the extracellular loops of  $OmpC$ , consequently restricting its target-cell range to a subpopulation of *E. coli* strains.

### ***CDI<sup>UPEC536</sup> targets OmpC loops 4 and 5***

We have now established that  $CDI^{UPEC536}$  targets the extracellular loops of both  $OmpC$  and  $OmpF$ . Given that  $OmpF$  is relatively conserved amongst *E. coli*, we hypothesize that  $OmpC$  is driving the target range of  $CDI^{UPEC536}$ . Therefore, we wanted to define which extracellular loops in  $OmpC$  serve as the binding-epitope for  $CdiA^{UPEC536}$ . Given that the major differences between a sensitive and resistant allele are in loops 4 and 5, we reasoned these loops are targeted (Fig. 5B). To test this, we grafted loop 4 or 5 from a sensitive allele onto a resistant  $OmpC$  and asked if this sensitized the chimeric  $OmpC$  to  $CDI^{UPEC536}$ . We chose CFT073 as the resistant allele to test, and grafted loop 4 or loop 5 from the sensitive F11 or UPEC536 alleles, respectively. CFT073 and UPEC536 have the same loop 4, but have considerable differences in loop 5 (Fig. 7A). Similarly, CFT073 and F11 have similar loop 5 residues, but CFT073 has an 8 residue insertion in loop 4 (Fig. 7A). JCM158 *ompC* target strains were complemented with the chimeric *ompC* alleles on plasmid pTrc99a and incubated with  $CDI^+$  or  $CDI^-$  UPEC536 inhibitor cells. Targets expressing the chimeric *ompC* alleles showed no growth defect against  $CDI^-$  inhibitors, and as predicted  $OmpC^{CFT073}$  with either  $L5^{UPEC536}$  or  $L4^{F11}$  are fully sensitive to  $CDI^{UPEC536}$  (Fig. 7B). Interestingly, these results demonstrate that having a sensitive L4 is not sufficient for target cells to be

susceptible to CDI<sup>UPEC536</sup>, as the resistant CFT073 allele and sensitive UPEC536 allele share the same L4 sequence. These results indicate that OmpC loops 4 and 5 are important determinants for target cell susceptibility to CDI<sup>UPEC536</sup>.

Crystal structure analysis of various *E. coli* OmpC proteins reveals that loops 4 and 5 are close enough to interact with one another (Fig. 5C)<sup>229,231</sup>. Polymorphisms in these loops amongst the crystalized alleles have a drastic effect on the positioning of the loops in the structure (Fig. 5C). With this information, along with the results from the loop-swap experiments performed above, we postulate that the interaction between loops 4 and 5 in OmpC is important for CDI<sup>UPEC536</sup>. OmpC<sup>EC93</sup> is the most sensitive allele to CDI<sup>UPEC536</sup> that we have identified thus far with target cells losing ~3.5 logs of viability (Fig. 8C). There are only 5 residues that differ in OmpC<sup>EC869</sup> compared to OmpC<sup>EC93</sup> (Fig. 8A), yet target cells expressing OmpC<sup>EC869</sup> only lose 0.5-1 log of viability (Fig. 8C). Further analysis reveals that one of these differences, the G188E change in loop 4 of OmpC<sup>EC869</sup> has the potential to influence the interaction between loops 4 and 5. The crystal structure of OmpC<sup>UTI2</sup>, an allele that is resistant to CDI<sup>UPEC536</sup>, shows that residue D231 in L5 in all 3 alleles comes in close proximity to E188 in L4 (Fig. 8B)<sup>231</sup>. When we made OmpC<sup>EC869</sup> with residue E188 changed to a glycine, it became fully sensitive to CDI<sup>UPEC536</sup> (Fig. 8C). Additionally, western blot analysis reveals there are similar levels of OmpC in each target cell (Fig. 8D). Taken all together, these data suggest our hypothesis is correct; the interaction between loops 4 and 5 in OmpC is important for CDI<sup>UPEC536</sup> and likely make up at least part of the binding-epitope for CdiA<sup>UPEC536</sup>.

***OmpC/F heterotrimers support CDI<sup>UPEC536</sup>***

The fact that both OmpC and OmpF are required for CDI<sup>UPEC536</sup> and target-cell binding is puzzling. There are three models that support this observation. First, one porin could act as the receptor and the other a translocator (translocator model), similar to how colicins transit through the outer membrane. Colicin-bound receptor recruits a translocator to facilitate colicin entry into the periplasm<sup>51</sup>. In fact, most group A colicins use OmpF as their translocator<sup>51</sup>. Second, CdiA<sup>UPEC536</sup> could bind to OmpC and OmpF homotrimers that are next to one another in the membrane (double-receptor model). This seems feasible as they are the two most abundant outer-membrane proteins in *E. coli*<sup>232</sup>. Lastly, OmpC and OmpF are 60% homologous and have been shown to form heterotrimers *in vivo*<sup>233</sup>. The “heterotrimer model” suggests a 2:1 or 1:2 ratio of OmpC to OmpF could act as the receptor protein. The fact that there are examples of OmpC and OmpF alleles from different species that do not support CDI<sup>UPEC536</sup> suggests CdiA<sup>UPEC536</sup> is interacting with a specific epitope on both OmpC and OmpF. Unfortunately, this does not rule out any of these models, it is just perhaps suggestive of the latter two.

To test the heterotrimer model, we tried to covalently fuse 3 OmpF proteins together by linking 3 *ompF* genes on plasmid pTrc99a with 7 or 16 residue linkers between the monomers. Consequently, only homotrimers of OmpF or OmpC would be present in the outer membrane of target cells. If these cells are resistant to CDI<sup>UPEC536</sup>, this would suggest heterotrimers act as the receptor protein. Unfortunately, all efforts put into making this cell line failed, as the covalently linked trimer could not properly insert into the outer membrane (Fig. 9B & 9D). However, we were able to successfully link 2 OmpF proteins together with both the 7 (short) and 16 (long) residue linker sequences connecting the two monomers.



To test if the OmpF dimers are folding and inserting into the outer membrane properly, *ompF* cells were complemented with empty vector, monomeric or dimeric OmpF on a pTrc99a plasmid and incubated with 1 nM of colicin E5. This is a group A colicin that requires OmpF to enter cells and subsequently cleave a subset of tRNAs to cause cell death<sup>51</sup>. Addition of colicin E5 to cells expressing OmpF dimers were inhibited to the same level as WT cells or cells expressing OmpF monomers (Fig. 9C). Additionally, each cell line grew at the same rate and to the same level when no colicin was added, suggesting there is no growth defect from expressing dimeric OmpF (Fig. 9C). Furthermore, a western blot reveals full-length OmpF dimer and no degradation products, assuring the dimer did not get processed into monomeric OmpF (Fig. 9D). However, OmpF naturally forms trimers in the membrane and it seems unlikely that the dimeric form is in the membrane on its own. Rather, dimeric OmpF is most likely forming a heterotrimer with OmpC to be able to insert properly and be functional in the membrane.

If this is the case, cells expressing dimeric OmpF but lacking *ompC* should not be able to insert OmpF into the outer membrane, rendering them resistant to colicin E5. To test this, *ompF ompC* cells were complemented with empty vector, monomeric OmpF or dimeric OmpF on plasmid pTrc99a and subjected to 1 nM colicin E5. Only WT cells and cells expressing monomeric OmpF were inhibited by the colicin. All other cells grew to the same level, indicating OmpF is not present on their surface (Fig. 9A). Furthermore, western blot analysis again showed full-length OmpF dimers present in these cells (Fig 9B). This illustrates that dimeric OmpF cannot insert into the membrane without the presence of OmpC. Taken all together, these results demonstrate that an OmpF dimer and OmpC monomer can form functional heterotrimers in the outer membrane of *E. coli* cells.

Now we asked if  $\text{CDI}^{\text{UPEC536}}$  can utilize OmpC/F heterotrimers as a receptor. The same  $\text{ompC}^+ \text{ompF}^-$  cells complemented with monomeric or dimeric OmpF were competed against EPI100 cells expressing  $\text{CDI}^{\text{UPEC536}}$  or no CDI genes. Incredibly,  $\text{CDI}^{\text{UPEC536}}$  inhibits target cells expressing OmpF dimers to the same level as WT targets, or targets expressing OmpF monomers (Fig. 10A). Furthermore, each target strain grew to equivalent levels against CDI inhibitors (Fig. 10A). Next,  $\text{CDI}^{\text{UPEC536}}$  was assayed for its ability to bind cells containing OmpC/F heterotrimers. The targets used in the competition assay were transformed with pCH450::DsRed and incubated with MC4100 GFP<sup>+</sup> cells expressing  $\text{CDI}^{\text{EC93}}$ ,  $\text{CDI}^{\text{UPEC536}}$ , or no CDI genes. Consistent with the competition results,  $\text{CDI}^{\text{UPEC536}}$  could bind to target cells that contained OmpC/F heterotrimers (Fig. 10B). Thus, OmpC/F heterotrimers form a functional receptor for  $\text{CDI}^{\text{UPEC536}}$ .

## Discussion

The work presented here establishes OmpC and OmpF as novel receptor proteins for  $\text{CDI}^{\text{UPEC536}}$ . Both porins are required for CDI and cell-cell binding, and changes in the extracellular loops of each protein can affect target cell susceptibility. This suggests  $\text{CdiA}^{\text{UPEC536}}$  binds to both OmpC and OmpF on target cells. We demonstrate that  $\text{CDI}^{\text{UPEC536}}$  can target heterotrimers of OmpC/F, supporting Nikaido's original observation that functional OmpC/F heterotrimers exist *in vivo*<sup>233</sup>. Unfortunately this does not prove heterotrimers act as the exclusive receptor because we only showed they are capable and not required to support CDI. However, if heterotrimers are the relevant targets for  $\text{CdiA}^{\text{UPEC536}}$ , suitable environmental conditions would be required for effective CDI because *ompC* and *ompF* expression fluctuates based on media osmolarity. OmpF is preferentially expressed in

low osmolarity or high cAMP levels, and while OmpC dominates in high osmolarity. We only tested a 2:1 OmpF to OmpC ratio, but if this is the relevant ratio to form a functional receptor, it seems environmental conditions, and therefore levels of each porin present could affect the efficiency of CDI.

Furthermore, this work demonstrates that BamA is not the universal receptor for all *E. coli* CDI systems. In fact, BamA might be the outlier. The amount of variability seen in the extracellular loops of OmpC could conceivably support discrete binding sites for multiple CdiA proteins. In support of this model, the CDI system found in *Enterobacter cloacae* ATCC 13047 also uses both OmpC and OmpF as receptor proteins, and seems to target a different epitope on OmpC than CDI<sup>UPEC536</sup> (data not shown). An alignment of CdiA from EC93 and UPEC536 shows two regions of variability, the greatest one clustered around amino acids 1300-1600 (Fig. 1A). Given these two CDI systems target distinct receptor proteins, it seems conceivable that this region corresponds to the receptor-binding domain of CdiA. Identifying this domain is enticing as it could lead to the ability to engineer CDI systems capable of targeting specific species or even strains of bacteria. This has promising implications for the design of antibiotics specific to the pathogen present in a host.

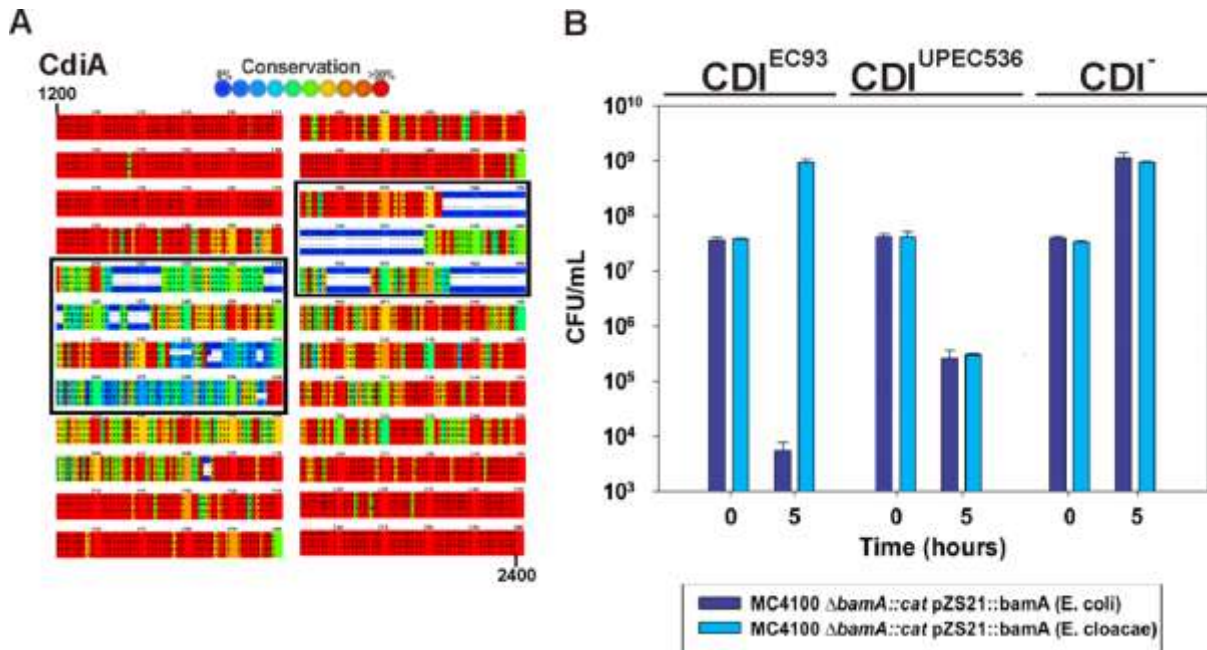
CDI<sup>EC93</sup> was shown to target loops 6 and 7 of BamA<sup>*E. coli*</sup>, limiting its target range to strictly *E. coli*<sup>118</sup>. OmpC and OmpF vary amongst strains of *E. coli* with extracellular loops 4 and 5 in OmpC showing the greatest amount of polymorphism. We illustrated that these two loops are targeted by CDI<sup>UPEC536</sup>, limiting its target range even further than CDI<sup>EC93</sup> to specific strains of *E. coli*. Interestingly, all five resistant *ompC* alleles tested are from uropathogenic *E. coli* (UPEC). OmpC and OmpF are two of the most positively selected proteins in *E. coli*, but also specifically in UPEC strains<sup>230,234</sup>. Given that UPEC536 shares

the same niche as these other urinary tract infection (UTI) strains, it seems unusual for CdiA<sup>UPEC536</sup> to target such polymorphic receptor proteins. In fact, we sequenced the *ompC* allele from 52 human UPEC isolates, and 46 cat or dog UPEC isolates and found 33 unique OmpC proteins. Of these 33, 19 were resistant to CDI<sup>UPEC536</sup>. This raises questions about the function of CDI<sup>UPEC536</sup> as a competition system in nature. Given its selective target range and the ability to proficiently target its own OmpC, this suggests CDI<sup>UPEC536</sup> is optimized to deliver into sibling cells as a mechanism of kin selection.

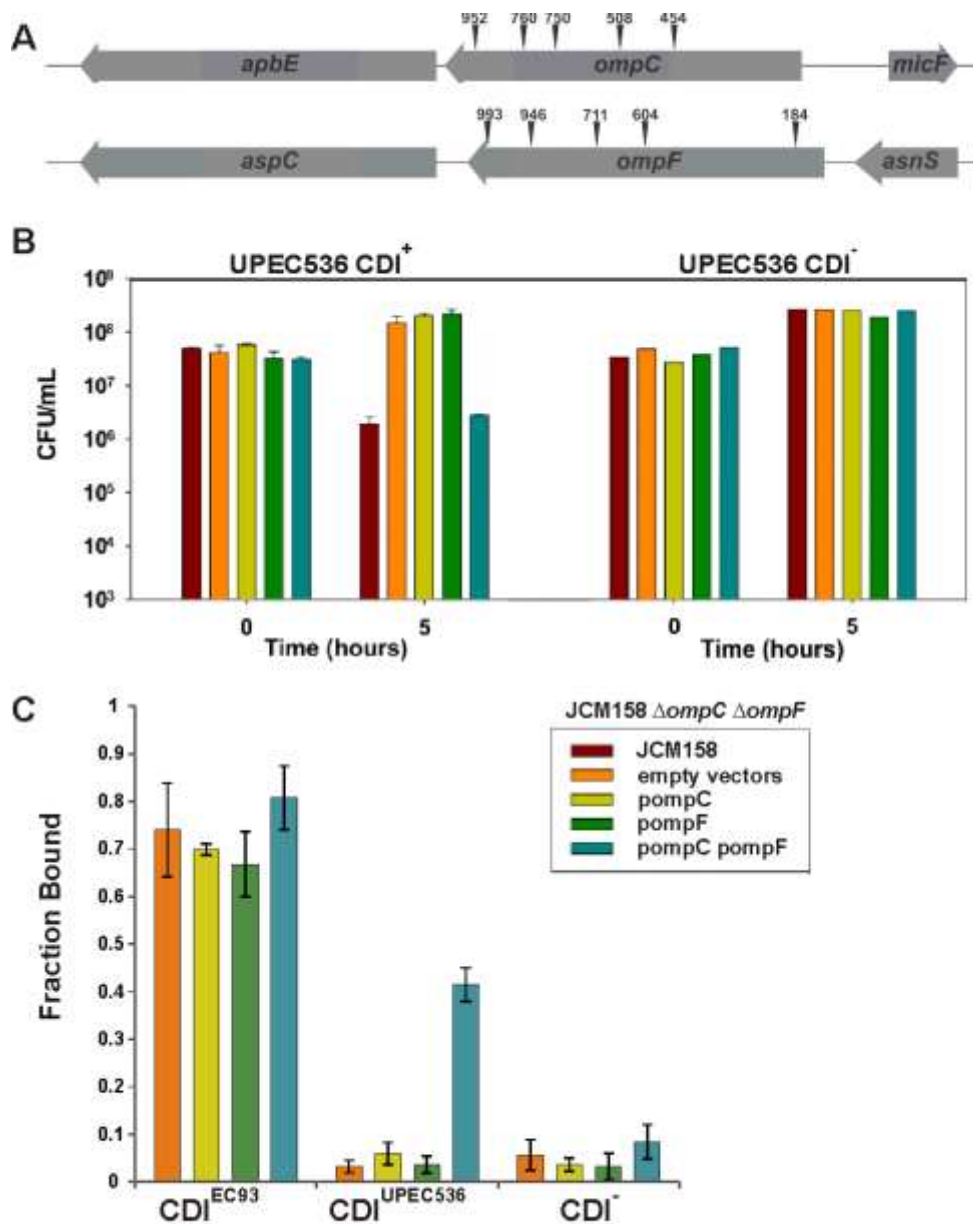
One idea, which was proposed by Ruhe *et al.*, is that CDI systems can be used to regulate community behavior. When bacteria decide to switch to the pathogenic state, it requires the group to collectively turn on the proper genes, which are often located in pathogenicity islands (PAIs)<sup>235</sup>. This is going to come at a cost to the cell as it must use protein machinery and ATP to make and the resources needed to invade a host. If there are siblings in the population that have failed to turn on their pathogenic genes (“cheaters”) this can prevent the group from efficiently invading the host<sup>236</sup>. CDI<sup>UPEC536</sup> is located in pathogenicity island-II in UPEC536. In fact, CDI systems are typically found in PAIs. While nothing is known about the regulation of CDI<sup>UPEC536</sup>, one could imagine a situation where it is under the same regulation as the pathogenic genes found in the island. If a cell has neglected to start expressing those genes, it will not be protected from sibling cells delivering CdiA-CT because it is not expressing its cognate immunity protein, CdiI. This would allow CDI<sup>+</sup> cells to pick out “cheaters” in the population when the group has decided to turn pathogenic and invade a host<sup>237</sup>. While this is completely speculative, it is supported by the presence of CDI systems in PAIs, as well as the restricted target range seen now in CDI<sup>EC93</sup> and CDI<sup>UPEC536</sup>. Furthermore, with the evidence of OmpC and OmpF being under positive selection in UPEC

strains, and CdiA<sup>UPEC536</sup> existing exclusively in UPEC strains, perhaps CDI is a method of kin selection within UTI strains.

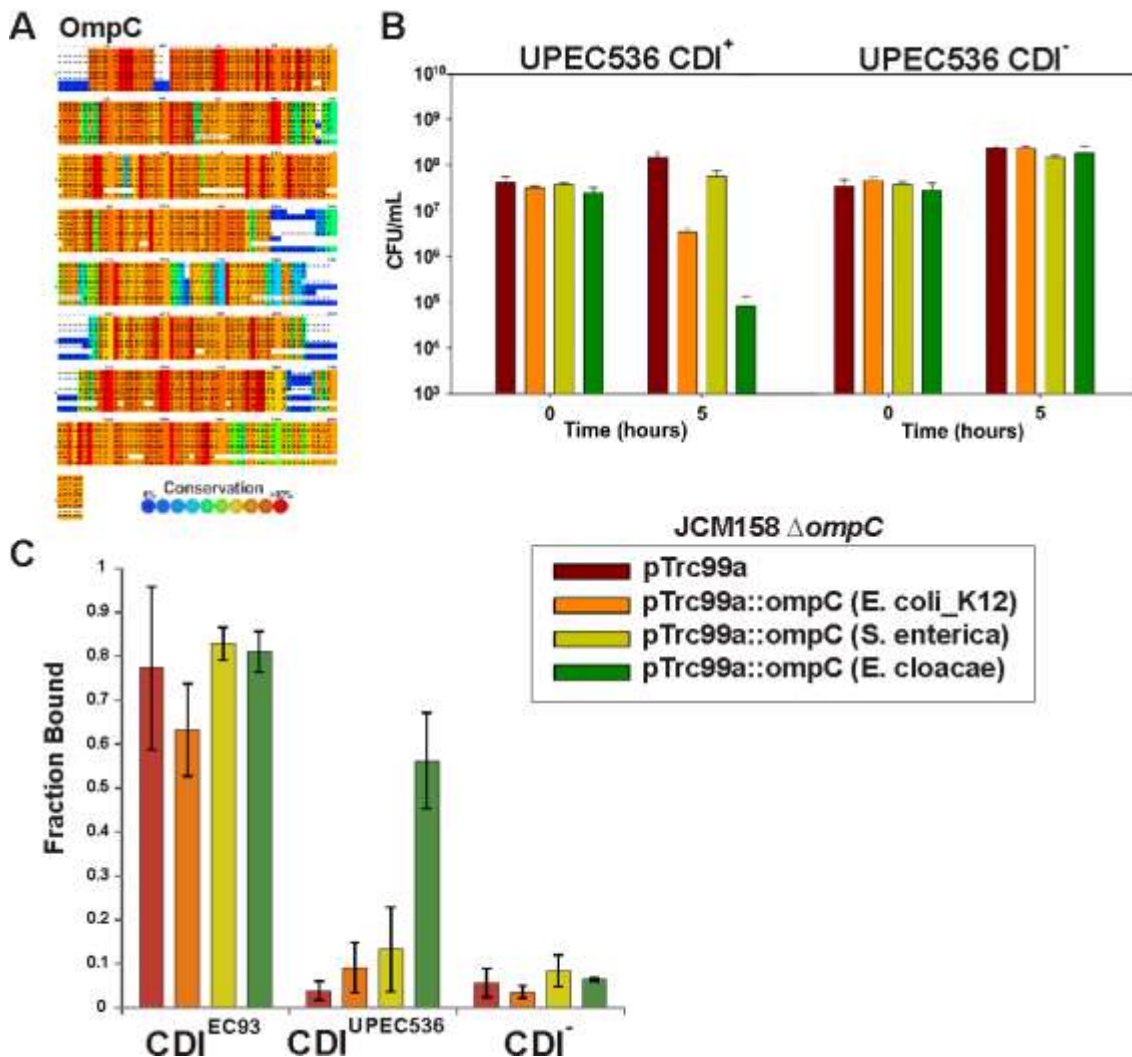
**Figure 1.  $CDI^{UPEC536}$  is not restricted by  $BamA^{E. coli}$ .** **A)** An alignment of residues 1200-2400 of CdiA from EC93 and UPEC536. Full length CdiA is 3132 and 3201 residues for EC93 and UPEC536, respectively. Boxed regions highlight the regions of variability seen between the two proteins. Alignment was performed by PRALINE Multiple Sequence Alignment. **B)** MC4100  $\Delta bamA::cat$  target cells expressing either  $BamA^{E. coli}$  or  $BamA^{E. cloacae}$  were co-cultured with EPI100 inhibitor cells expressing  $CDI^{EC93}$ ,  $CDI^{UPEC536}$ , or  $CDI^{-}$  cosmids for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.



**Figure 2. OmpC and OmpF are required for CDI<sup>UPEC536</sup>.** **A)** Open reading frames of *ompC* and *ompF* in their genomic context. Arrows indicate 5 unique transposon insertions in *ompC* and *ompF* that confer resistance to CDI<sup>UPEC536</sup>. **B)** JCM158  $\Delta ompC \Delta ompF$  target cells expressing the indicated constructs were co-cultured with UPEC536 CDI<sup>+</sup> or CDI<sup>-</sup> inhibitor cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **C)** The same target cells used in the competition experiments were labelled red and incubated with MC4100 GFP<sup>+</sup> cells carrying CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or CDI cosmids for 10 minutes at a 5:1 inhibitor to target ratio and analyzed by flow-cytometry.



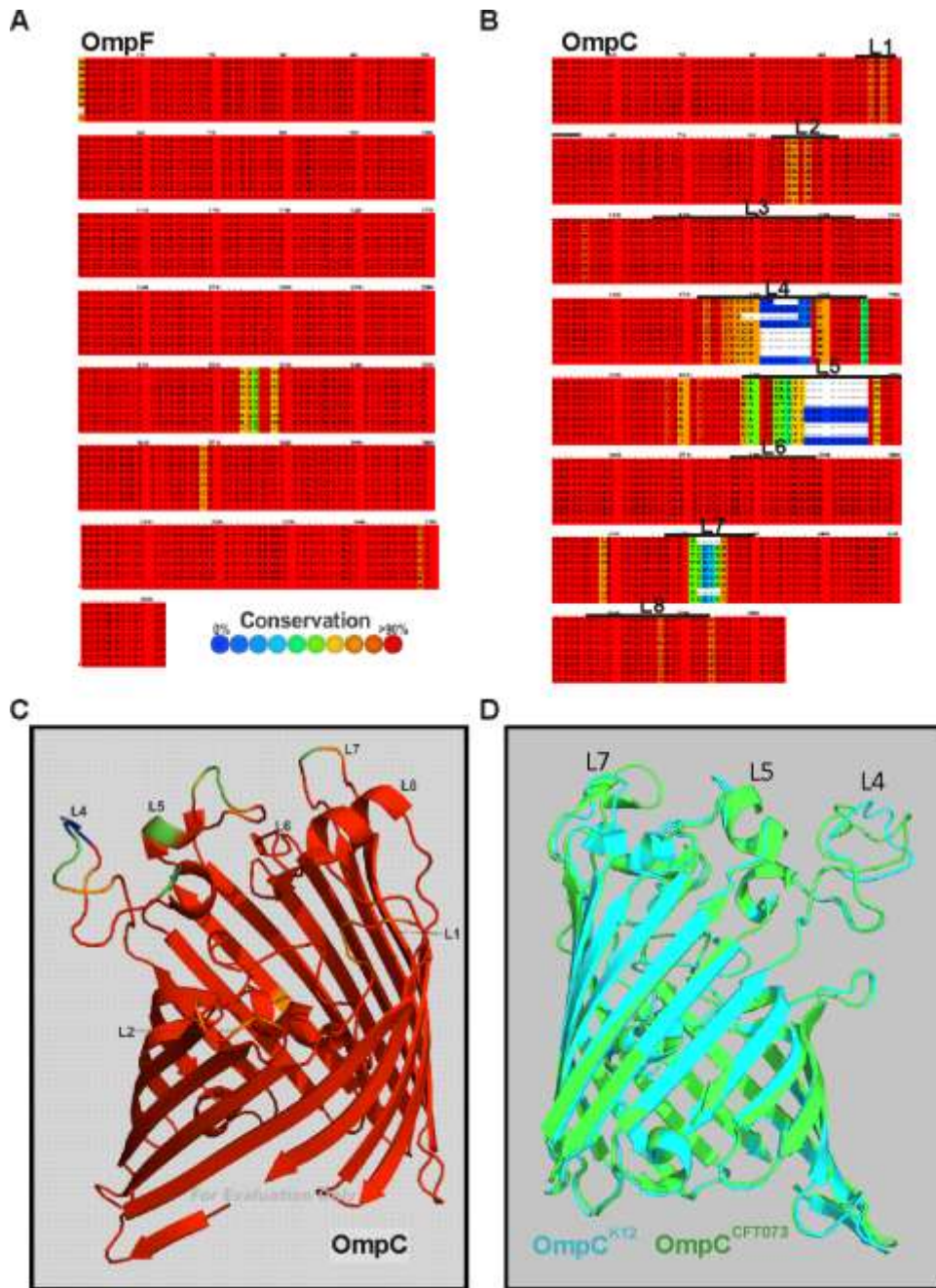
**Figure 3. Heterologous *ompC* alleles show varying levels of target cell susceptibility.** **A)** A heat map alignment of OmpC from various proteobacteria performed by PRALINE Multiple Sequence Alignment. **B)** JCM158  $\Delta ompC$  target cells expressing the annotated constructs were co-cultured with UPEC536 CDI<sup>+</sup> or CDI<sup>-</sup> inhibitor cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **C)** The same target cells used in the competition experiments were labelled red and incubated with MC4100 GFP<sup>+</sup> cells carrying CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or CDI<sup>-</sup> cosmids for 10 minutes at a 5:1 inhibitor to target ratio and analyzed by flow-cytometry.



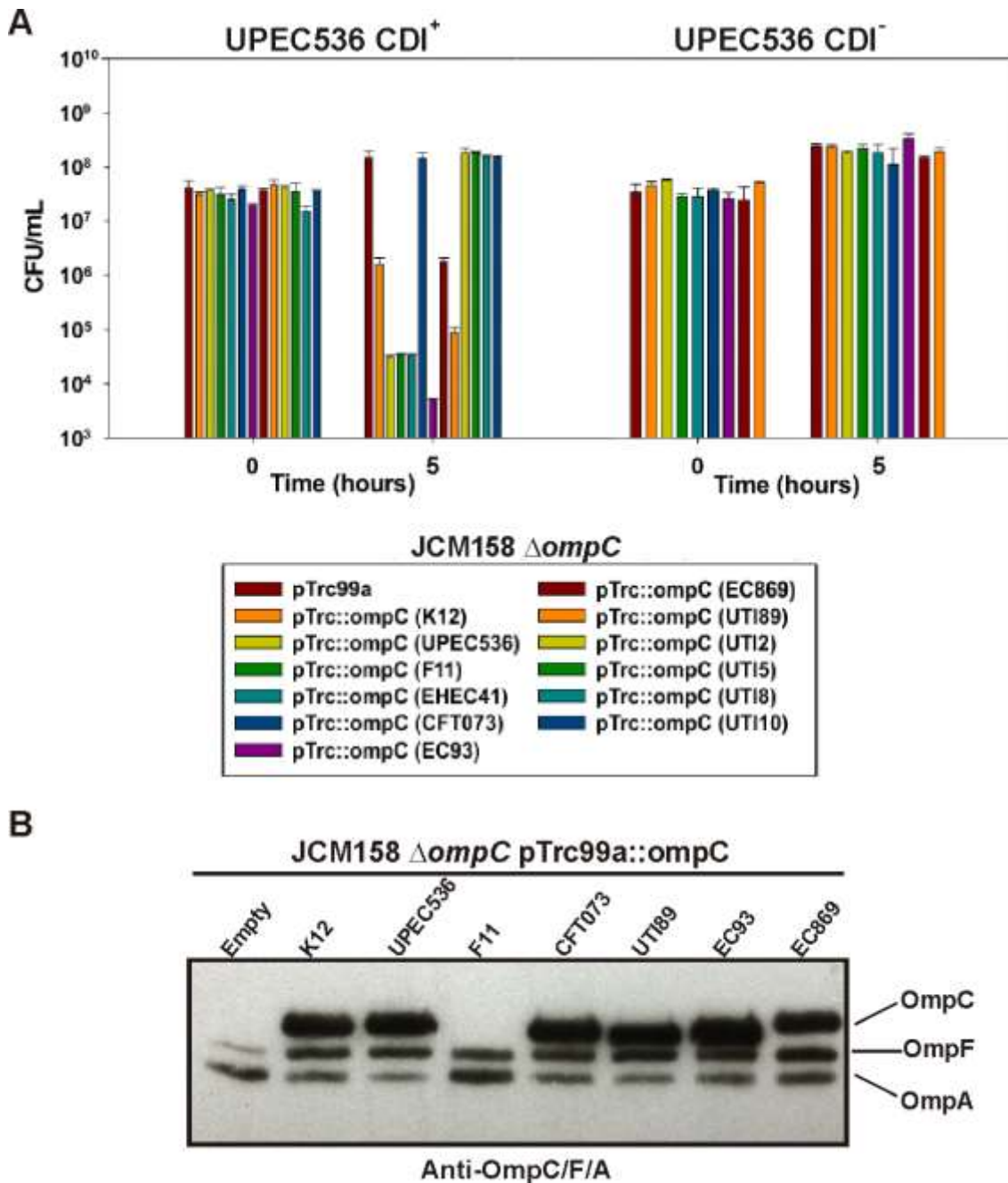




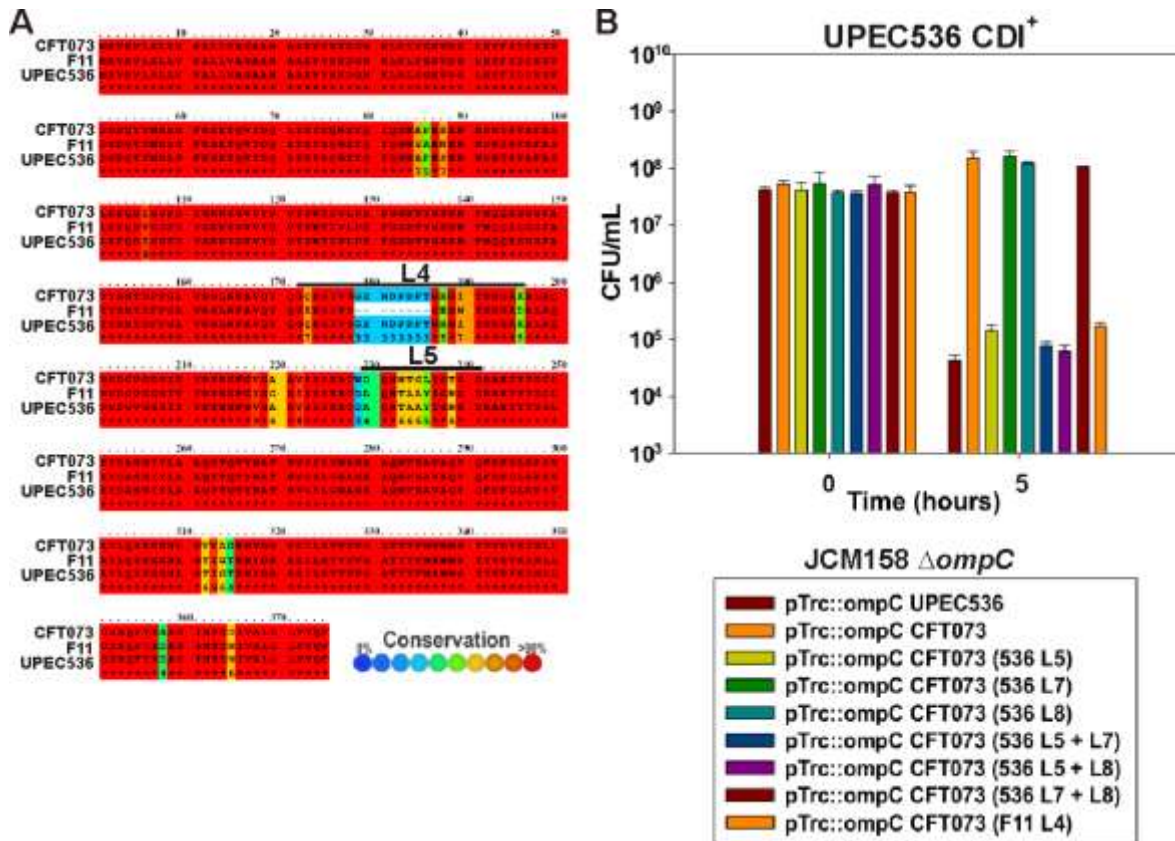
**Figure 5. OmpC and OmpF vary amongst strains of *E. coli*.** **A)** A heat map alignment of OmpF from various *E. coli* strains performed by PRALINE Multiple Sequence Alignment. **B)** OmpC from the same *E. coli* strains as A) were aligned using PRALINE Multiple Sequence Alignment. Black lines designate the 8 loops that connect the 16  $\beta$ -strands of an OmpC monomer. **C)** The heat map depicted in B) was grafted onto the crystal structure of *E. coli* K12 OmpC in monomeric form. Extracellular loops 1-2 and 4-8 are labelled. **D)** An overlay of *E. coli* K12 OmpC in light blue, and CFT073 OmpC in green.



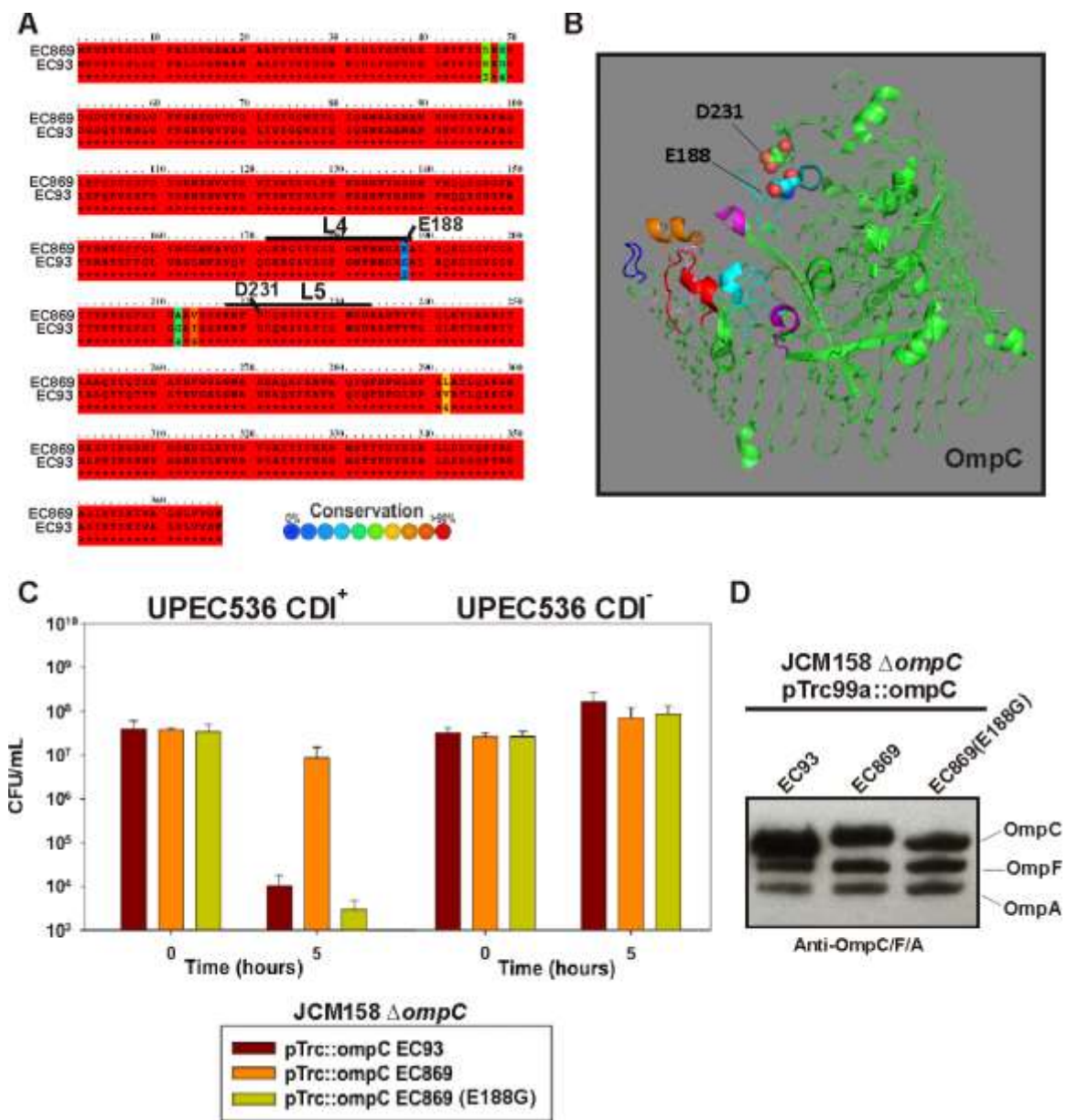
**Figure 6. *OmpC*<sup>*E.coli*</sup> polymorphism restricts *CDI*<sup>UPEC536</sup>.** **A)** JCM158  $\Delta ompC$  target cells expressing *ompC* alleles from a variety of *E. coli* strains were co-cultured with UPEC536 *CDI*<sup>+</sup> or *CDI*<sup>-</sup> inhibitor cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **B)** Western blot analysis of protein extracted from target cells in A). The antibody used was raised to *E. coli* OmpC but recognizes OmpC, OmpF, and OmpA.



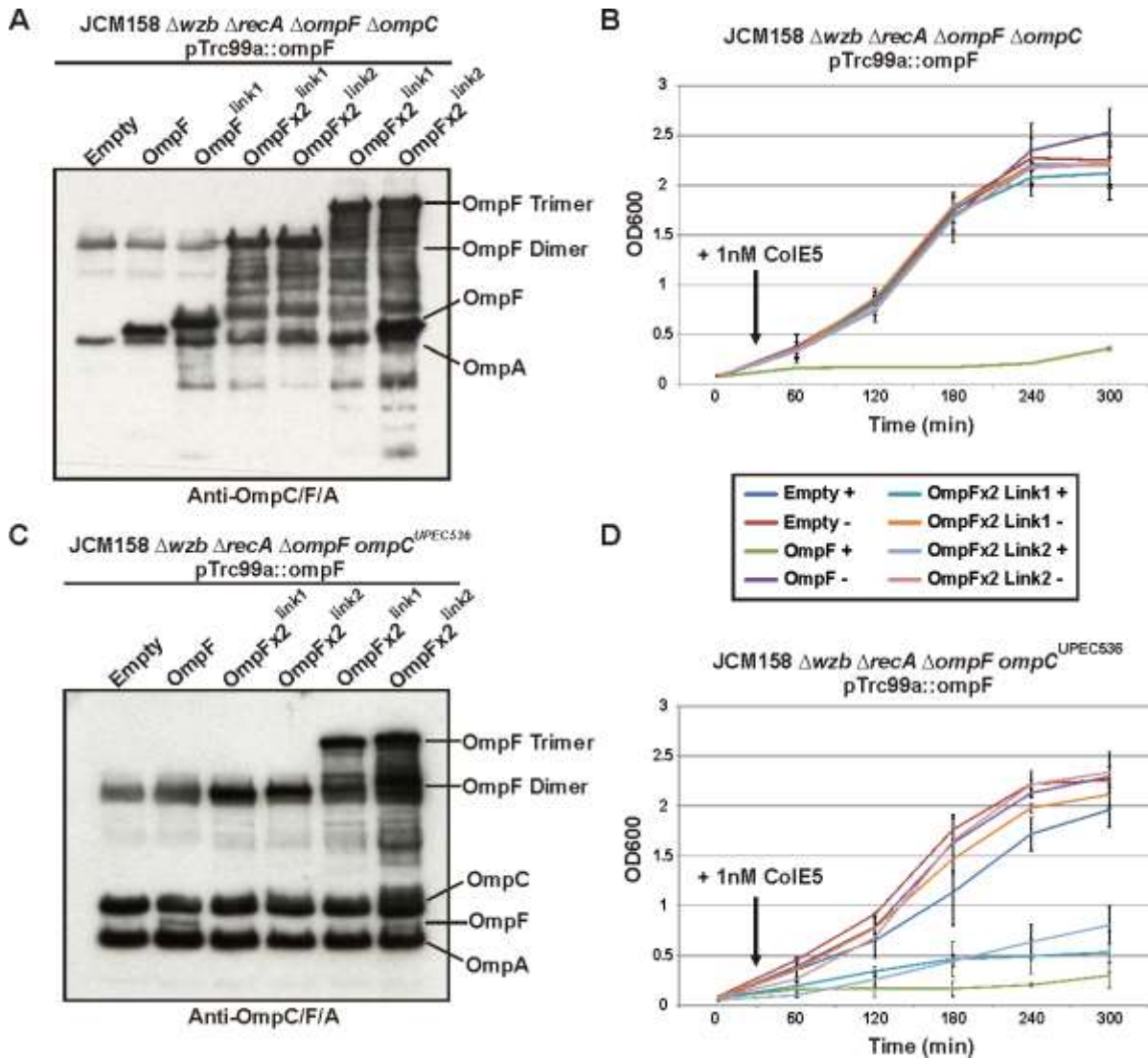
**Figure 7. CDI<sup>UPEC536</sup> targets *OmpC* extracellular loops 4 and 5.** **A)** A heat map alignment of *OmpC* from *E. coli* strains CFT073, F11, and UPEC536 performed by PRALINE Multiple Sequence Alignment. Loops 4 and 5 are designated by black bars. **B)** JCM158  $\Delta ompC$  target cells expressing the indicated *ompC* alleles were co-cultured with UPEC536 CDI<sup>+</sup> or CDI<sup>-</sup> inhibitor cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.



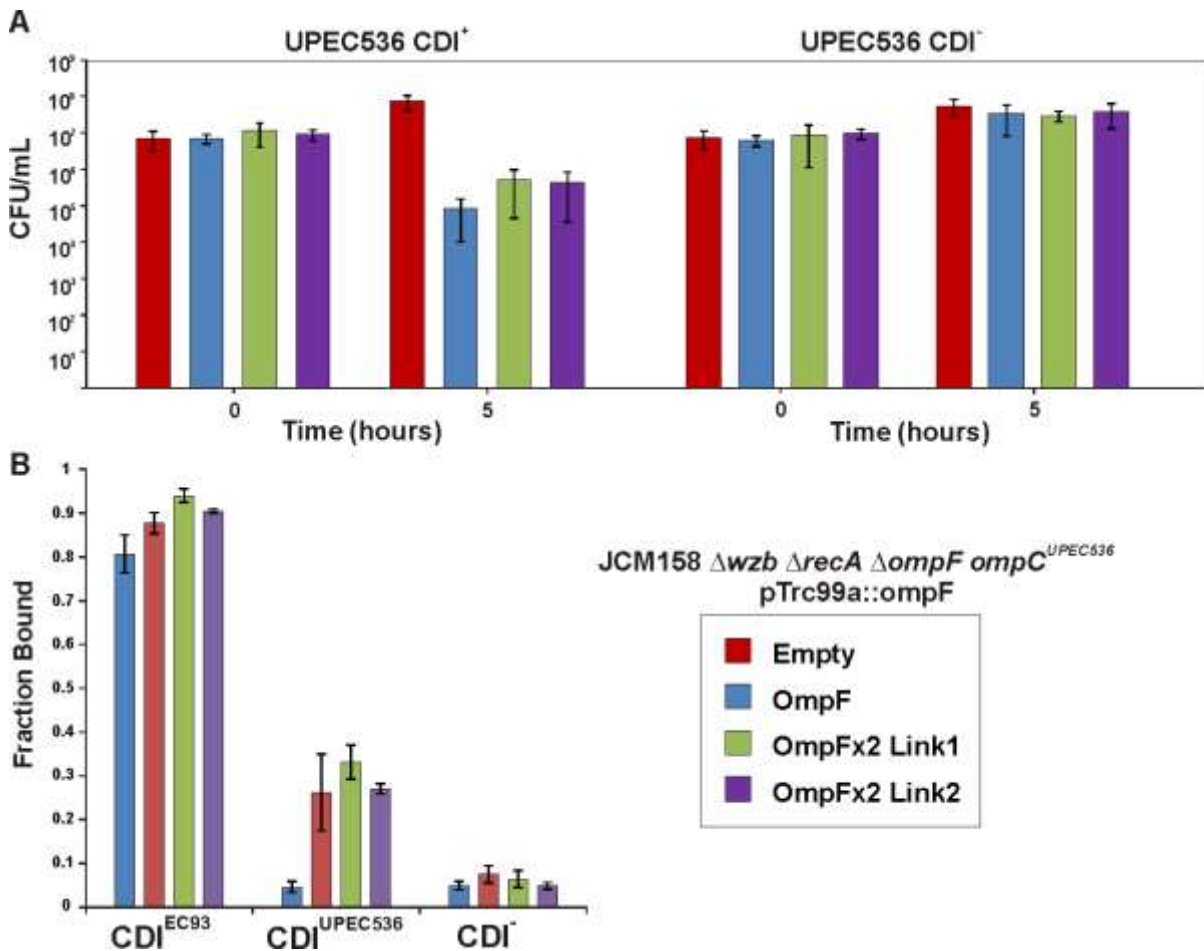
**Figure 8. The interaction between OmpC extracellular loops 4 and 5 is important for CDI<sup>UPEC536</sup>.** **A)** A heat map alignment of OmpC from *E. coli* strains EC869 and EC93 performed by PRALINE Multiple Sequence Alignment. Loops 4 and 5 are designated by black bars. **B)** Crystal structure of OmpC from *E. coli* strain UTI2 in homotrimer form. Residue E188 found in UTI2 and EC869 and residue D231 found in UTI2, EC869, and EC93 are depicted in stick form. **C)** JCM158  $\Delta ompC$  target cells expressing the annotated *ompC* alleles were co-cultured with UPEC536 CDI<sup>+</sup> or CDI<sup>-</sup> cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **D)** Western blot analysis of protein extracted from target cells in C). The antibody was raised to *E. coli* OmpC but recognizes OmpC, OmpF, and OmpA.



**Figure 9. Formation of functional OmpC/F heterotrimers *in vivo*.** **A)** Western blot analysis of the indicated cells. The antibody was raised to OmpC but recognizes OmpC/F/A. Monomeric, dimeric, and trimeric OmpF are labeled. **B)** Growth curve analysis of the designated cells with or without the addition of 1 nM of colicin E5 after 30 minutes of growth. **C) D)** Same analysis as A) and B) but cells are *ompC*<sup>+</sup>.



**Figure 10. OmpC/F heterotrimers support  $CDI^{UPEC536}$ .** **A)** The indicated target cells were co-cultured with UPEC536  $CDI^+$  or  $CDI^-$  inhibitor cells for 5 hours in LB broth at a 10:1 inhibitor to target ratio. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **B)** The same target cells used in the competition experiments were labelled red and incubated with MC4100 GFP<sup>+</sup> cells carrying  $CDI^{EC93}$ ,  $CDI^{UPEC536}$ , or  $CDI^-$  cosmids for 10 minutes at a 5:1 inhibitor to target ratio and analyzed by flow-cytometry.



**Chapter 4. Contact-dependent growth inhibition (CDI) systems  
deploy a toxic ribosomal RNase**



**Note:** This research was originally published in *Structure*. I collaborated on this project with Robert Morse in the laboratory of Celia Goulding at the University of California, Irvine. Figures 1 and 2 are solely his work. Furthermore, the majority of the writing in this chapter was done by Christopher Hayes and Celia Goulding.

**Beck CM**, Morse RP, Cunningham DA, Low DA, Goulding CW, Hayes CS. “The CdiA protein of *Enterobacter cloacae* deploys a toxic ribosomal RNase.” *Structure*. 2014 May 6;22(5):707-18.

## Abstract

Contact-dependent growth inhibition (CDI) is a form of inter-bacterial competition. CDI<sup>+</sup> cells export large CdiA effectors that carry a variety of C-terminal toxin domains (CdiA-CT). CdiA-CT toxins are specifically neutralized by cognate CdiI immunity proteins to protect toxin-producing cells from auto-inhibition. Here, we use high-resolution structure determination to elucidate the activity of a unique CDI toxin from *Enterobacter cloacae* (ECL). The CdiA-CT<sup>ECL</sup> structure is similar to the C-terminal nuclease domain of colicin E3, which cleaves 16S ribosomal RNA to disrupt protein synthesis. In accord with this structural homology, we show that CdiA-CT<sup>ECL</sup> uses the same nuclease activity to inhibit bacterial growth. Surprisingly, although colicin E3 and CdiA<sup>ECL</sup> carry equivalent toxin domains, the corresponding immunity proteins are unrelated in sequence, structure and toxin-binding site. Together, these findings reveal unexpected diversity amongst 16S rRNases and suggest that these nucleases are robust and versatile payloads for a variety of toxin-delivery platforms.

## Introduction

Bacterial genomes and plasmids encode a variety of peptide and protein toxins that mediate inter-bacterial competition. Colicins were the first such toxins to be identified and characterized from strains of *Escherichia coli*. Subsequently, it became apparent that other bacteria release similar toxins, which are now collectively termed bacteriocins<sup>238</sup>.

Bacteriocins are diffusible proteins that parasitize cell-envelope proteins to enter and kill bacteria. These toxins are composed of three domains, each responsible for a distinct step in the cell killing pathway. The central domain binds specific receptors on the surface of susceptible bacteria. The N-terminal domain mediates translocation across the cell envelope, and the C-terminal domain carries the bacteriocidal activity. This modular structure allows for delivery of diverse C-terminal toxins using conserved translocation and receptor-binding domains. For example, colicins E2 through E9 share virtually identical N-terminal domains but carry different C-terminal toxins with DNase<sup>91</sup>, ribosomal RNase<sup>100,239</sup> or tRNA anticodon nuclease activities<sup>101</sup>. Bacteriocin genes are always closely linked to immunity genes, which encode small proteins that bind the toxin domain and neutralize its toxicity. Thus, cells that harbor bacteriocinogenic plasmids are protected from toxin activity, but they may still be susceptible to the bacteriocins produced from other plasmids. Several different bacteriocin/immunity types are typically present in a given environment<sup>14,15</sup>, and these plasmids are predicted to have a significant impact on bacterial population structures<sup>240,241</sup>.

Research over the past decade has uncovered bacterial competition systems that require direct cell-to-cell contact for toxin delivery<sup>112,115,137,242,243</sup>. There are at least two pathways – mediated by type V and type VI secretion systems – for contact-dependent toxin delivery between Gram-negative bacteria<sup>167,244</sup>. The type V mechanism was the first to be

identified and this phenomenon was termed "CDI" for contact-dependent growth inhibition<sup>112</sup>. CDI is mediated by the CdiB/CdiA family of two-partner secretion proteins. CdiB is a predicted  $\beta$ -barrel protein that resides in the outer membrane and is required for export of CdiA effectors. CdiA proteins are very large (250 – 600 kDa) and are thought to extend from the inhibitor cell to interact with neighboring target bacteria. Although CdiA and bacteriocins are unrelated, these effector proteins share a number of general features. Like bacteriocins, CdiA proteins bind to specific receptors on the surface of target bacteria and these interactions determine the target-cell range<sup>118,126</sup>. Additionally, CDI toxin activity is carried at the extreme C-terminus of CdiA, and some portion of this CdiA-CT region is translocated into target bacteria<sup>115,116,122</sup>. CDI loci also encode CdiI immunity proteins, which bind and inactivate CdiA-CTs to protect toxin-producing cells from auto-inhibition. Finally, CDI systems deploy a variety of distinct toxin domains, which have different biochemical activities. Remarkably, functional chimeric effector proteins can be produced experimentally by fusing different toxins onto CdiA at a conserved VENN peptide motif, which demarcates the CdiA-CT region<sup>115</sup>. There is also substantial evidence that *cdiA-CT/cdiI* gene pairs are exchanged horizontally between bacteria<sup>121</sup>, suggesting that modularity is exploited to diversify toxin/immunity pairs. In fact, bacteria collectively share a large repository of toxin/immunity gene pairs and these modules appear to be exchanged between a number of different toxin-delivery systems<sup>121,172,245,246</sup>. For example, at least two CdiA proteins carry toxin domains that resemble bacteriocin nucleases. CdiA<sup>Dd3937</sup> from *Dickeya dadantii* 3937 carries a CT domain with 35% identity to pyocin S3<sup>115</sup>, and the C-terminal region of CdiA<sup>K96243</sup> from *Burkholderia pseudomallei* K96243 is 49% identical to colicin E5. Biochemical analyses have confirmed that each CDI toxin has the same nuclease activity as

the corresponding bacteriocin<sup>115,120</sup>. Together, these observations suggest that CDI loci integrate toxin/immunity gene pairs from diverse sources and that this diversity contributes to inter-strain competition.

In an effort to understand CDI toxin/immunity diversity and uncover new toxin activities, we have initiated structural studies of CdiA-CT/CdiI pairs from various bacteria. Here, we describe the structure and function of the CDI toxin/immunity protein pair from *Enterobacter cloacae* ATCC 13047 (ECL). The CdiA-CT<sup>ECL</sup> toxin shares no significant sequence identity with proteins of known function, but the three-dimensional structure of CdiA-CT<sup>ECL</sup> resembles the C-terminal nuclease domain of colicin E3. In accord with this structural homology, CdiA-CT<sup>ECL</sup> cleaves 16S rRNA at the same site as colicin E3 and this nuclease activity is responsible for the inhibition of bacterial growth. In contrast, CdiI<sup>ECL</sup> does not resemble the colicin E3 immunity protein (ImE3), and the two immunity proteins bind to different sites on their respective cognate toxin domains. Inspection of other CdiA proteins from *Erwinia chrysanthemi* EC16 (Uniprot: P94772), *Enterobacter hormaechei* ATCC 49162 (F5S237) and *Pseudomonas viridiflava* UASWS0038 (K6CF79) has revealed that their toxin domains share a common nuclease motif with colicin E3<sup>104</sup>. Analysis of CdiA-CT<sup>EC16</sup> from *Erwinia chrysanthemi* EC16 confirms that this toxin has 16S rRNase activity and demonstrates that the associated CdiI<sup>EC16</sup> immunity protein is specific to CdiA-CT<sup>EC16</sup> and does not provide protection against the CdiA-CT<sup>ECL</sup> nuclease. Together, these observations indicate that 16S rRNase toxins are more diverse and widespread than previously recognized.

## Materials and Methods

**Bacterial strains and plasmids used in this study**

<b>Strain or plasmid</b>	<b>Description</b>	<b>Reference</b>
<b>Strains<sup>a</sup></b>		
BL21-Gold(DE3)	<i>E. coli</i> B, F <sup>-</sup> <i>ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>dcm</i> <sup>+</sup> <i>gal</i> l(DE3) <i>endA</i> Hte, Tet <sup>R</sup>	Stratagene
X90	F <sup>'</sup> <i>lacI</i> <sup>q</sup> <i>lac'</i> <i>pro'</i> /ara Δ( <i>lac-pro</i> ) <i>nal1 argE</i> (amb) <i>rif</i> <sup>r</sup> <i>thi-1</i> , Rif <sup>R</sup>	177
EPI100	F <sup>-</sup> <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> ) φ80 <i>dlacZ</i> Δ <i>M15</i> Δ <i>lacXcZ</i> Δ <i>M15</i> Δ <i>lacX recA1 endA1 araD139</i> Δ( <i>ara, leu</i> )7697 <i>galU galK</i> λ <sup>-</sup> <i>rpsL nupG</i> , Str <sup>R</sup>	Epicentre
<i>Enterobacter cloacae</i> ATCC 13047	Type strain, Amp <sup>R</sup>	ATCC
DY378	W3110 <i>lcI857</i> Δ( <i>cro-bio</i> )	247
CH1944	X90 (DE3) Δ <i>rna</i> , Rif <sup>R</sup>	248
CH2016	X90 (DE3) Δ <i>rna</i> Δ <i>slyD::kan</i> , Kan <sup>R</sup>	248
CH10013	Spontaneous rifampicin-resistant derivative of <i>E. coli</i> JCM158, Rif <sup>R</sup>	This study
CH10420	CH1944 Δ <i>rsmE::kan</i> , Kan <sup>R</sup>	This study
CH11196	<i>E. cloacae</i> Δ <i>vasK1::kan</i> , Amp <sup>R</sup> Kan <sup>R</sup>	This study
CH11247	<i>E. cloacae</i> <i>spc-araC</i> <sup>Eco</sup> :: <i>cdiB</i> , Amp <sup>R</sup> Spc <sup>R</sup>	This study
CH11248	<i>E. cloacae</i> <i>spc-araC</i> <sup>Eco</sup> :: <i>cdiB</i> Δ <i>araBAD::kan</i> , Amp <sup>R</sup> Spc <sup>R</sup> Kan <sup>R</sup>	This study
CH11249	<i>E. cloacae</i> <i>spc-araC</i> <sup>Eco</sup> :: <i>cdiB</i> Δ <i>araBAD</i> , Amp <sup>R</sup> Spc <sup>R</sup>	This study
CH11250	<i>E. cloacae</i> <i>spc-araC</i> <sup>Eco</sup> :: <i>cdiB</i> Δ <i>araBAD</i> Δ <i>vasK1::kan</i> , Amp <sup>R</sup> Spc <sup>R</sup> Kan <sup>R</sup>	This study
CH11275	<i>E. cloacae</i> Δ <i>wzb::kan</i> , Amp <sup>R</sup> Kan <sup>R</sup>	This study
CH11432	<i>E. cloacae</i> Δ <i>wzb</i> , Amp <sup>R</sup>	This study
CH11445	<i>E. cloacae</i> Δ <i>wzb</i> Δ <i>cdiA::spc</i> , Amp <sup>R</sup> Spc <sup>R</sup>	This study
CH11475	<i>E. cloacae</i> Δ <i>wzb</i> Δ <i>cdiA1::kan</i> , Amp <sup>R</sup> Kan <sup>R</sup>	This study
<b>Plasmids</b>		
pTrc99a	IPTG-inducible expression plasmid, Amp <sup>R</sup>	GE Healthcare
pCH450	pACYC184 derivative with <i>E. coli</i> <i>araBAD</i> promoter for arabinose-inducible expression, Tet <sup>R</sup>	249
pKOBEG	L-arabinose inducible expression phage 1 <i>red</i> genes, temperature-sensitive replication origin, Cm <sup>R</sup>	250
pCP20	Heat-inducible expression of FLP recombinase, Cm <sup>R</sup> Amp <sup>R</sup>	225

pWEB-TNC	Cosmid cloning vector, Amp <sup>R</sup> Cm <sup>R</sup>	Epicentre
pKAN	pBluescript with kanamycin-resistance cassette ligated to SmaI site, Amp <sup>R</sup> Kan <sup>R</sup>	251
pSPM	pBluescript with spectinomycin-resistance cassette ligated between BamHI and EcoRI sites, Amp <sup>R</sup> Spc <sup>R</sup>	252
pCH8001	pET21KS:: <i>yqcFG</i> , expression construct containing the <i>Bacillus subtilis</i> 168 <i>yqcFG</i> operon, Amp <sup>R</sup>	lab collection
pCH10163	Cosmid pCdiA-CT/ <i>pheS</i> * that carries a <i>kan-pheS</i> * cassette in place of the <i>E. coli</i> EC93 <i>cdiA-CT/cdiI</i> coding sequence. Used for allelic exchange and counter-selection. Cm <sup>R</sup> Kan <sup>R</sup>	122
pET21K:: <i>cdiA-CT/cdiI</i> <sup>ECL</sup>	Overproduces CdiA-CT/CdiI <sup>ECL</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21K:: <i>cdiA-CT(D203A)/cdiI</i> <sup>ECL</sup>	Overproduces CdiA-CT(D203A)/CdiI <sup>ECL</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21K:: <i>cdiA-CT(D205A)/cdiI</i> <sup>ECL</sup>	Overproduces CdiA-CT(D205A)/CdiI <sup>ECL</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21K:: <i>cdiA-CT(H207A)/cdiI</i> <sup>ECL</sup>	Overproduces CdiA-CT(H207A)/CdiI <sup>ECL</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21K:: <i>cdiA-CT(K214A)/cdiI</i> <sup>ECL</sup>	Overproduces CdiA-CT(K214A)/CdiI <sup>ECL</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21S:: <i>cdiA-CT/cdiI</i> <sup>EC16</sup>	Overproduces CdiA-CT/CdiI <sup>EC16</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21K:: <i>cdiI</i> <sup>EC16</sup>	Overproduces CdiI <sup>EC16</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pCH450:: <i>cdiA-CT</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiA-CT</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(D203A)</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiA-CT(D203A)</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(D205A)</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiA-CT(D205A)</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(H207A)</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiA-CT(H207A)</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(K214A)</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiA-CT(K214A)</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT</i> <sup>EC16</sup>	Arabinose-inducible expression of <i>cdiA-CT</i> <sup>EC16</sup> , Tet <sup>R</sup>	This study
pTrc99a:: <i>cdiI</i> <sup>ECL</sup>	IPTG-inducible expression of <i>cdiI</i> <sup>ECL</sup> , Amp <sup>R</sup>	This study
pTrcKX:: <i>cdiI</i> <sup>EC16</sup>	IPTG-inducible expression of <i>cdiI</i> <sup>EC16</sup> , Amp <sup>R</sup>	This study
pCH450:: <i>cdiI</i> <sup>ECL</sup>	Arabinose-inducible expression of <i>cdiI</i> <sup>ECL</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiI</i> <sup>ECL</sup> -	Arabinose-inducible expression of <i>cdiI</i> <sup>ECL</sup> - <i>his</i> <sub>6</sub>	This study

<i>his<sub>6</sub></i>	for purification of CdiI <sup>ECL</sup> -His <sub>6</sub> , Tet <sup>R</sup>	
C16 pCH450KX:: <i>cdiI</i> <sup>E</sup>	Arabinose-inducible expression of <i>cdiI</i> <sup>ECL16</sup> , Tet <sup>R</sup>	This study
pCH10445	Constitutive expression of chimeric <i>cdiA</i> <sup>ECL93</sup> - <i>CT</i> <sup>ECL</sup> and <i>cdiI</i> <sup>ECL</sup> genes, Cm <sup>R</sup>	This study
pCH10508	Derivative of pCH10445 carrying the H207A mutation in the <i>cdiA-CT</i> <sup>ECL</sup> coding sequence, Cm <sup>R</sup>	This study
pCH11050	pKAN- $\Delta$ <i>vasKI</i> <sup>ECL</sup> , Amp <sup>R</sup> Kan <sup>R</sup>	This study
pCH11506	pKAN- $\Delta$ <i>wzb</i> <sup>ECL</sup> , Amp <sup>R</sup> Kan <sup>R</sup>	This study
pCH11507	pSPM- $\Delta$ <i>cdiA</i> <sup>ECL</sup> , Amp <sup>R</sup> Spc <sup>R</sup>	This study
pCH11508	pKAN- $\Delta$ <i>cdiAI</i> <sup>ECL</sup> , Amp <sup>R</sup> Kan <sup>R</sup>	This study

<sup>a</sup>All bacterial strains are *Escherichia coli* unless otherwise noted.

### ***Bacterial strains and growth conditions***

All bacterial strains and plasmids used in this study are listed in the table above. Bacteria were grown in LB media or LB-agar with antibiotics at the following concentrations: ampicillin (Amp) 150 µg/mL; kanamycin (Kan) 50 µg/mL; rifampicin (Rif) 200 µg/mL; spectinomycin (Spc) 100 µg/mL; and tetracycline (Tet) 10 µg/mL. The  $\Delta$ *rsmE::kan* mutation<sup>184</sup> was transferred into *E. coli* strain CH1944 by bacteriophage P1-mediated transduction. *E. cloacae* genes were deleted using the same protocol as described for *E. coli*<sup>251</sup>. DNA sequences located upstream and downstream of target genes were amplified and cloned into plasmid pKAN<sup>251</sup> or pSPM<sup>252</sup> to flank kanamycin- or spectinomycin-resistance cassettes, respectively. The resulting plasmids were linearized by restriction endonuclease digestion and electroporated into *E. cloacae* cells expressing the phage  $\lambda$  Red proteins from plasmid pKOBEG<sup>250,253</sup>.

The *wzb* upstream region was amplified with oligonucleotides (restriction sites underlined) **ECL-wzb(KO)-Sac** (5' - TTT GAG CTC CTG CAG GCG CTG ATG C)/**ECL-wzb(KO)-Bam** (5' - TTT GGA TCC GCG GAT TAC CAG GTA TG), and the downstream region with **ECL-wzb(KO)-Eco** (5' - TTT GAA TTC CAG CAG GGA TAA CAA TGA



C)/**ECL-wzb(KO)-Kpn** (5′ - TTT GGT ACC CCC TTT CGG CAG CAC). The kanamycin-resistance cassette was removed from the resulting strain CH11275 using FLP recombinase<sup>185,225</sup> to generate strain CH11432. The *cdiA*<sup>ECL</sup> upstream region was amplified with **ECL-cdiA(KO)-Sac** (5′ - TTT GAG CTC GTT TTC AGC GCC CGG TTG)/**ECL-cdiA(KO)-Bam** (5′ - TTT GGA TCC TGT CCC TTA GAA AGC CAC TG), and the downstream region with **ECL-cdiA(KO)-Eco** (5′ - GCG GAA TTC ACA GGA TTA AGG ACT AGT G)/**ECL-cdiA(KO)-Kpn** (5′ - CTC GGT ACC GAG CTC TGT TTC AGT TG). The resulting construct was used to generate strain CH11445. The *cdiI*<sup>ECL</sup> downstream region was amplified with primers **ECL-cdiI(KO)-Eco** (5′ - ACA GAA TTC ACT ACC GAG CCG GGA GC)/**ECL-cdiI(KO)-Kpn** (5′ - CGG GGT ACC TAC GCG TAG TGA ATT ATC), and used with the *cdiA*<sup>ECL</sup> upstream homology fragment to generate the  $\Delta cdiAI::kan$  deletion construct. This construct was used to delete the *cdiI*<sup>ECL</sup> gene from strain CH11445, thereby generating immunity-deficient strain CH11475 to serve as target bacteria for competition experiments.

Inducible *E. cloacae* inhibitor cells (strain CH11250) were generated in several steps. The *E. coli araBAD* promoter was first recombined upstream of the *cdiBAI* gene cluster in *E. cloacae*. A region containing the promoter, *araC* and a spectinomycin-resistance cassette was amplified from TGI1<sup>115</sup> with primers **ECL-spec-araC-cdiB(OE)-for** (5′ - AGT TGC GCG TTT TTT GAC AGG CAA GTC TGA AAA CGG ACT TTT AGC AAT GCT TGC ATA ATG) and **ECL-spec-araC-cdiB(OE)-rev** (5′ - CGA CGT TTT CCC TGA CAT GGT GAA TTC CTC CTG CTA GCC CAA AAA AAC GGG TAT GGA G). A region upstream of *cdiB*<sup>ECL</sup> was amplified with **ECL-spec-araC-cdiB(UP)-for** (5′ - CGC TTT CAG TTC CAG TTT GC) and **ECL-spec-araC-cdiB(UP)-rev** (5′ - GCC TGT CAA AAA ACG CGC); and

the downstream homology region amplified with **ECL-spec-araC-cdiB(DS)-for** (5′ - CCA TGT CAG GGA AAA CGT CGA TAC) and **ECL-spec-araC-cdiB(DS)-rev** (5′ - AAT TTT CCC TTC GAG GCC G). The three products were combined by OE-PCR and the resulting product was electroporated into *E. cloacae* to generate strain CH11247. The *araBAD*<sup>ECL</sup> operon was then deleted. Primers **ECL-araBAD(OE)-for** (5′ - GTT TTC TGA TGG AGC AAC ACC GAA TTG GAG CTC CAC CGC) and **ECL-araBAD(OE)-rev** (5′ - GTT TTA TGC TGG GTT TAT ATA CAG TCA AAA GCT GGG TAC CGG G) were used to amplify the kanamycin-resistance cassette from pCH70. Primers **ECL-araBAD(UP)-for** (5′ - TTT AAG CTT GAT TAT CCC TTC CCC ACG C) and **ECL-araBAD(UP)-rev** (5′ - GGT GTT GCT CCA TCA GAA AAC) were used to amplify a region upstream of *araB*; and primers **ECL-araBAD(DS)-for** (5′ - ACT GTA TAT AAA CCC AGC ATA AAA C) and **ECL-araBAD(DS)-rev** (5′ - TTT CTC GAG CAG GGG CGG TAA TAA ACC G) were used to amplify a region downstream of *araD*. The PCR products were combined with the kanamycin-resistance cassette by overlap extension-PCR (OE-PCR)<sup>254</sup> and the final product electroporated into CH11247 cell to generate strain CH11248. The kanamycin-resistance cassette was subsequently removed to produce strain CH11249. Finally, the *vasK1* gene was deleted from the inhibitor cells. The *vasK1* upstream region was amplified with primers **ECL-vasK1(KO)-Sac** (5′ - TTT GAG CTC GAA ATC GAC GCC GGT CTG)/**ECL-vasK1(KO)-Bam** (5′ - TTT GGA TCC TTT CCT TGC GGC AAT CCG); and the downstream region with primers **ECL-vasK1(KO)-Eco** (5′ - TTT GAA TTC CAA GGA CAG CCG TAT GAC)/**ECL-vasK1(KO)-Kpn** (5′ - TTT GGT ACC GAA TCG ACA TCA GCA TCT C). The resulting construct was electroporated into CH11249 cells followed by selection for kanamycin-resistant recombinants.

### ***Protein expression constructs***

Protein overproduction constructs were generated using plasmid pET21S<sup>115</sup> and its derivative pCH8001, which contains a KpnI restriction site immediately upstream of NcoI. The *cdiA-CT/cdiI<sup>ECL</sup>* coding sequence was amplified from *E. cloacae* genomic DNA using primers **ECL-CT-Kpn/Nco-for** (5' - GGG GGT ACC ATG GCT GAG AAT AAC TCG CTG GC) and **ECL-cdiI-Nhe-rev** (5' - CTC TGC TAG CGT TGT TAA GAC TAT GAT AAA AAT C). The PCR product was digested with KpnI/NheI and ligated to KpnI/SpeI-digested pCH8001 to generate plasmid pET21K::*cdiA-CT/cdiI<sup>ECL</sup>*. Point mutations were introduced into *cdiA-CT<sup>ECL</sup>* using megaprimer PCR<sup>187</sup>. Megaprimers were first generated using mutagenic forward primers: **ECL-CT(D203A)-for** (5' - CTA CAT TAC CCC CGC TAT TGA TAG TCA TAA), **ECL-CT(D205A)-for** (5' - CAT TAC CCC CGA TAT TGC TAG TCA TAA TGT TAC), **ECL-CT(H207A)-for** (5' - CCC GAT ATT GAT AGT GCT AAT GTT ACT AAT GG), and **ECL-CT(K214A)-for** (5' - GTT ACT AAT GGC TGG GCG ATG TTC AAT AGT AAA GG) in conjunction with primer **ECL-cdiI-Nhe-rev**. These products were used in subsequent reactions with **ECL-CT-Kpn/Nco-for** to generate complete *cdiA-CT/cdiI<sup>ECL</sup>* modules, which were ligated into pCH8001. The *cdiA-CT/cdiI<sup>EC16</sup>* region was amplified from *Erwinia chrysanthemi* EC16 genomic DNA using primers **EC16-CT-Nco-for** (5' - AGG CCA TGG TGG AGA ATA ACC TTT TGT CTG C) and **EC16-virA-Spe-rev** (5' - TCA TAC TAG TTT GGG AAT CAA GGA TCT CAT ATC G). The resulting product was digested with NcoI/SpeI and ligated to plasmid pET21S. The *cdiI<sup>EC16</sup>* gene was amplified with primers **EC16-virA-Kpn-for** (5' - TTT GGT ACC ATG CAA

GAG CCA GTA AG) and **EC16-virA-Spe-rev**, digested with KpnI/SpeI and ligated into pCH8001.

Arabinose-inducible *cdiA-CT* expression constructs were generated using plasmid pCH450<sup>249</sup>. The *cdiA-CT*<sup>ECL</sup> and *cdiA-CT*<sup>EC16</sup> coding sequences were amplified using **ECL-CT-Kpn/Nco-for** with **ECL-CT-Xho-rev** (5' - TCC CTC GAG ACC ACT AGT CCT TAA TCC), and **EC16-CT-Nco-for** with **EC16-CT-Xho-rev** (5' - TTT CTC GAG TTA TTT TAC CTT TAA ATC), respectively. The products were digested with NcoI/XhoI and ligated to plasmid pCH450. For *in vivo* immunity assays, compatible *cdiI* expression constructs were generated using plasmid pTrc99A derivatives. The *cdiI*<sup>ECL</sup> gene was amplified with **ECL-cdiI-Eco-for** (5' - GAT TAA GGA ATT CTG GTA TGT TTG G) and **ECL-cdiI-Sac-rev** (5' - AGT GAG CTC TGT TTC AGT TGT TAA GAC) and ligated to pTrc99A using EcoRI and SacI restriction sites. The *cdiI*<sup>ECL</sup> gene was also amplified with **ECL-cdiI-Eco-for** and **pET-Pst** (5' - CGG CTG CAG CAG CCA ACT CAG TGG) and ligated to plasmid pCH450 for the overproduction and purification of CdiI<sup>ECL</sup>-His<sub>6</sub>. The *cdiI*<sup>EC16</sup> gene was amplified with **EC16-virA-Kpn-for** and **EC16-virA-Xho-rev** (5' - TTT CTC GAG TTA TTG GGA ATC AAG GAT C) and ligated to plasmid pTrc-*rhsI*<sub>B</sub><sup>252</sup> using KpnI and XhoI restriction sites. The *cdiI*<sup>ECL</sup> and *cdiI*<sup>EC16</sup> genes were also subcloned into plasmid pCH450 to test immunity function in *E. cloacae* growth competitions.

The EC93-ECL chimeric CDI system was constructed by allelic exchange of the counter-selectable *pheS\** marker from cosmid pCH10163 as described<sup>122</sup>. The *cdiA-CT/cdiI*<sup>ECL</sup> coding region was amplified with primers **ECL-CT-chim-for** (5' - TCG CTG GCA CTG GTT GCC AGA GGT TG) and **ECL-cdiI-chim-rev** (5' - TTA GTT GTT AAG ACT ATG ATA AAA ATC). An upstream *cdiA*<sup>EC93</sup> homology region was amplified with

primers **DL1527** (5′ - GAA CAT CCT GGC ATG AGC G) and **ECL-chim(OE)-rev** (5′ - CAA CCT CTG GCA ACC AGT GCC AGC GAA TTA TTC TCA ACC GAG TTC CTA CCT G), and a downstream homology region amplified with primers **ECL-chim(OE)-for** (5′ - GAT TTT TAT CAT AGT CTT AAC AAC TAA CCC AAA GGT TAG ACA CCA GAC C) and **DL2368** (5′ - GTT GGT AGT GGT GGT GCT G). The three PCR products were combined by OE-PCR using oligonucleotides **DL1527** and **DL2368**. The final DNA product (100 ng) was electroporated together with pCH10163 (300 ng) into *E. coli* strain DY378 cells<sup>247</sup> and recombinants selected on yeast extract glucose-agar supplemented with 33 µg/mL chloramphenicol and 10 mM D/L-*p*-chlorophenylalanine.

### ***Protein purification and crystallography***

The CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup>-His<sub>6</sub> complex was overproduced from plasmid pET21K::*cdiA-CT/cdiI<sup>ECL</sup>* in either *E. coli* BL21 (DE3) Gold (Stratagene) or *E. coli* CH2016 cells induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Selenomethionine (SeMet) labeled complex was produced from cells grown in M9 minimal medium supplemented with leucine, isoleucine and valine at 50 mg/L; phenylalanine, lysine and threonine at 100 mg/L; and SeMet at 75 mg/L) as previously described<sup>255</sup>. Cells were washed with resuspension buffer [20 mM sodium phosphate (pH 7.0), 200 mM NaCl], then broken by sonication on ice in resuspension buffer supplemented with 10 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride. Unbroken cells and debris were removed by centrifugation at 18,000 5g for 30 min followed by filtration through a 0.45 µm filter. Clarified lysates were loaded onto a Ni<sup>2+</sup> charged HiTrap column (GE Healthcare) and washed with resuspension buffer containing 10 mM imidazole. The CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup>-

His<sub>6</sub> complex was eluted with a linear gradient of imidazole (10 – 500 mM) in resuspension buffer. Fractions were collected, combined, and concentrated to ~500 µL in a centrifugal concentrator. Complexes were further purified by gel filtration on a Superdex 200 column equilibrated with 20 mM Tris-HCl (pH 7.4), 150 mM NaCl using an AKTA FPLC. SeMet-labeled complex was concentrated to 9 mg/mL in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl for crystallization trials.

Crystals of SeMet-labeled CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup> complex were grown over 1 month at room temperature by hanging drop-vapor diffusion with a reservoir containing 1.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M Bis Tris (pH 5.1) and 1% (wt/vol) PEG 3350. The hanging drop contained a 2:1 (vol:vol) ratio of protein to reservoir solutions. The SeMet-labeled complex crystallized in space group P4<sub>1</sub>22 with unit cell dimensions of 85.64 Å 5 85.64 Å 5 75.17 Å and one complex per asymmetric unit. Crystals were mounted and data collected under cryo-conditions with the addition of 40% (vol/vol) glycerol as cryoprotectant to the reservoir condition. A Se-MAD dataset was collected at 70K at the Se-absorption edge (0.9759 Å), inflection (0.9794 Å) and remote (1.377 Å). Data reduction for each wavelength was carried out with the HKL2000 suite<sup>256</sup>, resulting in three 100% complete datasets with combined significant anomalous differences up to 4.6 Å resolution. Experimental phasing and initial model building was performed using AutoSol and Autobuild in PHENIX<sup>257</sup>, in which six Se atoms were located per asymmetric unit. A high-resolution dataset was collected with a second P4<sub>1</sub>22 Se-Met-labeled crystal with a slightly smaller unit cell (85.25 Å 5 85.25 Å 5 74.91 Å) at 70K at 1.00 Å. Data reduction was carried out with the HKL2000 suite, resulting in a 100% complete dataset up to 2.4 Å. The structure was solved via molecular replacement utilizing phenix.automr, and the final model obtained through iterative manual building in

Coot and refinement with phenix.refine<sup>258-260</sup>. The final model includes residues 160 – 235 of CdiA-CT<sup>ECL</sup> (numbered from Ala1 of the AENN motif) and residues 1 – 141 of CdiI<sup>ECL</sup> with a final R<sub>work</sub>/R<sub>free</sub> (%) 18.3/23.7. CdiA-CT<sup>ECL</sup> residues Leu160 – Lys163, His188 and loop L4 (residues Ser206 – Asn211) were modeled as alanines due to lack of observable side chain density. Additionally, residues Glu72, Lys73, Glu99 and His102 of CdiI<sup>ECL</sup> were modeled as alanines. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.pdb.org) as PDB ID code 4NTQ.

### ***Toxin-immunity binding and RNase assays***

CdiA-CT/CdiI-His<sub>6</sub> complexes were denatured in binding buffer [20 mM sodium phosphate (pH 7.0, 150 mM NaCl, 10 mM β-mercaptoethanol] supplemented with 6 M guanidine-HCl, and CdiA-CT proteins isolated in the void volume during Ni<sup>2+</sup>-affinity chromatography<sup>115</sup>. Toxins were refolded by dialysis against binding buffer. CdiI-His<sub>6</sub> proteins were purified by Ni<sup>2+</sup>-affinity chromatography under non-denaturing conditions as described<sup>120</sup>. The isolated toxins and immunity proteins were dialyzed against reaction buffer and quantified by absorbance at 280 nm. Purified CdiA-CT and CdiI-His<sub>6</sub> protein were mixed at 10 μM final concentration in binding buffer and binding interactions assessed by co-purification during Ni<sup>2+</sup>-affinity chromatography as described<sup>115,120</sup>. *In vitro* nuclease activity assays were conducted in binding buffer supplemented with 10 mM MgCl<sub>2</sub>. Ribosomes were isolated from S30 lysates of *E. coli* strain CH1944 by centrifugation at 100,000 x g for 4 hr at 4 °C. Ribosomes (10 μM) were incubated with purified CdiA-CT (1 μM) in binding buffer supplemented 10 mM MgCl<sub>2</sub> for 1 hr at 37 °C. Where indicated, CdiI-His<sub>6</sub> immunity protein was also included at 1 μM final concentration or 15μM of antibiotic.

Reactions were analyzed by denaturing electrophoresis on 50% urea – 6% polyacrylamide gels in 15 Tris-borate-EDTA buffer. Gels were transferred to nylon membrane and hybridized to 5'-radiolabeled oligonucleotide (5' - TAA GGA GGT GAT CCA ACC GCA), which is complementary to the 3'-end of *E. coli* 16S rRNA. For primer extension analysis, ribosomes were isolated from *E. coli* CH10420 cells that lack the RsmE methyltransferase. Reactions were quenched with guanidinium isothiocyanate-phenol and rRNA extracted as described<sup>248</sup>. Primer extension was performed as previously described<sup>261</sup> using 5'-radiolabeled oligonucleotide (5' - GGT TCC CCT ACG GTT ACC TTG) as a primer.

CdiA-CT toxicity and *in vivo* nuclease activity was assessed in *E. coli* strain X90 (Table S2). Arabinose-inducible pCH450::*cdiA-CT* expression constructs were introduced together with IPTG-inducible pTrc::*cdiI* plasmids and transformants selected on LB-agar supplemented with Tet, Amp, 0.4% D-glucose and 1 mM IPTG. The resulting strains were grown to mid-log phase in LB media with Tet, Amp, 0.4% D-glucose and 1 mM IPTG, then cells collected by centrifugation and resuspended in fresh LB supplemented with Tet, Amp, 1 mM IPTG and 0.2% L-arabinose. Cell growth was monitored by measuring the optical density at 600 nm (OD<sub>600</sub>) as a function of time. Culture samples were harvested into an equal volume of ice-cold methanol, cells were collected by centrifugation and frozen at –80 °C. RNA was isolated from the frozen cell pellets using guanidine isothiocyanate-phenol and 10 µg of total RNA used for northern blot analysis as described<sup>248</sup>.

### ***Growth competitions***

*E. cloacae* inhibitor cells (CH11250 or CH11196) were co-cultured with target *E. cloacae*  $\Delta$ *cdiAI* cells (CH11475) at a 10:1 ratio on LB-agar with 0.2% L-arabinose. Co-



cultures were incubated at 37 °C for 4 hr and then cells harvested into 15 mL M9 salts for serial dilution and colony forming unit (CFU) enumeration on selective media. Immunity function was evaluated through expression of *cdiI* genes in targets from pCH450 constructs. Cross-species competitions were performed under the same conditions using *E. coli* strain CH10013 as targets. Immunity function was assessed in *E. coli* targets through *cdi* expression from pTrc constructs.

The chimeric EC93-ECL CDI system was expressed in *E. coli* EPI100 cells from cosmids pCH10445 and pCH10508. The CDI<sup>+</sup> inhibitors were mixed at a 1:1 ratio with *E. coli* CH10013 target cells in LB media and incubated at 37 °C with shaking in a baffled flask for three hours. Samples were taken periodically for enumeration of viable target cells as CFU/mL. *E. coli* EPI100 carrying cosmid pWEB-TNC (Epicentre) was used as a mock (CDI<sup>-</sup>) inhibitor and immunity function was assessed using plasmid pTrc expression constructs. Co-culture samples were harvested into an equal volume of ice-cold methanol for subsequent RNA analysis as described above.

## Results

### *Crystallization and structure of the CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup> complex*

In a previous study, we used structural analysis to elucidate the activities of CDI toxins from *E. coli* EC869 and *Burkholderia pseudomallei* 1026b<sup>122</sup>. Because the CDI toxin/immunity pair from *E. cloacae* ATCC 13047 shares no obvious sequence homology with proteins of known function, we followed a similar structure-based approach to characterize this system. The CdiA-CT<sup>ECL</sup> region is demarcated by the AENN peptide motif and corresponds to residues Ala3087 to Asp3321 of full-length CdiA<sup>ECL</sup>. We co-expressed

CdiA-CT<sup>ECL</sup> with His<sub>6</sub>-tagged CdiI<sup>ECL</sup> and purified the complex to near homogeneity. The N-terminal region of CdiA-CT<sup>ECL</sup> underwent significant degradation during crystallization, presumably because this region is disordered. Similar N-terminal degradation has been observed with other CdiA-CTs<sup>122</sup>. The CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup> complex crystallized in space group P4<sub>1</sub>22 with one heterodimeric complex per asymmetric unit. The structure was solved by selenium multiple wavelength anomalous dispersion (Se-MAD) phasing to 2.4 Å resolution. The final refined model contains CdiA-CT<sup>ECL</sup> residues 160 – 235 (numbered from Ala1 of the AENN motif) and CdiI<sup>ECL</sup> residues 1 – 145. In addition, 62 well-resolved water molecules are included in the final model resulting in R<sub>work</sub>/R<sub>free</sub> of 18.3/23.7.

The resolved C-terminal domain of CdiA-CT<sup>ECL</sup> consists of an N-terminal  $\alpha$ -helix followed by a twisted five-stranded antiparallel  $\beta$ -sheet (Fig. 1A). The domain contains two long loops, L2 and L4, which connect  $\beta$ 1 to  $\beta$ 2 and  $\beta$ 3 to  $\beta$ 4, respectively (Fig. 1A). Weak electron density was observed for loop L4, likely due to its flexibility, and thus Ser206 – Asn211 were modeled as alanine residues. The CdiI<sup>ECL</sup> immunity protein comprises three- and four-stranded antiparallel  $\beta$ -sheets, forming a  $\beta$ -sandwich that is decorated with three  $\alpha$ -helices (Fig. 1A). The toxin and immunity protein interface is elaborate and mediated by a series of hydrogen-bonds (H-bond), electrostatic and hydrophobic interactions. CdiA-CT<sup>ECL</sup> residues within loops L2 – L6 form H-bonds and ion-pair interactions with CdiI<sup>ECL</sup> residues in loops L1', L2' and L3' and the edge of the  $\beta$ -sandwich ( $\beta$ 3',  $\beta$ 5' and  $\beta$ 6') (Fig. 1B). Notably, Arg222 and the C-terminal Asp235 residue of CdiA-CT<sup>ECL</sup> interact extensively with the immunity protein. A water-mediated network of H-bonds also contributes to the interface, resulting in more than 20 ion-pair/H-bond interactions between toxin and immunity proteins. In addition, there is a hydrophobic interface of approximately 300 Å<sup>2</sup> consisting of Ile178,

Val192, Tyr199 and Phe216 from CdiA-CT<sup>ECL</sup>, and Phe76, Phe78, Val95 and Phe97 from CdiI<sup>ECL</sup> (Fig. 1C). Overall, the CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup> complex has an interface of 1399 Å<sup>2</sup>, burying 27.6 and 17.1% of the solvent-accessible surface areas of the toxin and immunity proteins, respectively.

### ***CdiA-CT<sup>ECL</sup> is structurally homologous to the nuclease domain of colicin E3***

CdiA-CT<sup>ECL</sup> shares no structural homology with previously characterized CDI toxins from *E. coli* EC869 and *B. pseudomallei* 1026b<sup>120,122</sup>. Searches for structural homologues using the DALI server<sup>262</sup> revealed that CdiA-CT<sup>ECL</sup> is similar to the C-terminal nuclease domain of colicin E3 (ColE3-CT). Colicin E3 is a plasmid-encoded bacteriocin found in some *E. coli* strains, and its nuclease domain cleaves 16S rRNA between residues A1493 and G1494 (*E. coli* numbering) to interfere with protein synthesis<sup>105,108</sup>. The CdiA-CT<sup>ECL</sup> and ColE3-CT domains share a twisted antiparallel β-sheet and superimpose with an rmsd of 2.1 Å over 76 α-carbons, corresponding to a Z-score of 4.8, whereas the sequence identity between the two domains is approximately 18% (Figs. 2A & S2). Residues Asp510, His513 and Glu517 of colicin E3 are thought to function directly in catalysis<sup>104,105,263</sup>, and CdiA-CT<sup>ECL</sup> residues Asp203, Asp205 and Lys214 superimpose upon the colicin E3 active-site residues (Figs. 2B & S2). Together, these structural similarities suggest that CdiA-CT<sup>ECL</sup> may also share 16S rRNA nuclease activity with colicin E3.

Although the CdiA-CT<sup>ECL</sup> and ColE3-CT toxin domains are structurally similar, the corresponding immunity proteins are not related to one another in either primary or tertiary structure (Figs. S3A & S3B). The colicin E3 immunity protein (ImE3) is significantly smaller than CdiI<sup>ECL</sup> (~9.9 versus 16.9 kDa), and the two proteins have different folds (Fig.

S3B). A DALI search reveals that CdiI<sup>ECL</sup> is most similar to the Whirly family of single-stranded DNA binding proteins<sup>264</sup>. The closest structural homologues are two proteins of unknown function from cyanobacteria (PDB ID codes: 2IT9 and 2NVN), which superimpose onto CdiI<sup>ECL</sup> with rmsd of 3.6 – 4.0 Å over 120 – 122 α-carbons (Fig. S3C). The two immunity proteins also bind to different sites on their cognate toxin domains. ImE3 binds to an 'exosite' that leaves the colicin E3 active site exposed<sup>103</sup>, whereas CdiI<sup>ECL</sup> binds directly over the predicted active site (Fig. 2C). Structural alignment of the complexes reveals that the immunity binding occurs at distinct non-overlapping positions (Fig. 2D). Interestingly, ColE3-CT contains a C-terminal extension not found in CdiA-CT<sup>ECL</sup> (Fig. 2A). This C-terminal tail forms a short α-helix in some ColE3-CT structures, and this structure would likely interfere with CdiI<sup>ECL</sup> binding were it present in CdiA-CT<sup>ECL</sup>. Similarly, the orientation of loop L2 differs considerably between the toxins (Fig. 2A), and these loops could block the binding of non-cognate immunity proteins (Fig. 2D). Despite these differences, each immunity protein is predicted to prevent its cognate toxin from entering the ribosome A site (Figs. S4A & S4B), and therefore toxin inactivation is fundamentally the same for both systems.

### ***CdiA-CT<sup>ECL</sup> cleaves 16S rRNA in vivo and inhibits cell growth***

To test CdiA-CT<sup>ECL</sup> for nuclease activity, we cloned the *cdiA-CT<sup>ECL</sup>* coding sequence under the control of the arabinose-inducible P<sub>BAD</sub> promoter and expressed the toxin in *E. coli* cells. Cells carrying this construct do not grow in the presence of L-arabinose, but control cells with the empty plasmid vector grow unimpeded (Fig. 3A). Northern blot analysis revealed that the 3'-end of 16S rRNA is cleaved in cells that express *cdiA-CT<sup>ECL</sup>* but not in

control cells (Fig. 3B). We next tested whether this toxin activity is neutralized by CdiI<sup>ECL</sup>. We cloned the *cdiI*<sup>ECL</sup> immunity gene under control of an IPTG-inducible promoter and introduced the resulting plasmid into cells that harbor the arabinose-inducible toxin construct. Cells that co-express *cdiA-CT*<sup>ECL</sup> and *cdiI*<sup>ECL</sup> grow at the same rate as control cells carrying the empty vector (Fig. 3A), indicating that the immunity protein is fully protective. Furthermore, no 16S rRNA cleavage is detected in cells that co-express toxin and immunity proteins (Fig. 3B), strongly suggesting that the nuclease activity is responsible for growth inhibition.

Kleanthous and colleagues have predicted that a different CdiA-CT from *Erwinia chrysanthemi* EC16 also possesses colicin E3-like activity<sup>104</sup>. CdiA-CT<sup>EC16</sup> appears to have the same catalytic motif as ColeE3-CT<sup>104,265</sup>, but does not share significant sequence identity with CdiA-CT<sup>ECL</sup> (Fig. S2B). The CdiI<sup>EC16</sup> and CdiI<sup>ECL</sup> immunity proteins also appear to be unrelated based on primary sequence. We expressed *cdiA-CT*<sup>EC16</sup> in *E. coli* cells and found that this toxin also inhibits cell growth and induces 16S rRNA cleavage (Figs. 3A & 3B). Both growth inhibition and nuclease activities are blocked by co-expression of *cdiI*<sup>EC16</sup>, but not *cdiI*<sup>ECL</sup> (Figs. 3A & 3B). Similarly, expression of *cdiI*<sup>EC16</sup> does not protect against the toxic effects of *cdiA-CT*<sup>ECL</sup> expression (Figs. 3A & 3B). Thus, although the CDI toxins appear to share biochemical activities, only cognate immunity proteins are able to block toxicity.

#### ***CdiA-CT*<sup>ECL</sup> cleaves 16S rRNA in vitro**

In principle, CdiA-CT<sup>ECL</sup> and CdiA-CT<sup>EC16</sup> could induce an endogenous nuclease that is actually responsible for 16S rRNA cleavage. Therefore, we tested purified CdiA-CTs for

nuclease activity *in vitro*. Each CdiA-CT/CdiI-His<sub>6</sub> pair was first purified as a complex, and then the toxins were isolated from immunity proteins using Ni<sup>2+</sup>-affinity chromatography under denaturing conditions. Purified CdiA-CTs were refolded by dialysis against non-denaturing buffer prior to activity assays. Before testing for RNase activity, we first confirmed that each refolded toxin is able to re-bind its cognate immunity protein. We mixed CdiI-His<sub>6</sub> with either cognate or non-cognate CdiA-CT and subjected the mixtures to Ni<sup>2+</sup>-affinity chromatography under non-denaturing conditions. Each CdiA-CT co-purified with its cognate immunity protein (Fig. S5A), indicating that the toxins can re-establish specific binding interactions after denaturation and refolding. We next treated ribosomes with the CdiA-CTs and analyzed the reactions by northern blot hybridization. Purified CdiA-CT<sup>ECL</sup> and CdiA-CT<sup>EC16</sup> both cleave a 3'-fragment from 16S rRNA, and the activity of each toxin is blocked by equimolar cognate CdiI protein (Fig. 4).

We next used primer extension analysis to determine whether the CDI toxins cleave 16S rRNA at the same position as colicin E3. We generated an oligonucleotide that hybridizes to residues C1501 – C1521 of *E. coli* 16S rRNA (Fig. 5A) and used it as a primer in reverse transcription reactions to screen for cleavage sites. Residue U1498 of 16S rRNA is methylated at the N3 position (Fig. 5A) and this modified base is predicted to interfere with reverse transcription. Therefore, we repeated the *in vitro* nuclease reactions using ribosomes isolated from an *E. coli*  $\Delta rsmE::kan$  mutant, which lacks the U1498 methyltransferase<sup>266</sup>. Analysis of these nuclease reactions shows a strong primer-extension arrest corresponding to residue G1494 (Figs. 5A & 5B). This primer extension product is not observed when ribosomes are mock-treated with buffer, nor when the reactions contain equimolar cognate CdiI protein (Fig. 5b). These data are consistent with CdiA-CT-mediated cleavage of the

phosphodiester bond linking residues A1493 and G1494 (Fig. 5A). Thus, CdiA-CT<sup>ECL</sup> and CdiA-CT<sup>ECL16</sup> both appear to cleave 16S rRNA at the same site as colicin E3.

Colicin E3 binds near the A site of 70S ribosomes<sup>103</sup>. We reasoned that CdiA-CT<sup>ECL</sup> should also bind in the A site if it cleaves at the same position as colicin E3. To test this, we incubated 70S ribosomes with the A site acting antibiotics streptomycin and gentamicin and asked if they could block CdiA-CT<sup>ECL</sup> activity. As a specificity control, erythromycin, which is an antibiotic that functions at the exit tunnel, was included. The Northern blot analysis to the reactions show that CdiA-CT<sup>ECL</sup> activity is blocked by the addition of A site acting antibiotics, but not erythromycin (Fig S1), indicating CdiA-CT<sup>ECL</sup> likely binds to the A site of ribosomes.

#### ***Mutational analysis of the CdiA-CT<sup>ECL</sup> active site***

The side-chains of CdiA-CT<sup>ECL</sup> Asp203 and Lys214 overlay with active-site residues Asp510 and Glu517 (respectively) of colicin E3 (Figs. 2B & S2). However, because loop L4 is not well resolved in the CdiA-CT<sup>ECL</sup> structure, it is difficult to unambiguously identify a catalytic residue corresponding to His513 of colicin E3. Therefore, we mutated CdiA-CT<sup>ECL</sup> residues Asp203, Asp205, His207 and Lys214 individually to alanine and tested the resulting proteins for toxicity *in vivo* and 16S rRNase activity *in vitro*. CdiA-CT<sup>ECL</sup> variants containing Asp203Ala, His207Ala or Lys214Ala mutations have no effect on *E. coli* cell growth (Fig. 6A), suggesting that nuclease activity is disrupted. The Asp205Ala variant shows a delayed inhibition phenotype, in which cell growth is arrested ~90 min after toxin expression is induced (Fig. 6A). Comparable results were obtained with *in vitro* reactions using purified toxin variants. CdiA-CT<sup>ECL</sup> carrying the Asp203Ala, His207Ala and Lys214Ala mutations

have no detectable rRNase activities *in vitro*, whereas the Asp205Ala variant exhibits lower activity than the wild-type enzyme (Fig. 6B). We note that all the CdiA-CT<sup>ECL</sup> variants appear to be folded properly, because each protein efficiently re-binds cognate CdiI<sup>ECL</sup> immunity protein (Fig. S5B). Together, these experiments indicate that Asp203, His207 and Lys214 are required for toxin activity and could function in catalysis, whereas Asp205 plays an important yet non-essential role.

### ***CdiA-CT<sup>ECL</sup> is delivered into target bacteria during CDI***

We next asked whether the *E. cloacae* CDI system is expressed and deployed for competition. We reasoned that *E. cloacae* mutants lacking the immunity gene should be susceptible to inhibition. We deleted the *cdiA*<sup>ECL</sup> and *cdiI*<sup>ECL</sup> genes and tested the resulting double-mutant strain in competition co-cultures with wild-type *E. cloacae* cells. We found that the *E. cloacae*  $\Delta cdiA^{\text{ECL}} \Delta cdiI^{\text{ECL}}$  mutants are not inhibited by wild-type cells in either liquid or solid media (Fig. 7A and data not shown), suggesting that the CDI<sup>ECL</sup> system is not functional or may not be expressed under laboratory conditions. Therefore, we introduced the *E. coli* *araBAD* promoter upstream of the *E. cloacae* *cdi* locus to allow inducible expression of the system with arabinose and used this inducible inhibitor strain for competitions on solid growth media. When the CDI<sup>ECL</sup> system is induced, the growth of target cells is suppressed approximately 16-fold compared to co-cultures with *E. cloacae* cells carrying the wild-type *cdi* locus (Fig. 7A). Moreover, target cell growth is restored if they carry a plasmid-borne copy of the *cdiI*<sup>ECL</sup> immunity gene, but the non-cognate *cdiI*<sup>EC16</sup> gene provides no protection (Fig. 7A). We also tested the inducible *E. cloacae* inhibitor cells in co-cultures with *E. coli* target cells. Because *E. cloacae* ATCC 13047 uses one of its type VI secretion systems to



inhibit *E. coli*<sup>118</sup>, we first deleted the *vasK1* gene (ECL\_01536) from the *E. cloacae* inhibitor strain to inactivate type VI secretion. Remarkably, *E. coli* cells are more sensitive to growth inhibition, with viable target cell counts reduced ~100-fold after four hrs of co-culture (Fig. 7B). Notably, *E. coli* cell growth is unaffected during co-culture with *E. cloacae* containing the wild-type *cdi* locus (Fig. 7B), again indicating that the CDI<sup>ECL</sup> system is not expressed on lab media. Moreover, *E. coli* targets are protected by plasmid-borne *cdiI*<sup>ECL</sup>, but not *cdiI*<sup>EC16</sup> (Fig. 7B), confirming that growth inhibition is due to CdiA-CT<sup>ECL</sup> toxin activity.

Many CdiA-CT toxins are modular and can be exchanged between different CdiA proteins to generate functional effector molecules<sup>115,116,120,122</sup>. To test whether the CdiA-CT/CdiI<sup>ECL</sup> toxin/immunity protein complex is functional in the context of another CDI system, we replaced the *cdiA-CT/cdiI*<sup>EC93</sup> region of the *E. coli* EC93 CDI system with the *E. cloacae* toxin/immunity coding sequences. This fusion produces a chimeric CdiA protein with CdiA-CT<sup>ECL</sup> grafted onto CdiA<sup>EC93</sup> at the VENN peptide motif. *E. coli* cells expressing the CdiA<sup>EC93</sup>-CT<sup>ECL</sup> chimeric effector protein are potent inhibitors. Viable *E. coli* target cells are reduced ~10<sup>4</sup>-fold after three hrs of co-culture in broth with CdiA<sup>EC93</sup>-CT<sup>ECL</sup> inhibitors (Fig. 8A). Again, target cells that carry the *cdiI*<sup>ECL</sup> immunity gene are not inhibited and grow to the same level as cells cultured with mock-inhibitor cells that lack a CDI system (Fig. 8A). However, target cells expressing non-cognate *cdiI*<sup>EC16</sup> are inhibited to the same extent as cells that carry no immunity gene (Fig. 8A). Because the inhibition effect is so profound in these co-culture experiments, we asked whether toxin-damaged ribosomes could be detected in the target cells. We isolated total RNA from each competition co-culture and performed northern blot analysis to assay for RNase activity. Cleaved 16S rRNA is readily detected when the target cells lack immunity or express non-cognate *cdiI*<sup>EC16</sup> immunity (Fig. 8B). This

nuclease activity is not observed when target cells carry the cognate *cdiI*<sup>ECL</sup> gene (Fig. 8B). We also generated and tested a chimeric CdiA<sup>EC93</sup>-CT<sup>ECL</sup> effector carrying the His207Ala active site mutation in the CdiA-CT<sup>ECL</sup> toxin domain. Cells expressing this mutant effector do not inhibit *E. coli* targets, and no 16S rRNA cleavage is detected in the competition co-culture (Figs. 8A & 8B). Together, these results demonstrate that the CdiA-CT<sup>ECL</sup> toxin is delivered into target bacteria during CDI and that 16S RNase activity is solely responsible for growth inhibition.

## Discussion

These studies demonstrate that the C-terminal toxin domain of CdiA<sup>ECL</sup> adopts a structure that is similar to the C-terminal nuclease domain of colicin E3. In accord with the structural homology, the CdiA-CT<sup>ECL</sup> toxin domain has the same 16S rRNase activity as colicin E3, and this activity is required for growth inhibition. A catalytic mechanism of colicin E3 activity has been proposed by Ramakrishnan, Kleantous and their colleagues based on the structure of ColE3-CT bound to the ribosome<sup>105</sup>. Their model postulates that Glu517 of colicin E3 acts a general base to abstract a proton from the 2'-OH of 16S rRNA residue A1493. The resulting alkoxide subsequently attacks the phosphodiester linking A1493 and G1494 to cleave the 16S rRNA chain. The charged imidazole side-chain of His513 is thought to stabilize the transition state as well as donate a proton to the 5'-OH leaving group after cleavage. Colicin E3 residues Asp510 and Glu515 are within hydrogen-bonding distance of His513 and may promote its protonation<sup>105</sup>. Although ColE3-CT and CdiA-CT<sup>ECL</sup> have a similar arrangement of predicted active-site residues, there are also differences that may be functionally significant. For example, CdiA-CT<sup>ECL</sup> loop 4, which

contains the predicted active-site residues, is longer and displaced relative to the corresponding loop in colicin E3. Loop 4 also interacts directly with CdiI<sup>ECL</sup>, so it is possible that this region is positioned akin to colicin E3 in the absence of immunity protein. Flexibility of this region could also explain why the side-chains for loop L4 (Ser206 – Asn211) are not resolved in the final model. Superimposition of nuclease domains reveals that the side-chains for residues Asp203, Asp205 and Lys214 of CdiA-CT<sup>ECL</sup> are in the same relative positions as Asp510, His513 and Glu517 of colicin E3, respectively. Mutagenesis of these CdiA-CT<sup>ECL</sup> residues indicates that they are important for nuclease activity, but it is unclear whether they play the same catalytic roles as in colicin E3. For example, Lys214 in CdiA-CT<sup>ECL</sup> is unlikely to function as a generalized base as proposed for Glu517 of ColE3-CT, especially as nearby residues Asp203 and Asp205 in CdiA-CT<sup>ECL</sup> should favor side-chain protonation. Lysine residues are often found in the active sites of nucleases, where they typically function to position the scissile phosphodiester or stabilize pentavalent transition states<sup>267-269</sup>. Additionally, residue His207 is also critical for CdiA-CT<sup>ECL</sup> activity. A more definitive model of CdiA-CT<sup>ECL</sup> catalysis must likely await a structure of the toxin bound to the ribosome.

The colicin E3 nuclease domain makes a series of specific contacts with the ribosome A site that are presumably important for substrate binding. The residues mediating the ribosome contact are highly conserved amongst other E3-like enzymes such as colicins E4, E6 and cloacin DF, but in most instances, the CdiA-CT<sup>ECL</sup> nuclease domain does not share this homology. In particular, ColE3-CT loop L2 (linking  $\beta$ 1 to  $\beta$ 2) makes a number of specific contacts with the ribosome A site. Arg495 and Gln489 bind the nucleobase and phosphate of A1493, Lys496 interacts with C518 and Lys494 holds G530 in the *syn*

conformation through a bridging water molecule<sup>105</sup>. Not one of these residues is shared with CdiA-CT<sup>ECL</sup> (see Fig. S2A). In fact, loop L2 of CdiA-CT<sup>ECL</sup> is significantly displaced compared to the corresponding loop in ColE3-CT. This displacement may result from the binding of CdiI<sup>ECL</sup>, which would clash with loop L2 if it were in the ColE3-CT conformation. In the absence of CdiI<sup>ECL</sup>, it is possible that loop L2 adopts a conformation similar to that of the ColE3-CT domain, but L2 sequence divergence suggests that each loop makes distinct contacts with the ribosome. Similarly, colicin E3 residues that contact ribosomal protein S12 are not conserved within CdiA-CT<sup>ECL</sup>. Colicin E3 residues Tyr460 – Tyr464 form an intriguing pseudo- $\beta$ -sheet interaction with S12, with the side-chains of His462, Asp463 and Tyr464 making specific hydrogen-bond contacts with S12<sup>105</sup>. The corresponding region of CdiA-CT<sup>ECL</sup> was degraded during crystallization, precluding direct comparison with colicin E3. However, the relevant primary sequences reveal no obvious homology, suggesting that this region makes unique contacts if it does indeed interact with ribosomal protein S12.

The other striking difference between the two systems is the immunity protein. CdiI<sup>ECL</sup> and ImE3 are significantly different in size, share less than 12% sequence identity, and bind to distinct, non-overlapping sites on their nuclease domains. Moreover, the two immunity proteins have different tertiary structures and folds. Structural homology searches reveal that the Whirly family of single-stranded DNA-binding proteins are most similar to CdiI<sup>ECL</sup>. This homology is relatively weak ( $Z$ -scores 4.1 to 4.3 and rmsd  $\sim$ 4.0 Å), but these proteins all share the same topology. The fact that the immunity proteins for colicin E3 and CdiA-CT<sup>ECL</sup> toxins are unrelated in both primary sequence and tertiary structure suggests that these toxin-immunity pairs have independent origins. Toxin-immunity protein pairs are

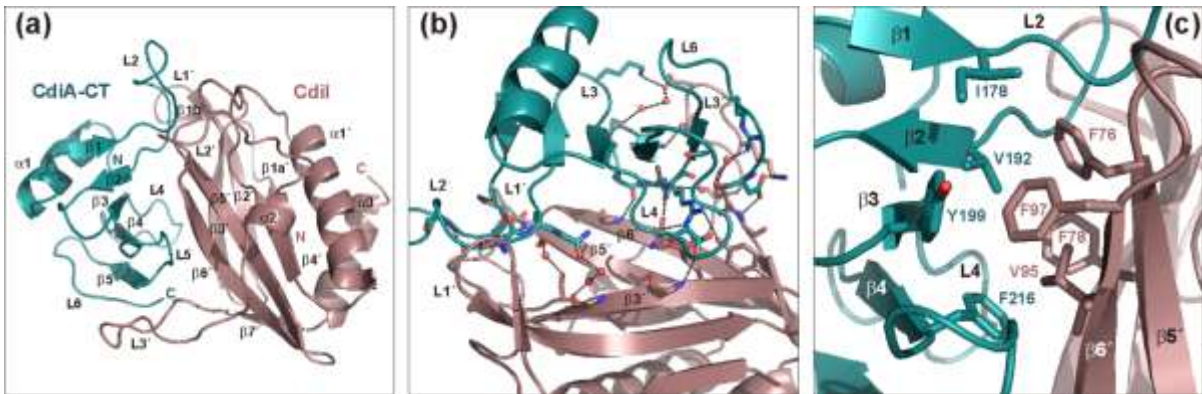
thought to evolve through initial changes in the immunity protein, followed by compensatory changes in the toxin that restore high binding affinity between the two proteins<sup>270,271</sup>. In general, there are fewer functional constraints to impede the mutational drift of immunity genes, which often only need to encode proteins that bind to toxins. By contrast, toxins are often enzymes that must retain the ability to bind substrates and catalyze reactions. These general assumptions are largely supported by analyses of CDI toxin-immunity families, in which immunity proteins typically share less sequence identity than the toxins<sup>167</sup>. Thus, it is formally possible that ImE3 and CdiI<sup>ECL</sup> arose from a common ancestor, but the different folds of these proteins make this possibility less likely. Moreover, in the overwhelming majority of cases, cognate toxin/immunity gene pairs are closely linked and therefore must presumably co-evolve as a unit. Therefore, it seems unlikely that the *cdiI*<sup>ECL</sup> immunity gene evolved from an *imE3* homologue. Based on this reasoning, we speculate that the CdiA-CT<sup>ECL</sup>/CdiI<sup>ECL</sup> protein pair evolved from a lineage that is distinct from ColeE3-CT/ImE3. In this model, the structural and enzymatic similarities between CdiA-CT<sup>ECL</sup> and ColeE3-CT are the result of convergent evolution to target a highly conserved and critical structure in bacterial cells.

CDI systems are modular and related *cdiA-CT/cdiI* gene pairs are often found in different bacterial species. For example, *Enterobacter* sp. R4-368 contains a predicted CDI system that encodes a toxin (Uniprot: R9VTZ7) with 79% sequence identity (over 101 C-terminal residues) to CdiA-CT<sup>ECL</sup> and an immunity protein (Uniprot: R9VPD7) with 61% sequence identity to CdiI<sup>ECL</sup>. Several other enterobacteria (*Yersinia aldovae*, *Photorhabdus luminescens*, *Photorhabdus asymbiotica* and *Xenorhabdus bovienii*) carry "orphan" *cdiA-CT/cdiI* modules that are also clearly related to the *cdiA-CT*<sup>ECL</sup>/*cdiI*<sup>ECL</sup> pair. Orphan modules

represent the displaced 3'-ends of *cdiA* genes together with the linked immunity reading frame. These gene pairs are often found in tandem arrays downstream of *cdiBAI* loci <sup>121</sup>. Additionally, the CdiA-CT<sup>ECL</sup> toxin domain is also found at the C-terminus of two other large proteins that are unrelated to CdiA. *Streptomyces* sp. HGB0020 encodes a YD-repeat protein (Uniprot: S2YVI0) with the toxin domain, and *Amycolatopsis orientalis* HCCB10007 contains an uncharacterized protein of ~110 kDa (Uniprot: R4T474) that also contains a C-terminal domain related to the CdiA-CT<sup>ECL</sup> nuclease. We have recently shown that YD-repeat proteins from both Gram-negative and Gram-positive bacteria deliver protein toxins into other bacteria in a contact-dependent manner <sup>252</sup>, strongly suggesting that the *Streptomyces* YD-repeat protein deploys a 16S rRNase domain to inhibit neighboring cells. Consistent with this model, this YD-repeat gene is immediately followed by an unannotated reading frame that encodes a small protein sharing 36% sequence identity with CdiI<sup>ECL</sup>. Presumably, the 110 kDa protein from *Amycolatopsis* also functions in intercellular competition. The *Amycolatopsis* genomic neighborhood contains genes encoding VgrG and PAAR-domain proteins, which are often associated with antibacterial type VI secretion systems <sup>153,244,252</sup>. Thus, the 110 kDa protein from *Amycolatopsis* may represent a new type of type VI effector protein. Finally, many strains of *Neisseria gonorrhoeae* and *Neisseria meningitidis* carry predicted toxins related to the CdiA-CT<sup>ECL</sup> nuclease domain. The encoding genes are all found associated with *mafABI* loci, which are thought to encode adhesins that bind glycolipids on host mammalian cells <sup>272</sup>. However, *mafB* genes encode variable C-terminal domains that are related to known toxins, strongly suggesting that they represent yet another contact-dependent toxin delivery system. Together, these observations

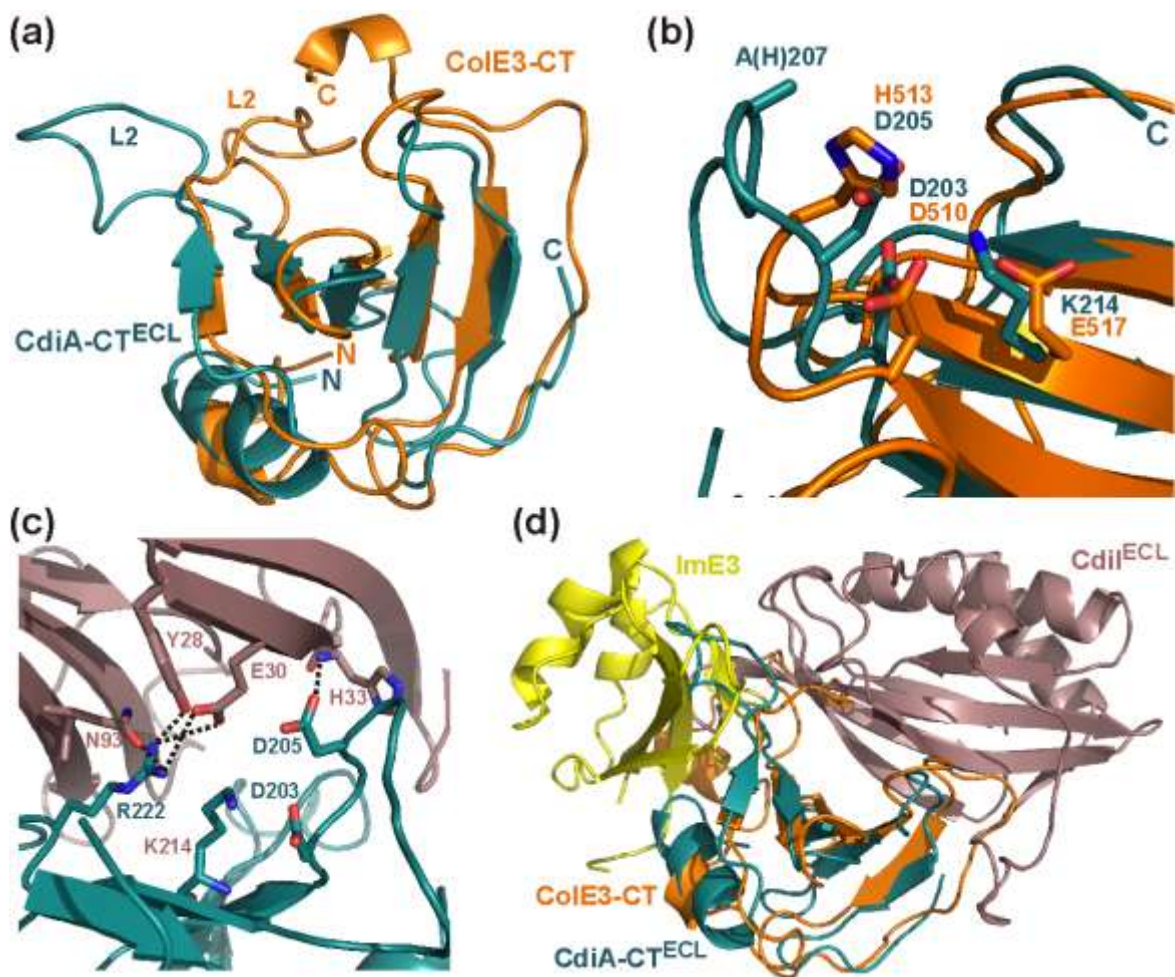
strongly suggest that sequences toxin/immunity gene pairs are horizontally exchanged between bacteria and fused to various toxin-delivery platforms.

**Figure 1. Structure of the CdiA-CT/CdiI<sup>ECL</sup> complex.** **A)** The CdiA-CT<sup>ECL</sup> toxin (teal) and CdiI<sup>ECL</sup> immunity protein (salmon pink) are depicted in ribbon representation with secondary structure elements. The amino and carboxyl termini are indicated by N and C, respectively. CdiI<sup>ECL</sup> elements are denoted with a prime symbol (') to differentiate them from the toxin secondary structure elements. **B)** The CdiA-CT/CdiI<sup>ECL</sup> interface is mediated by an extensive network of ion-pair and hydrogen-bond interactions. Residues are indicated in one-letter code and rendered as stick representations. Water molecules are depicted as red spheres and interacting bonds as dotted lines. **C)** The CdiA-CT/CdiI<sup>ECL</sup> interface also contains hydrophobic interactions mediated by the indicated residues.

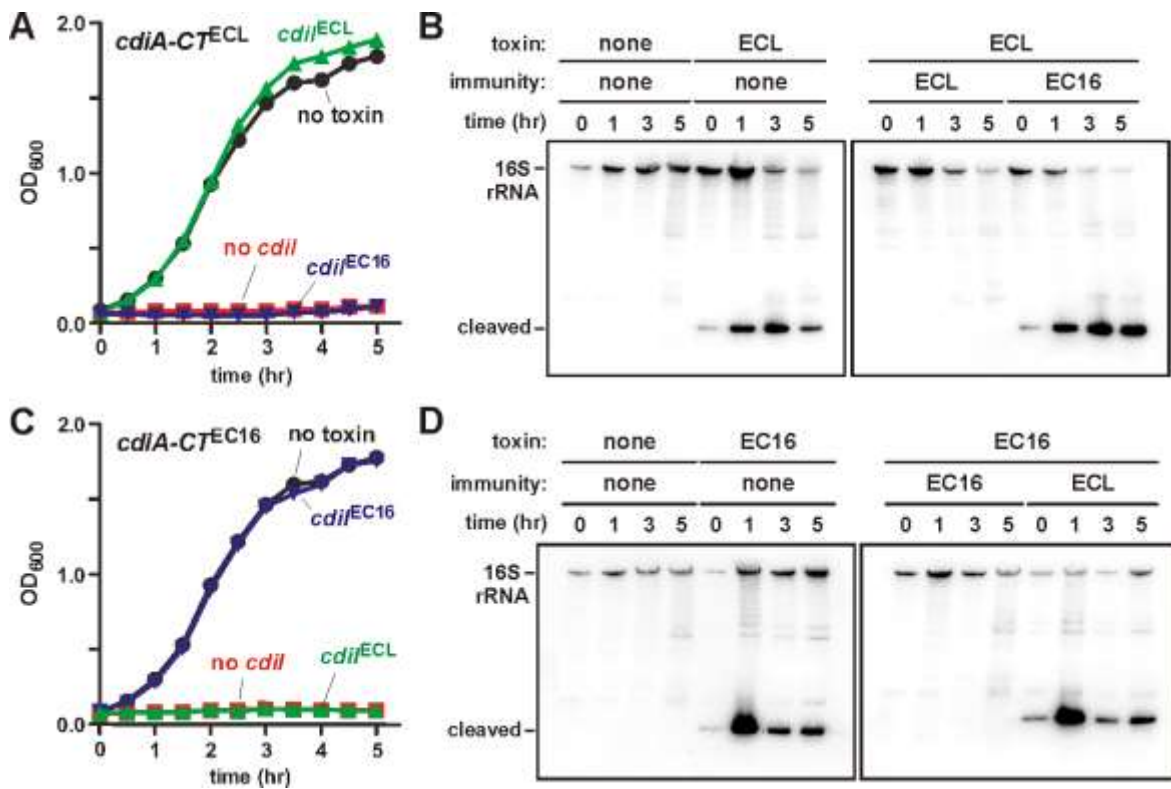




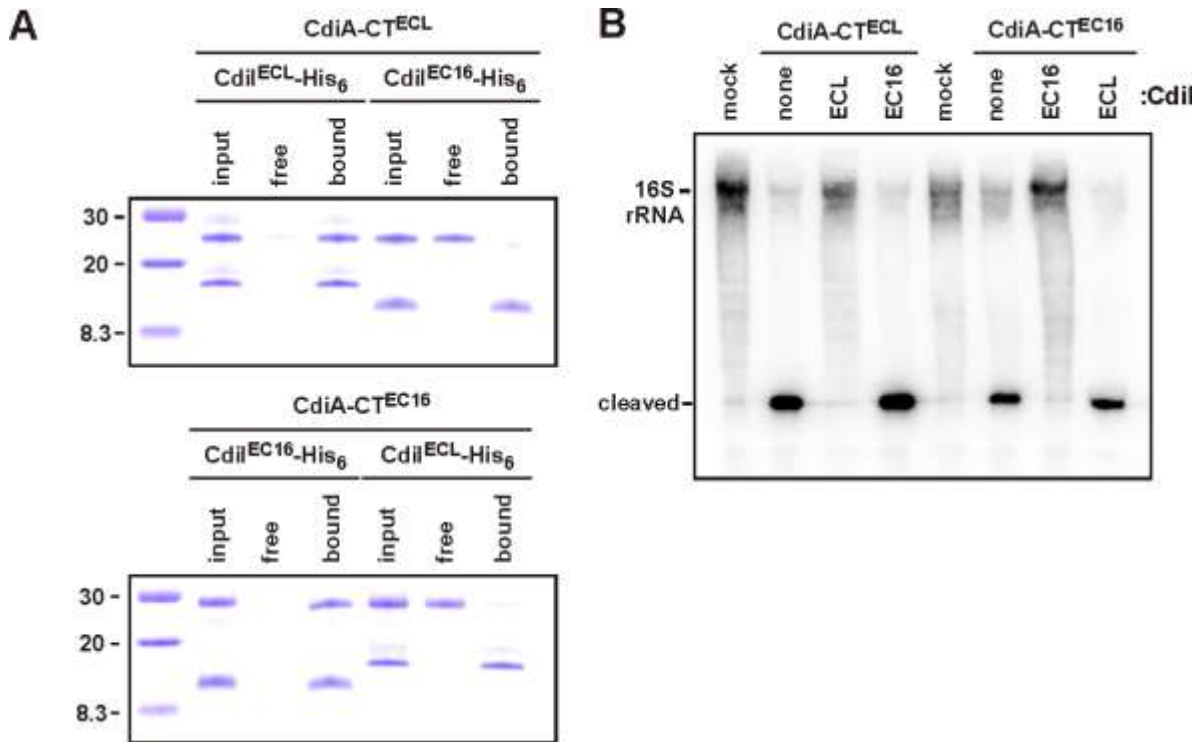
**Figure 2. CdiA-CT<sup>ECL</sup> share structural similarities with the nuclease domain of colicin E3.** **A)** Superimposition of CdiA-CT<sup>ECL</sup> (teal) and the C-terminal nuclease domain of colicin E3 (ColE3-CT, orange) (PDB ID: 2B5U). The toxin domains superimpose with an rmsd of 2.1 Å. **B)** Colicin E3 residues Asp510, His513 and E517 are involved in catalysis and superimpose with residues Asp203, Asp205 and Lys214 of CdiA-CT<sup>ECL</sup>. His207 of CdiA-CT<sup>ECL</sup> is located within disordered loop L4 and is modeled as an alanine residue. Residues are indicated in one-letter code and rendered as stick representations. **C)** The predicted CdiA-CT<sup>ECL</sup> active site is occluded by bound CdiI<sup>ECL</sup>. Interacting bonds are represented by black dotted lines. **D)** Superimposition of CdiA-CT/CdiI<sup>ECL</sup> with the ColE3-CT/ImE3 complex. Ribbon representations of CdiA-CT<sup>ECL</sup> (teal), CdiI<sup>ECL</sup> (salmon pink), ColE3-CT (orange) and ImE3 (yellow) are depicted. Also see Figures S2 and S3.



**Figure 3. CdiA-CT toxin activity *in vivo*.** **A)** *cdiA-CT<sup>ECL</sup>* and *cdiA-CT<sup>EC16</sup>* expression was induced at 0 hr and cell growth monitored by measuring the culture optical density at 600 nm (OD<sub>600</sub>). Red curves are from cells that lack an immunity gene; and the green and blue curves represent cells that co-express *cdiI<sup>ECL</sup>* and *cdiI<sup>EC16</sup>*, respectively. The black curve shows cell growth in the absence of toxin expression. **B)** Total RNA was isolated from the cells in panel A and analyzed by northern blot using a probe to the 3'-end of *E. coli* 16S rRNA. Toxin and immunity genes are indicated as **ECL** for *E. cloacae* and **EC16** for *E. chrysanthemi* EC16. The migration positions of full-length and cleaved 16S rRNA are indicated. Also see Figure S4.

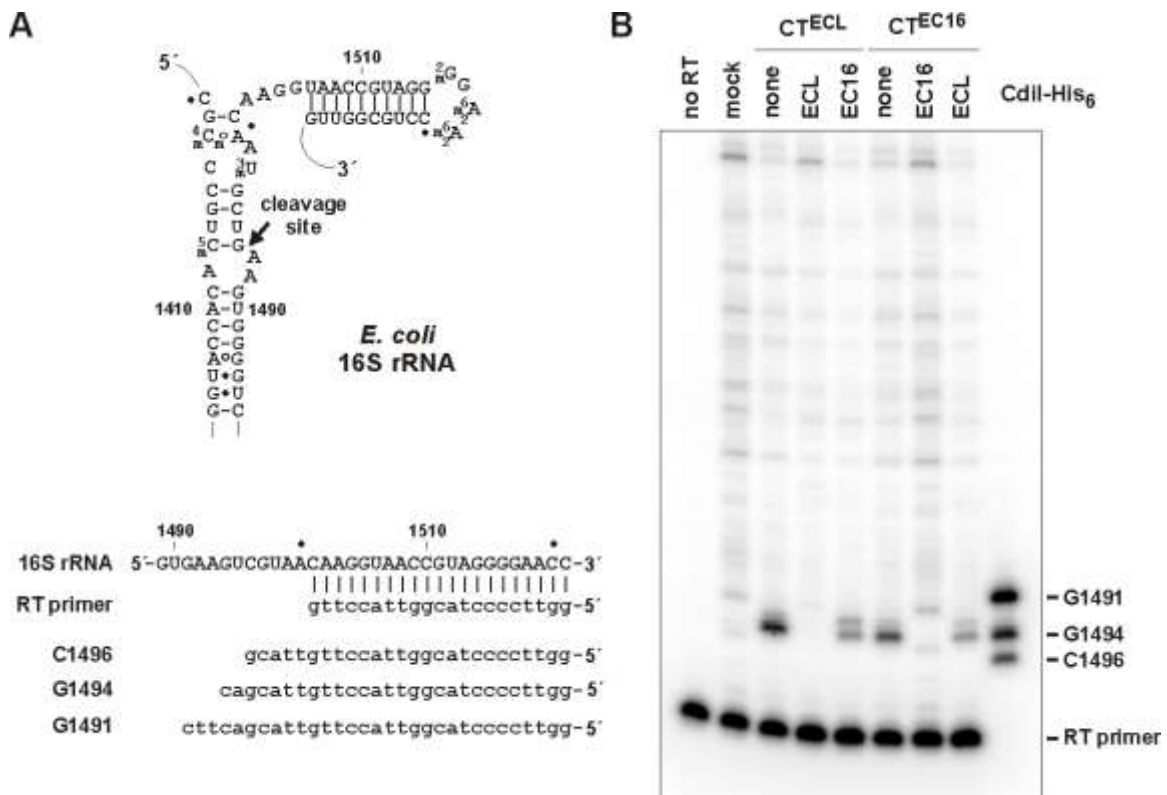


**Figure 4. CdiA-CT toxin activity *in vitro*.** **A)** Purified CdiA-CTs were mixed with equimolar CdiI<sup>ECL</sup>-his6 or CdiI<sup>EC16</sup>-his6 and put over a Ni<sup>2+</sup> column (input). The unbound (free) and bound (bound) fractions were collected and analyzed by SDS-PAGE. **B)** Isolated *E. coli* ribosomes were treated with purified CdiA-CT<sup>ECL</sup> and CdiA-CT<sup>EC16</sup>, and the reactions analyzed by northern blot using a probe to the 3'-end of 16S rRNA. Where indicated, CdiI-His<sub>6</sub><sup>ECL</sup> (**ECL**) or CdiI-His<sub>6</sub><sup>EC16</sup> (**EC16**) was added at an equimolar ratio to the toxin. Mock reactions are ribosome samples that were treated with buffer. The migration positions of full-length and cleaved 16S rRNA are indicated.

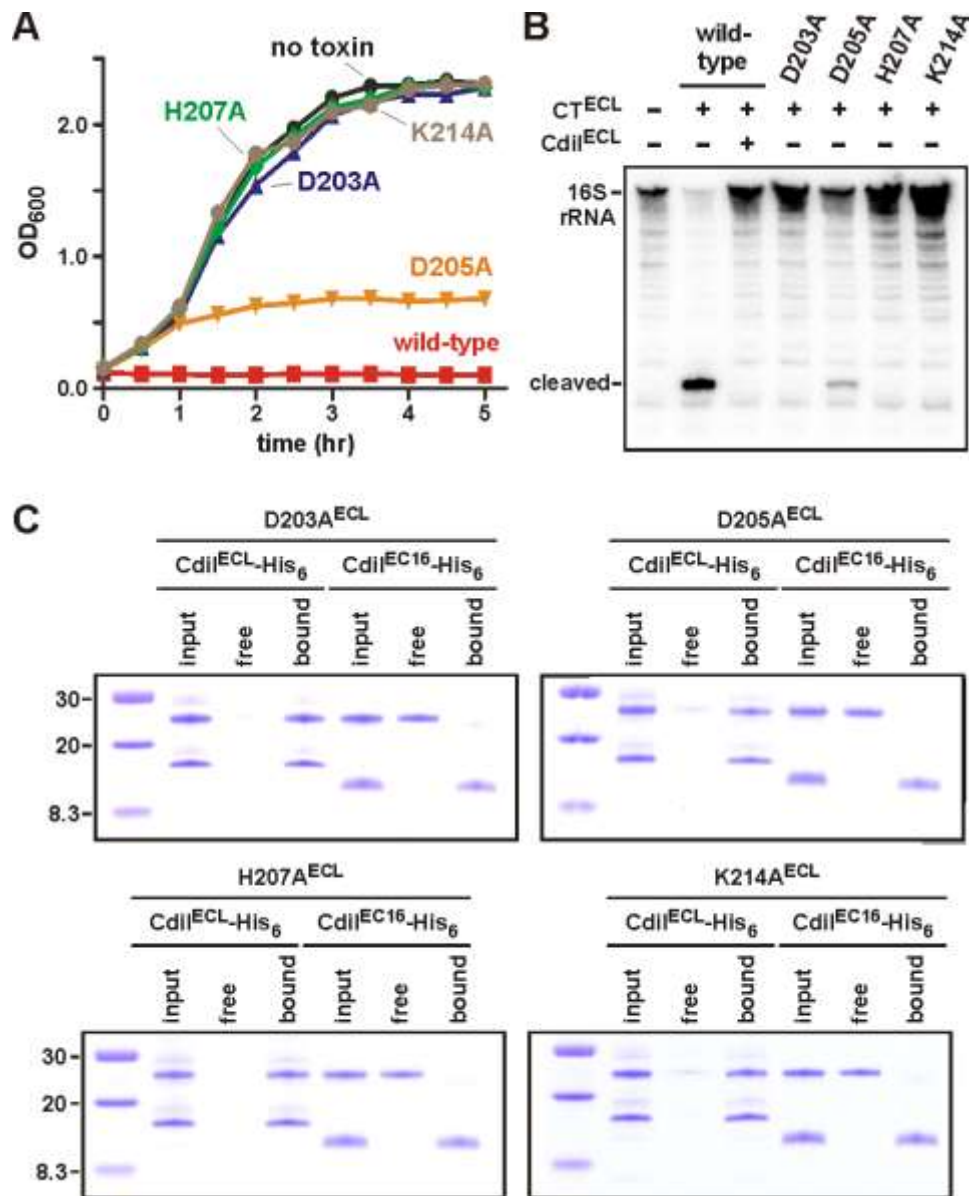


**Figure 5. CdiA-CT<sup>ECL</sup> and CdiA-CT<sup>EC16</sup> cleave 16S rRNA between A1393 and G1394.**

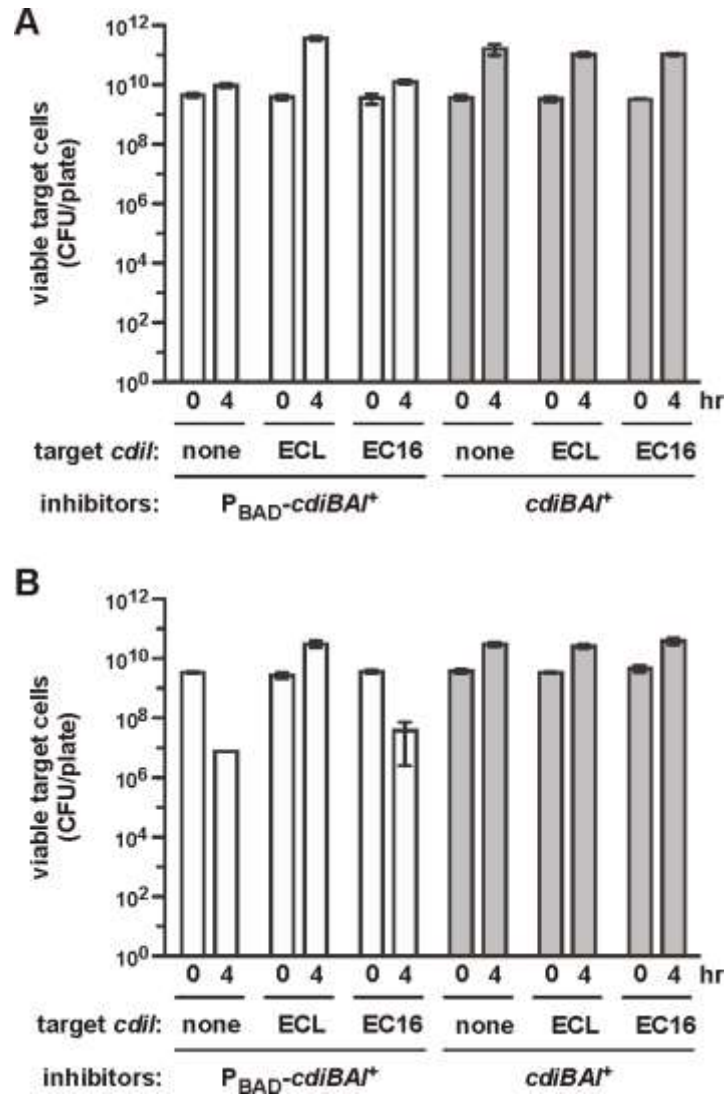
**A)** Nucleotide sequence and secondary structure of the 30S subunit decoding center. The sequence of the reverse transcription (**RT**) primer is shown in heteroduplex with its complementary sequence in 16S rRNA. Oligonucleotides **C1496**, **G1494** and **G1491** were used as gel-migration standards. The 16S rRNA cleavage site is indicated by an arrow. **B)** Ribosomal RNA was extracted from the indicated *in vitro* nuclease reactions and hybridized to radiolabeled **RT primer** for primer extension analysis using reverse transcriptase. Reactions were resolved on a denaturing gel and visualized by phosphorimaging. The migration positions of oligonucleotides **C1496**, **G1494** and **G1491** are indicated.



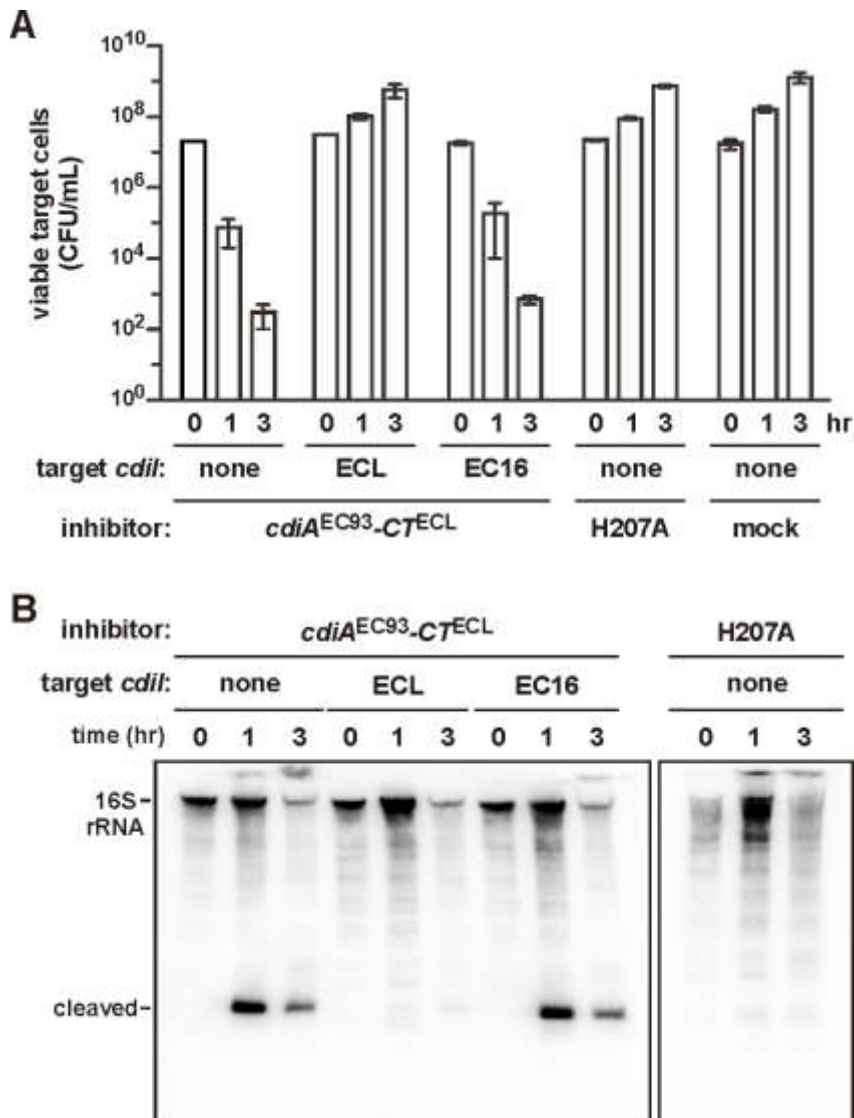
**Figure 6. Mutagenesis of predicted active-site residues in CdiA-CT<sup>ECL</sup>.** **A)** Expression of *cdiA-CT<sup>ECL</sup>* and the indicated mutated variants were induced at 0 hr with L-arabinose and cell growth monitored by measuring the optical density of the culture at 600 nm (OD<sub>600</sub>). The black curve shows the growth of a control culture without toxin expression. **B)** Isolated *E. coli* ribosomes were treated with purified CdiA-CT<sup>ECL</sup> toxins and RNA extracted for northern blot analysis. Where indicated, purified CdiI-His<sub>6</sub><sup>ECL</sup> was included in the reaction. The gel-migration positions of full-length and cleaved 16S rRNA are indicated. **C)** See Figure S5.



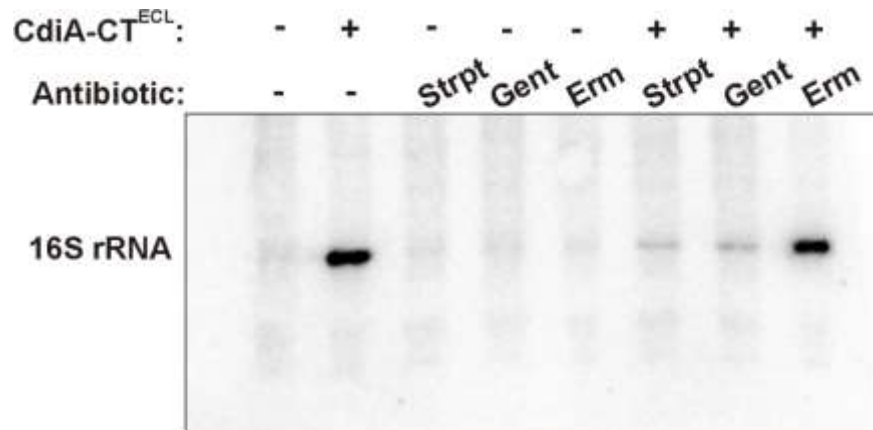
**Figure 7. Intercellular competitions with *E. cloacae* inhibitor cells.** **A)** Intra-species competition. *E. cloacae* inhibitor cells were co-cultured with *E. cloacae*  $\Delta cdiA^{ECL} \Delta cdiI^{ECL}$  target cells on LB-agar supplemented with arabinose. Where indicated, target cells were provided with plasmid-borne *cdiI*<sup>ECL</sup> or *cdiI*<sup>EC16</sup> immunity genes. Total viable target cells were determined as colony forming units. Open bars correspond to competitions with arabinose-inducible inhibitor cells ( $P_{BAD}$ -*cdiBAI*<sup>+</sup>), and grey bars correspond to competitions with inhibitors that carry the wild-type locus (*cdiBAI*<sup>+</sup>). **B)** Inter-species competition. *E. cloacae* inhibitor cells were co-cultured with *E. coli* target cells on LB-agar supplemented with arabinose. Where indicated, the target cells were provided with plasmid-borne *cdiI*<sup>ECL</sup> or *cdiI*<sup>EC16</sup> immunity genes. Viable target cells were determined as colony forming units (CFU), and data are reported as the mean  $\pm$  SEM for two independent experiments.



**Figure 8. The CdiA-CT<sup>ECL</sup> toxin domain is modular.** **A)** *E. coli* target cells were co-cultured with inhibitor cells that express chimeric CdiA<sup>EC93</sup>-CT<sup>ECL</sup> in broth. Where indicated, target cells were provided with plasmid-borne copies of the *cdiA*<sup>ECL</sup> or *cdiA*<sup>EC16</sup> immunity genes. The inhibitors labeled **H207A** express CdiA<sup>EC93</sup>-CT<sup>ECL</sup> containing the His207Ala mutation in the toxin domain. Mock inhibitors lack the CDI system. Viable target cell counts were determined as CFU/mL, and data are reported as the mean  $\pm$  SEM for two independent experiments. **B)** Total RNA was isolated from the co-culture experiments described in panel A and analyzed by northern blot using a probe to the 3'-end of 16S rRNA.

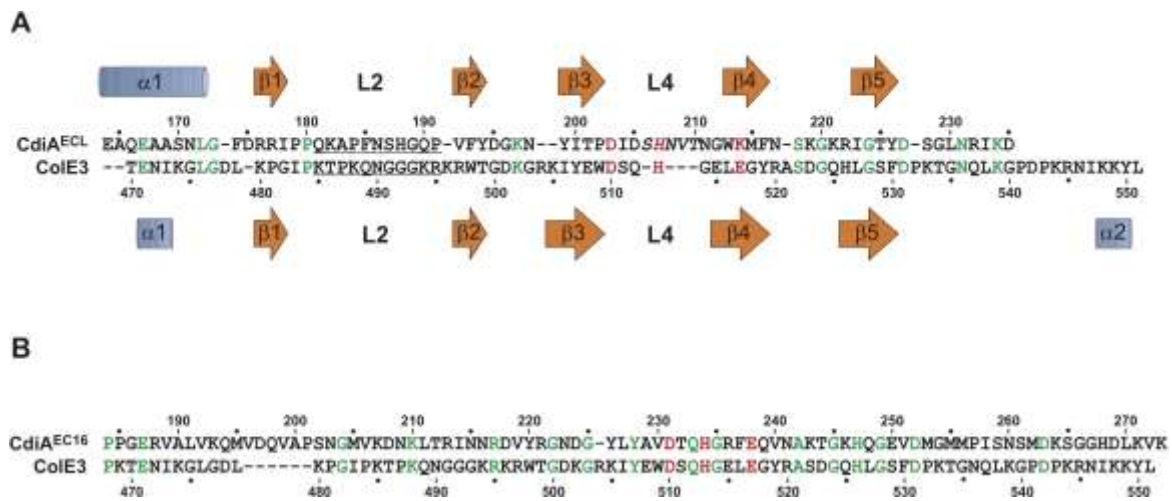


**Figure S1. A-site antibiotics block CdiA-CT<sup>ECL</sup> activity *in vitro*.** 1uM CdiA-CT<sup>ECL</sup> was incubated with 10uM of ribosomes with or without 15uM of Streptomycin (Strpt), Gentamicin (Gent), or Erythromycin (Erm) for 1 hour at 37° C. The reactions were analyzed by Northern blot to the 3'-end of 16S rRNA.

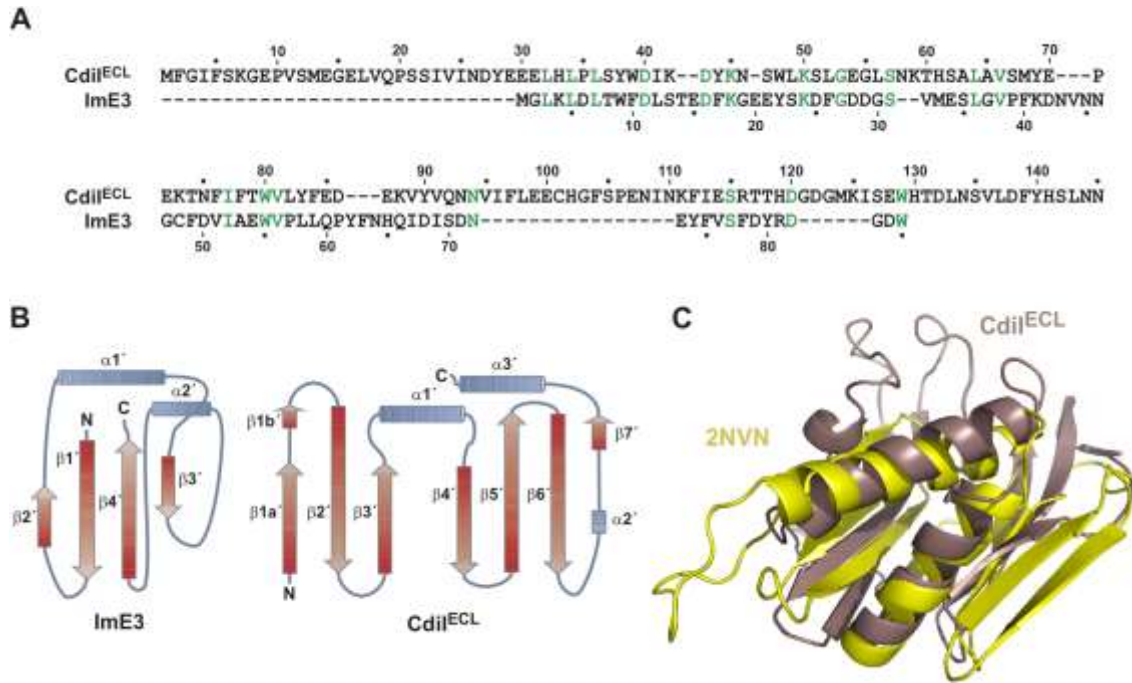




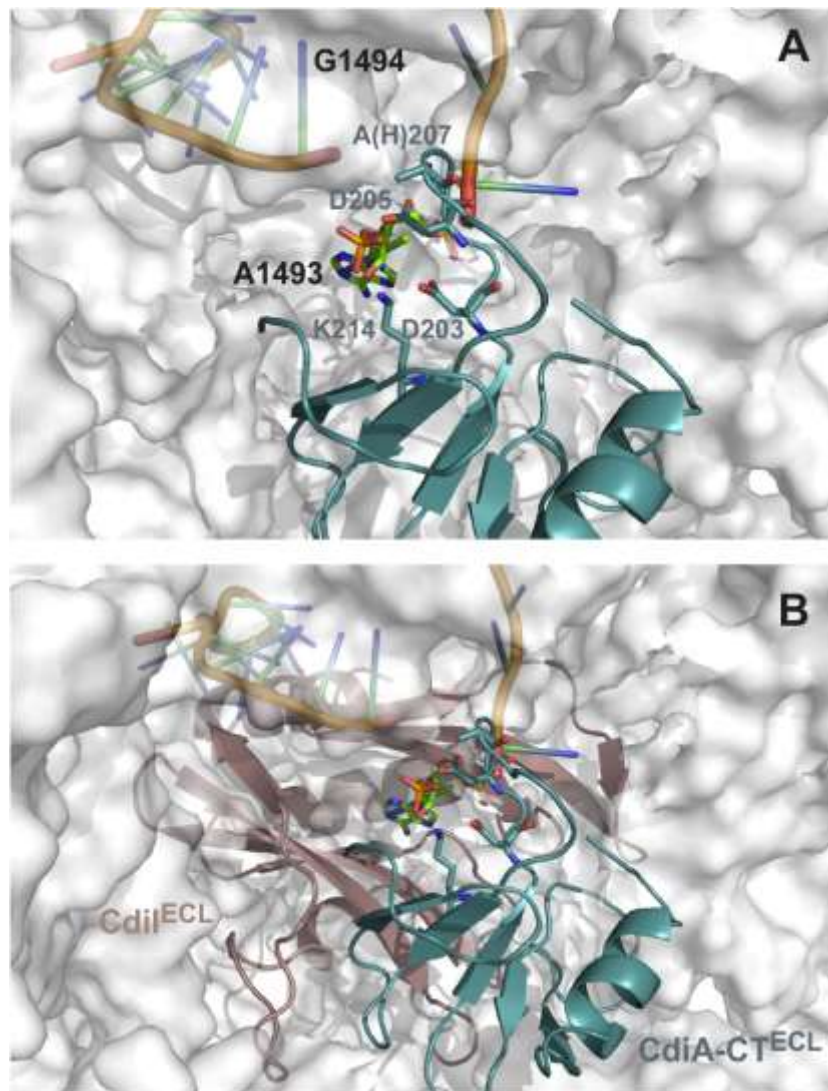
**Figure S2. Alignment of CDI toxins with the nuclease domain of colicin E3.** **A)** The sequences of CdiA-CT<sup>ECL</sup> and CoIE3-CT were aligned based on the superimposed structures of the nuclease domains. Secondary structure elements are indicated as blue  $\alpha$ -helices and orange  $\beta$ -sheets. The underlined residues correspond to loop L2 regions, which adopt different positions in the two domains. Italicized residues in CdiA-CT<sup>ECL</sup> correspond to the unresolved loop L4 region. Identical residues are rendered in green and the predicted active-site residues are shown in red. **B)** Alignment of CdiA-CT<sup>EC16</sup> and CoIE3-CT sequences. The toxin sequences were aligned using Clustal-W. Identical residues are rendered in green and the predicted active-site residues are shown in red.



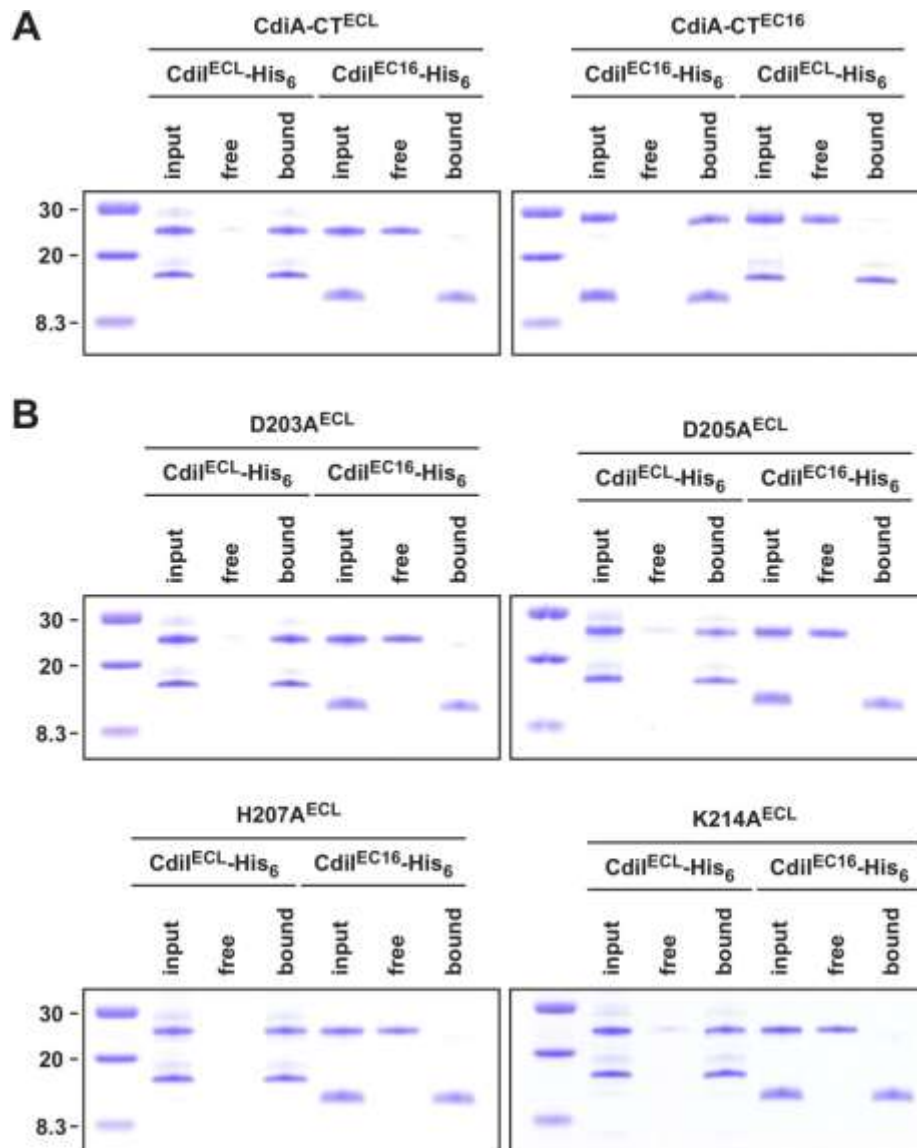
**Figure S3. CdiI<sup>ECL</sup> and ImE3 immunity proteins are unrelated.** **A)** Alignment of CdiI<sup>ECL</sup> and ImE3 immunity proteins sequences using Clustal-W. The proteins share 11.7% identity, with identical residues are rendered in green. **B)** Topologies of CdiI<sup>ECL</sup> and ImE3 immunity proteins. The N- and C-termini are indicated for each immunity protein as are the secondary structure elements. **C)** Superimposition of CdiI<sup>ECL</sup> and Uniprot entry Q31MH7 from *Synechococcus elongatus* PCC 7942 (PDB: 2NVN). The proteins overlay with rmsd 3.6 Å over 121  $\alpha$ -carbons and a Z-score of 4.1.



**Figure S4. Modeling of CdiA-CT<sup>ECL</sup> on the ribosome.** **A)** Superimposition of CdiA-CT<sup>ECL</sup> (teal ribbon) onto the structure of ColE3-CT bound to the ribosome in the post-cleavage state (PDB: 1JCH). The surface of the ribosome is rendered in grey, and the cleaved 16S rRNA strand is shown in ribbon representation. The ColE3-CT domain has been removed for clarity. Putative CdiA-CT<sup>ECL</sup> active-site residues and 16S rRNA residue A1493 are shown in stick representation with nitrogen, oxygen and phosphate atoms colored blue, red and orange, respectively. **B)** CdiI<sup>ECL</sup> (salmon pink) is superimposed upon the model to illustrate steric clashes that should prevent the complex from binding the ribosome A site.



**Figure S5. CdiA-CT/CdiI binding interactions.** A) SDS-PAGE analysis of CDI toxin/immunity protein binding. Isolated CdiA-CT toxins and CdiI-His<sub>6</sub> immunity proteins were mixed and subjected to Ni<sup>2+</sup>-affinity chromatography under non-denaturing conditions. Lanes labeled **input** represent CdiA-CT/CdiI-His<sub>6</sub> mixtures prior to chromatography, lanes labeled **free** are proteins that fail to bind Ni<sup>2+</sup>-resin, and lanes labeled **bound** are samples that were eluted with imidazole. B) CdiA-CT<sup>ECL</sup> toxin variants retain specific binding interactions with CdiI-His<sub>6</sub><sup>ECL</sup>. Mutated toxins were mixed with cognate or non-cognate CdiI-His<sub>6</sub> and processed as described for panel A. The migration positions of molecular weight standards are indicated and masses given in kDa.



**Chapter 5. Exploring the role of CysK in the activation of a contact-dependent growth inhibition (CDI) toxin**

**Note:** I collaborated on this project with Robert Morse and Parker Johnson from the laboratory of Celia Goulding at the University of California, Irvine. The crystal structure of CysK/CdiA-CT(H178A) was completed by Robert Morse and the structure of CysK/CdiA-CT/CdiI was done by Parker Johnson. Other than the two structures, this work is my own.

## Abstract

Contact-dependent growth inhibition (CDI) is a bacterial competition system found throughout proteobacteria. It is comprised of three genes, *cdiBAI*, where CdiB exports the large effector protein CdiA to the surface of CDI<sup>+</sup> cells. The C-terminal toxin domain of CdiA (CdiA-CT) gets delivered into target cells to cause growth inhibition while CdiI binds to CdiA-CT to neutralize its lethal activity. CdiA-CTs are polymorphic and encode for a variety of toxic activities. CdiA-CT from UPEC536 (CdiA-CT<sup>UPEC536</sup>) demonstrates general tRNase activity that requires the protein co-factor, CysK, for activity both *in vitro* and *in vivo*. Here, we show the crystal structure of the CysK/CdiA-CT/CdiI<sup>UPEC536</sup> ternary complex, supporting previous studies that CdiA-CT<sup>UPEC536</sup> binds in the active site of CysK. We identify the active-site residues of CdiA-CT<sup>UPEC536</sup> and demonstrate that a variety of distant *cysK* alleles can still bind to and activate CdiA-CT<sup>UPEC536</sup>. We identify a distant homolog of CdiA-CT<sup>UPEC536</sup> from the Gram-positive species *Ruminococcus lactaris* that does not require CysK and consequently has unique tRNase activity. Lastly, we show this distant homolog can be delivered into Gram-negative target cells via the CDI pathway. Taken all together, our results suggest CysK acts as a chaperone for CdiA-CT and reveal the qualities it contains to allow it to serve as a sufficient co-factor in the CDI pathway.

## Introduction

Contact-dependent growth inhibition (CDI) is a bacterial competition system that requires physical contact for delivery of a toxic effector protein into target cells<sup>112</sup>. CDI systems are widespread throughout proteobacteria and are often found in pathogenicity islands. A three gene locus, *cdiBAI*, mediates competition, where *cdiB/A* encodes for a two-partner secretion system<sup>112,114</sup>. CdiB is an outer membrane protein that exports CdiA onto the surface of cells. CdiA is a 200-600 kDa hemagglutinin repeat protein that is predicted to extend several hundred angstroms from CDI<sup>+</sup> cells<sup>220</sup>. The C-terminal domain of CdiA (CdiA-CT) encodes a toxin<sup>121</sup>. When CdiA comes in contact with a receptor protein on target cells, through an unknown mechanism, CdiA-CT gets delivered into the cytoplasm, resulting in growth inhibition<sup>116</sup>. *cdiI* encodes for an immunity protein that binds to and inactivates CdiA-CT, thus protecting CDI<sup>+</sup> cells from its own toxin<sup>122</sup>.

CdiA is highly conserved amongst different *Escherichia coli* strains until the last 200 amino acids where the toxic domain is demarcated by a conserved VENN peptide motif<sup>121</sup>. Consistent with the polymorphism seen in CdiA-CTs, they have distinct toxic activities<sup>121</sup>. Thus far, CdiA-CTs have been shown to be nucleases or pore-forming proteins<sup>167</sup>. The nucleases target tRNA or DNA to stop protein synthesis and replication, respectively<sup>122</sup>. The pore-formers dissipate proton-motive force upon insertion into the inner membrane<sup>125</sup>. With varying CdiA-CT sequence and activity, it is conceivable that each CdiI protein is polymorphic, as only a cognate immunity protein will inactivate toxin activity<sup>115</sup>. Likewise, *in vitro* experiments demonstrate that immunity proteins bind specifically to their cognate CdiA-CTs and block nuclease activity<sup>115</sup>. Non-cognate immunity proteins do not bind and



therefore will not block activity<sup>115</sup>. Ultimately, CDI systems contain a repertoire of toxins, thus allowing for diverse competition in the environment.

In 2013, Diner *et al.* showed that CdiA-CT from uropathogenic *E. coli* 536 (CdiA-CT<sup>UPEC536</sup>) is a general tRNase that cleaves in the single-stranded anti-codon loop of tRNA<sup>123</sup>. Surprisingly, this toxin requires a protein co-factor, CysK, which is necessary for tRNase activity both *in vitro* and *in vivo*<sup>123</sup>. CysK is a metabolic enzyme involved in the last step of cysteine biosynthesis. CysE takes serine and acetylates it to form O-acetylserine (OAS), and then CysK replaces the acetate with sulfide to form cysteine<sup>273</sup>. In *E. coli*, and closely related species, the C-terminus of CysE binds in the active site of CysK, forming the cysteine synthase complex<sup>274</sup>. CysM is an isomer of CysK that shares 39% sequence homology but does not bind CysE. CysM can use both sulfide and thiosulfate as the sulfur donor while CysK can only use sulfide<sup>273</sup>. It is thought that CysM is expressed under anaerobic conditions, while CysK dominates under aerobic conditions. The accumulation of OAS will result in spontaneous cyclisation to form N-acetylserine (NAS), which does not act as a substrate for CysK<sup>275</sup>. Instead, NAS serves as a co-factor for CysB, which will stimulate expression of *cysK*; thus accumulation of OAS will lead to increased expression of *cysK*<sup>275</sup>.

Diner *et al.* showed that CysK, but not CysM physically binds to CdiA-CT<sup>UPEC536</sup><sup>123</sup>. Furthermore, the absence of CysE does not affect target cell susceptibility to CDI<sup>UPEC536</sup> (data not shown). This suggests the physical binding of CysK to the toxin, and nothing to do with cysteine biosynthesis, is what activates it for tRNase activity. Interestingly, CdiI<sup>UPEC536</sup> can still interact with the toxin when it is bound to CysK, suggesting CdiI<sup>UPEC536</sup> and CysK have discrete binding sites<sup>123</sup>. Furthermore, the C-termini of CysE and CdiA-CT<sup>UPEC536</sup> are similar, the former ending in GDGI, the latter in GYGI. Diner *et al.* demonstrated that residues GYGI

in the toxin are necessary, but not sufficient for binding to CysK (unpublished data).

Additionally, they showed that incubating CysK with its substrate, OAS, could compete with toxin binding while NAS could not<sup>123</sup>. Taken together, this suggests CdiA-CT<sup>UPEC536</sup> binds in a similar manner as CysE, in the active site of CysK.

Here we show the crystal structure of *E. coli* CysK bound to CdiA-CT/CdiI<sup>UPEC536</sup>. The structure demonstrates that CdiA-CT<sup>UPEC536</sup> does in fact bind in the active site of CysK and CdiI<sup>UPEC536</sup> shows no interactions with CysK. We perform mutational analysis to identify the active-site residues required for tRNase activity *in vitro* and *in vivo*. We show that a cohort of CysK alleles from different species can still bind to and activate CdiA-CT<sup>UPEC536</sup>, demonstrating the importance of exploiting a conserved co-factor. Lastly, we look at a distant homolog of CdiA-CT<sup>UPEC536</sup>, which does not require CysK for activity, for some insights into the evolution of the toxin. From here forward, CdiA-CT<sup>UPEC536</sup> will be referred to as CdiA-CT or simply, CT.

## Materials and Methods

<i>Strain or plasmid</i>	<i>Description</i>	<i>Reference</i>
<i>Strains<sup>a</sup></i>		
BL21-Gold(DE3)	<i>E. coli</i> B, F <sup>-</sup> <i>ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>dcm</i> <sup>+</sup> <i>gal</i> 1(DE3) <i>endA</i> Hte, Tet <sup>R</sup>	Stratagene
X90	F <sup>-</sup> <i>lacI<sup>f</sup> lac<sup>'</sup> pro<sup>'</sup>/ara</i> Δ( <i>lac-pro</i> ) <i>nall argE</i> (amb) <i>rif<sup>f</sup> thi-1</i> , Rif <sup>R</sup>	177
EPI100	F <sup>-</sup> <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> ) φ80 <i>dlacZ</i> Δ <i>M15</i> Δ <i>lacXcZ</i> Δ <i>M15</i> Δ <i>lacX recA1 endA1 araD139</i> Δ( <i>ara, leu</i> )7697 <i>galU galK</i> ÷ <sup>-</sup> <i>rpsL nupG</i> , Str <sup>R</sup>	Epicentre
DY378	W3110 <i>lcI857</i> Δ( <i>cro-bio</i> )	247
CH1944	X90 (DE3) Δ <i>rna</i> , Rif <sup>R</sup>	248
CH2016	X90 (DE3) Δ <i>rna</i> Δ <i>slyD::kan</i> , Kan <sup>R</sup>	248

CH10801	JCM158Rif <sup>R</sup> $\Delta$ cysK	This study
CH8602	X90 $\Delta$ cysK	123
<b>Plasmids</b>		
pTrc99a	IPTG-inducible expression plasmid, Amp <sup>R</sup>	GE Healthcare
pCH450	pACYC184 derivative with <i>E. coli</i> araBAD promoter for arabinose-inducible expression, Tet <sup>R</sup>	249
pCP20	Heat-inducible expression of FLP recombinase, Cm <sup>R</sup> Amp <sup>R</sup>	225
pWEB-TNC	Cosmid cloning vector, Amp <sup>R</sup> Cm <sup>R</sup>	Epicentre
pCH10163	Cosmid pCdiA-CT/pheS* that carries a kan-pheS* cassette in place of the <i>E. coli</i> EC93 cdiA-CT/cdiI coding sequence. Used for allelic exchange and counter-selection. Cm <sup>R</sup> Kan <sup>R</sup>	122
pET21P:: <i>cdiA-CT/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	123
pET21P:: <i>cdiA-CT(D155A)/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT(D155A)/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21P:: <i>cdiA-CT(W176A)/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT(W176A)/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21P:: <i>cdiA-CT(H178A)/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT(H178A)/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21P:: <i>cdiA-CT(E181A)/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT(E181A)/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pET21P:: <i>cdiA-CT(T185I)/cdiI</i> <sup>UPEC536</sup>	Overproduces CdiA-CT(T185I)/CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pBAD:: <i>cdiI</i> <sup>UPEC536</sup>	Overproduces CdiI <sup>UPEC536</sup> -His <sub>6</sub> under control of araBAD promoter, Amp <sup>R</sup>	123
pCH450:: <i>cdiA-CT</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	224
pCH450:: <i>cdiA-CT(N149D)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(N149AD)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(K152A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(K152A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(D155A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(D155A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-</i>	Arabinose-inducible expression of <i>cdiA-</i>	This study

<i>CT(W176A)</i> <sup>UPEC536</sup>	<i>CT(W176A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	
pCH450:: <i>cdiA-CT(H178A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(H178A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	224
pCH450:: <i>cdiA-CT(E181A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(E181A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(T185A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(T185A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(T185I)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(T185I)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(R187A)</i> <sup>UPEC536</sup>	Arabinose-inducible expression of <i>cdiA-CT(R187A)</i> <sup>UPEC536</sup> , Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(T185I)</i> <sup>EC93o1</sup>	Arabinose-inducible expression of <i>cdiA-CT(T185I)</i> <sup>EC93o1</sup> , Tet <sup>R</sup>	This study
pET21P:: <i>cysK</i>	Overproduces CysK-His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	123
pET21P:: <i>CT/Imm.</i> <sup>R.lactaris</sup>	Overproduces <i>CT/Imm.</i> <sup>R.lactaris</sup> -His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pCH450:: <i>CT/Imm.</i> <sup>R.lactaris</sup> /DAS	Arabinose-inducible expression of <i>CT/Imm.</i> <sup>R.lactaris</sup> /DAS, Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT/cdiI</i> <sup>UPEC536</sup> /DAS	Arabinose-inducible expression of <i>cdiA-CT/cdiI</i> <sup>UPEC536</sup> /DAS, Tet <sup>R</sup>	123
pCH450:: <i>cdiA-CT(R.lactP139-D143)/cdiI</i> <sup>UPEC536</sup> /DAS	Arabinose-inducible expression of <i>cdiA-CT(R.lactP139-D143)/cdiI</i> <sup>UPEC536</sup> /DAS, Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(R.lactP139-D143,K127-G134)/cdiI</i> <sup>UPEC536</sup> /DAS	Arabinose-inducible expression of <i>cdiA-CT(R.lactP139-D143,K127-G134)/cdiI</i> <sup>UPEC536</sup> /DAS, Tet <sup>R</sup>	This study
pCH450:: <i>cdiA-CT(R.lactN138-L149)/cdiI</i> <sup>UPEC536</sup> /DAS	Arabinose-inducible expression of <i>cdiA-CT(R.lactN138-L149)/cdiI</i> <sup>UPEC536</sup> /DAS, Tet <sup>R</sup>	This study
pTrc99a:: <i>cdiI</i> <sup>UPEC536</sup>	Expresses <i>CdiI</i> <sup>UPEC536</sup>	This study
pTrc99aKX:: <i>imm.</i> <sup>R.lactaris</sup>	Expresses <i>Imm.</i> <sup>R.lactaris</sup>	This study
pDAL660Δ1-39	Expresses <i>CDI</i> <sup>EC93</sup> , Amp <sup>R</sup>	112
pDAL866	Expresses <i>CDI</i> <sup>UPEC536</sup> , Amp <sup>R</sup> Cm <sup>R</sup>	116
CH12389	Constitutive expression of chimeric <i>cdiA</i> <sup>EC93</sup> - <i>NT</i> <sup>UPEC536</sup> <i>CT</i> <sup>R.lactaris</sup> and <i>imm.</i> <sup>R.lactaris</sup> genes, Cm <sup>R</sup>	This study
CH12390	Constitutive expression of chimeric <i>cdiA</i> <sup>EC93</sup> - <i>NT</i> <sup>EC3006</sup> <i>CT</i> <sup>R.lactaris</sup> and <i>imm.</i> <sup>R.lactaris</sup> genes, Cm <sup>R</sup>	This study
pTrc99a:: <i>cysK-his6</i>	Constitutive expression of CysK-his6, Amp <sup>R</sup>	
pTrc99a:: <i>cysK</i> <sup>D.dada</sup>	Constitutive expression of CysK-his6,	This study

<i>niii</i> -his6	Amp <sup>R</sup>	
pTrc99a::cysK <sup>E.cloac</sup> <i>ae</i> -his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK <sup>B.subtili</sup> <i>s</i> -his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK <sup>H.influe</sup> <i>nza</i> -his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK <sup>N.lacta</sup> <i>mica</i> -his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(K42 A)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(Q14 3,F144A)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(R10 1E,K102E)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(K10 5E)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(K22 1E)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(K22 6A)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(Q22 8E)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK(K28 3E)-his6	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study
pTrc99a::cysK( $\Delta$ 30 7-324)	Constitutive expression of CysK-his6, Amp <sup>R</sup>	This study

### ***Bacterial strains and growth conditions***

All bacterial strains and plasmids used in this study are listed the table above.

Bacteria were grown in LB media or LB-agar with antibiotics at the following concentrations: ampicillin (Amp) 150  $\mu$ g/mL; kanamycin (Kan) 50  $\mu$ g/mL; rifampicin (Rif) 200  $\mu$ g/mL; spectinomycin (Spc) 100  $\mu$ g/mL; chloramphenicol (Cm) 100  $\mu$ g/mL and tetracycline (Tet) 10  $\mu$ g/mL. The  $\Delta$ cysK::kan mutation was transferred from the Keio collection into *E. coli* strain JCM158 by bacteriophage P1-mediated transduction<sup>226</sup>. The Kan<sup>R</sup> cassette was removed via pCP20<sup>225</sup>.

### *Plasmid constructs*

Protein overproduction constructs were generated using plasmid pET21S<sup>115</sup>. pET21P::*cdiA-CT/cdiI* constructs were made by amplifying *cdiA-CT/cdiI* with primers **536-CT-Nco-for** (5' - AGA CCA TGG TTG AGA ATA ATG CGC TGA G) and **536-cdiI-Xho-rev** (5' - GAT CTC GAG TAC AAT TAT CTG ATT GAT TTT T), digested with NcoI/XhoI, and ligated into digested pET21P. Point mutations were introduced into *cdiA-CT* using megaprimer PCR<sup>187</sup>. Megaprimers were first generated using mutagenic reverse primers: **536-CT(D155A)-rev** (5' - GAG AGT TCC AAT AAT AGC ATG ATC TTT CAG - 3'), **536-CT(W176A)-rev** (5' - CCT GCA TAT GAT CCG CAT ATC CTC CAT TC - 3'), **536-CT(E181A)-rev** (5' - GCG TAT TTT GCA TTG CCT GCA TAT GAT CC - 3'), **536-CT(T185I)-rev** (5' - TTC TTA ATC CTC TGA GGA TAT TTT GCA TTT CCT GC - 3') in conjunction with **536-cdiI-Xho-rev**. These products were used in subsequent reactions with **536-CT-Nco-for** to generate complete *cdiA-CT/cdiI* modules, which were ligated into pET21P. The *CT/immI*<sup>R.lactaris</sup> region was generated by IDT, digested with NcoI/SpeI and ligated into pET21P and p6520<sup>123</sup>.

*cdiA-CT(R.lactP139-D143)/cdiI/DAS* was made with megaprimer 5' - CAT GCG GAT ACG TTG AAA AAC CCC AAC CTC TCA GAC GTC AAC AAT CCT GAA GCT C - 3' *cdiA-CT(R.lactP139-D143,K127-G134)/cdiI/DAS* with 5' - CGC TCA GAG GAT TAA GAA AGA TCA AAA AGG GAT TAG AGG GGA CGT TGA AAA ACC CCA AC - 3' off of the P139-D143 construct. *cdiA-CT(R.lactN138-L149)/cdiI/DAS* was generated via OE-PCR using oligos 5' - AAT CCA AAC CTC TCT GAT GTA GAT AGA GCA TTA TTG CAG GCT GCG TAT GGC - 3' and 5' - CTA CAT CAG AGA GGT TTG GAT TTT TTA GTG AAC CTT CCA ACG TAT CCG CAT G - 3'. The *immI*<sup>R.lactaris</sup> gene was amplified with

primers **R.lact-imm-Kpn-for** (5' - TTT GGT ACC ATG CAA GAC AAG AGA AAA ATA AAA G - 3') and **R.lact-imm-Xho-rev** (5' - TTT CTC GAG ATC ACA TTA TTT TTT TGG ATA AAG TAT CTA TC - 3'), digested with KpnI/XhoI and ligated into pTrc99aKX.

Arabinose-inducible *cdiA-CT* expression constructs were generated using plasmid pCH450<sup>249</sup>. The *cdiA-CT* coding sequence was amplified using **536-CT-Nco-for** with **536-CT-Xho-rev** (5' - TTA CTC GAG GTA ATC ATA TTC CAT A), digested with NcoI/XhoI and ligated into pCH450. D155A, W176A, E181A, and T185I were amplified from the pET constructs while **N149D**, **K152A**, **T185A**, and **R187A** were generated using the megaprimers 5' - GAT AAC ACT ATA AAA GAT GCT CTG AAA GAT C - 3', 5' - CTA TAA AAA ATG CTC TGG CAG ATC ATG ATA TTA T - 3', 5' - CTT AAT CCT CTG AGC GCA TTT TGC ATT TCC TGC - 3' and 5' - ATG CAA AAT ACG CTC GCA GGA TTA AGA AAT C - 3', respectively.

pTrc99a::*cysK-his6* constructs were made by amplifying gDNA from the various species, digested with NcoI/SpeI (XhoI for *N. lactamica*) and ligated into pTrc99a. Primer pairs 5' - CGG GCC ATG GGT AAG ATC TAC GAA GAC/5' - GGA ACT AGT CTG CTG CAG TTC CTG TTC GG for *D. dadantii*, 5' - GGC CAT GGG TAA GAT TTA TGA AGA CAA C/5' - GCT GAC TAG TCT GTT GCA GTT CTT TCT CGG for *E. cloacae*, 5' - GTC GAC CAT GGT ACG TGT AGC AAA CTC C/5' - GCT TAC TAG TAT CGA ATT GGT ACA GCG GCG for *B. subtilis*, 5' - ATA CCA TGG CAA TTT ATG CAG AC/5' - TTT ACT AGT TCC CTC AAT CCC TTC AAA C for *H. influenza*, and 5' - AGA CCA TGG AAA TTG CAA ACA GCA TCA CC / 5' - AAA CTC GAG CGC CAA ATC GGC AAA CAG GGG CG –for *N. lactamica* were used for PCR. Each mutant CysK was generated using megaprimer PCR with oligos **K42A** 5' - CCA GCT TCA GCG TTG CGT GCC GTA TCG

GTG CC, Q143A F144A 5' - GCT GCT GCA AGA AGC CAG CAA TCC GGC AAA CC, **R101E K102E** 5' - CGC TTT CAG CAG CTC TTC GCG TTC AAT ACT CAT G , **K105E** 5' - TTG CAC CTA ACG CTT CCA GCA GCT TGC GGC G, **K221E** 5' - GCA GGT GAA GAG ATT GAA CCT GGC CCG CAT A, **K226A** 5' - AAA CCT GGC CCG CAT GCA ATT CAG GGT ATT GG, **Q228E** 5' - CCG CAT AAA ATT GAG GGT ATT GGC GCT G, **K283E** 5' - GTT GCC GCG GCG TTG GAA CTA CAA GAA GAT G. CysK  $\Delta$ 307-324 was generated with **CysK-Nco-for** and 5' - AAA CTC GAG TTA ATA ACG CTC ACC CGA TGA - 3', digested with Nco/Xho, and ligated into pTrc99a.

The EC93-536-R.lactaris and EC93-3006-R.lactaris chimeric CDI systems were constructed by allelic exchange of the counter-selectable *pheS\** marker from cosmid pCH10163 as described<sup>122</sup>. Primer pairs 5' - GAG CAT CAG TTA ATG TAT TAT CTA CTT TTT CAT TTT GCT GCT TAG GAT C/CDI204 for NT<sup>UPEC536</sup> or 5' - GAG CAT CAG TTA ATG TAT TAT CTA CTT TCA GAA CTT CTA TCT TAC TGG CC /CDI204 for NT<sup>EC3006</sup>, 5' - AAA GTA GAT AAT ACA TTA ACT GAT GCT C/ 5' - GGT CTG GTG TCT AAC CTT TGG GAT CAC ATT ATT TTT TTG GAT AAA GTA TCT ATC and DL1663/ CDI205 were used to generate the 3 PCR products. CDI205 and 205 were used in OE-PCR to combine the 3 PCR products. The final DNA product (100 ng) was electroporated together with pCH10163 (300 ng) into *E. coli* strain DY378 cells<sup>247</sup> and recombinants selected on yeast extract glucose-agar supplemented with 33  $\mu$ g/mL chloramphenicol and 10 mM D/L-*p*-chlorophenylalanine.

### ***Protein purification and crystallography***



The CdiA-CT/CdiI-His<sub>6</sub> complex was overproduced from plasmid pET21P::*cdiA-CT/cdiI* in either *E. coli* BL21 (DE3) Gold (Stratagene) or *E. coli* CH2016 cells induced with 1 mM isopropyl b-D-thiogalactopyranoside (IPTG). Selenomethionine (SeMet) labeled complex was produced from cells grown in M9 minimal medium supplemented with leucine, isoleucine and valine at 50 mg/L; phenylalanine, lysine and threonine at 100 mg/L; and SeMet at 75 mg/L) as previously described<sup>255</sup>. Cells were washed with resuspension buffer [20 mM sodium phosphate (pH 7.0), 200 mM NaCl], then broken by sonication on ice in resuspension buffer supplemented with 10 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride. Unbroken cells and debris were removed by centrifugation at 18,000 *g* for 30 min followed by filtration through a 0.45 µm filter. Clarified lysates were loaded onto a Ni<sup>2+</sup> charged HiTrap column (GE Healthcare) and washed with resuspension buffer containing 10 mM imidazole. The CdiA-CT/CdiI-His<sub>6</sub> complex was eluted with a linear gradient of imidazole (10 – 500 mM) in resuspension buffer. Fractions were collected, combined, and concentrated to ~500 µL in a centrifugal concentrator. Complexes were further purified by gel filtration on a Superdex 200 column equilibrated with 20 mM Tris-HCl (pH 7.4), 150 mM NaCl using an AKTA FPLC. SeMet-labeled complex was concentrated to 9 mg/mL in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl for crystallization trials.

The CdiA-CT(H178A)/CysK binary complex crystallized in space group P4<sub>1</sub> with two CdiA-CT/CysK complexes per asymmetric unit. The initial phases were determined by molecular replacement by molecular replacement using the structure of CysK as a search model (pdb ID 1OAS), and the CdiA-CT<sup>UPEC536</sup> molecules were manually built. The final model contains CdiA-CT<sup>UPEC536</sup> and CysK residues 127-227 and 2-314, respectively, and 110

water molecules resulting in an  $R_{\text{work}}/R_{\text{free}}$  (%) of 20.0/22.4. Additionally, both CysK molecules in the asymmetric unit contain a pyridoxal 5'-phosphate (PLP) bound in a Schiff base linkage to Lys42.

The CdiA-CT/CdiI/CysK ternary complex was solved to 3.0 Å resolution by MR using the UPEC536 CdiA-CT/CysK binary complex as search model, and the structure of the CdiI protein was manually built. The proteins crystallized with four complexes per asymmetric unit in space group  $P2_1$ . The final complex model contains 531 residues per complex and includes Leu133-Ile228 (91 residues) of CdiA-CT, Ile2-Val128 (127 residues) of CdiI, and Gly2-Asp314 (313 residues) of CysK, with 110 water molecules (per asymmetric unit) as well as a PLP molecule linked to Lys42 in the active site of each CysK monomer, resulting in a final  $R_{\text{work}}/R_{\text{free}}$  of 19.65/24.5.

### ***Western blot analysis***

Cells were grown to mid-log in 2mL liquid LB-Amp with aeration and harvested for protein extraction with urea lysis buffer. Equal amounts of protein were subjected to 10% acrylamide SDS-PAGE for 1 hr at 100V. The gel was transferred onto nitrocellulose membrane, and incubated with 1:25,000 of anti-his6 antibody, a kind gift from David Low.

### ***In vitro binding and RNase assays***

CdiA-CT/CdiI-His<sub>6</sub> complexes were denatured in binding buffer [20 mM sodium phosphate pH 7.0, 150 mM NaCl] supplemented with 6 M guanidine-HCl, and CdiA-CT proteins isolated in the void volume during Ni<sup>2+</sup>-affinity chromatography<sup>115</sup>. Toxins were refolded by dialysis against binding buffer. CdiI-His<sub>6</sub> and CysK-His<sub>6</sub> proteins were purified

by Ni<sup>2+</sup>-affinity chromatography under non-denaturing conditions as described<sup>120</sup>. Imm.<sup>R.lactaris</sup>-His<sub>6</sub> was refolded after the denaturing prep. The isolated toxins and immunity proteins were dialyzed against binding buffer and quantified by absorbance at 280 nm. Purified CdiA-CT, CysK-His<sub>6</sub> and CdiI-His<sub>6</sub> proteins were mixed at 10 μM final concentration in binding buffer and binding interactions assessed by co-purification during Ni<sup>2+</sup>-affinity chromatography as described<sup>115,120</sup>. *In vitro* nuclease activity assays were conducted in reaction buffer [20mM Tris pH7.5, 150mM NaCl, MgCl<sub>2</sub>]. S30 lysates of *E. coli* strain CH1944 were made by centrifugation at 30,000 x g for 30 min at 4 °C. X90 total RNA was made by guanidine isothiocyanate-phenol extraction as previously described<sup>121</sup> and reactions were carried out for 1 hr at 37 °C. Reactions were analyzed by denaturing electrophoresis on 50% urea – 6% polyacrylamide gels in 15 Tris-borate-EDTA buffer. Gels were transferred to nylon membrane and hybridized to 5'-radiolabeled oligonucleotides to hybridize various tRNAs.

### ***In vivo toxicity assays***

CdiA-CT toxicity and *in vivo* nuclease activity was assessed in *E. coli* strain X90. Arabinose-inducible pCH450::*cdiA-CT/cdiI/DAS* expression constructs were introduced together with IPTG-inducible pTrc::*cdiI* plasmids and transformants selected on LB-agar supplemented with Tet, Amp, 0.4% D-glucose and 1 mM IPTG. The resulting strains were grown to mid-log phase in LB media with Tet, Amp, 0.4% D-glucose and 1 mM IPTG, then cells collected by centrifugation and resuspended in fresh LB supplemented with Tet, Amp, 1 mM IPTG and 0.2% L-arabinose. Cell growth was monitored by measuring the optical density at 600 nm (OD<sub>600</sub>) as a function of time. Culture samples were harvested into an

equal volume of ice-cold methanol, cells were collected by centrifugation and frozen at  $-80^{\circ}\text{C}$ . RNA was isolated from the frozen cell pellets using guanidine isothiocyanate-phenol and 10  $\mu\text{g}$  of total RNA used for northern blot analysis as described <sup>248</sup>.

X90 or X90  $\Delta\text{cysK}$  cells were made TSS-competent, 100 ng of pDNA was transformed and cells were recovered with 0.4% glucose for 1 hr at  $37^{\circ}\text{C}$ . Cells were plated on LB-agar plates supplemented with Tet, Amp, 0.4% D-glucose or L-arabinose.

### ***Growth competitions***

*E. coli* EPI100 inhibitor cells carrying the indicated CDI cosmids were co-cultured with the designated target cells at a 1:1 ratio in LB broth, shaking at 215rpm at  $37^{\circ}\text{C}$  for 3 hrs. Cells were harvested at time 0 and 3 hours into M9 salts for serial dilution and colony forming units/mL (CFU/mL) enumeration on selective media. CysK complementation and immunity function was evaluated through expression of *cysK* or *cdiI* genes in targets from plasmid pTrc99a.

## **Results**

### **Characterization of CdiA-CT**

#### ***Structure of the CysK/CdiA-CT/CdiI complex***

CdiA-CT/CdiI-his6 and CysK-his6 were purified separately under native conditions and combined for co-crystallization. The ternary complex was solved to  $3.0 \text{ \AA}$  resolution with four complexes per asymmetric unit in space group  $P2_1$ . The final complex model contains 531 residues per complex and includes Leu133-Ile228 (91 residues) of CdiA-CT, Ile2-Val128 (127 residues) of CdiI, and Gly2-Asp314 (313 residues) of CysK, with 110 water

molecules (per asymmetric unit) as well as a PLP molecule linked to Lys42 in the active site of each CysK monomer, resulting in a final R<sub>work</sub>/R<sub>free</sub> of 19.65/24.5. The cysteine synthase complex is made up of CysE as a dimer of homotrimers bound to a maximum of two molecules of CysK in the form of a dimer<sup>276</sup>. The CysK/CdiA-CT/CdiI complex shows CysK as a dimer (Fig. 1A), however it is unclear what ratio the CysK/CdiA-CT/CdiI complex is *in vivo*.

The structure confirms previous analysis that the extreme C-terminus of CdiA-CT binds in the active site of CysK (Fig. 1B) and that CdiI binds to a discrete position on CdiA-CT (Fig. 1A). The fold of CdiA-CT does not provide any obvious clefts for tRNA docking (Fig. 2A). Furthermore, an electrostatic map does not reveal any dramatic pockets of positive charge that would facilitate interactions with the phosphate backbone of tRNA (Fig. 2B). We have previously reported that an H178A mutation renders the toxin inactive, but it can still bind to CdiI and CysK<sup>123</sup>. However, the location of this residue on the structure still does not give significant insight into identifying the active site (Fig. 2B & 2C). Of course, the presence of CdiI could be changing the conformation of CdiA-CT to an inactive state. To rule this out, CdiA-CT(H178A) was co-crystallized with CysK to look for conformational changes between the CdiI-bound and unbound CT. The protein complex crystallized in space group P4<sub>1</sub> with two CdiA-CT/CysK complexes per asymmetric unit. The final model contains CdiA-CT and CysK residues 127-227 and 2-314, respectively, and 110 water molecules resulting in an R<sub>work</sub>/R<sub>free</sub> (%) of 20.0/22.4. Additionally, both CysK molecules in the asymmetric unit contain a pyridoxal 5'-phosphate (PLP) bound in a Schiff base linkage to Lys42. An overlay of CdiA-CT/CysK with CdiA-CT(H178A)/CysK shows CdiA-CT has no significant change in conformation or fold in the presence or absence of CdiI (Fig. 2A).

While CdiA-CT(H178A) is technically not the active form of the toxin, it is doubtful that this single-residue change, especially given its location on CdiA-CT, could drastically change the conformation of the CT. Taken all together, the toxin conformation remains the same with or without immunity bound. Unfortunately, this still leaves us perplexed as to where the active site of the toxin is located.

### ***Identifying active-site residues in CdiA-CT***

To help reveal some of the residues involved in tRNA binding or cleavage, and therefore the location of the active site, we looked at an alignment of CdiA-CT homologs. We reasoned that conserved residues would be involved with enzymatic activity, while non-conserved residues have drifted because their function is to maintain the fold of the toxin. An alignment of 7 homologs to CdiA-CT reveals numerous conserved residues, and therefore does not disclose information about the active site (Fig. 3A). We then looked at the mechanism of how other proteins bind to tRNA, especially to the anti-codon loop. It has been demonstrated that some tRNA synthetases stabilize the interaction with their cognate tRNA by flipping out a base in the single-stranded sections of the molecule (3' acceptor stem or anti-codon loop) and stacking it with an aromatic residue, thus allowing for pi-stacking<sup>277</sup>. Closer analysis of the alignment performed reveals a conserved aromatic residue at position 176, which maps to close proximity with H178A on the structure (Fig. 2A & 2B).

Many endonucleases coordinate divalent cations to help stabilize negative transition states<sup>278</sup>. We reasoned that some negatively charged residues should be required for CdiA-CT tRNase activity since its activity *in vitro* is dependent on a divalent cation (data not shown). Furthermore, we predicted some positively charged residues might be involved in

electrostatic interactions to help bind negatively charged tRNA. Lastly, we considered residues capable of hydrogen bonding as this interaction would also assist in binding to tRNA. This led us to mutate the following residues to alanine and test their ability to cleave tRNA. W176 was chosen based on its possible role in stacking a base from the anti-codon loop of bound tRNA. D155 and E181 were selected for possible cation coordination. N149, K152, and R187 were chosen for their potential to bind tRNA and T185 was chosen for its ability to form hydrogen bonds. Furthermore, each of these residues is highly conserved (Fig. 3A) and map to a location on the toxin that would allow them access to a tRNA substrate (Fig. 3B).

#### ***In vivo analysis of active-site residues***

Each mutant was cloned into plasmid pCH450 which contains an araC promoter that is induced with arabinose and suppressed by glucose. Equivalent amounts of plasmid DNA were transformed into *E. coli* or *E. coli cysK* cells and plated on LB-glucose, LB, or LB-arabinose plates for increasing levels of induction. Firstly, WT toxin is able to kill cells even when suppressed with glucose, but unable to kill *cysK* cells under any condition, as expected (Fig. 4). Mutants D155A, H178A, and E181A were unable to kill cells under any condition while W176A showed partial toxicity to WT cells (Fig. 4). Interestingly, T185A showed WT levels of killing, but T185I has a decrease in activity, suggesting the bulky side chain of isoleucine can disrupt substrate binding (Fig. 4). Lastly, N149D, K152A, and R187A each killed cells to WT level, indicating they are not involved in tRNase activity (Fig. 4).

#### ***In vitro analysis of active-site residues***

To assess whether each of these mutants can still fold properly, we purified each toxin and evaluated its ability to bind CysK *in vitro*. Each CdiA-CT was purified by denaturing it away from Ni<sup>2+</sup>-bound CdiI-his6. The toxins were then refolded and incubated with CysK-his6. “Input” refers to what was loaded onto a Ni<sup>2+</sup> column, “unbound” represents what did not bind to the column, and “bound” is what eluted off the column. Each mutant toxin tested is able to bind CysK, suggesting they can re-fold properly (Fig. 5A).

Next, these mutants were tested for their ability to cut tRNA *in vitro*. 1uM of each purified toxin was incubated with 1uM of CysK-his6 and 5ug of guanidine isothiocyanate-phenol extracted total RNA. 1uM of CdiI was included where indicated. Ethidium bromide and Northern blot analysis to Ile<sup>1</sup> tRNA shows D155A, H178A, and E181A have abolished activity, while W176A and T185I have decreased activity compared to the WT enzyme (Fig. 5B). These results are consistent with the *in vivo* toxicity assay. Taken all together, these data demonstrate that residues D155, W176, H178, E181, and T185 each play a role in enzymatic activity or substrate binding.

## **Characterization of CysK**

### ***Mutational analysis of CysK***

Now that we have established a potential active site for CdiA-CT, we wanted to identify the role of CysK in the complex. We reasoned CysK could be directly involved in substrate binding or even enzymatic activity. We used genetic selection to reveal mutations in CysK that correlated with abolished tRNase activity. However, we only found mutations that impeded the ability of CysK to bind CdiA-CT. Therefore, we took a similar approach as the one used to identify active-site residues in CdiA-CT. We looked for conserved CysK residues



that had the potential to be involved in substrate binding or enzymatic activity (Fig. 6A). Each CysK mutant was cloned into pTrc99a and co-transformed into *E. coli cysK* cells with either pCH450::CdiA-CT or pCH450::CdiA-CT(H178A) and selected on LB-glucose plates. We found that even with suppressed toxin expression, each mutant was still able to activate CdiA-CT and kill cells (Fig. 6B).

Each pTrc99a::cysK-his6 mutant was also expressed in JCM158 *cysK* target cells and competed against EPI100 cells expressing no CDI genes, CDI<sup>UPEC536</sup> or a chimeric CDI system where CdiA-CT/CdiI<sup>UPEC536</sup> is deployed by CdiA<sup>EC93</sup>. We reasoned this would be a more sensitive method of testing the ability of each mutant allele to activate CdiA-CT as much less toxin is delivered into target cells than the amount present when expressed from plasmid pCH450. Targets were tested against both CDI<sup>+</sup> inhibitors because CdiA<sup>UPEC536</sup> delivers CdiA-CT into target cells with lower efficiency. We hoped using this inhibitor would be an even more sensitive method to test for CdiA-CT activation. Western blot analysis using anti-his6 antibody showed relatively equal expression of each mutant, as compared to WT (data not shown). Furthermore, each target grew to full capacity in the competitions against CDI mock inhibitors, demonstrating expression of these mutant alleles does not affect cell growth (Fig. 6C). In competition against the two CDI<sup>+</sup> inhibitors, each mutant was inhibited to WT levels, except for R101E K102E and  $\Delta$ 307-324 (Fig. 6C). Unfortunately, after deeper analysis of these mutants, it was discovered that neither support growth in minimal media in the absence of *cysM*. This suggests these mutants do not have functional enzymatic activity, most likely due to aberrant folding, and therefore should not be studied further. These results suggest CysK might not play a role in enzymatic activity or

substrate binding. Alternately, perhaps a single-residue change is not enough to potentiate a phenotype that would reveal itself in our toxicity assays.

### ***Distant CysK alleles activate CdiA-CT***

We have not been able to find small changes in CysK that influence CdiA-CT activity. Therefore, we decided to make big changes and ask if they have any consequence on toxin activation. We reasoned if we made such changes but found a conserved region that was required for activation, this might direct us to the region of CysK involved in tRNase activity. Instead of making drastic changes to *E. coli* CysK and hoping it would still be able to fold properly, we used CysK alleles from different species of bacteria. Figure 7A shows an alignment of the species chosen. Each allele was cloned into pTrc99a and subjected to the same co-transformation toxicity assay as described above. Remarkably, each CysK allele was able to activate CdiA-CT (7B). CysK<sup>N.lactamica</sup> was the only allele that required arabinose induction of the CT to kill cells (Fig. 7B). Thus, even distant alleles of CysK are still able to activate CdiA-CT *in vivo*.

### ***In vitro analysis of distant CysK alleles***

To test whether these CysK alleles could bind to the toxin, the same protein-binding assay was used as described earlier. We found CysK<sup>H.influenza</sup> and CysK<sup>N.lactamica</sup> are the only two alleles with reduced binding to the CT (Fig. 7C). This is only somewhat reflective of the *in vivo* transformation assay results, as CysK<sup>H.influ</sup> seemed to activate CdiA-CT to WT levels (Fig. 7B).

To examine if the distant CysK alleles could alter tRNase activity, perhaps by changing the cleavage site or specificity of the substrate, RNA from *in vitro* reactions was analyzed. 0.1uM, 1uM, or 10uM of CysK-his6 was incubated with 1uM CT and 5ug of guanidine isothiocyanate-phenol extracted total RNA. 1uM of CdiI was added where indicated. The reactions were subjected to gel electrophoresis and imaged by ethidium bromide staining or Arg<sup>2</sup> radiolabeled probe. There was no change in cleavage pattern, indicating each CysK allele still supports CdiA-CT to cut in the anti-codon loop of all tRNAs (Fig. 8). Furthermore, the results are consistent with the protein-binding data in that reactions containing CysK<sup>H.influ</sup> or CysK<sup>N.lact</sup> show a decrease in tRNase activity (Fig. 8).

### ***Distant CysK alleles support CDI***

Competition analysis was used as a more sensitive method to look for a decrease in CdiA-CT activation. Again, JCM158 *cysK* target cells were complemented with pTrc99a::cysK-his6 constructs and competed against EPI100 cells expressing no CDI genes, CDI<sup>UPEC536</sup> or CDI<sup>EC93</sup> deploying CdiA-CT/CdiI<sup>UPEC536</sup>. Western blot analysis to his6 illustrates that each allele of CysK is expressed to similar levels as CysK<sup>E.coli</sup> (Fig. 9). Interestingly, this competition analysis only somewhat mirrors the *in vitro* results. Each target strain grew 2 logs against the CDI mock control, indicating expression of the foreign alleles does not affect target cell growth (Fig. 9). Strikingly, target cells carrying CysK<sup>E.cloacae</sup>, *B.subtilis*, or *N.lactamica* are inhibited 1 log less than targets expressing the rest of the CysK alleles (Fig. 9). It is unclear if this is physiologically relevant, or can give us insight into the role of CysK in toxin activation. Considering each distant allele of CysK is still able to activate

CdiA-CT, these data again suggest that the physical binding of CysK is the mechanism of toxin activation.

### **Insights from a distant homolog of CdiA-CT**

#### ***CT<sup>R.lactaris</sup> does not require CysK in vivo***

When looking at an alignment of CdiA-CT to distant homologs, it becomes apparent the toxins found in Gram-positive species contain a 9 residue insertion between residues K197 and N198 in CdiA-CT<sup>UPEC536</sup> (Fig. 10A). This insertion maps to a loop connecting two  $\alpha$ -helices that protrude out from the CysK active site (Fig. 10B). Given the above results, we now postulate the role of CysK is to simply bind to and stabilize the CT so it can perform its enzymatic activity. We hypothesized that the insertion found in the toxins from Gram-positive species might stabilize the toxin to the point where it would no longer require CysK for activity. To test this, we looked at *in vivo* toxicity of a homolog from the Gram-positive species, *Ruminococcus lactaris* (CT<sup>R.lactaris</sup>). We transformed pCH450::CT/I<sup>R.lactaris</sup>-DAS into *cysK*<sup>+</sup> or *cysK*<sup>-</sup> *E. coli* cells and selected on LB-glucose or LB-arabinose plates. Upon arabinose induction, the DAS-tagged immunity protein will deliver the CT/I-DAS complex to the ClpXP degradation machinery. The machinery will selectively degrade the DAS-tagged immunity protein, leaving the CT free in the cytoplasm of cells to perform its toxic function. First, each construct was able to grow under glucose conditions, indicating the transformation was efficient (Fig. 10C). Second, upon arabinose induction, pCH450::CdiA-CT/CdiI-DAS killed WT cells but not *cysK*<sup>-</sup> cells, as expected (Fig. 10C). However, pCH450::CT/I<sup>R.lactaris</sup>-DAS killed both WT and *cysK*<sup>-</sup> cells when induced with arabinose.

These results demonstrate that the CdiA-CT homolog from *R. lactaris* does not require CysK for activity *in vivo*.

### ***CT<sup>R.lactaris</sup> does not require or bind CysK in vitro***

Next, we asked if CT<sup>R.lactaris</sup> requires CysK for *in vitro* activity. To test this, we incubated 1uM of each toxin in the presence or absence of 1uM CysK with 5ug of guanidine isothiocyanate-phenol extracted total RNA. 1uM of immunity protein was added where indicated. Not surprisingly, CT<sup>R.lactaris</sup> shows the same level of activity, independent of CysK (Fig. 11A). We were curious if CT<sup>R.lactaris</sup> could still bind to CysK. Its C-terminal tail ends in PYGGIN, so it was not likely. Indeed, when purified CT<sup>R.lactaris</sup> was incubated with CysK-his6, no binding was observed (Fig. 11B), consistent with the *in vitro* results. Thus, CT<sup>R.lactaris</sup> does not require CysK for *in vitro* tRNase activity.

### ***CT<sup>R.lactaris</sup> acts as a specific tRNase in vitro***

As previously mentioned, CdiA-CT is a general tRNase that cleaves in the anticodon loop, and it behaves as such in the assay performed above (Fig. 11A). Strikingly, even though CT<sup>R.lactaris</sup> seems to cleave in the same position as CdiA-CT, it has diminished tRNase activity *in vitro* (Fig. 11A). However, the *in vivo* toxicity assay showed CT<sup>R.lactaris</sup> to be just as lethal as CdiA-CT (Fig. 10C). We hypothesized that CT<sup>R.lactaris</sup> is cutting only a subset of tRNAs, targeting all tRNAs but less efficiently, or requires a different cofactor to be completely active.

To test the dependence on a distinct cofactor, we performed *in vitro* reactions in the presence or absence of S30 cell extract. S30 extract contains all elements of a cytoplasm, so

it would presumably contain any cofactors that were present in the *in vivo* toxicity assay. CdiA-CT activity was enhanced greatly by the addition of S30, likely because of the addition of CysK found in the extract (Fig. 11C). Conversely, CT<sup>R.lactaris</sup> activity was not enhanced in the presence of S30, suggesting it does not use a cofactor found in S30 cell extract (Fig. 11C).

To test if it cuts all tRNAs, but less efficiently, Northern blot analysis to the *in vitro* reactions was performed. This reveals cleaved tRNA<sup>Ala1B</sup> in reactions containing CdiA-CT but not CT<sup>R.lactaris</sup> (Fig. 11A). Notably, the toxin was denatured during the purification process. However, we made sure it re-folded properly by demonstrating it could re-bind to its cognate immunity protein (Fig. 12C). Taken all together, these results indicate that CT<sup>R.lactaris</sup> does not utilize a co-factor or act as a general tRNase *in vitro*, suggesting it targets a specific subpopulation of tRNA species.

### ***CT<sup>R.lactaris</sup> acts as a specific tRNase in vivo***

We next asked if CT<sup>R.lactaris</sup> has the same activity *in vivo* as it does *in vitro*. To test this, we took RNA samples from cells expressing CdiA-CT or CT<sup>R.lactaris</sup> with their cognate or noncognate immunity protein. Upon guanidine isothiocyanate-phenol extraction, RNA was subjected to Northern blot analysis. Cleaved tRNA<sup>Ala1B</sup> was observed in cells expressing CdiA-CT, but not in cells expressing CT<sup>R.lactaris</sup> (Fig. 12B). This result supports the hypothesis that CT<sup>R.lactaris</sup> is a specific tRNase both *in vitro* and *in vivo*. More work will have to be done to identify the specific tRNA(s) targeted by CT<sup>R.lactaris</sup>.

### ***Chimeric CdiA-CT<sup>UPEC536</sup>/CT<sup>R.lactaris</sup> shows specific tRNase activity***

The point of studying CT<sup>R.lactaris</sup> was to gain some insight into the requirement of CysK for CdiA-CT activation. We have now established that CT<sup>R.lactaris</sup> does not bind to or require CysK *in vivo* or *in vitro*, while CdiA-CT does. Since the main difference between these two toxins is a 9 residue insertion between two  $\alpha$ -helices, we reasoned this insertion is what drives the CysK-independent activity seen in CT<sup>R.lactaris</sup>. To test this, we engineered 3 chimeric proteins (UPEC536/R.lact CT1, CT2, and CT3) to try to stabilize CdiA-CT to render it CysK-independent. Figure 13A shows the sequences of the 3 chimeric constructs that were cloned into pCH450::CdiA-CT/I-DAS and assayed for their toxicity in *cysK*<sup>+</sup> and *cysK*<sup>-</sup> *E. coli* cells. Unfortunately, the toxicity of the first 2 chimeras is CysK-dependent (data not shown) and CT3 has a defect in toxicity in WT cells, suggesting aberrant folding (Figure 13B). As a longshot, we asked if the chimeric proteins have altered tRNase activity. To do this, we purified RNA from *cysK*<sup>+</sup> cells expressing each of these constructs and performed a Northern blot to look at tRNA<sup>Thr2</sup>. Remarkably, UPEC536/R.lact CT1 no longer shows characteristic general tRNase activity (Fig. 13C) even though it inhibited the growth of *E. coli* cells to the same levels as the WT construct (Fig. 13B). This suggests the insertion of residues NPNLSD at position K197 in CdiA-CT and removal of N189 (Fig. 13A), has turned CdiA-CT into a specific tRNase, perhaps cutting the same subset of tRNAs as CT<sup>R.lactaris</sup>. Further analyses will have to be performed to prove this hypothesis.

***Immunity<sup>R.lactaris</sup> displays partial protection of CdiA-CT<sup>UPEC536</sup> activity***

As expected, the immunity proteins of CdiA-CT and CT<sup>R.lactaris</sup> are similar. An alignment displays 18% sequence identity (Fig. 10A). It is reasonable to believe that these immunity proteins might offer some cross-protection for their noncognate toxin. The *in vitro* reactions

discussed above were performed with both cognate and noncognate immunity proteins. The results show Imm.<sup>R.lactaris</sup> providing partial protection of CdiA-CT activity (Fig. 11A). Furthermore, *in vivo* analysis also reveals almost full protection from both growth inhibition (data not shown) and tRNase activity when expressed inside *E. coli* cells (Fig. 12B). Consistent with the *in vitro* and *in vivo* results, Immunity<sup>R.lactaris</sup>-his6 is able to partially bind to CdiA-CT (Fig. 12C).

### ***CT<sup>R.lactaris</sup> can be delivered via CDI***

As mentioned previously *R. lactaris* is a Gram-positive species, which contain different types of contact-dependent competition systems than those found in Gram-negative species. Homologs of CT<sup>R.lactaris</sup> in Gram-positives can be found in rearrangement hotspot (Rhs) proteins, a competition system that also requires direct cell-to-cell contact for delivery of a toxic domain derived from its C-terminus (Rhs-CT). Rhs proteins are involved in intra-species competition and thus Rhs-CTs must cross the membrane of Gram-positive cells. We asked if CT<sup>R.lactaris</sup> could cross the outer and inner membranes of Gram-negative bacteria via the CDI pathway. CdiA-CTs can be further broken down into an N- and C-terminal domain (CdiA-CT<sup>NT</sup> and CdiA-CT<sup>CT</sup>). CdiA-CT<sup>CT</sup> encodes for the toxic domain, while CdiA-CT<sup>NT</sup> is what directs the toxin to a specific inner-membrane protein to be delivered to the cytoplasm (data not shown). We reasoned CT<sup>R.lactaris</sup> would not have the proper domain at its N-terminus to be directed through the inner membrane of Gram-negative bacteria. Therefore, we fused CT<sup>R.lactaris</sup> to the N-terminal domain of CdiA-CT<sup>UPEC536</sup> or CdiA-CT<sup>EC3006</sup> and then fused the subsequent chimeric CdiA-CTs to CdiA<sup>EC93</sup>. The resulting chimeric CdiA proteins are depicted in Figure 14A. EPI100 inhibitor cells expressing CDI<sup>UPEC536</sup>, CDI<sup>UPEC536/R.lactaris</sup>,



CDI<sup>EC3006/R.lactaris</sup>, or CDI cosmids were co-cultured with target cells expressing no immunity, cognate, or noncognate immunity protein. Strikingly, the inhibitors carrying the two chimeric CdiA constructs are able to deploy CT<sup>R.lactaris</sup> into target cells (Fig. 14B). This is demonstrated by targets losing viability against CDI<sup>EC3006/R.lactaris</sup> inhibitors and only growing ~0.25 logs against CDI<sup>UPEC536/R.lactaris</sup> inhibitors. Furthermore, expression of cognate immunity protein rescues this growth inhibition. Moreover, these results further support the ability of Imm.<sup>R.lactaris</sup> to partially block CdiA-CT activity.

## Discussion

CdiA-CT is the only known example of a CDI toxin that requires a cofactor for activation. In fact to this date, no bacteriocin or type VI effector protein has been shown to require a cofactor<sup>51,128</sup>. Many questions have arisen since its discovery. First, why does CdiA-CT even require a CysK? One idea is that the binding event protects CdiA-CT from degradation in target cells, prolonging its enzymatic activity. However, there are numerous examples of CDI toxins that do not require a cofactor and kill target cells with the same efficiency as CdiA-CT<sup>117,120</sup>. Nonetheless, CdiA-CT does require a cofactor in its CDI pathway, and there are many reasons CysK is a worthy protein to exploit. Firstly, we demonstrated that CysK from a cohort of distantly related species was still able to bind to and activate CdiA-CT. Considering CDI systems are widespread amongst proteobacteria, it is ideal to utilize a cofactor that is conserved. Another preferred characteristic of a cofactor in the CDI pathway would be that it is essential, this way target cells cannot lose its function and become resistant to CdiA-CT activity. CysK is not essential, as it has an analog, CysM, which is also able to produce cysteine from OAS<sup>273</sup>. However, the two enzymes can use

different substrates and we have observed that CysK is more efficient than CysM at cysteine metabolism (unpublished data). Therefore, under some physiological conditions, CysK could potentially identify as an essential protein. This could even suggest that CDI systems delivering CdiA-CT<sup>UPEC536</sup> have evolved to be expressed or utilized most efficiently under such conditions.

Next, what is CysK's role in tRNase activity? Does it provide any residues that are involved in electron transfer, stabilization of intermediates, or coordination of a metal ion? Or rather, does it simply bind to and stabilize the toxin? With help from the crystal structure and homologs we were able to identify active-site residues of CdiA-CT. While it is unclear where a tRNA species would bind to the CysK/CdiA-CT complex, we propose residue W176 flips out a base in the anti-codon loop of bound tRNA which allows residue D155 or E181 to act as a general base and strip a proton from the 2' hydroxyl of the tRNA. This would lead to auto-attack of the backbone where H178 could offer a proton to the 5'-OH leaving group. Lastly, we propose residue T185 is involved in substrate binding. This threonine is found mid-way through an  $\alpha$ -helix with its sidechain exposed to the periphery. When we changed this threonine to an isoleucine we saw a decrease in tRNase activity. Considering this is in close proximity to the proposed active site, we postulate that tRNA makes contact with this threonine and the bulky side chain of isoleucine, or lack of ability to H-bond, hinders tRNA docking. This is of course all completely speculative and a resolved structure of tRNA bound to the complex will be required to determine the precise mechanism of cleavage.

Unfortunately, we were unable to concretely identify the role of CysK in the complex. Genetic selection, making forced mutations, and examining alleles from distant species failed to find mutations in CysK that affected toxin activation without reducing its ability to bind

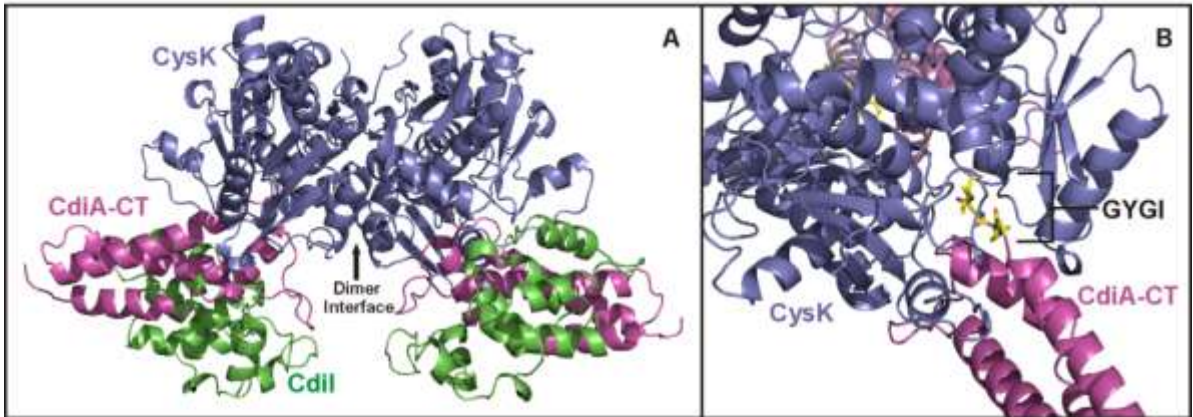
the CT. Thus, we postulate that CysK acts to stabilize the toxin which allows it to properly bind and cleave tRNA.

We also looked at a distant homolog of CdiA-CT for some insights on the role of CysK in the complex. We established CT<sup>R.lactaris</sup> does not require CysK for activity both *in vivo* and *in vitro*. Furthermore, we found it has a slightly altered activity than CdiA-CT in that it acts as a specific tRNase, although still cutting in the anti-codon loop. We credit this specificity to residues N139-D143 in CT<sup>R.lactaris</sup> because when we grafted these residues onto CdiA-CT, it became a specific tRNase. Reasonably, this insertion is located distal from the active site, suggesting its role in substrate binding and not enzymatic activity. Perhaps then, the role of CysK is to stabilize CdiA-CT in a way that allows the complex to bind tRNAs non-discriminately. The implications of being able to cut a broad range of tRNAs versus a narrow range is unknown. Finally, we demonstrated that a toxin from a competition system used by Gram-positive species could be delivered via the CDI pathway into Gram-negative target cells. This demonstrates the adaptability of CDI systems, and their potential to deliver a vast repertoire of toxins.

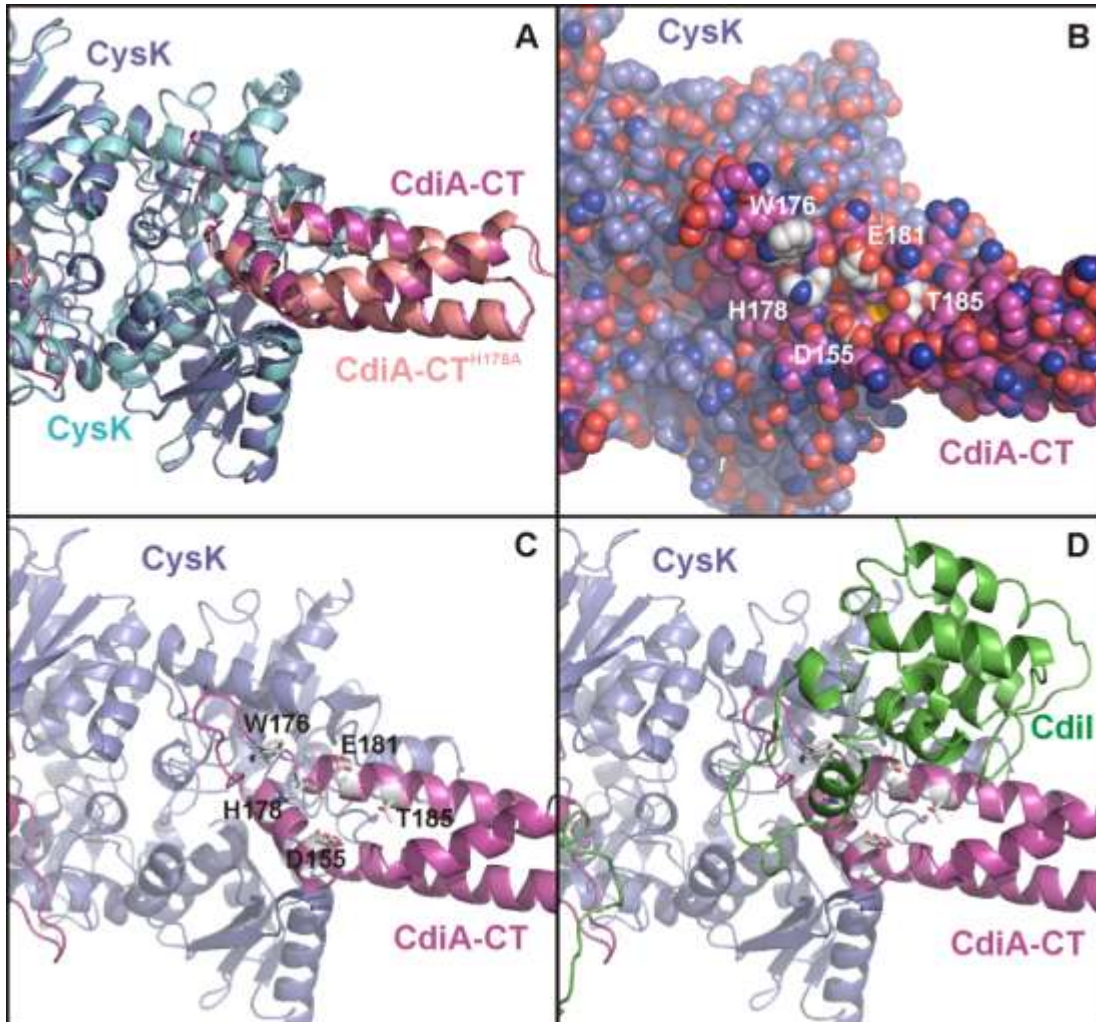
Lastly, does delivered CdiA-CT affect cysteine metabolism? Cysteine metabolism is under strict regulation in the cell where increased amounts of OAS will upregulate CysK expression<sup>275</sup>. With CdiA-CT binding in the active site of CysK, it could potentially displace CysE or prevent OAS from entry. This would decrease cysteine production and ultimately alter the expression of the metabolic enzymes. Furthermore, since the toxin can bind to CysK in the presence of immunity protein, even sibling cells can be affected. With the recent finding that CDI systems have an extremely narrow target range and are mostly limited to intra-species and even intra-strain delivery, this could possibly be a way to regulate CysK

activity amongst kin<sup>118</sup>. Extensive efforts were put into trying to try to find an effect on cysteine biosynthesis in cells upon delivery of CdiA-CT, but none seemed to be relevant under natural conditions. For example, we could block CysK activity *in vivo* by overexpressing CdiA-CT inside cells, but when we tried to block the enzyme by delivering CdiA-CT via the CDI pathway, we no longer saw a loss of CysK activity (data not shown). However, perhaps under certain environmental conditions, there could be enough CdiA-CT delivery into cells, or scarce enough amounts of CysK, where CDI would in fact affect cysteine metabolism. In addition, the 9 residue insertion that presumably makes CT<sup>R.lactaris</sup> CysK-independent is only found in Gram-positive species which do not contain CDI systems. This further supports the idea that CDI systems are involved in communication rather than competition.

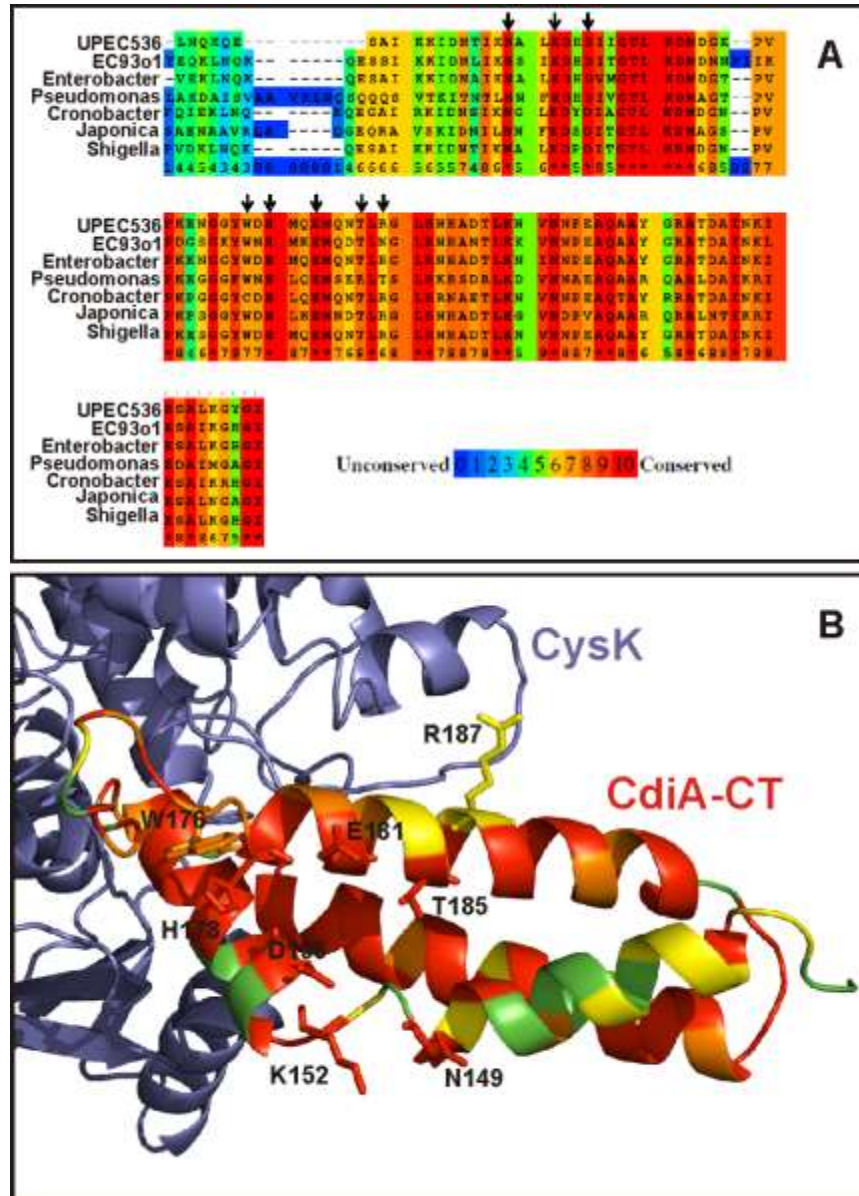
**Figure 1. Structure of CysK/CdiA-CT/CdiI ternary complex.** **A)** CdiA-CT is depicted in pink, CysK in purple, and CdiI in green, each in ribbon form. The model presented shows CysK in its dimer form, with hydrophobic interactions facilitating binding at the described interface. PLP is not shown. **B)** CysK in purple and CdiA-CT in pink with its C-terminal GYGI amino acids (yellow) sticking into the active site of CysK. PLP is not shown.



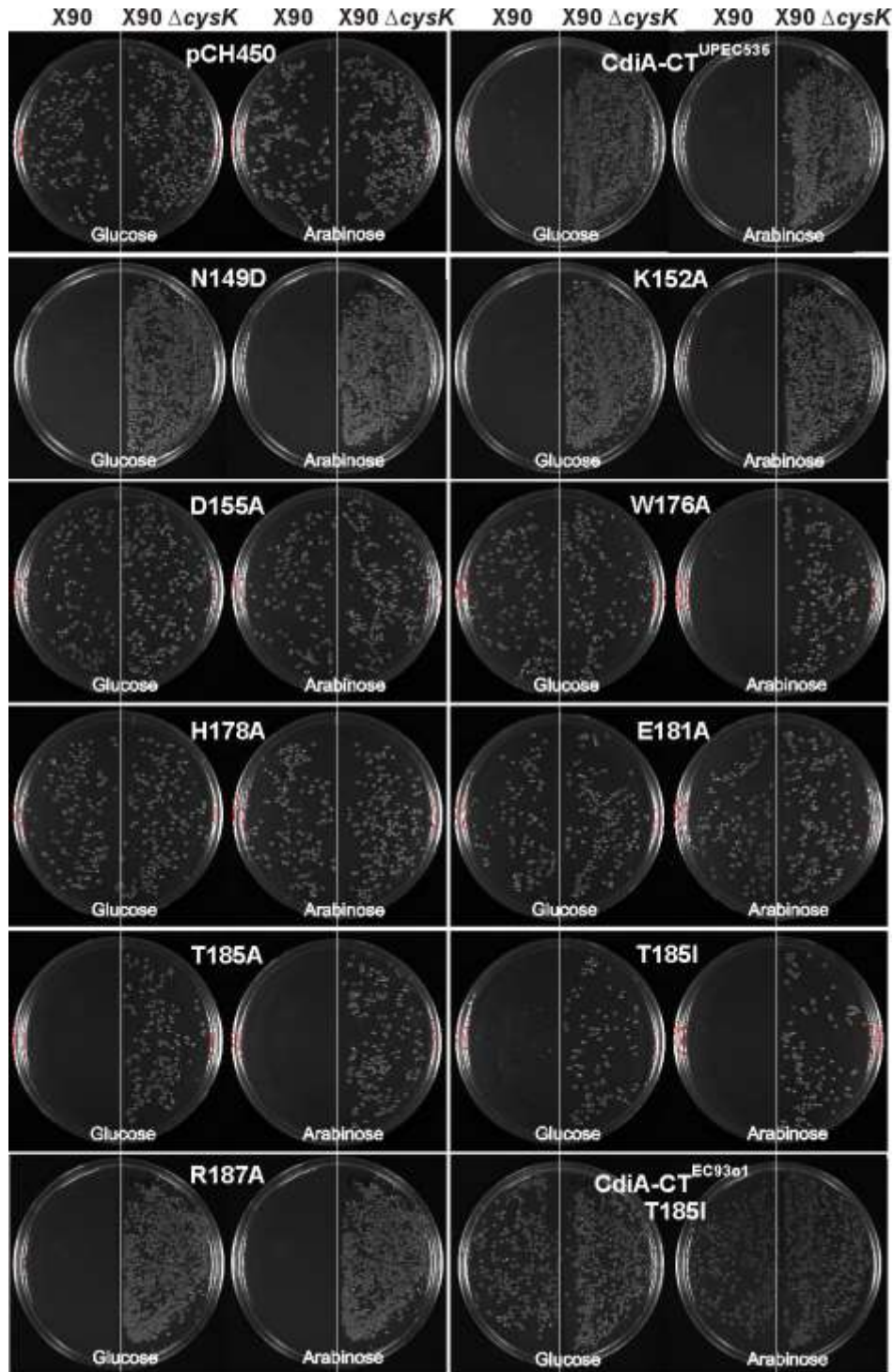
**Figure 2. Proposed active-site residues of CdiA-CT .** **A)** An overlay of CysK/CdiA-CT(H178A) in light blue and light pink with CysK/CdiA-CT in purple and pink. PLP is not shown. **B)** CysK in purple and CdiA-CT in pink with the proposed active-site residues in white. The molecule is in sphere form to demonstrate electrostatic potential. **C)** The same molecule as B), but in ribbon form with the active-site residues shown as sticks. **D)** The same molecule as C), but CdiI has been added to show it blocks the proposed active-site residues.



**Figure 3. Homologs of CdiA-CT show conserved active-site residues.** **A)** An alignment of CdiA-CT to homologous toxins found in the species stated below. Black arrows point to conserved residues that were mutated for analysis. This is a heat map alignment performed by PRALINE Multiple Sequence Alignment. **B)** The heat map in A) is grafted onto the crystal structure of CdiA-CT with CysK in purple. The residues pointed to in A) are shown in stick form.

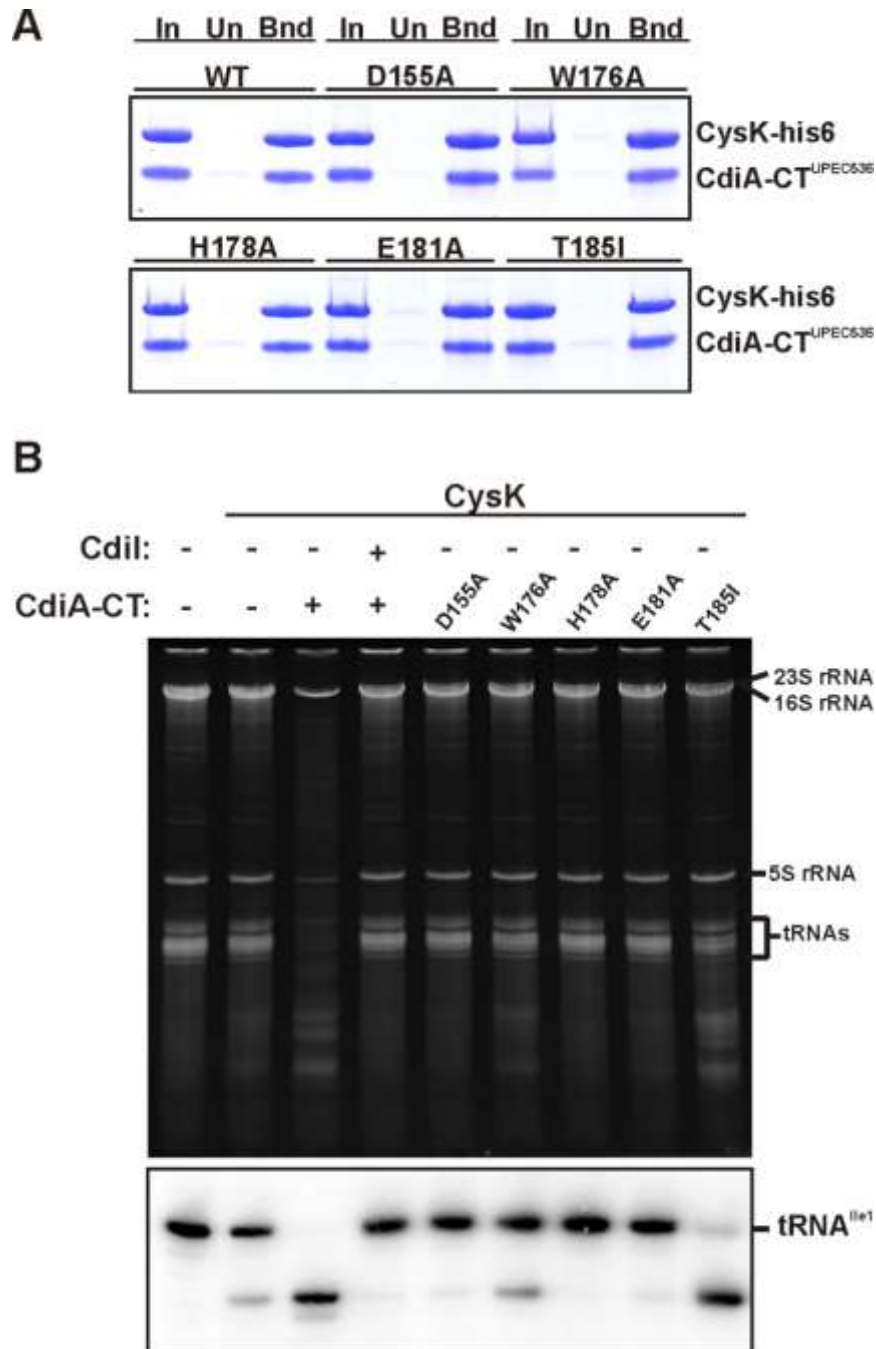


**Figure 4. Mutational analysis of CdiA-CT active-site residues *in vivo*.** Arabinose-inducible plasmid, pCH450, with the specified *cdiA-CT* alleles, was transformed into *E. coli* X90 or X90  $\Delta$ *cysK* cells and selected on LB-glucose or LB-arabinose plates supplemented with tetracycline. The last panel is a homolog of CdiA-CT from *E. coli* EC93. The T185I mutation has a greater effect on this toxin.

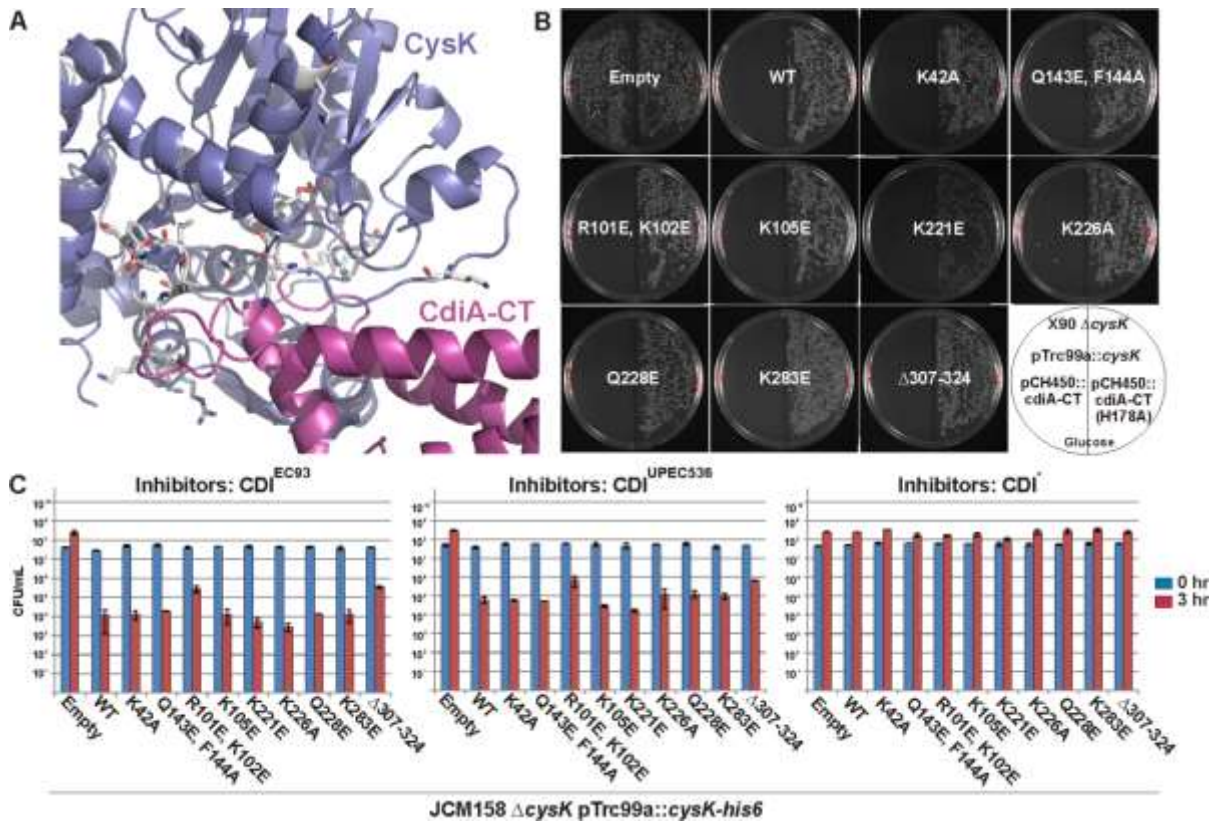




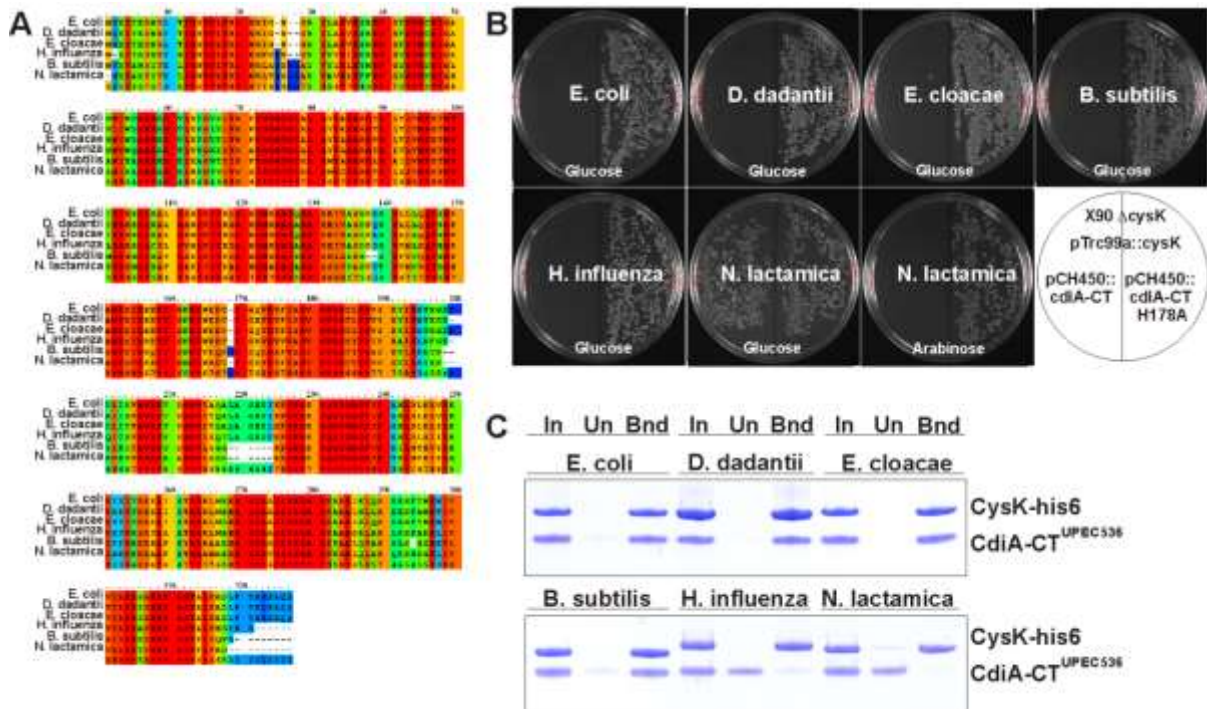
**Figure 5. Mutational analysis of CdiA-CT active-site residues *in vitro*.** **A)** Purified CdiA-CT mutants were mixed with equimolar CysK-his6 and put over a Ni<sup>2+</sup> column (In). The unbound (Un) and bound (Bnd) fractions were collected and analyzed by SDS-PAGE. **B)** 1uM of CdiA-CT, CdiI, and CysK were added to 8ug of total RNA where indicated. RNA was analyzed by ethidium bromide stain and Northern blot using a probe to tRNA<sup>Ile1</sup>.



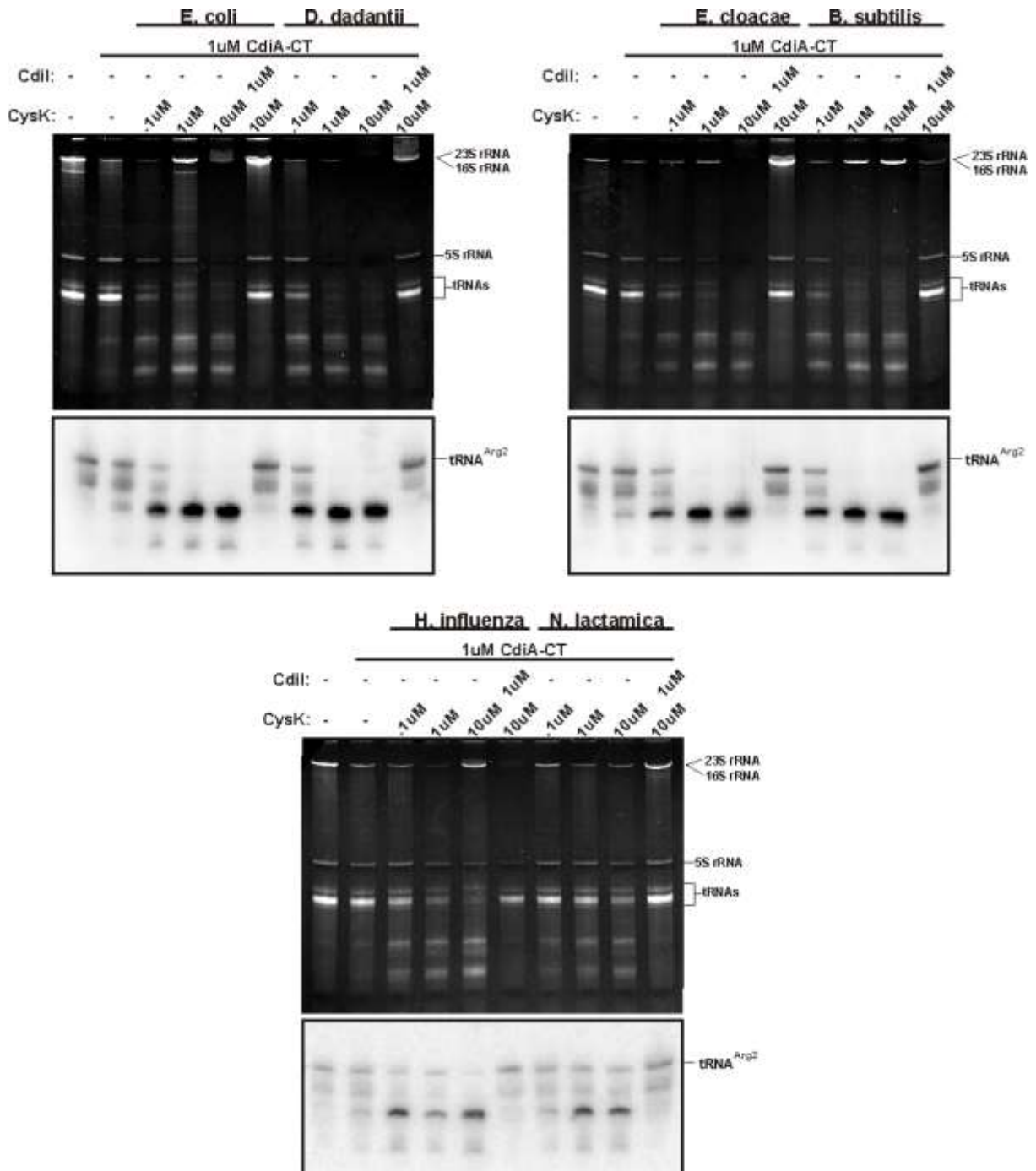
**Figure 6. Mutational analysis of CysK *in vivo*.** **A)** CdiA-CT in pink and CysK in purple with each mutated CysK residue shown in stick form in white. **B)** *E. coli* X90  $\Delta$ cysK was co-transformed with the specified pTrc99a::cysK construct and either pCH450::CdiA-CT or CdiA-CT(H178A) and selected on LB-glucose plates supplemented with tetracycline and ampicillin. **C)** The annotated target cells were co-cultured with CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or CDI<sup>I</sup> inhibitors at a 1:1 ratio in liquid broth for 3 hours. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.



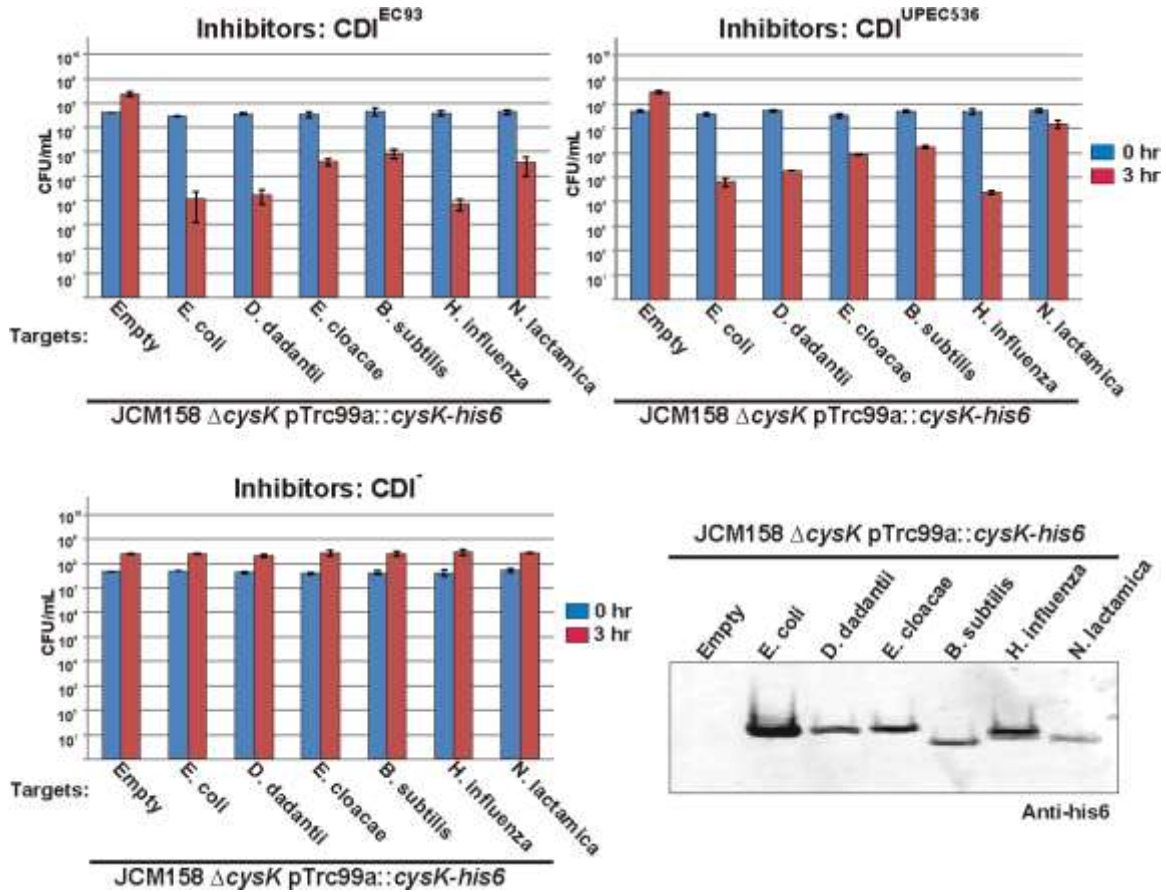
**Figure 7. Distant CysK alleles activate and bind CdiA-CT.** **A)** A heat map alignment of CysK from the species designated using PRALINE Multiple Sequence Alignment **B)** *E. coli* X90  $\Delta$ cysK was co-transformed with the specified pTrc99a::cysK construct and either pCH450::CdiA-CT or CdiA-CT(H178A) and selected on LB-glucose plates supplemented with tetracycline and ampicilin. The last panels shows pTrc99a::CysK<sup>N.lactamica</sup> required arabinose induction to activate CdiA-CT. **C)** Purified CdiA-CT was mixed with the annotated CysK-his6 at equimolar ratio and put over a Ni<sup>2+</sup> column (In). The unbound (Un) and bound (Bnd) fractions were collected and analyzed by SDS-PAGE.



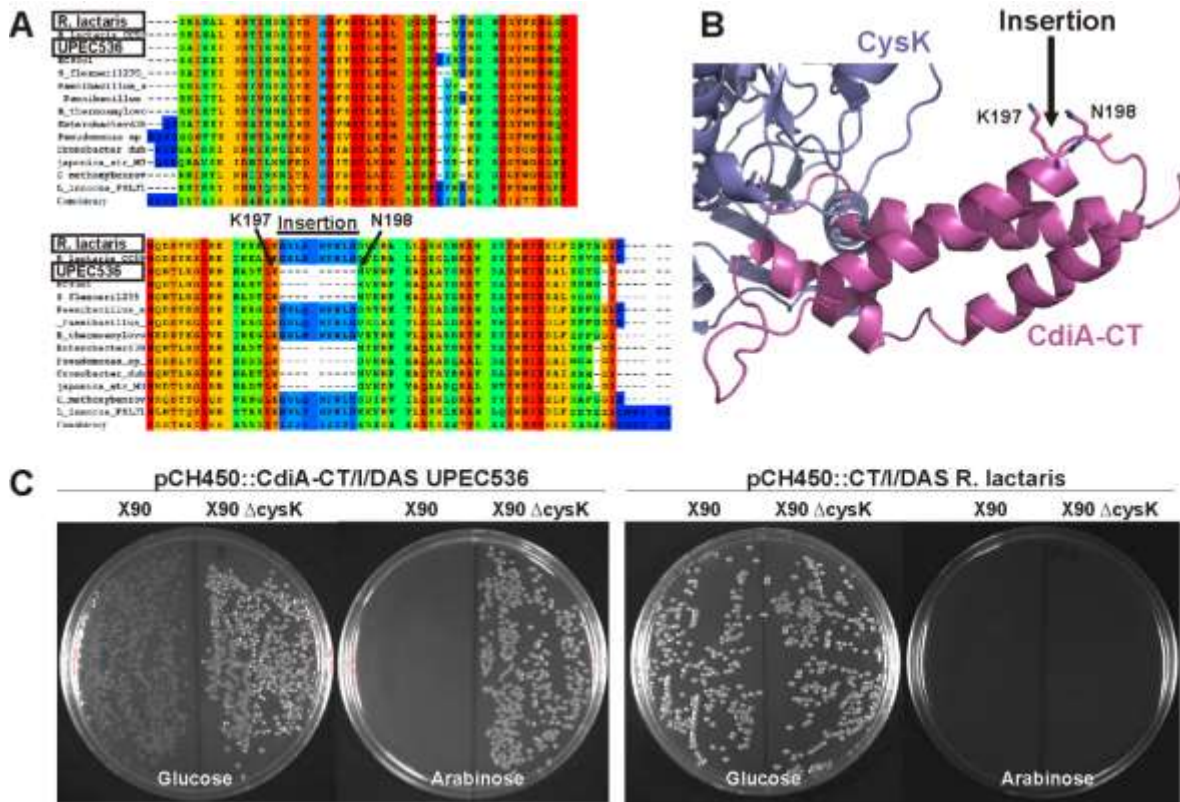
**Figure 8. Distant CysK alleles activate CdiA-CT *in vitro*.** Purified CdiA-CT was incubated with increasing amounts of the indicated CysK and 8ug of total RNA. The reactions were analyzed by ethidium bromide stain and Northern blot to tRNA<sup>Arg2</sup>.



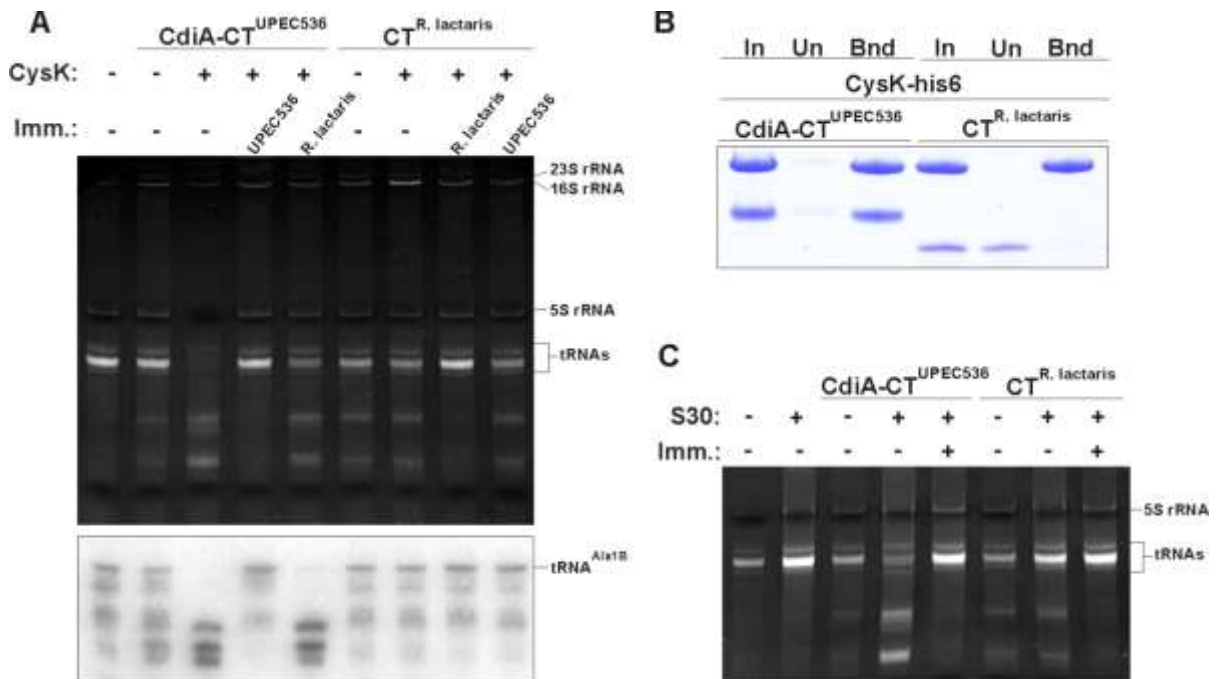
**Figure 9. Distant CysK alleles support CDI.** The indicated target cells were co-cultured with CDI<sup>EC93</sup>, CDI<sup>UPEC536</sup>, or CDI inhibitors at a 1:1 ratio in liquid broth for 3 hours. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. Western blot analysis of CysK-his6 expression levels in target cells using anti-his6 antibody.



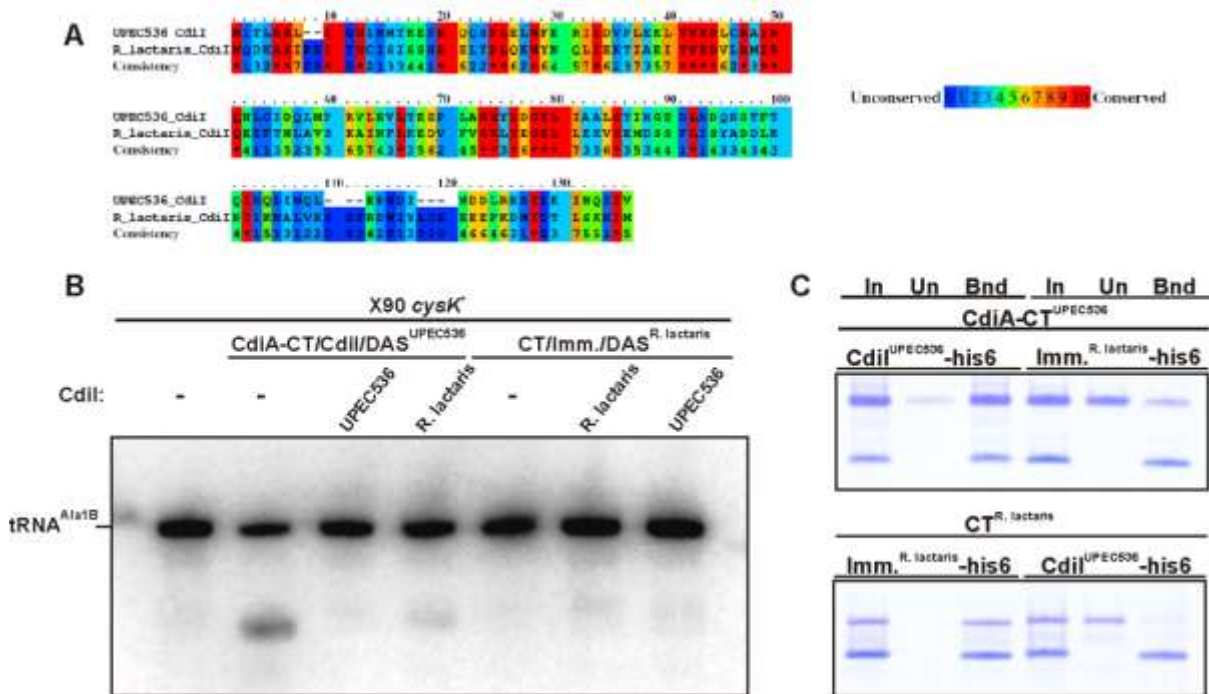
**Figure 10. CT<sup>R.lactaris</sup> does not require CysK *in vivo*.** **A)** An alignment of distant homologs to CdiA-CT using PRALINE Multiple Sequence Alignment and shown as a heat map with conserved residues in red and divergent residues in cool colors. The 9 residue insertion found in Gram-positive species is highlighted with a black bar. The insertion is between residues K197 and N198 in CdiA-CT. **B)** Structure of CysK in purple and CdiA-CT in pink to show the insertion site between two  $\alpha$ -helices in CdiA-CT. **C)** The two arabinose-inducible constructs were transformed into *E. coli* X90 or X90  $\Delta$ cysK cells and selected on LB-glucose or LB-arabinose plates supplemented with tetracycline.



**Figure 11.  $CT^{R.lactaris}$  does not require or bind CysK *in vitro*.** **A)** 1uM of toxin was incubated with 8ug of total RNA and equimolar CysK and/or Immunity where indicated. Reactions were examined via ethidium bromide stain and Northern blot analysis to  $tRNA^{Ala1B}$ . **B)** Equimolar ratios of toxin and CysK-his6 were incubated with  $Ni^{2+}$  beads (In). The fractions that did (Bnd) and did not bind (Un) the  $Ni^{2+}$  beads were analyzed by SDS-Page. **C)** 1uM of toxin was incubated with 8ug of total RNA with or without S30 extract. 1uM of Immunity protein was added where indicated. RNA in the reactions was observed via ethidium bromide stain.

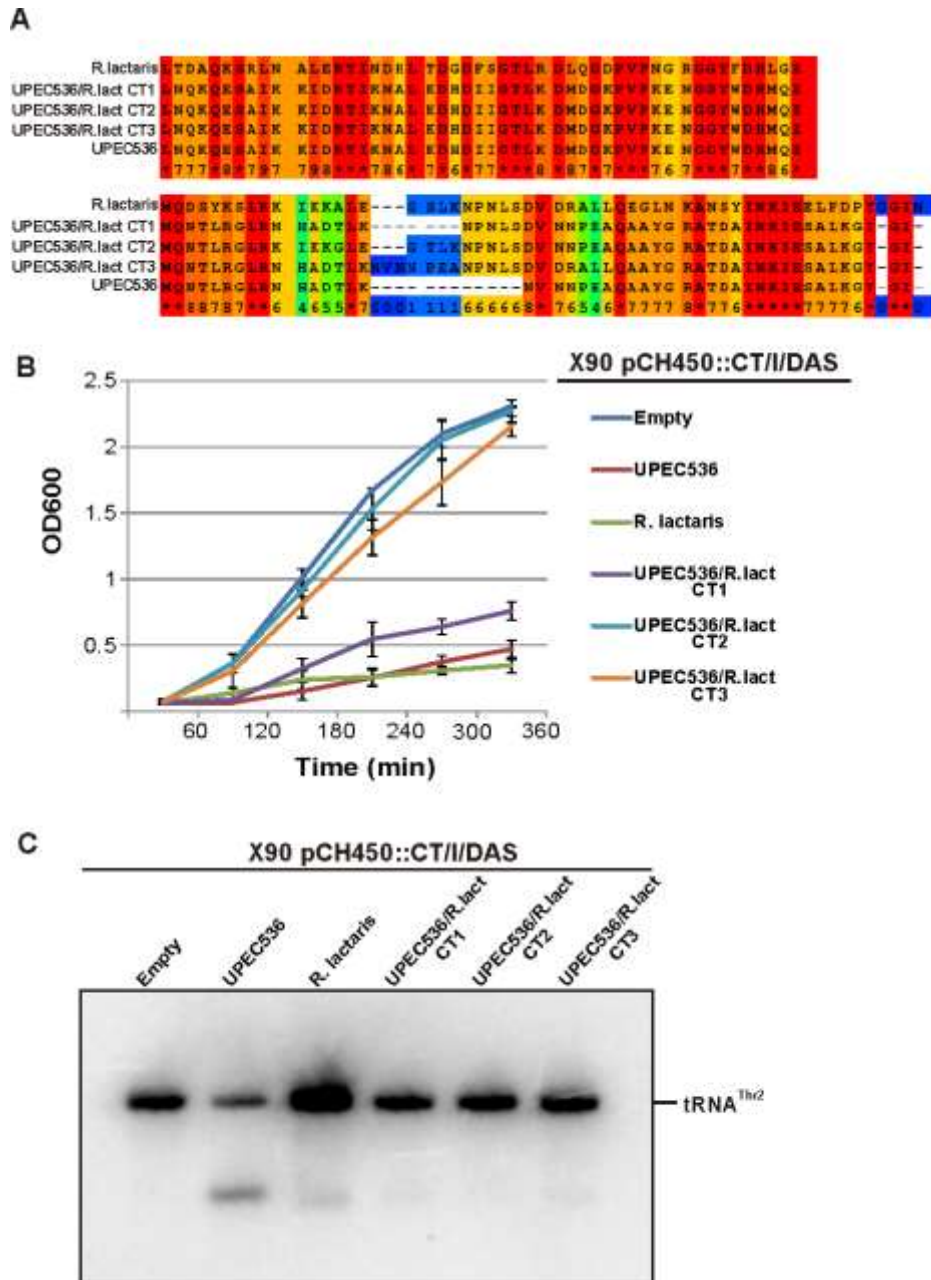


**Figure 12. Analysis of CdiI<sup>UPEC536</sup> and Immunity<sup>R.lactaris</sup>.** A) PRALINE Multiple Sequence Alignment of CdiI<sup>UPEC536</sup> and Immunity<sup>R.lactaris</sup> and shown as a heat map. B) *E. coli* X90 *cysK*<sup>+</sup> cells carrying the annotated toxin and immunity constructs were induced for both toxin and immunity expression for 3 hours. RNA was collected and analyzed via Northern blot to tRNA<sup>Ala1B</sup>. C) Equimolar ratios of the indicated purified toxin and immunity-his6 proteins were mixed with Ni<sup>2+</sup> beads and assayed for their ability to bind one another. “In” represents what was added to the beads, “Un” denotes what did not bind, and “Bnd” is what eluted off the Ni<sup>2+</sup> beads.

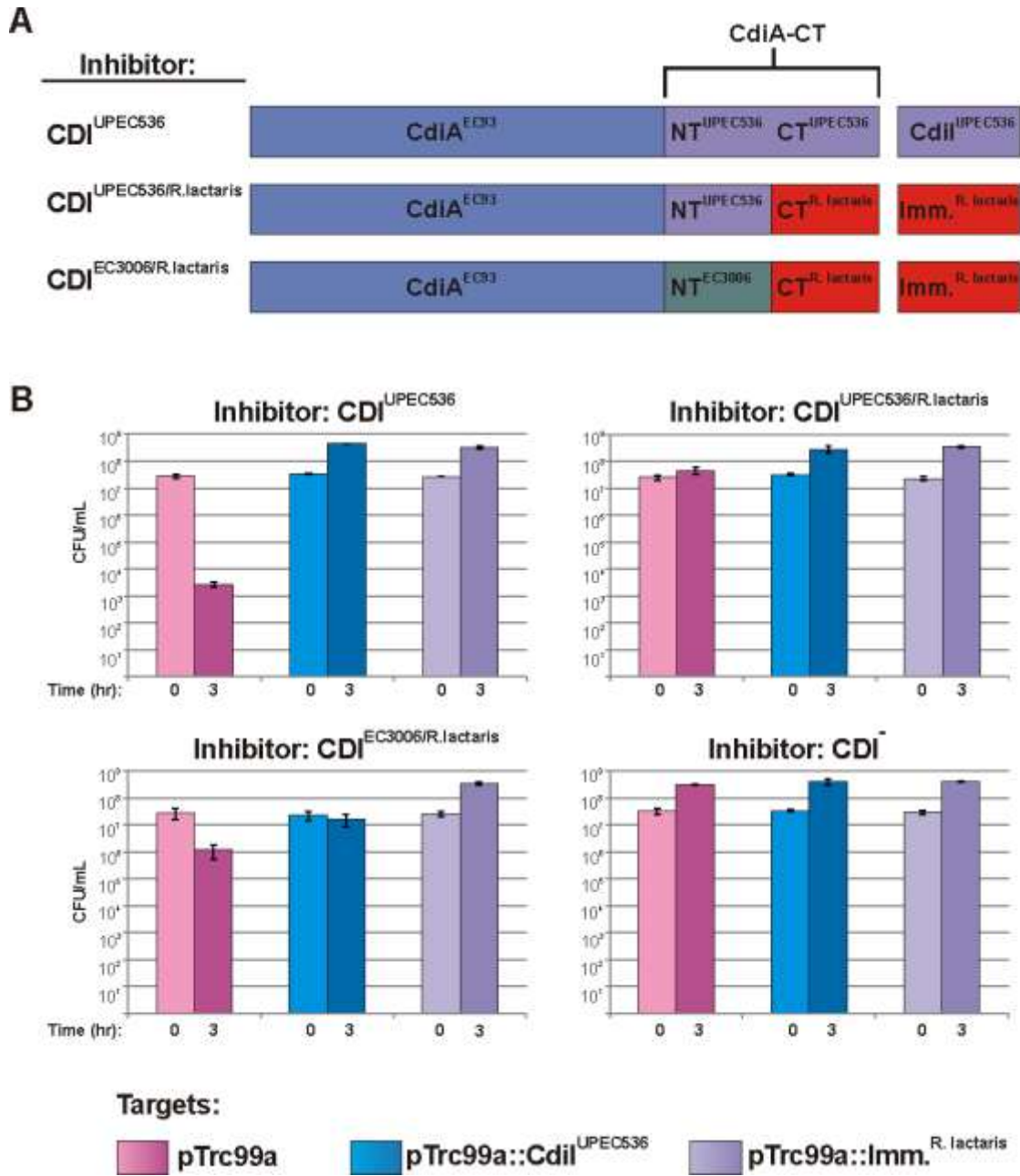




**Figure 13. Chimeric CdiA-CT<sup>UPEC536</sup>/CT<sup>R.lactaris</sup> shows specific tRNase activity. A)** PRALINE Multiple Sequence Alignment of CdiA-CT, CT<sup>R.lactaris</sup>, and the 3 chimeric toxins generated. They are shown as a heat map with unconserved residues in cool colors. **B)** *E. coli* X90 *cysK*<sup>+</sup> cells carrying the indicated arabinose-inducible toxin construct were induced after 30 minutes of growth. **C)** At 3 hours, RNA was collected and analyzed via Northern blot to tRNA<sup>Thr2</sup>.



**Figure 14. CT<sup>R.lactaris</sup> is delivered into target cells via the CDI pathway. A)** Chimeric CdiA proteins present in the designated inhibitor cells. CdiA-CT is composed of an N- and C-terminal domain. Figure is not drawn to scale. **B)** *E. coli* JCM158 target cells with the indicated immunity constructs were co-cultured with each inhibitor cell at a 1:1 ratio for 3 hours in liquid broth. Growth inhibition was assessed by quantifying the number of viable target cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.



**Chapter 6. Characterizing Type VI secretion in *Enterobacter cloacae***

**Note:** Other members of the Hayes lab participated on this project. Fellow graduate student Sonya Donato contributed to the work presented in Figures 6, 9, and 12 and will be finishing the remaining pieces of the project. Undergraduate students David Cunningham, Jeff Kim, Donghoon (Alex) Kang, and Ian Singleton and incoming Master's student Kelsy Siegel helped with competition experiments. David also played a significant role in starting this project.

## Abstract

The type VI secretion system (T6SS) is the most recently identified secretion system found in one-fourth of sequenced Gram-negative bacteria. It forms a dynamic syringe-like structure that delivers toxic effector proteins into neighboring cells. Its role in bacterial competition and virulence has been demonstrated in a limited number of species. Here we show *Enterobacter cloacae* ATCC13047 is a great model organism to study T6S. It contains a functional T6SS that is constitutively expressed under laboratory conditions and is involved in inter- and intra-species competition. There are at least 13 components required to build the T6S organelle, including a PAAR domain-containing protein. We find at least one of the PAAR-containing rearrangement hot spot (Rhs) proteins found in *E. cloacae* is required for T6 assembly and function. Furthermore, we demonstrate that full-length Rhs, not just its PAAR domain, is required for T6S. Our work raises questions about the function of PAAR-containing proteins and suggests Rhs proteins have a vital structural role in T6 assembly in *E. cloacae*.

## Introduction

The type VI secretion system (T6SS) is a membrane-spanning apparatus that injects toxic effector molecules into neighboring prokaryotic or eukaryotic cells<sup>127</sup>. It is found in 25% of sequenced proteobacteria where strains can have up to six type VI loci present in their genome<sup>135</sup>. There are 13 core components required to make a functioning T6 organelle<sup>136,138</sup>. With the exception of Hcp, VgrG, ClpV, and VasK, the remaining core components are named TssX for Type six secretion gene X. The organelle is made up of an inner tube capped by spike proteins that gets encapsulated by a sheath and attached to the cell envelope<sup>151</sup>. The inner tube is made up of hemolysin-containing protein (Hcp) that form monohexameric rings that stack on top of one another<sup>149,150</sup>. Valine-glycine repeat protein G (VgrG) trimers cap this tube and a PAAR (proline, alanine, alanine, arginine) motif-containing protein binds to the C-terminus of VgrG to further sharpen the tube spike<sup>153</sup>. The sheath is made up of alternating TssB/C units<sup>151</sup>. This entire tubule assembles onto a baseplate and is anchored to a membrane-associated complex made up of three proteins, including VasK<sup>151</sup>. The capped tubule and baseplate are structurally and functional homologous to the tail-spike of T4 phage<sup>140</sup>. Similar to this contractile phage, upon contraction of the sheath, the inner tube is propelled forward out of the cell where the tube spike is thought to help penetrate target cells<sup>152</sup>.

There are numerous toxins that can be deployed by the type VI secretion system. Small effector molecules (10-20kDa) have been demonstrated to bind inside the Hcp hexameric structure, becoming part of the inner tube upon assembly<sup>164</sup>. The C-terminus of some VgrG proteins contain a toxic domain, thus receiving the name “evolved” or “extended” VgrGs<sup>133</sup>. Likewise, PAAR proteins can have toxins on the N- or C-terminal end of the PAAR domain that interacts with VgrG<sup>153</sup>. PAAR domains are also found in the N-terminus of

rearrangement hotspot (Rhs) proteins. Rhs proteins have a core of YD-repeats with a C-terminal toxin domain that is demarcated by conserved DPXGL motifs<sup>121</sup>. However, the reported PAAR-containing proteins that are required to support T6 activity are only ~95 residues while Rhs proteins are ~1400 residues<sup>153</sup>. Nevertheless, Rhs proteins are deployed by the T6SS, presumably by their PAAR domain interacting with a VgrG protein. There are also medium sized effectors that are too big to fit inside an Hcp tube, thus their mechanism of delivery is unknown. The small and medium sized effectors typically act in the periplasm where they target the cell wall or other membrane components<sup>128</sup>. Rhs proteins almost always encode nucleases that act in the cytoplasm, and PAAR proteins contain a mix of periplasmic- and cytoplasmic-acting toxins<sup>153</sup>. Ultimately, there are numerous diverse toxins deployed by the T6SS.

Despite the large variety of bacteria that contain T6SSs, functional T6S has only been experimentally shown in a handful of species. The field is still trying to understand the mechanism of T6 assembly, firing, and delivery of effectors into their proper compartments in target cells. Here we demonstrate that *Enterobacter cloacae* ATCC13047 (ECL) is a great model organism to study T6 secretion. We show it contains a functional T6S locus that is involved in inter- and intra-species competition. We establish that the two Rhs proteins found in ECL are functional and utilize specific VgrG proteins to deploy into target cells. We demonstrate that at least one Rhs protein is required for a functional T6 apparatus. Surprisingly, we find that full-length Rhs protein, and not just its PAAR domain, is required for functional T6 secretion in ECL. Furthermore, we demonstrate the assembly of the T6 organelle, and not just function is dependent on full-length Rhs. This suggests Rhs proteins

are more involved in the T6 assembly process than previously believed, and do more than just sharpen the tip of VgrG proteins.

## Materials and Methods

### *Bacterial strains and plasmids used in this study*

<i>Strains or plasmids</i>	<i>Description<sup>a</sup></i>	<i>Reference</i>
<b>Strain</b>		
X90	F' <i>lacI<sup>d</sup> lac' pro' araΔ(lac-pro) nalI argE(amb) rif<sup>R</sup> thi-1</i> , Rif <sup>R</sup>	177
CH6733	<i>Salmonella enterica</i> serovar <i>typhimurium</i> with genomically-encoded chloramphenicol resistance cassette flanked by Roth universal primers, Cm <sup>R</sup>	David Low
CH8160	<i>Alcaligenes faecalis</i> Rif <sup>R</sup> ATCC 35655	This study
CH8161	<i>Citrobacter freundii</i> Rif <sup>R</sup> ATCC 8090	118
CH8162	<i>Enterobacter aerogenes</i> Rif <sup>R</sup> ATCC 13048	118
CH7839	<i>Enterobacter cloacae</i> ATCC 13047	ATCC
CH8163	<i>Enterobacter cloacae</i> Rif <sup>R</sup> ATCC 13047	118
CH11641	<i>Serratia marcescens</i> Rif <sup>R</sup> ATCC 8100	118
CH8330	<i>Proteus mirabilis</i> Rif <sup>R</sup>	Dylan Peterson
CH11196	ECL $\Delta vasK1::kan$	This study
CH11396	ECL $\Delta vasK1$	This study
CH11615	ECL $\Delta vasK1$ Rif <sup>R</sup>	This study
CH12037	ECL $\Delta vasK2::kan$	This study
CH12038	ECL $\Delta vasK1 \Delta vasK2::kan$	This study
CH12384	ECL $\Delta vgrG1::kan$	This study
CH11436	ECL $\Delta vgrG2::kan$	This study
CH12414	ECL $\Delta vgrG2 \Delta vgrG1::kan$	This study
CH11199	ECL $\Delta hcp3::spec$	This study
CH11191	ECL $\Delta clpV1::kan$	This study
CH11202	ECL $\Delta tae4::spec$	This study
CH11204	ECL $\Delta tae4 \Delta tai4::kan$	This study
CH11205	ECL $\Delta tae4 \Delta tai4$	This study
CH11685	ECL $\Delta tae4 \Delta tai4$ Rif <sup>R</sup>	This study
CH11876	ECL $\Delta 01553::kan$	This study
CH11895	ECL $\Delta 01553 \Delta 01554::spec$	This study
CH11904	ECL $\Delta 01553 \Delta 01554::spec$ Rif <sup>R</sup>	This study
CH11178	ECL $\Delta rhs1::kan$	This study



CH11179	ECL $\Delta rhs1$	This study
CH11530	ECL $\Delta rhs1 rhs2 1-206::kan$	This study
CH11645	ECL $\Delta rhs1 rhs2 1-857::kan$	This study
CH11531	ECL $\Delta rhs1 rhs2 1-1161::kan$	This study
CH12486	ECL $\Delta rhs1 rhs2 1-1277::kan$	This study
CH11181	ECL $\Delta rhs1 \Delta rhsI1::kan$	This study
CH11182	ECL $\Delta rhs1 \Delta rhsI1$	This study
CH11185	ECL $\Delta rhs1 \Delta rhsI1 Rif^R$	This study
CH12415	ECL $\Delta rhs1 \Delta vgrG1::kan$	This study
CH12416	ECL $\Delta rhs1 \Delta vgrG2::kan$	This study
CH11186	ECL $\Delta rhs2::kan$	This study
CH11733	ECL $\Delta rhs2$	This study
CH12497	ECL $\Delta rhs2 rhs1 1-250::kan$	This study
CH12226	ECL $\Delta rhs2 rhs1 1-966::kan$	This study
CH11749	ECL $\Delta rhs2 rhs1 1-1323::kan$	This study
CH12492	ECL $\Delta rhs2 rhs2 1-1330::kan$	This study
CH11188	ECL $\Delta rhs2 \Delta rhsI2::spec$	This study
CH11223	ECL $\Delta rhs2 \Delta rhsI2::spec Rif^R$	This study
CH11748	ECL $\Delta rhs2 \Delta rhsI1::kan$	This study
CH11903	ECL $\Delta rhs2 \Delta rhs1$	This study
CH12385	ECL $\Delta rhs2 \Delta vgrG1::kan$	This study
CH12386	ECL $\Delta rhs2 \Delta vgrG2::kan$	This study
CH11221	ECL $\Delta rhs1 \Delta rhsI1 \Delta rhs2 \Delta rhsI2::spec$	This study
CH11472	ECL $\Delta hcp1::kan$	This study
CH11652	ECL $\Delta hcp2::kan$	This study
CH11653	ECL $\Delta hcp4::kan$	This study
CH11654	ECL $\Delta hcp5::kan$	This study
CH12482	ECL $tssC-GFP::kan$	This study
CH12549	ECL $\Delta rhs2 \Delta rhs1 tssC-GFP::kan$	This study
CH12031	ECL $spec\_araC\_pBAD\_T6SS-2$	This study
CH12039	ECL $\Delta vasK1 spec\_araC\_pBAD\_T6SS-2$	This study
CH11498	ECL $\Delta tep1$	This study
CH11500	ECL $\Delta tep1 \Delta tep1::kan$	This study
CH11670	ECL $\Delta tep2$	This study
CH11659	ECL $\Delta tep2 \Delta tep2::kan$	This study
CH11209	ECL $\Delta tep4$	This study
CH11211	ECL $\Delta tep4 \Delta tep4::kan$	This study
CH11675	ECL $\Delta tep5$	This study
CH11660	ECL $\Delta tep5 \Delta tep5::kan$	This study
CH11807	ECL $\Delta 01556$	
CH11830	ECL $\Delta 01556 \Delta 01557::kan$	This study
CH1347	ECL $\Delta 03144$	This study

CH11902	ECL $\Delta 03144 \Delta 03145::kan$	This study
CH11922	ECL $\Delta 02217 \Delta 02217imm. \Delta 04194$	This study
CH11943	ECL $\Delta 02217 \Delta 02217imm. \Delta 04194 \Delta 04194imm.::kan$	This study
<b>Plasmids</b>		
pTrc99a	IPTG-inducible expression plasmid, Amp <sup>R</sup>	GE Healthcare
pTrc99aCm	Cm <sup>R</sup> IPTG-inducible expression plasmid, Amp <sup>R</sup>	This study
pCP20	Heat-inducible expression of FLP recombinase, Cm <sup>R</sup> Amp <sup>R</sup>	225
pKOBEG	L-arabinose inducible expression phage I <i>red</i> genes, temperature-sensitive replication origin, Cm <sup>R</sup>	250
pKAN	pBluescript with kanamycin-resistance cassette ligated to SmaI site, Amp <sup>R</sup> Kan <sup>R</sup>	251
pSPM	pBluescript with spectinomycin-resistance cassette ligated between BamHI and EcoRI sites, Amp <sup>R</sup> Spc <sup>R</sup>	252
pCH450	pACYC184 derivative with <i>E. coli araBAD</i> promoter for arabinose-inducible expression, Tet <sup>R</sup>	249
pCH450K	pACYC184 derivative with <i>E. coli araBAD</i> promoter for arabinose-inducible expression with KpnI site, Tet <sup>R</sup>	This study
pZS21kan	pZS31 derivative carrying an ampicillin-resistance cassette, Kan <sup>R</sup>	118
pET21P:: <i>rhs1</i> (82-467)	Overproduces Rhs1(82-467)-His <sub>6</sub> under control of phage T7 promoter, Amp <sup>R</sup>	This study
pZS21kan:: <i>tai4</i>	Constitutive expression of Tai4, Kan <sup>R</sup>	This study
pZS21kan:: <i>rhsI1</i>	Constitutive expression of RhsI1, Kan <sup>R</sup>	This study
pTrc99aCm:: <i>0155</i> <sub>4</sub>	Constitutive expression of 01554, Cm <sup>R</sup>	This study
pTrc99aCm:: <i>rhsI2</i>	Constitutive expression of RhsI2, Cm <sup>R</sup>	This study
pCH450:: <i>vasK1</i>	Arabinose-inducible expression of VasK1, Tet <sup>R</sup>	This study
pCH450:: <i>hcp3</i>	Arabinose-inducible expression of Hcp3, Tet <sup>R</sup>	This study
pCH450K:: <i>clpV1</i>	Arabinose-inducible expression of ClpV1, Tet <sup>R</sup>	This study
pCH450K:: <i>hcp1</i>	Arabinose-inducible expression of Hcp1, Tet <sup>R</sup>	This study
pCH450K:: <i>hcp2</i>	Arabinose-inducible expression of Hcp2, Tet <sup>R</sup>	This study
pCH450K:: <i>hcp4</i>	Arabinose-inducible expression of Hcp4, Tet <sup>R</sup>	This study

pCH450:: <i>hcp5</i>	Arabinose-inducible expression of Hcp5, Tet <sup>R</sup>	This study
pCH450K:: <i>rhs1(1-230)</i>	Arabinose-inducible expression of Rhs1(1-230), Tet <sup>R</sup>	This study
pCH450K:: <i>rhs1(94-230)</i>	Arabinose-inducible expression of Rhs1(94-230), Tet <sup>R</sup>	This study
pCH450K:: <i>rhs1(1-467)</i>	Arabinose-inducible expression of Rhs1(1-467), Tet <sup>R</sup>	This study
pCH450:: <i>rhs1(82-467)</i>	Arabinose-inducible expression of Rhs1(82-467), Tet <sup>R</sup>	This study
pCH450K:: <i>rhs2(80-163)</i>	Arabinose-inducible expression of Rhs2(80-163), Tet <sup>R</sup>	This study
pCH450K:: <i>02217-02217imm.</i>	Arabinose-inducible expression of 02217-02217Imm., Tet <sup>R</sup>	This study
pCH450K:: <i>03143-03145</i>	Arabinose-inducible expression of 03143-03145, Tet <sup>R</sup>	This study
pCH450K:: <i>tep1-tip1-hcp1</i>	Arabinose-inducible expression of Tep1-Tip1-Hcp1, Tet <sup>R</sup>	This study
pCH450K:: <i>hcp2-tep2-tip2</i>	Arabinose-inducible expression of Hcp2-Tep2-Tip2, Tet <sup>R</sup>	This study
pCH450:: <i>hcp3-tae4-tai4</i>	Arabinose-inducible expression of Hcp3-Tae4-Tai4, Tet <sup>R</sup>	This study
pCH450:: <i>hcp4-tep4-tip4</i>	Arabinose-inducible expression of Hcp4-Tep4-Tip4, Tet <sup>R</sup>	This study
pCH450:: <i>tep5-tip5-hcp5</i>	Arabinose-inducible expression of Tep5-Tip5-Hcp5, Tet <sup>R</sup>	This study

<sup>a</sup>Abbreviations: Amp<sup>R</sup>, ampicillin resistant; Cm<sup>R</sup>, chloramphenicol resistant; Kan<sup>R</sup>, kanamycin resistant; Rif<sup>R</sup>, rifampicin resistant; Str<sup>R</sup>, streptomycin resistant; Tet<sup>R</sup>, tetracycline resistant

#### ***Oligonucleotides used in this study***

<i>tae4-up</i>	5' - GGG GAG CTC CCC AGC CAG GTA ATA TG - 3' 5' - CCA CTA GTT CTA GAG CGG CTT GTT TCT CCT TGA AAA G - 3'	This study
<i>tae4-ds</i>	5' - GCT TAT CGA TAC CGT CGA CAC CTT CTG GAG CCT GAA ATG - 3' 5' - TCT GAT AAT GAC CAG GCT CGG TAC C - 3'	This study
<i>tai4-ds</i>	5' - GCT TAT CGA TAC CGT CGA CTA GTA AAG ATG AAA TCG GC - 3' 5' - ATA GGT ACC GTC ACT TCG ATG CGG - 3'	This study
<i>01553-up</i>	5' - CAA GAG CTC CGG GAT GGT TGC C - 3' 5' - ATT GGA TCC GTC CTG TTA CCA GTC - 3'	This study

<i>01553</i> -ds	5' - AGG CTC GAG ACA TT CAA TTA TTA GG - 3' (XhoI) 5' - AAC GGT ACC TGG CGA TAA ACC CGC - 3'	This study
<i>01554</i> -ds	5' - CCA ACT CGA GTT AAA TAG GAA ACG - 3' (XhoI) 5' - CCA GGT ACC AAA GTG CTG TGT GC - 3'	This study
<i>clpV1</i> -up	5' - TTT GAG CTC GGT CAA ACT GAT CGC CAA C - 3' 5' - TTT GGA TCC GTT TCC ATT AAC AAT GAT TCG G - 3'	This study
<i>clpV1</i> -ds	5' - TTT GAA TTC CAA GAG AAC TGT AAA AGA TGA C - 3' 5' - TTT GGT ACC GCC TCG TTA ATC AGC TCC - 3'	This study
<i>vasK1</i> -up	5'- TTT GAG CTC GAA ATC GAC GCC GGT CTG -3' 5'- TTT GGA TCC TTT CCT TGC GGC AAT CCG -3'	This study
<i>vasK1</i> -ds	5'- TTT GAA TTC CAA GGA CAG CCG TAT GAC - 3' 5'- TTT GGT ACC GAA TCG ACA TCA GCA TCT C - 3'	This study
<i>vasK2</i> -up	5' - TTT GAG CTC GGC AAC CGC CTG ACA C - 3' 5' - CCA CTA GTT CTA GAG CGG CTT CCG TAG TCT TCG GTG C - 3'	This study
<i>vasK2</i> -ds	5' - GCT TAT CGA TAC CGT CGA CGG ACA GTA CGG AAA GCA G - 3' 5' - TTT GGT ACC GCC GAG CCA TTC - 3'	This study
<i>vgrG1</i> -up	5' - GTT GAG CTC GTT ATG GAT GTC ATT TTG TCA ATC 5' - AAT GGA TCC GAG CAT AAT CGT TAT TCC GTA ATG	This study
<i>vgrG1</i> -ds	5' - AAG GAA TTC AAT AAG TAA ACG TAA TTA GAA AC 5' - CTT GGT ACC AGC AAA AGT TCG ATT TAT TCA AC	This study
<i>vgrG2</i> -up	5' - TTT GAG CTC CCC TTG CTA CGG CCA AAC -3' 5' - TTT GGA TCC TCG TTA TTC CAC TAT GGG C - 3'	This study
<i>vgrG2</i> -ds	5' - TTT CTC GAG GCT GGA GCG GTG CTT G - 3'(XhoI) 5' - TTT GGT ACC CGA GTC CAG ACA ATC AGG - 3'	This study
<i>hcp1</i> -up	5' - cagGAGCTCtagcgatagcatggacg 5' - ccactagtctagagcggcgatttgacatacaactcc	This study
<i>hcp1</i> -ds	5' - atgGAATTCcaacatactaaaacgtgatccc 5' - agcGGTACCaccgcgctccctgc	This study

<i>hcp2</i> -up	5' - catGAGCTCccctcttccctcgctc 5' - gggGGATCCcattctgaaagctccttttcag	This study
<i>hcp2</i> -ds	5' - attGAATTCcagacagatgatccgttgc 5' - agaGGTACCctgagtgaacgggttcatc	This study
<i>hcp3</i> -up	5' - TTT GAG CTC CCA GGT GCA GGA GAT TC - 3' 5' - CCA CTA GTT CTA GAG CGG CTA CTC TTC GTC GAT GAA C - 3'	This study
<i>hcp3</i> -ds	5' - GCT TAT CGA TAC CGT CGA CGT AGT GGG TCC GAA AGG G - 3' 5' - AAA GGT ACC TTC CAG AGT GTT ACA TGC - 3'	This study
<i>hcp4</i> -up	5' - cacGAGCTCtctctgatttccgctgc 5' - aggGGATCCcatagtctactcatcatcatgt	This study
<i>hcp4</i> -ds	5' - tagGAATTCagagcgttaattatgcgtactc 5' - catGGTACCctgcaggttttcatacacg	This study
<i>rhs1</i> -up	5' - TTT GAG CTC ATA CAC CCT CCA GGA AGG - 3' 5' - TTT GGA TCC GCC TTA CAC ATT CCG GTT G - 3'	This study
<i>rhs1</i> -ds	5' - GAA GAA TTC TGG CAA GAG GAT TAC TTA ATG - 3' 5' - TTT GGT ACC CAT CAT TAG TAA TGC AAA G - 3'	This study
<i>rhs11</i> -ds	5' - TTT GAA TTC GTT AAT GTT CGA CCA CAA TTG - 3' 5' - TTT GGT ACC CAT GCC GAT GTC GAT TAA AG - 3'	This study
<i>rhs2</i> -up	5'- TTT GAG CTC ACC CGC TCA ATG TCA GAA C - 3' 5'- TTT GAA TCC CCC TGG TGT TAA TGG TGG -3'	This study
<i>rhs2</i> -ds	5'- TTT GAA TTC CAA TGA ATA TGC TGA ATG TGA G -3' 5'- TTT GGT ACC ACT TCG TCA TTA TCA TCT GC -3'	This study
<i>rhs12</i> -ds	5'- TTT CTC GAG ATC CCG GAA ACC ATT AAG -3' (XhoI) 5'- TTT GGT ACC CAT CCT CTG CAT GAA GAT G - 3'	This study
<i>rhs2(206)</i> -up	5' - GCC GAG CTC GGC GCA TCC TGC CTT GGC - 3' 5' - CAG GGA TCC TGC CCA GCT ACT TAG AGA GCG C - 3'	This study

<i>rhs2(857)</i> -up	5' - TTT GAG CTC TGC TGA GTG CCG TGA TC - 3' 5' - TTT GGA TCC ACT AGC TTT CGA TAC CCA GCG C - 3'	This study
<i>rhs2(1161)</i> -up	5' - TAC GAG CTC GAA GGG CGT CTG CTG AAG C - 3' 5' - TGT GGA TCC AGT AAA TCT AAC CGC TGC TCT GG - 3'	This study
<i>rhs2(1277)</i> -up	5' - TAC GAG CTC GAA GGG CGT CTG CTG AAG C - 3' 5' - AAC GCG GCC GCT TAC GCG GAG AGT CCC CAC (NotI)	This study
<i>rhs1(250)</i> -up	5' - GCG GAG CTC CTA ACC TGG CGG GTG 5' - GAG GGA TCC TTA CCC CAG TGC CAG C	This study
<i>rhs1(966)</i> -up	5' - GGT GAG CTC CCG CTG GGA CAG C 5' - GCT GGA TCC TTA CCC GCT GCC GTA G	This study
<i>rhs1(1323)</i> -up	5' - TTT GAG CTC ACA GAA GTG ATC AGC CAG - 3' 5' - TTT GGA TCC CTA TAT TCG GGT TAG ACT ATT AGC - 3'	This study
<i>rhs1(1330)</i> -up	5' - TAA GAG CTC GCG GGA AGA ACG GG - 3' 5' - CCT GCG GCC GCG AGT CCC CAC GGA TCA ATC - 3' (NotI)	This study
<i>tep1</i> -up	5' - gttGAGCTCctgacggcaccacc 5' - agaGGATCCcataacgtatccatactgttttatgg	This study
<i>tep1</i> -ds	5' - gatGAATTCgctatctataacgacttaaaggataaatag 5' - aaaGGTACCagggatcacggttttagtatggtg	This study
<i>tip1</i> -ds	5' - attGAATTCgctttgataagtagcattgagttta 5' - aaaGGTACCagggatcacggttttagtatggtg	This study
<i>tep2</i> -up	5' - acaGAGCTCtcaactcactgaaaaggagc 5' - gcaGGATCCcatctgtctgccaagaatc	This study
<i>tep2</i> -ds	5' - gcaGAATTCaggtctactgatgaaacc 5' - atcCTCGAGtgcaaaaaataacgttgactcatc (XhoI)	This study
<i>tip2</i> -ds	5' - aacGAATTCatattgattagcgcattgtgaaag 5' - atcCTCGAGtgcaaaaaataacgttgactcatc	This study
<i>tep4</i> -up	5' - CAC GAG CTC GGT AGA TGT TCA TGA TC - 3' 5' - CCA CTA GTT CTA GAG CGG AAT TAA CGC TCT GCC CAG C - 3'	This study
<i>tep4</i> -ds	5' - GCT TAT CGA TAC CGT CGA CGG TTT TTA ACT GGC GAG TC - 3' 5' - CCC GGT ACC TAG CGT TGA TGA TCA G - 3'	This study
<i>tip4</i> -ds	5' - GCT TAT CGA TAC CGT CGA CTC ATA AAC	This study

	GCT ATT TAC CG -3' 5' - CCA CCG GTA CCG ACA GCG CAA GG - 3'	
<i>tep5-up</i>	5' - ttaGAGCTCgtcatatgacagttccattaagt 5' - gacGGATCCcatacacacctcatagccatt	This study
<i>tep5-ds</i>	5' - catGAATTCgaccggatattgtaactgaac 5' - cacCTCGAGccgtagctcgctcattc (XhoI)	This study
<i>tip5-ds</i>	5' - tggGAATTCgattaataaggaggaaaatgatggc 5' - cacCTCGAGccgtagctcgctcattc	This study
<i>02217-up</i>	5' - CAA GAG CTC GCG TGA GCA TGC GAC - 3' 5' - TTA GGA TCC GGT AAT TAG TAA ATT G - 3'	This study
<i>02217-ds</i>	5' - TAA CTC GAG AAG CGG GGC TGT ACA AAT G - 3' (XhoI) 5' - TTG AGA ATG GTA CCA GAA AAG CCC - 3'	This study
<i>02217imm.-ds</i>	5' - CTT CTC GAG TAT AAT TCT CAA TTC TC - 3' (XhoI) 5' - GGT CGG TAC CGC ACT CTG AAG CGC - 3'	This study
<i>04194-up</i>	5' - GCC GAG CTC CCG CCG CAA CTG C 5' - GTT GGA TCC CAG CGG GTG AAC AAC AAC	This study
<i>04194-ds</i>	5' - TAC GAA TTC TGA CAG TGA ATG TTG AAG CG 5' - TAT GGT ACC CTG CAA AAA GCC CCT ACC	This study
<i>04194imm.-ds</i>	5' - GCA GAA TTG GTA GTG TTC AAT ATA AGC CCC G 5' - TCC GGT ACC CTG GAA CTG AAG CAG GC	This study
<i>03144-up</i>	5' - GGT GAG CTC CGC ATA TGT GTT TAA GG - 3' 5' - CAT GGA TCC CTC TAC TTT ATA TGG - 3'	This study
<i>03144-ds</i>	5' - AGA GAA TTC ATG AGG TTG TTA AAT AA - 3' 5' - TCC GGT ACC TTT GCT TAA AGG G - 3'	This study
<i>03145-ds</i>	5' - CAA GAA TTC AGG AAA AAA TTG ATT TTA - 3' 5' - TTA GGT ACC TCG ATC CTT GCC G - 3'	This study
<i>01556-up</i>	5' - ACC GAG CTC TGG CTA ATC AGC GAA TG - 3' 5' - ACT GGA TCC CTC ATT TAA TCG ATT CG - 3'	This study
<i>01556-ds</i>	5' - GAA GAA TTC CCT GAC GAG TTT TGA G - 3' 5' - CAA GGT ACC GTC TGC TTA ATT TCG - 3'	This study
<i>01557-ds</i>	5' - ATC GAA TTC AAG CAG ACG GTG AC - 3' 5' - TGC GGT ACC GTC TGC CCC TGG - 3'	This study
<i>Tai4-EcoRI- for</i>	5' - TTT GAA TTC TTC TGG AGC CTG AAA TGA AAA AG - 3'	This study
<i>Tai4-BamHI- rev</i>	5' - TTT GGA TCC CTA CTT TGA GGA TTT GAG TGG - 3'	This study
<i>01554-KpnI-</i>	5' - GAG GGT ACC ATG AAA TCG TTC TTA TCA	This study

for	GGC - 3'	
01554-XhoI-rev	5' - ATA CTC GAG CTA TTT AAC CGG AGT TGG TG - 3'	This study
VasK1-NcoI-for	5' - TTT CCA TGG TGA CGA CTC TTC TTT C - 3'	This study
VasK1-XhoI-rev	5' - TTT CTC GAG TTA TGG GCA TGA GAA ACG - 3'	This study
Hcp1-KpnI-for	5' - TTT GGT ACC ATG TCA AAT CCG GCT TAT TTG - 3'	This study
Hcp1-XhoI-rev	5' - TTT CTC GAG TTT AGT ATG TTG CTC GCT C - 3'	This study
Hcp2-KpnI-for	5' - TTT GGT ACC ATG GCT ATA CCC GCA TAT C - 3'	This study
Hcp2-XhoI-rev	5' - TTT CTC GAG ATC ATC TGT CTG CCC AAG - 3'	This study
Hcp3-NcoI-for	5' - ATA CCA TGG CTA TTG ATA TGT TTC - 3'	This study
Hcp3-XhoI-rev	5' - TTT CTC GAG CAC TAC TAT TAT GCT TCT TTG - 3'	This study
Hcp4-KpnI-for	5' - GAG GGT ACC ATG GCA ATT CCT GTA TAT CTT TTC - 3'	This study
Hcp4-XhoI-rev	5' - TTT CTC GAG AAT TAA CGC TCT GCC CAG - 3'	
Hcp5-NcoI-for	5' - TTT CCA TGG CTG TAC CGG TCC - 3'	This study
Hcp5-XhoI-rev	5' - TTT CTC GAG GCA CTT CAA ACC GTA GC - 3'	This study
RhsI1-EcoRI-for	5' - TTT GAA TTC GGC AAG AGG ATT ACT TAA TG - 3'	This study
RhsI1-BamHI-rev	5' - TTT GGA TCC CGA ACA TTA ACA TAT TAA ATC G - 3'	This study
RhsI2-KpnI-for	5' - TTT GGT ACC ATG AGT GAA ATT GAA GAG TC - 3'	This study
RhsI2-XhoI-rev	5' - TTT CTC GAG GTA TCC TAG CCA TAA AAA TAA TC - 3'	This study
02217-KpnI-for	5' - GCG GGT ACC ATG GGT TAC CGT ACC G	This study
02217Imm.-XhoI-for	5' - TGT CTC GAG CAT TAC CAG CGA AGC TC	This study
03143-KpnI-for	5' - CTC GGT ACC ATG CAG CAA TAC ACC AAC	This study
03145-XhoI-for	5' - CTT CTC GAG ATG CTA TCG AAA ATA CAA C	This study
Rhs1-KpnI-for	5' - GTG GGT ACC ATG AGC GAT AAC AAC GCG GCC - 3'	This study
Rhs1(94)-	5' - GCG GGT ACC ATG GGT GAT GCG CTG TTC C	This study



KpnI-for		
Rhs1(82)- NcoI-for	5' - TGG CCA TGG TTA CTG ACG ATA TCA G - 3'	This study
Rhs1(230)- XhoI-rev	5' - CAG CTC GAG TTA GCC GAT GAT CAC ATT CG	This study
Rhs1(467)- XhoI-rev	5' - AAA CTC GAG TTA GGC AGC GGT TAC GCG	This study
Rhs2(80)- KpnI-for	5' - CG GGT ACC ATG GCC GAT GCC GGG G	This study
Rhs2(163)- XhoI-rev	5' - CTG CTC GAG TTA GCC GAT GAA GAC GTT GTG	This study

### ***Bacterial strains, plasmids, and growth conditions***

All bacterial strains and plasmids used in this study are listed in the table above. Bacteria were grown in LB media or LB-agar with antibiotics at the following concentrations: ampicillin (Amp) 150 µg/mL; kanamycin (Kan) 50 µg/mL; rifampicin (Rif) 200 µg/mL; spectinomycin (Spc) 100 µg/mL; and tetracycline (Tet) 10 µg/mL. The listed primer pairs were used to amplify ~400bp up or downstream of the designated gene. The upstream product was digested with SacI/BamHI, downstream product with EcoRI/KpnI (unless noted otherwise) and ligated into pKAN or pSPM to flank kanamycin- or spectinomycin-resistance cassettes, respectively. *tae4*, *tai4*, *vasK2*, *hcp3*, *tep4*, and *tip4* constructs were produced by OE-PCR. The resulting plasmids were linearized by restriction endonuclease digestion and electroporated into *E. cloacae* cells expressing the phage λ Red proteins from plasmid pKOBEG<sup>250,253</sup>. Kan cassettes were removed using pCP20<sup>225</sup>. For construction of the plasmids listed in the first table above, the primer pairs listed in the second table were used for PCR off *E. cloacae* gDNA. The PCR products were digested with the annotated enzymes and ligated into their respective plasmids.

### ***Growth competitions***

Donor and inhibitor cells were grown to late-log phase with aeration in liquid LB. Cells were spun down, concentrated with M9 salts, and mixed at a 1:1 ratio at OD600 17:17. 100uL of the mixture was plated onto LB-agar plates and incubated at 37°C for 4 hours. Cells were harvested with 1.5mLs of M9 salts, serially diluted, and plated on selective agar plates for colony counts.

### ***Western blots analysis***

Cells were grown in 2mLs of liquid LB supplemented with 0.2% arabinose where indicated to mid-log with aeration. Cells were lysed in urea lysis buffer or RIPA buffer. Supernatants were ethanol precipitated for at least 2 hours at -20°C. Equal amounts of protein were subjected to 10% acrylamide SDS-PAGE and transferred onto nitrocellulose or PVDF membrane and incubated with 1:10,000 anti-Rhs1 or 1:25,000 anti-Hcp3.

## **Results**

### ***T6SS-1 mediates inter-species competition***

ECL contains 2 T6SS loci that we have named T6SS-1 and -2 based on their chronological location in the genome (Fig. 1A & 1B). To determine if either of these loci is functional in bacterial competition, WT, *vasK1*<sup>-</sup> and *vasK2*<sup>-</sup> donor cells were assayed for their ability to inhibit the growth of *Escherichia coli* recipients. Competition data will be represented as the competitive index of donor over recipient cells unless noted otherwise. We found that WT and *vasK2*<sup>-</sup> cells kill *E. coli* 3-4 logs while *vasK1*<sup>-</sup> cells minimally hindered *E. coli* growth (Fig. 2A). We note this may be due to the slightly slower growth rate of *E. coli*

compared to ECL under the conditions tested (data not shown). These experiments demonstrate that T6SS-1, but not T6SS-2 is on and functional under laboratory conditions.

Next we wanted to see if T6SS-1 gave ECL a competitive advantage against other species of bacteria. ECL, *vasK1*<sup>-</sup> or *vasK2*<sup>-</sup> cells were competed against various proteobacteria. Only *Proteus mirabilis* was able to hinder the growth of ECL during co-culture, while ECL dominated all other species tested (Fig. 2A). We note that *P. mirabilis* has its own T6SS that may be facilitating this competitive advantage. Furthermore, it seems ECL fares better against *P. mirabilis* when it has a functional T6SS (Fig. 2A). These experiments further demonstrate that T6SS-1 mediates inter-species competition and is efficient to outcompete a variety of proteobacteria under laboratory conditions.

### ***Tae4*<sup>ECL</sup> and ECL01553 are deployed by T6SS-1**

Next we examined what toxins could be deployed by T6SS-1 to facilitate this competitive advantage. Tae4-Tai4 is a T6 effector-immunity (E-I) pair that has previously been described to function as an amidase, acting in the periplasm to cause cell lysis. ECL contains this E-I pair in T6SS-1, immediately downstream of the only Hcp in the locus (Hcp3) (Fig. 1A). Therefore, we assumed this effector is deployed by T6SS-1. To test this, we made *Δtae4Δtai4* recipient cells and asked how they fared in competition against various donor cells. Consistent with Tae4 serving as a functional amidase, *Δtae4Δtai4* cells were less fit than WT donors, and this decrease in fitness was dependent on *tae4*, *vask1*, *hcp3*, *clpV1* and at least one *vgrG* in donor cells (Fig. 3). Furthermore, expression of Tai4 restored their fitness defect (Fig. 3A), suggesting Tae4-Tai4 serves as a functional T6SS E-I pair that is deployed by T6SS-1.

Further examination of the T6SS-1 locus reveals ECL01553 has predicted lipase function. Mougous identified and characterized 5 classes of lipase proteins that act as effectors. ECL01553 has the characteristic GX SXG motif that is present in classes 1-4<sup>165</sup>. Consistent with other T6 effectors, which are found in bicistrons with cognate immunity genes, ECL10554 is found immediately downstream and has a predicted signal sequence that would direct it to the periplasm (Fig. 1A). Based on these observations, we hypothesized that ECL01553 encodes an effector that targets the phospholipid bilayer while ECL01554 encodes for an immunity protein capable of inactivating ECL01553. To test this, we used the same approach as described above and competed  $\Delta I0553 \Delta I0554$  recipients against various T6SS<sup>+</sup> and T6SS<sup>-</sup> donors. WT donors have a 3-log competitive advantage over these recipients, and this is abolished when donors lack *01553* or when recipients express *01554* (Fig. 4A). The fitness advantage was also dependent on donors expressing T6SS-1 genes (Fig. 4B & 4C). Thus, *01553-01554* also act as an effective T6SS-1 E-I pair that is able to severely inhibit the growth of recipient cells.

### ***Rhs1 and Rhs2 are functional effectors deployed by T6SS-1***

Of note, at the time of these experiments, it was unclear if Rhs proteins were delivered via the T6SS. It had been demonstrated in *Dickeya Dadantii* that donor cells required both Hcp and VgrG to inhibit  $\Delta rhsA \Delta rhsIA$  or  $\Delta rhsB \Delta rhsIB$  recipient cells<sup>252</sup>. However, it was thought that Rhs proteins get decorated by Hcp, capped by VgrG, and exported out of donor cells independent of the T6SS organelle. Furthermore, the interaction between the C-terminus of VgrG and PAAR domains was not discovered yet. This section is presented in a

manner where it is unclear how Rhs proteins are deployed, because at the time of these experiments, this was actually the case.

Rhs proteins have previously been reported as potent contributors to bacterial competition. The genome of ECL contains two Rhs proteins, named Rhs1 and Rhs2, which each reside upstream of genes encoding predicted cognate immunity proteins (Fig. 1A & 1C). To test if Rhs1 or Rhs2 are used in bacterial competition,  $\Delta rhs1 \Delta rhsI1$  or  $\Delta rhs2 \Delta rhsI2$  cells with or without plasmid-borne copies of cognate immunity proteins were competed against WT, and  $rhs1^-$  or  $rhs2^-$  donor cells. WT donors inhibited the growth of  $\Delta rhs1 \Delta rhsI1$  and  $\Delta rhs2 \Delta rhsI2$  recipients by 2 and 3 logs, respectively (Fig. 5A). Moreover, this inhibition was abrogated when target cells were expressing their cognate immunity protein, or when the cognate Rhs was missing in the donor cell (Fig. 5A). These results demonstrate that both Rhs1 and Rhs2 can be used in intra-species bacterial competition.

As previously mentioned, the toxic activity of Rhs proteins reside in the C-terminus (Rhs-CT)<sup>121</sup>. RhsI1 and RhsI2 are both predicted cytoplasmic proteins, so we reasoned Rhs1-CT and Rhs2-CT both act in the cytoplasm. Furthermore, bioinformatic analysis reveals Rhs1-CT contains an HNH nuclease domain, while Rhs2 is predicted to have DNase activity, so the substrates of each protein are located in the cytoplasm. In accord with these observations, expressing Rhs1-CT or Rhs2-CT in the cytoplasm of *E. coli* cells results in cell death while *E. coli* cells expressing cognate immunity proteins survive (data not shown). These data suggest Rhs1 and Rhs2 both mediate bacterial competition by targeting substrates in the cytoplasm of recipient cells.

Rhs1 is located directly downstream of T6SS-1 while Rhs2 resides as a distant site in the chromosome, unlinked to any other T6SS gene (Fig. 1A & 1C). We wanted to determine if

the entire T6SS-1 is required for Rhs1 or Rhs2 deployment, or if only VgrG and Hcp are required, like previously reported in *D. dadantii*. First, we confirmed that both *hcp3*<sup>-</sup> and *vgrG*<sup>-</sup> donor cells are not able to deliver Rhs1 or Rhs2 (Fig 5B & 7B). Next, we competed *vasKI*<sup>-</sup> and *clpVI*<sup>-</sup> donors against  $\Delta$ *rhs1* $\Delta$ *rhsI1* or  $\Delta$ *rhs2* $\Delta$ *rhsI2* recipient cells (Fig. 5B). VasK and ClpV are both essential components needed for proper T6SS assembly and disassembly, respectively. Donors lacking either essential component are no longer able to inhibit  $\Delta$ *rhs1* $\Delta$ *rhsI1* or  $\Delta$ *rhs2* $\Delta$ *rhsI2* recipient cells, but inhibition can be fully restored when *vasKI*<sup>-</sup> cells are complemented with plasmid-borne VasK1 and partially restored when *clpVI*<sup>-</sup> donors are complemented (Fig. 5B). Thus, we conclude that Rhs1 and Rhs2 both require the entire T6S apparatus, and not just VgrG and Hcp, for proper delivery into target cells.

ECL contains 5 Hcp alleles scattered throughout its genome, but only Hcp3 is linked to other T6 components. We were curious if any other Hcp protein was required for T6SS-1 function. To test this, we made individual knock outs of each *hcp* allele in ECL and competed these donors against *E. coli*,  $\Delta$ *rhs1* $\Delta$ *rhsI1*,  $\Delta$ *rhs2* $\Delta$ *rhsI2*, and  $\Delta$ 01553 $\Delta$ 01554 recipients to look for effects on inter-species, Rhs, and T6 effector delivery, respectively. We found that only *hcp3*<sup>-</sup> donors lost their ability to inhibit any of the recipient cells (Fig. 6A). Next we examined if we could complement *hcp3*<sup>-</sup> donors with the other 4 *hcp* alleles. However, the inhibitory phenotype of *hcp3*<sup>-</sup> donors could only be restored by expressing Hcp3, and not any of other *hcp* allele (Fig. 6B). An alignment of all 5 Hcp proteins shows that Hcp3 is the prevalent outlier of the group (Fig. 6C). Additionally, Hcp3 is the only allele present in T6SS-1 (Fig. 1A). These data suggest Hcp3 is not only required for T6S, but also the only Hcp in ECL capable of making a functional T6SS-1.

When examining Rhs deployment in *vgrG*<sup>-</sup> cells, we noticed that *vgrG1*<sup>-</sup> donors could inhibit  $\Delta rhs1\Delta rhs11$  recipients, but were unable to inhibit  $\Delta rhs2\Delta rhs12$  recipients (Fig. 7B). Conversely, *vgrG2*<sup>-</sup> donors showed the opposite phenotype where they were unable to inhibit  $\Delta rhs1\Delta rhs11$  recipients, but could inhibit  $\Delta rhs2\Delta rhs12$  recipients (Fig. 7B). These results demonstrate that Rhs1 requires VgrG2 for deployment, while Rhs2 requires VgrG1. A closer analysis of the two VgrG proteins reveals that the two proteins are 100% identical until their C-termini where VgrG1 extends an extra 106 residues (Fig. 7A). Furthermore, Rhs1 and Rhs2 proteins are divergent, with only 29% sequence homology. Therefore, we hypothesized Rhs1 and Rhs2 were interacting specifically with the C-terminus of VgrG1 and VgrG2, respectively. Before any subsequent experiments could be performed to prove our model, the laboratory of John Mekalanos (Schneider *et al.*) revealed the C-terminus of VgrG proteins bind to PAAR-containing proteins in *Vibrio cholera* and *E. coli*<sup>153</sup>. Rhs1 and Rhs2 each contain distinct PAAR domains at their N-termini, consistent with our model of Rhs1 and Rhs2 interacting with the divergent C-termini of VgrG1 and VgrG2 in ECL. Taken all together, these results demonstrate that Rhs1 and Rhs2 are T6SS-1 effectors that mediate intra-species competition via the toxic domains in their C-termini.

### ***Rhs proteins are required for T6S***

Schneider *et al.* also reported that at least one PAAR-containing protein is required in *V. cholera* and *Acinetobacter baylyi* for Hcp secretion and the ability to inhibit *E. coli* cells<sup>153</sup>. PAAR motifs coordinate zinc ions and use hydrophobic packing to bring PAAR-containing proteins to a sharp point<sup>153</sup>. Therefore, Schneider *et al.* proposed that PAAR proteins are a core component of the T6SS that act by sharpening the tip of the VgrG spike<sup>153</sup>. We had

noticed in our experiments that *E. coli* cells have the same fitness against *rhsI<sup>-</sup>rhs2<sup>-</sup>* donor cells as they do against *vasKI<sup>-</sup>* cells (Fig. 1A & 8A). We found this unusual because this suggests Rhs1 and Rhs2 are the only T6 effectors effective against *E. coli*. We had demonstrated Tae4 and 01553 are functional T6 effectors, and other T6 effectors have been shown to kill *E. coli* cells. If Rhs1 and Rhs 2 are the only T6S effectors mediating the inhibition of *E. coli*, then *E. coli* cells expressing both RhsI1 and RhsI2 should be resistant to WT ECL. However, this was not the case. ECL was still able to inhibit the growth of these cells, suggesting other T6 effectors are being deployed by T6SS-1 to inhibit the growth of *E. coli*, but in a manner that requires the presence of at least one Rhs protein in ECL (Fig. 8A). Of note, the same immunity expression constructs are able to protect  $\Delta r h s 1 \Delta r h s I 1 \Delta r h s 2 \Delta r h s I 2$  cells from inhibition against WT ECL, indicating the constructs are able to protect recipients from Rhs1 and Rhs2 activity (Fig, 8A). Furthermore, we showed that delivery of 01553 into recipient cells and proper Hcp3 secretions is dependent on at least one cognate Rhs/VgrG pair (Fig. 8B & 8C). Therefore, we conclude that the presence of either Rhs1 or Rhs2 in ECL is required for functional T6S.

The above results are not unreasonable considering both Rhs proteins have a PAAR domain, and Schneider *et al.* demonstrated the requirement of at least one PAAR-containing protein for proper T6 function<sup>153</sup>. However, ECL contains multiple additional PAAR-containing proteins encoded in its genome. It seems that in ECL, the PAAR domain required for T6SS-1 function is present in its Rhs proteins. It is conceivable that perhaps these other PAAR proteins are not expressed and therefore do not have the opportunity to participate in T6SS-1 function. To rule this out, we complemented *rhsI<sup>-</sup>rhs2<sup>-</sup>* donor cells with plasmid-borne PAAR proteins. However, these constructs were unable to restore T6SS-1 function



(Fig. 10A), again suggesting the PAAR domains in the Rhs proteins are required for T6 function in ECL. However, expression of the PAAR domains of Rhs1 or Rhs2 was also not sufficient to complement *rhs1<sup>-</sup> rhs2<sup>-</sup>* donor cells (Fig. 10A). Western blot analysis reveals at least some of the Rhs1 PAAR constructs are being expressed in donor cells. This led us to believe we did not have proper expression levels of the Rhs PAAR domains, and moved on to the subsequent experiments.

Interestingly, the PAAR containing proteins in *V. cholera* and *E. coli* are only ~95 residues and they are able to support T6 function, while Rhs1 and Rhs2 in ECL are ~1400 residues<sup>153</sup>. If it is only the PAAR motifs that are required for T6 function, then only the PAAR domain of Rhs should be required to restore T6 activity. To test this, we inserted stop codons in Rhs1 and Rhs2 after the predicted PAAR domains in each protein (Rhs-PAAR) (Fig. 9A). Surprisingly, we found these donor cells could not inhibit the growth of *E. coli* or  $\Delta I0553\Delta I0554$  recipients (Fig. 9B). We reasoned we might have hindered the fold of the PAAR protein by removing its C-terminus, so we inserted a stop codon after the YD-repeat motifs of each Rhs (Rhs-YD) (Fig. 9A). Again, surprisingly, these donor cells also failed to inhibit the growth of either recipient tested (Fig. 9B). Finally, when we insert a stop codon just upstream of the toxin domain (Rhs  $\Delta$ CT), but downstream of the conserved DPXGL motifs (Fig. 9A), are we able to restore full delivery of 01553 into recipients, and full *E. coli* inhibition (Fig. 9B). Interestingly, if we insert the stop codon in Rhs1 just 7 residues upstream of the toxin domain (Rhs DDPXL) and therefore remove some of the DPXGL motif (Fig. 9A), we lose some inhibitory activity (Fig. 9B). Consistent with the competition results, Hcp3 secretion is only fully restored in the mutants that contain stop codon insertions immediately upstream of the toxin domain (Fig. 9C).

It is still formerly possible that these truncated Rhs proteins are not being expressed, or we have disrupted their proper fold, therefore resulting in their destabilization and subsequent degradation. To rule this out, we extracted protein from each of the Rhs1 donor cells tested, and performed a western blot with antibodies raised to residues 82-467 of Rhs1. A prominent band corresponding to the correct size is present in Rhs1-PAAR cells, suggesting this protein is expressed and stable in these donor cells (Fig. 9D). No band was detected in Rhs-YD cells, but there is a high molecular weight band unique to  $\Delta$ Rhs2, Rhs1-DDPXL, and Rhs1- $\Delta$ CT samples (Fig. 9D). While this does not correspond to full-length protein, it suggests all three cell lines are expressing Rhs1 to the same level. Taken all together, these results demonstrate that full-length Rhs protein, and not just its PAAR domain, is required for functional T6S in ECL.

### ***Rhs proteins are required for T6 assembly***

We have demonstrated that *rhs1<sup>-</sup> rhs2<sup>-</sup>* donors have abolished T6S activity based on their inability to inhibit *E. coli* cells, deliver 01553, and secrete Hcp3 into the supernatant. Next we wanted to know if the T6SS apparatus was still able to assemble, just not fire in the absence of Rhs. Previous studies have reported the ability to track T6 firing events in real-time by fusing fluorescent proteins to TssC, one of the sheath proteins<sup>139</sup>. We took a similar approach to look at T6 assembly and firing in ECL. By fusing GFP to the C-terminus of TssC, we were able to see both punctate and dynamic green fluorescence (Fig. 11). We gather that the punctate spots represent seeding T6SSs and the dynamic fluorescence represents assembly and firing events because *hcp3<sup>-</sup>* and *vgrG<sup>-</sup>* cells only contain diffuse green fluorescence (data not shown). Surprisingly, *rhs1<sup>-</sup> rhs2<sup>-</sup>* cells contain neither punctate

nor dynamic green spots, indicating there is no assembly of the T6S organelle in *rhs*<sup>-</sup> cells (Fig. 11).

### ***T6SS-2 in nonfunctional***

We examined the genome of ECL for more T6S E-I pairs. Interestingly, we found that immediately up or downstream of the other 4 *hcp* alleles present in ECL, there are always two unannotated open reading frames. Further analysis of these ORFs reveals they encode for potential E-I pairs. Each effector was found in other T6SS<sup>+</sup> bacteria in a bicistron with its presumed immunity protein. Furthermore, the Hcp protein that is linked to each effector in ECL is also found in these other T6SS<sup>+</sup> strains, either linked or unlinked to the E-I pair. Therefore, we presumed these were E-I pairs that bound to their linked Hcp protein to be delivered via the T6SS. We named the E-I pairs Tep-Tip for type VI effector protein and type VI immunity protein. To test this, we deleted each Tep-Tip pair and asked if the parental strain could inhibit their growth. We found that each Tep-Tip deletion strain has no fitness defect against parental donors, thus concluding the effectors are not being deployed into recipient cells (Fig. 12A). However, if the Tep-Tip loci are not being expressed then this could explain the lack of inhibition seen. To rule this out, we cloned each Hcp-Tep-Tip or Tep-Tip-Hcp locus into an arabinose inducible plasmid, the same plasmid we used to complement other T6SS components, and asked if donors expressing the loci could inhibit the Tep-Tip deleted recipients. We tested all 4 pairs and saw no fitness defect in recipients (Fig. 12C and data not shown).

Next, we looked at the non-Rhs PAAR-containing proteins that we had identified earlier for any toxin motifs. We found 4 PAAR proteins that were either fused or linked to predicted

toxic domains. ECL02217 and 04194 share similar toxin domains and likewise have homologous immunity proteins located immediately downstream of each gene. To test for delivery of either of these toxin domains, we made recipient cells lacking the immunity proteins for each of these potential effectors. ECL03144 is located upstream of Rhs2 with transcription going in the opposite direction of Rhs2 and encodes for a PAAR-toxin with predicted pore-forming activity. Interestingly, ECL03143 also contains PAAR motifs. Lastly, although it does not contain a PAAR domain, ECL01556 is found in the T6SS-1 locus and homologous to Cas9 and we reasoned its predicted nuclease activity could be used in bacterial competition. However, these toxic domains do not seem to be deployed, as recipient cells grow to the same levels against WT or *toxin*<sup>-</sup> donors (Fig. 12B). Again, we tried to express each PAAR-toxin pair in donor cells, but there was still no fitness defect in recipient cells (data not shown).

We hypothesized these Hcp-linked and PAAR-containing toxins are normally deployed through T6SS-2, but this locus is not on under laboratory conditions. To test this, we put an arabinose inducible promoter in front of T6SS-2 and asked if it could make a functional T6SS. To do this, we competed *vaskI*<sup>-</sup> T6SS-2<sup>+</sup> cells with *E. coli*, *Salmonella enterica* and each of the potential E-I recipient strains tested above. We found that T6SS-2, even when turned on, could not inhibit the growth of any recipient strain tested (Fig. 13A & 13B and data not shown). Even when we forced expression of the potential E-I pairs in the T6SS-2<sup>+</sup> donor cells, there was still no inhibition seen (data not shown). Therefore we conclude that T6SS-2 is either non-functional, or we have not successfully forced the right levels of expression of each component needed to build the T6S apparatus.

Upon further analysis of the T6SS-2 locus, we found multiple issues that could contribute to its lack of activity. First, the VgrG present in the locus (VgrG3) has a stop codon 603 residues into the ORF, most likely rendering it non-functional. However, restoring the stop codon to a tyrosine and expressing this gene in T6SS-2<sup>+</sup> cells still does not support T6SS-2 activity (Fig. 13C). Second, there is a PAAR-containing protein (PAAR-01815) that is truncated after residue 63 by a transposon insertion. Bioinformatic analysis shows full-length PAAR-01815 is found in other closely related *Enterobacter* species with a WT VgrG3 allele. This suggests the C-terminus of VgrG3 binds to PAAR-01815 to build a functional T6SS. Lastly, *Enterobacter hormaechei* contains a similar T6SS-2 locus; however it contains an Hcp protein in the locus while ECL does not. Taken all together, T6SS-2 in ECL is non-functional due to missing or mutated core components. Further studies will have to be performed to test if the Hcp-linked effectors or PAAR-containing toxins in ECL are functional T6SS E-I pairs.

## **Discussion**

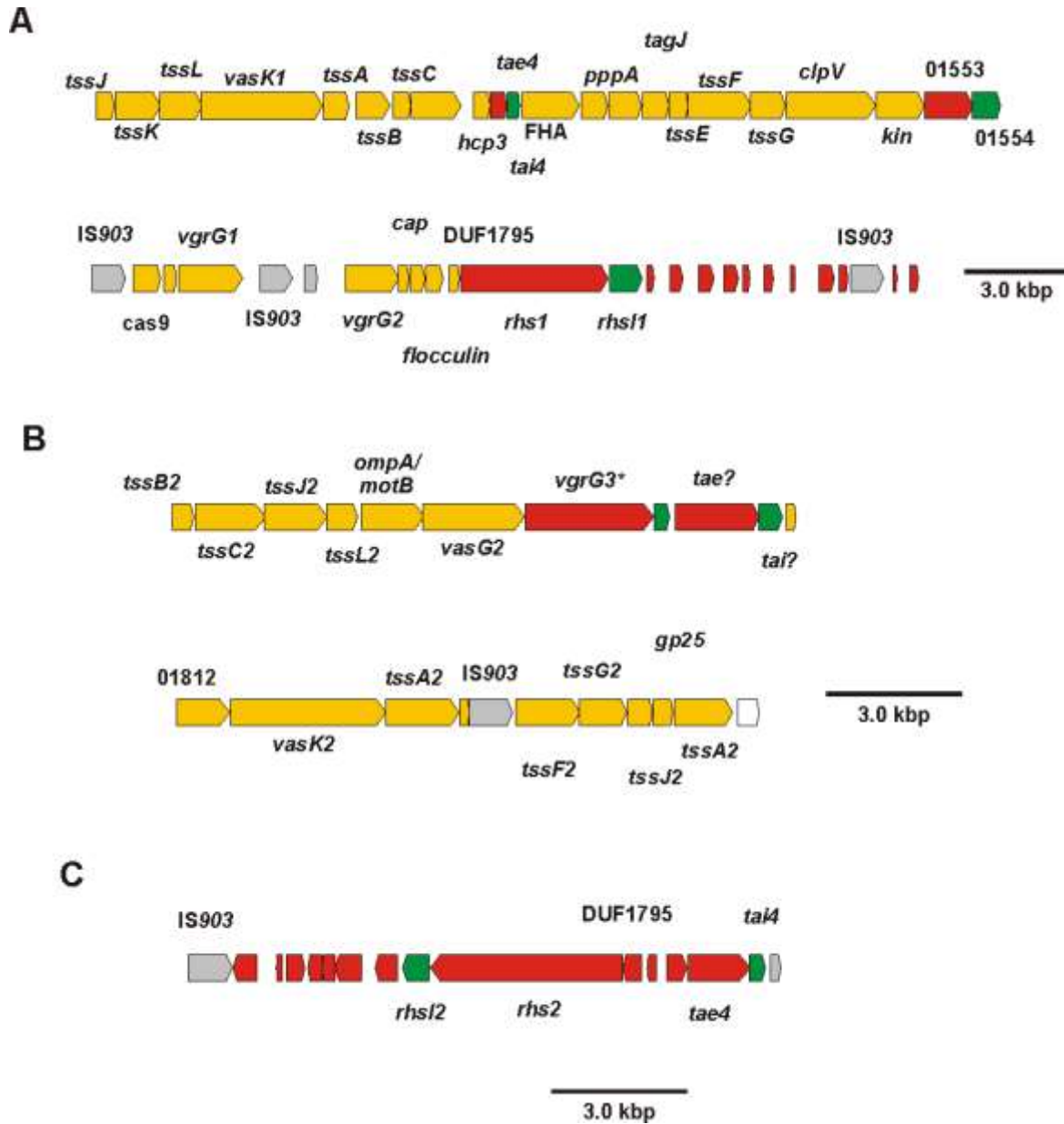
The work presented here demonstrates that ECL is a practical model organism to study T6S. We established that T6SS-1 plays a prominent role in both inter- and intra-species competition while T6SS-2 has some crucial defects that render it inactive. We showed that Tae4-Tai4 and 01553-01554 are both effective E-I pairs that likely contribute to the strong competitive phenotype of ECL. Furthermore, we demonstrated that Rhs1 and Rhs2 both contribute to bacterial competition, most likely by their C-terminal toxin domains targeting nucleic acid substrates in the cytoplasm of recipient cells. We determined that Rhs1 and Rhs2 are specific for VgrG2 and VgrG1, respectively, consistent with the divergent C-termini of the VgrG proteins interacting with the distinct PAAR domains present in Rhs1 and Rhs2.

Strikingly, we find that the PAAR domains of Rhs proteins are not sufficient to support T6S, and that only full-length Rhs protein can restore proper T6S. Most importantly, we demonstrate that not only T6 firing, but assembly of the apparatus is dependent on full-length Rhs.

As mentioned previously, T6S has been studied in a limited number of bacterial species. Based on the results presented here, it seems different bacterial species may have evolved distinct ways to assemble and fire their T6SSs. For example, in *V. cholera* and *A. baylyi*, the PAAR-containing proteins required to produce a functional T6SS are only ~95 residues while Rhs1 and Rhs2 in ECL are ~1400 residues<sup>153</sup>. Furthermore, in *Serratia marcescens*, *rhs*<sup>-</sup> donors still exhibit function T6S while *D. dadantii rhsA*<sup>-</sup> *rhsB*<sup>-</sup> cells have abrogated T6 activity (data not shown)<sup>279</sup>. Schneider *et al.* proposed that the role of PAAR-containing proteins is to sharpen the T6SS apparatus by interacting with the C-terminus of the spike protein, VgrG<sup>153</sup>. VgrG is also considered a baseplate component because it is necessary for Hcp tubule formation, which is in turn required for sheath polymerization<sup>151</sup>. We propose a model in which Rhs proteins in ECL and *D. dadantii* act as a seed for proper VgrG folding and subsequent Hcp tubule formation. The YD-repeat regions have the potential to act as a scaffold by encasing the spike tip or facilitating interactions with the membrane complex. Evidence for this is given by the recent crystal structure of an Rhs-repeat protein forming a cocoon around its toxin domain<sup>280</sup>. It seems the DPXGL motifs in Rhs also play some unidentified, perhaps structural role for the apparatus. With the presence of T6SSs in 25% of sequenced strains, and their prevalence in pathogenic species, it is important to understand the mechanism of toxin delivery. Identifying the differences between T6S in various species will lead to an improved understanding of these systems.

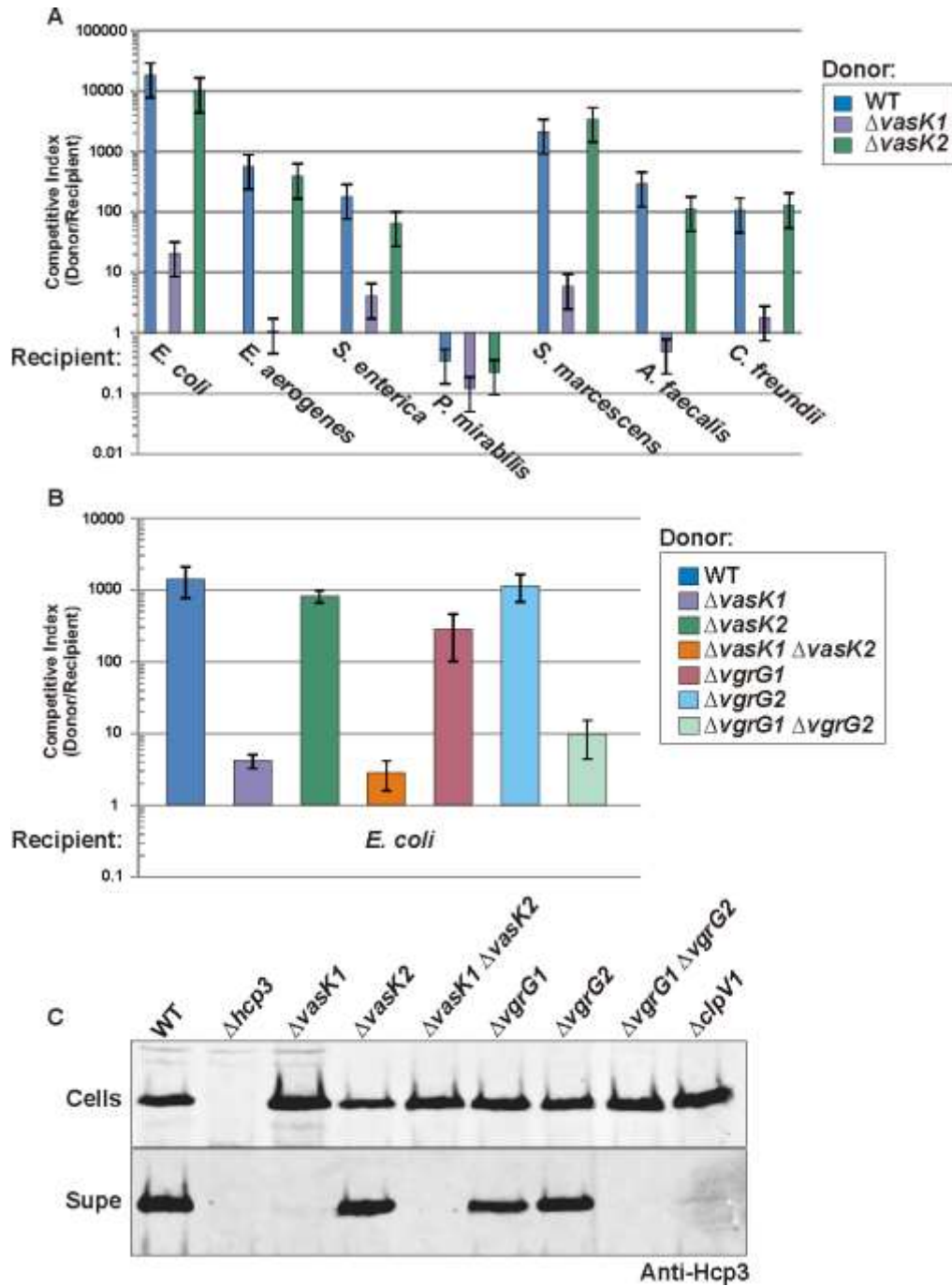
In addition, understanding the mechanism of effector delivery will help us identify the role of T6SSs in nature. It is unclear why so many T6SSs are present in the genome of some bacteria and if these systems facilitate a competitive advantage in specific habitats. Carrying independent T6S loci to deploy specific effectors comes at a cost, as compared to having one locus and regulating expression of E-I pairs based on the environmental conditions encountered. Nonetheless, some bacteria carry multiple loci and we still do not understand the advantage of this, or what excludes core components from different loci working together. We understand that VgrG and PAAR domains must evolve to work together, as do TssC/B, but the specificity of the remaining components remains unknown. Are VgrG proteins specific to Hcp or do they recognize the tubular structure that Hcp forms? Do TssC/B directly interact with Hcp or do they recognize the baseplate components? Can we swap components and direct differential assembly of T6SSs and therefore delivery of distinct effectors? Can each T6SS deliver into eukaryotic cells? These are all questions that remain to be answered, and the answers could give some insight into why some strains have so many loci and the contribution of each locus to bacterial competition and/or virulence. Further examination of T6SS-2 in ECL could give some insight into which core components interact with one another. Ultimately, clarifying the role of T6SSs in nature will help us understand how microbial communities are shaped and the potential consequences these communities have on an ecosystem.

**Figure 1. T6S and Rhs loci in *Enterobacter cloacae* ATCC 13047. A) T6SS-1 locus B) T6SS-2 locus C) Rhs2 locus. Open reading frames of T6SS genes are annotated in yellow. Red represents genes that contain toxic domains and green are the cognate immunity proteins located downstream of each toxin. IS elements are shown in grey.**

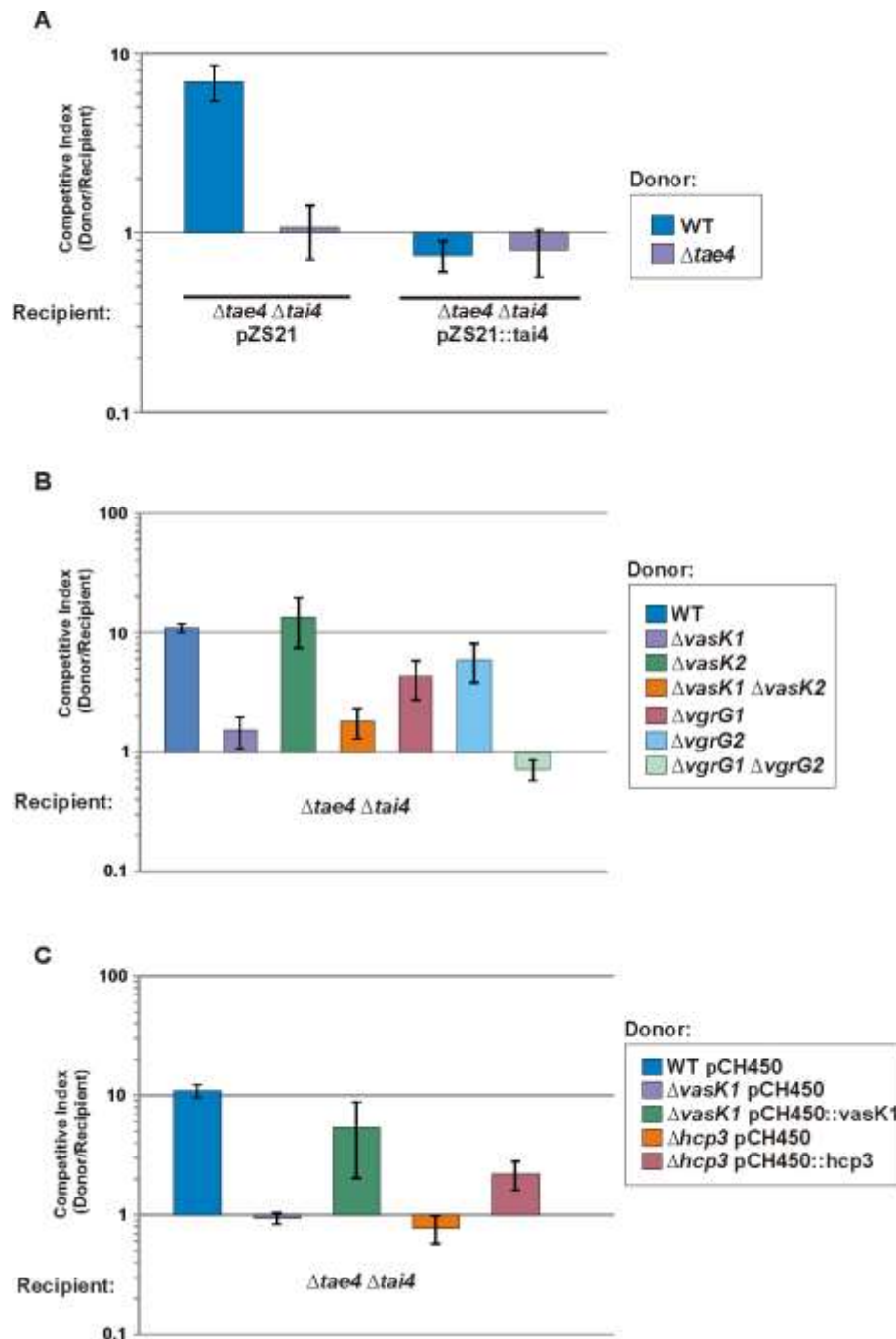




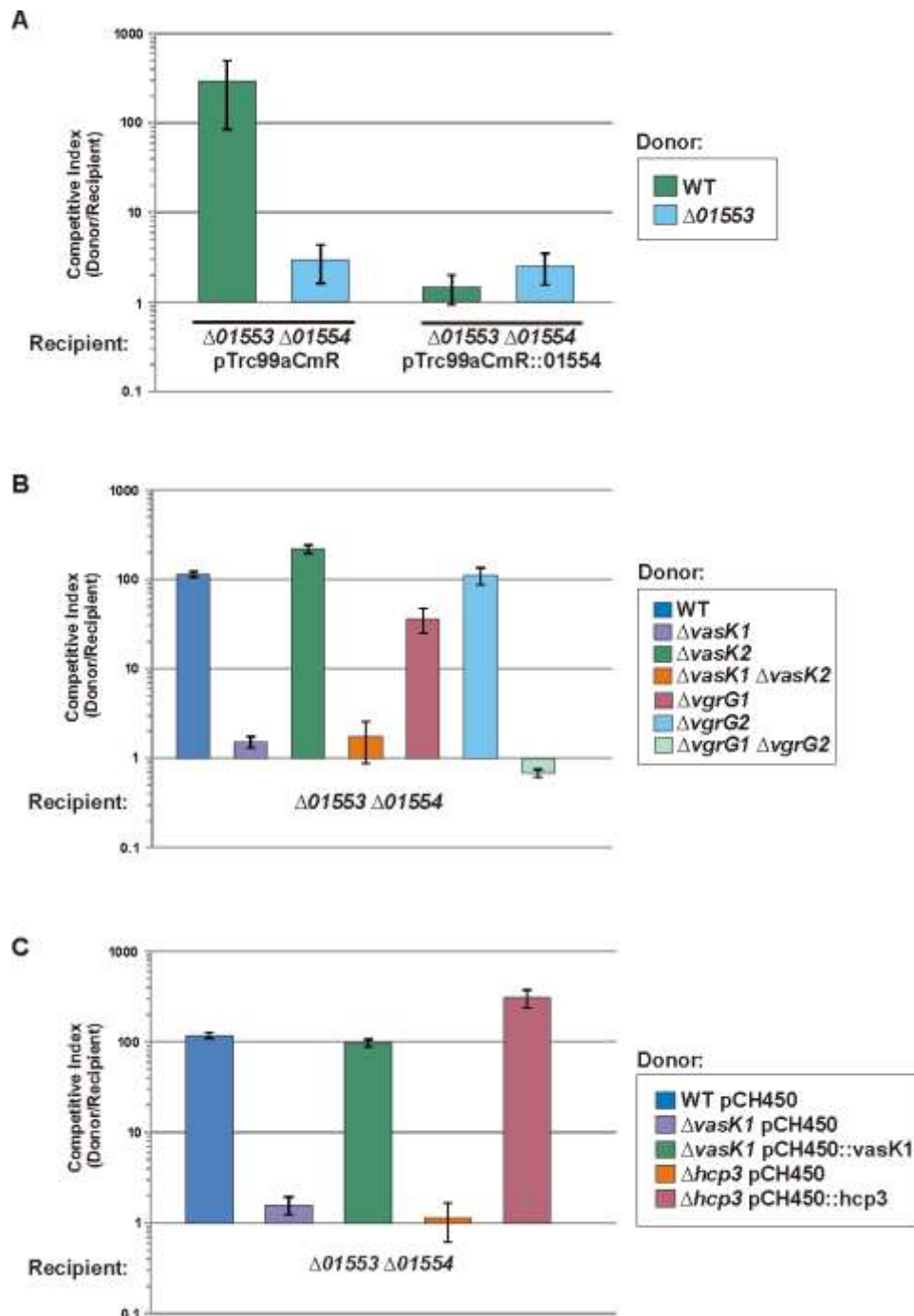
**Figure 2. T6SS-1 mediates inter-species competition.** **A)** Various species of bacteria were co-cultured with the annotated ECL donor cells at a 1:1 ratio under contact-promoting conditions for 4 hours. The results are reported as the mean  $\pm$  SEM for two independent experiments. **B)** *E. coli* recipients were co-cultured under the same conditions as A) with the specified ECL donor cells. The results are reported as the mean  $\pm$  SEM for three independent experiments. **C)** Western blot analysis of the supernatant and cytoplasm of the indicated ECL donor cells grown to mid-log in liquid LB.



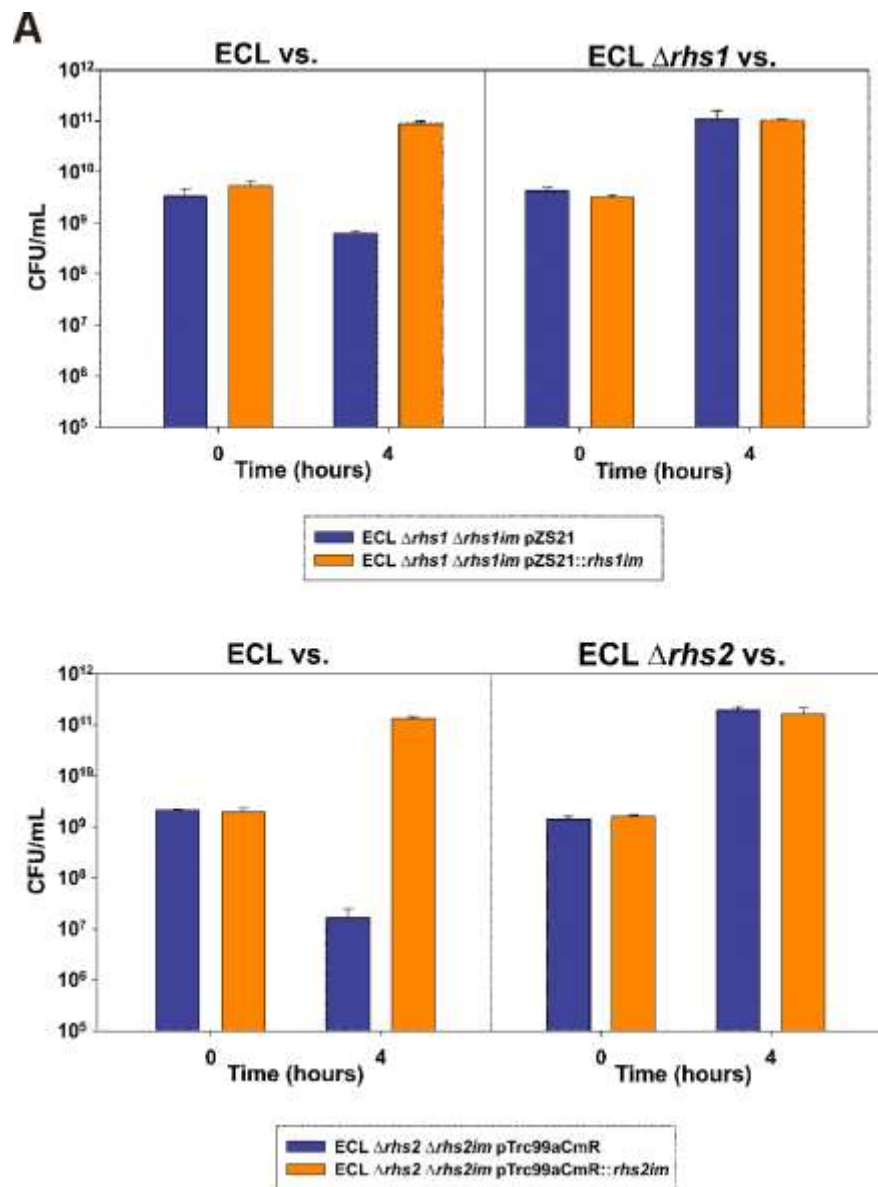
**Figure 3. *Tae4-Tai4* is a functional E-I pair deployed by T6SS-1. A) and B)** The indicated ECL recipient and donor cells were mixed at a 1:1 ratio under contact-promoting conditions for 4 hours. The results are reported as the mean  $\pm$  SEM for three independent experiments. **C)** ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with the indicated ECL recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The results are reported as the mean  $\pm$  SEM for three independent experiments.

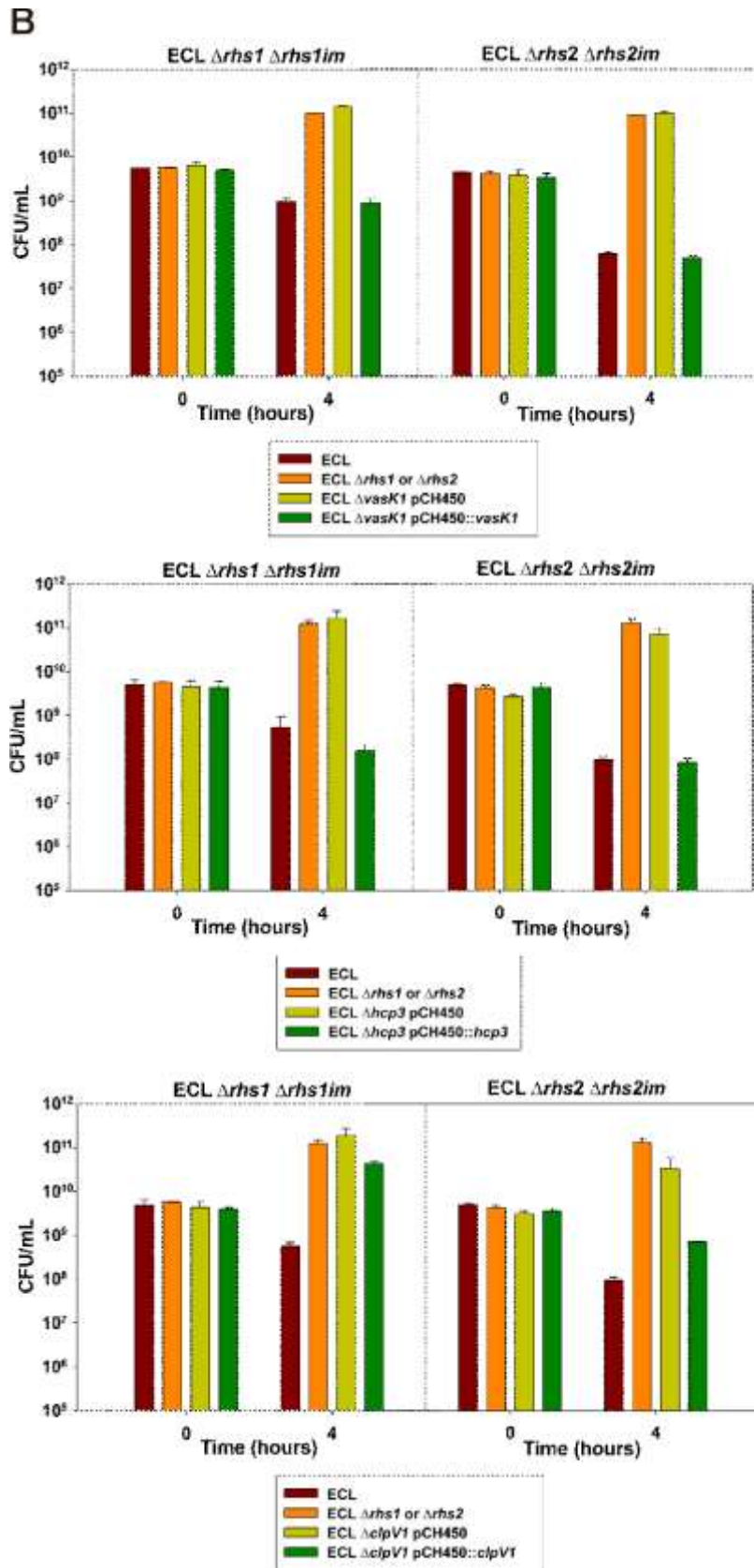


**Figure 4. ECL01553-01554 is a functional E-I pair deployed by T6SS-1. A) and B)** The specified ECL recipient and donor cells were mixed at a 1:1 ratio under contact-promoting conditions for 4 hours. The results are reported as the mean  $\pm$  SEM for three independent experiments. **C)** ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with the specified ECL recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The results are reported as the mean  $\pm$  SEM for three independent experiments.

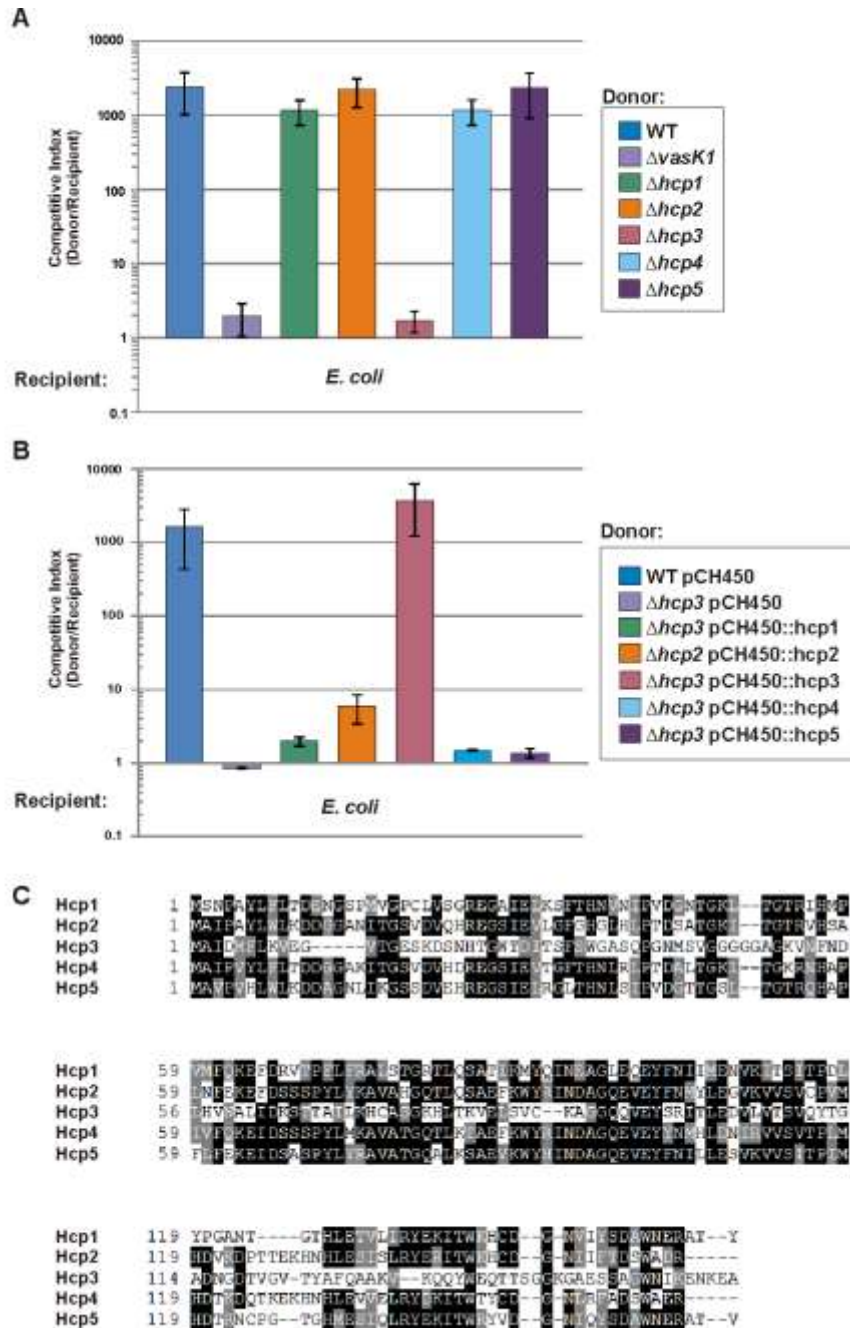


**Figure 5. Rhs 1 and Rhs2 are deployed by T6SS-1.** **A)** Donors and recipients are shown at the top and bottom of each graph, respectively. Cells were mixed at a 1:1 ratio under contact-promoting conditions for 4 hours. Growth inhibition was assessed by quantifying the number of viable recipient cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments. **B)** Recipient and donor cells are shown at the top and bottom of each graph, respectively. Donors were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with the annotated recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. Growth inhibition was assessed by quantifying the number of viable recipient cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.

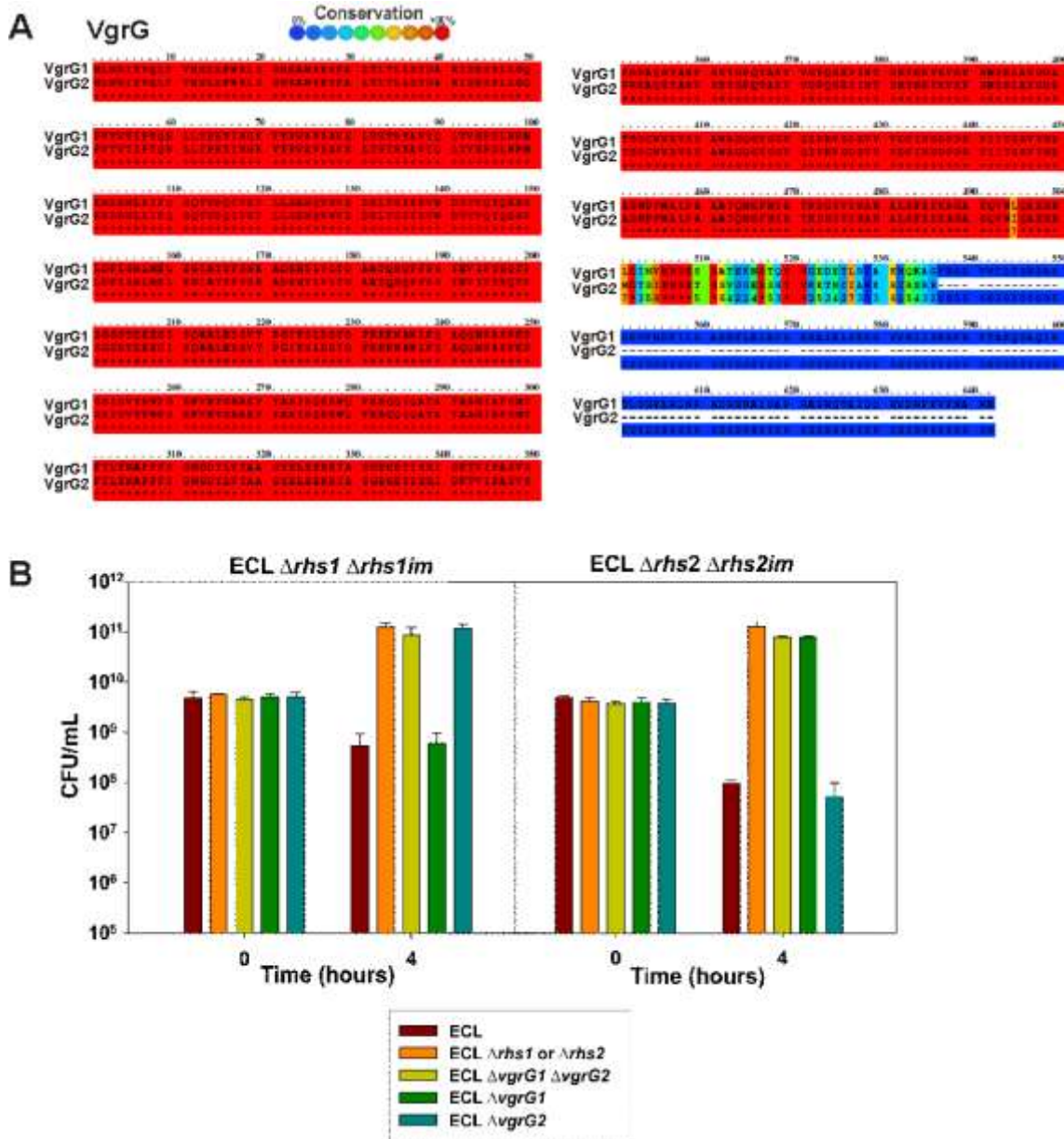




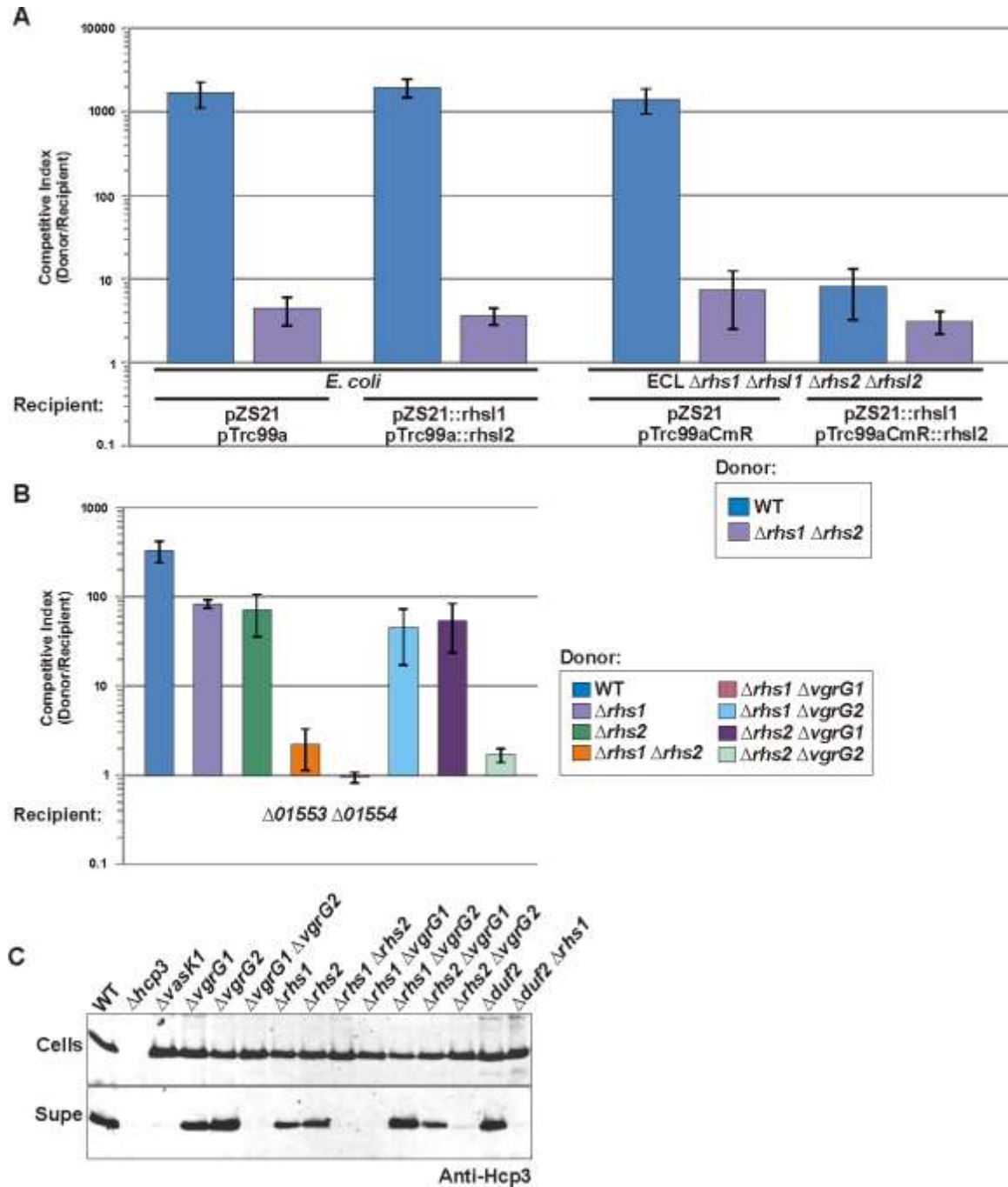
**Figure 6. T6SS-1 requires Hcp3.** A) ECL donor cells were mixed at a 1:1 ratio with *E. coli* recipients under contact-promoting conditions for 4 hours. The data are reported as the mean  $\pm$  SEM for three independent experiments. B) ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with *E. coli* recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The results are reported as the mean  $\pm$  SEM for three independent experiments. C) Alignment of the 5 Hcp proteins found in ECL. Shaded areas represent conserved residues between the Hcps. The alignment was performed by T-coffee and analysis by Box Shade.



**Figure 7. Divergent VgrG C-termini direct Rhs deployment.** A) PRALINE Multiple Sequence Alignment of VgrG1 and VgrG2 in ECL. B) Donors and recipients are shown at the bottom and top of each graph, respectively. Cells were mixed at a 1:1 ratio under contact-promoting conditions for 4 hours. Growth inhibition was assessed by quantifying the number of viable recipient cells as colony-forming units per milliliter (CFU/mL) and data are reported as the mean  $\pm$  SEM for two independent experiments.

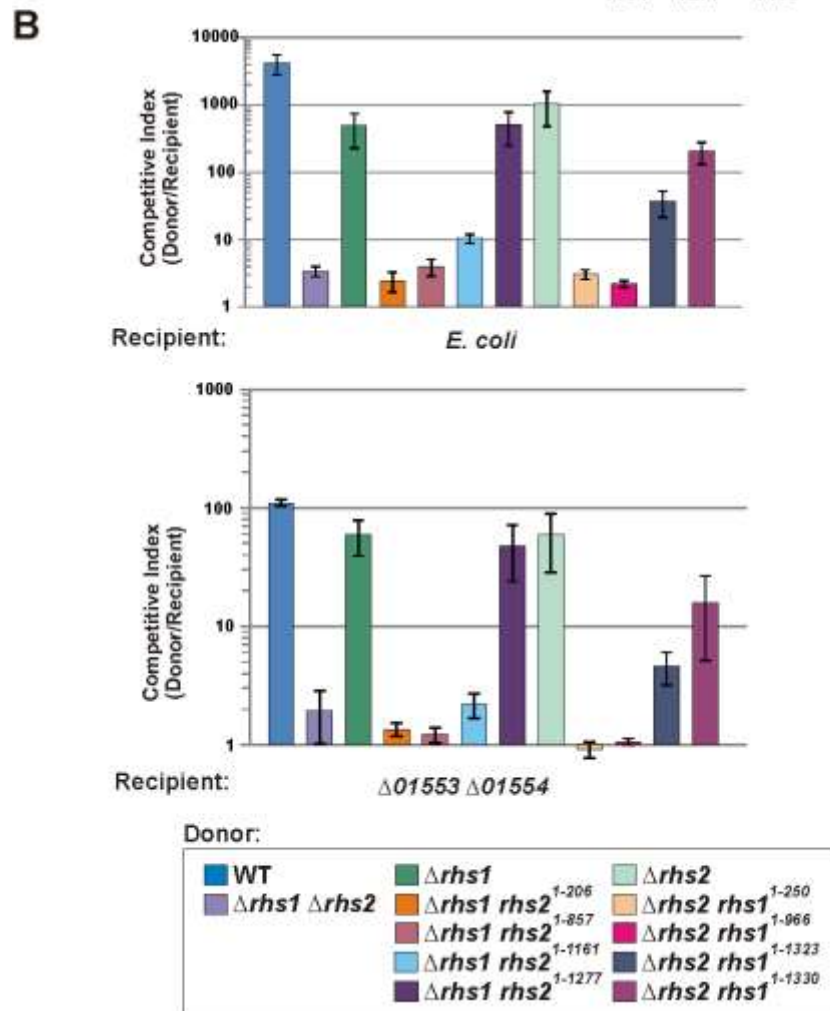
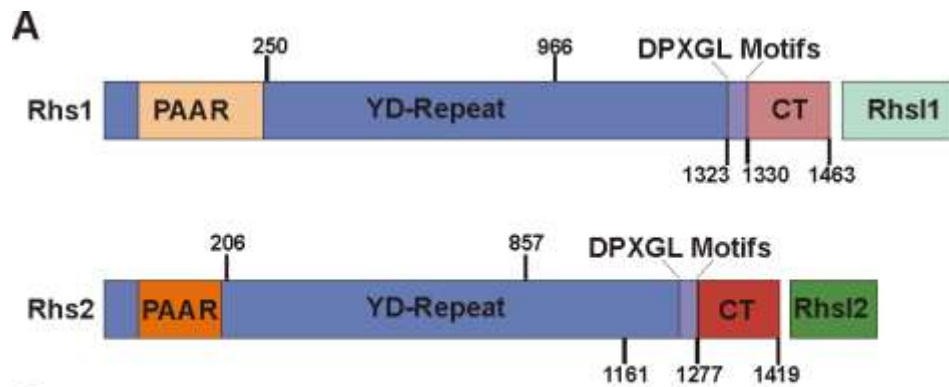


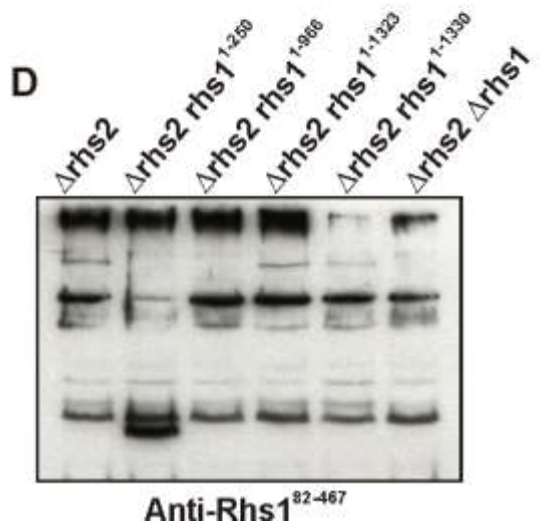
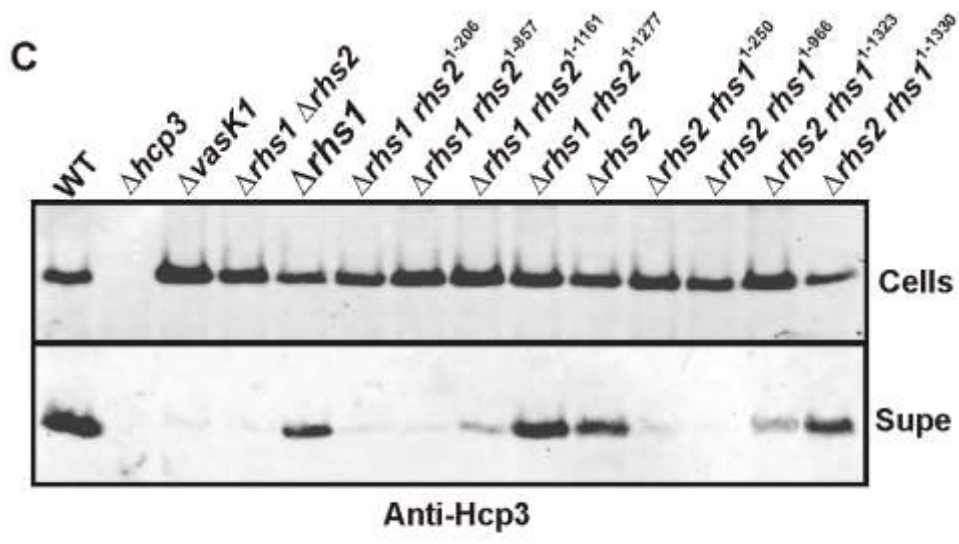
**Figure 8. Rhs proteins are required for T6SS-1 function.** A) and B) The indicated recipient and ECL donor cells were mixed at a 1:1 ratio for 4 hours under contact-promoting conditions. The data are reported as the mean  $\pm$  SEM for three independent experiments. C) Western blot analysis of the supernatant and cell lysate of donor cells grown to mid-log in liquid LB.



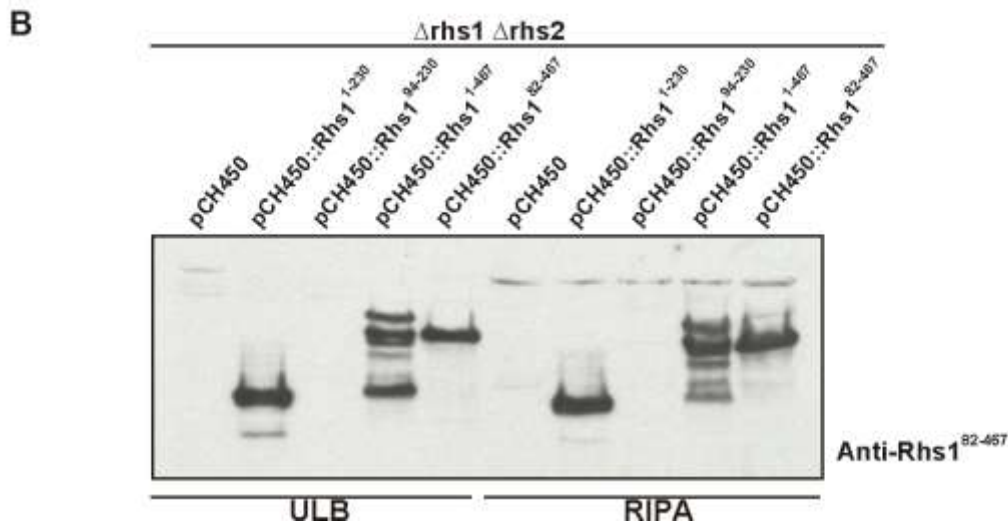
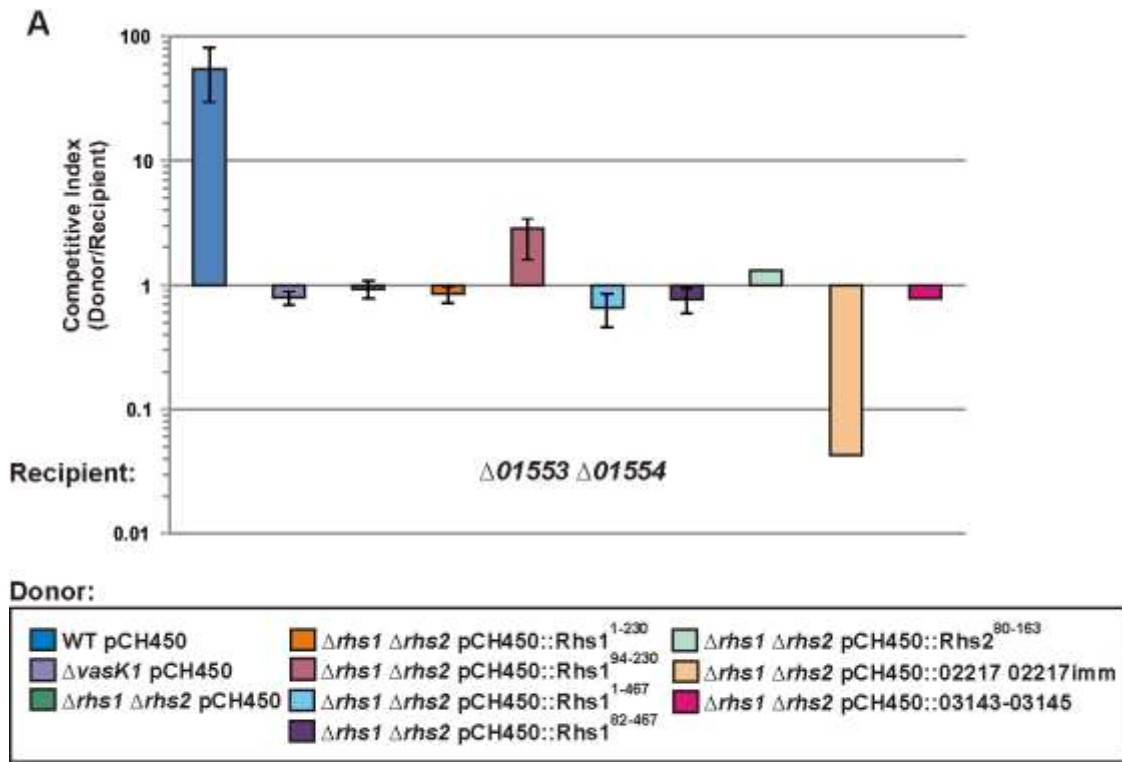


**Figure 9. Full-length Rhs proteins are required for T6SS-1 function.** A) Domains found in Rhs1 and Rhs 2. Stop codon insertions tested in B), C) and D) are labelled. Figure is not drawn to scale. B) The indicated recipient and ECL donor cells were mixed at a 1:1 ratio for 4 hours under contact-promoting conditions. The data are reported as the mean  $\pm$  SEM for three independent experiments. C) Western blot analysis of the supernatant and cell lysate of donor cells grown to mid-log in liquid LB. D) Western blot analysis of protein harvested from the specified donor cells using RIPA buffer.



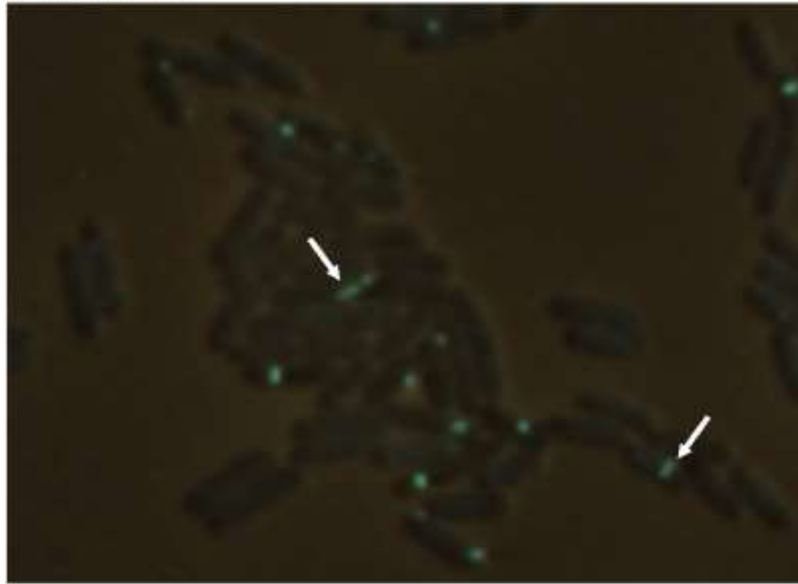


**Figure 10. PAAR-containing proteins do not restore T6SS-1 activity.** **A)** ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The first 7 bars are reported as the mean  $\pm$  SEM for two independent experiments. **B)** Western blot analysis of the cell lysate of specified donor cells grown to mid-log in liquid LB supplemented with 0.2% L-arabinose. Cells were lysed with either urea lysis buffer (ULB) or RIPA buffer.



**Figure 11. Rhs is required for T6SS-1 assembly.** The indicated cells were incubated on minimal media agar pads for 2 hours. Cells were viewed at 100X and captured every 15 seconds for 1.5 minutes for live-cell imaging. The images below represent one of the frames taken. White arrows are pointing to assembled T6SS-1.

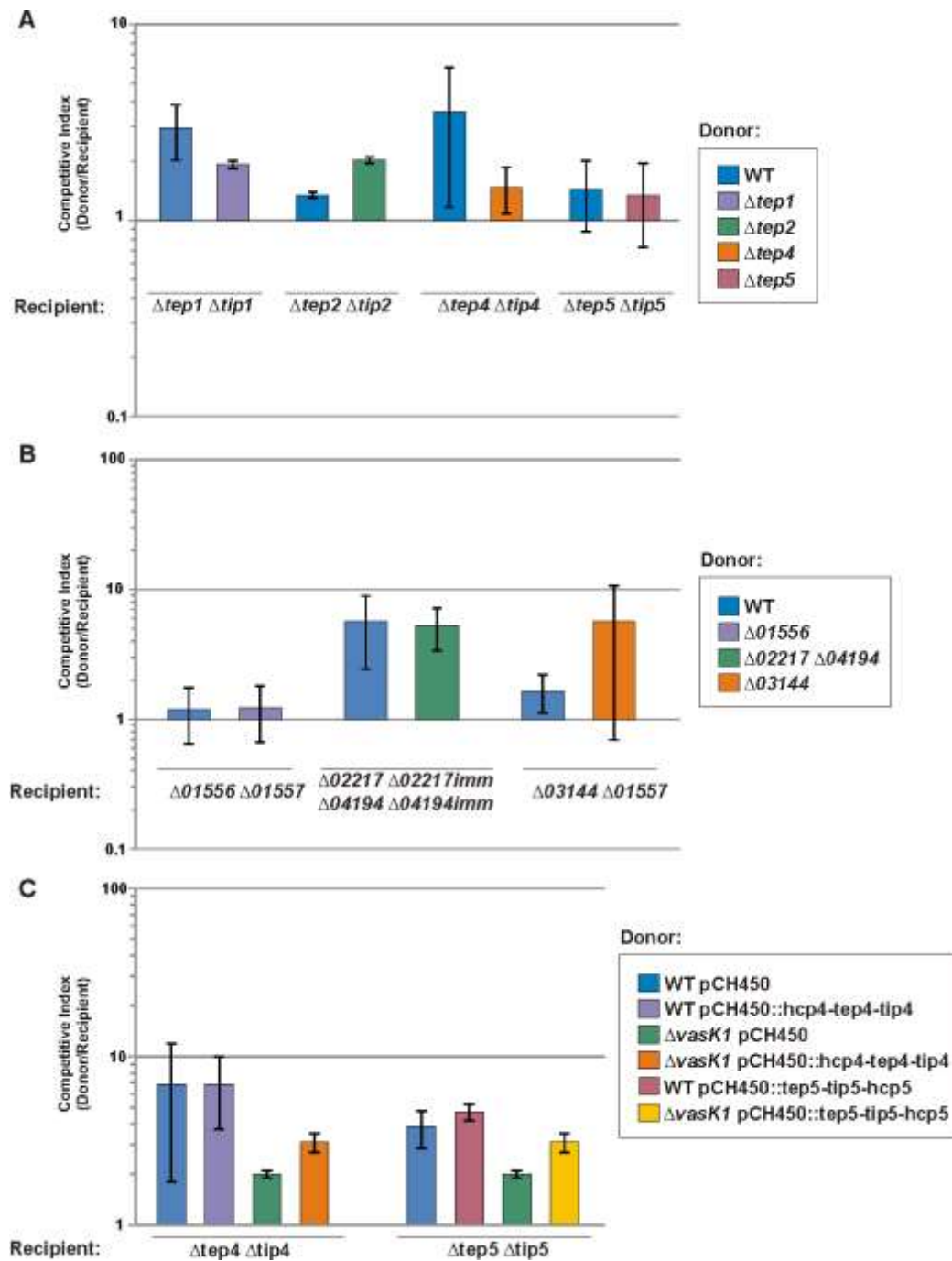
WT *tssC-GFP*



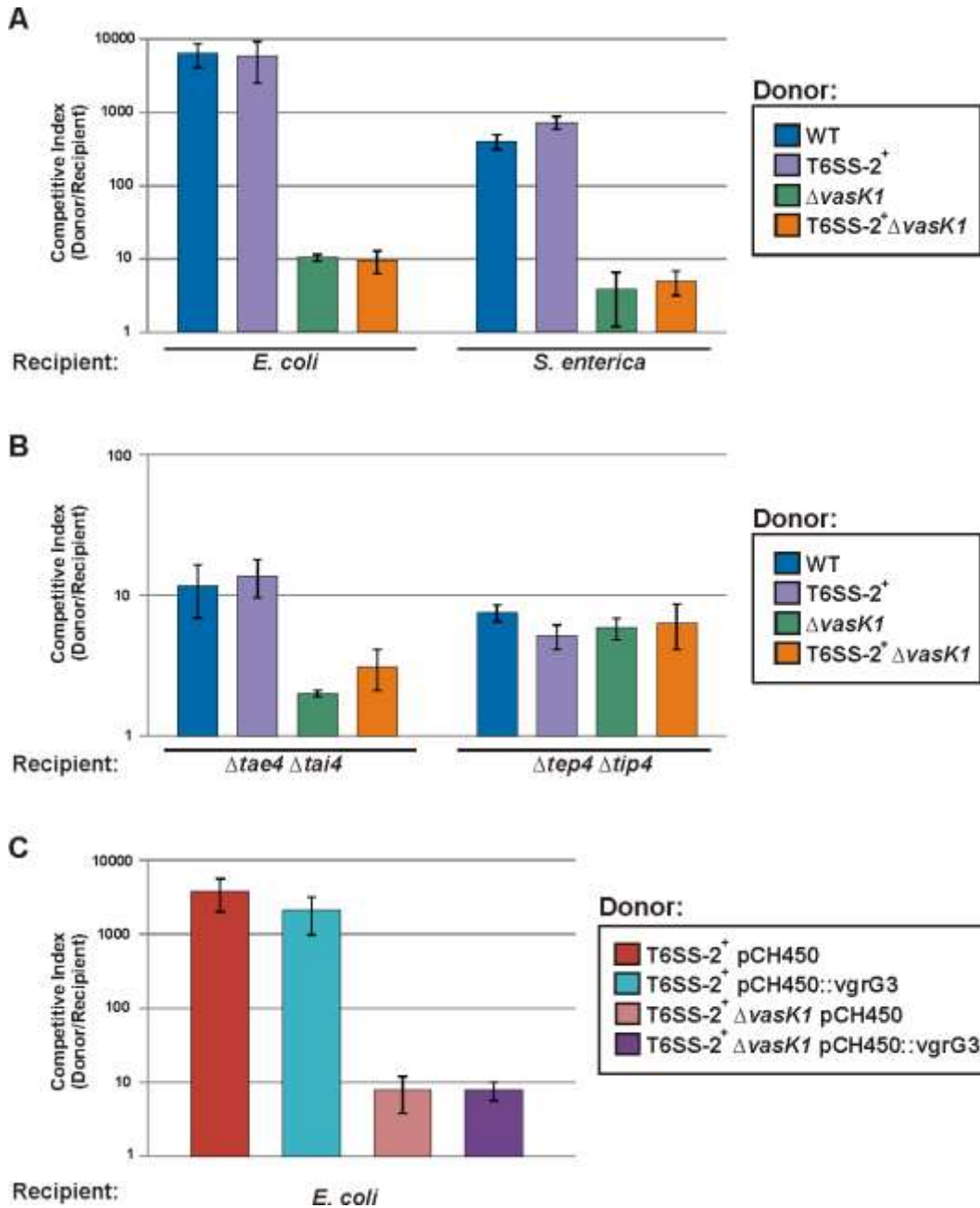
$\Delta$ *rhs1*  $\Delta$ *rhs2* *tssC-GFP*



**Figure 12. Hcp-linked E-I pairs and PAAR-containing toxins are not deployed.** A) and B) The indicated ECL donor and recipient cells were mixed at a 1:1 ratio for 4 hours under contact-promoting conditions. The data are reported as the mean  $\pm$  SEM for three independent experiments. C) ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The data are reported as the mean  $\pm$  SEM for three independent experiments.



**Figure 13. T6SS-2 is not functional. A) and B)** ECL donors were mixed at a 1:1 ratio with the indicated recipients for 4 hours under contact-promoting conditions. The data are reported as the mean  $\pm$  SEM for two independent experiments. **C)** ECL donor cells were grown in 0.2% L-arabinose, and then mixed at a 1:1 ratio under contact-promoting conditions with *E. coli* recipients on LB-agar supplemented with 0.4% L-arabinose for 4 hours. The data are reported as the mean  $\pm$  SEM for two independent experiments.



## **Chapter 7. Conclusion**

## Conclusion

Bacteria can be found in almost any ecological niche trying to flourish on limited resources. These harsh conditions have driven the evolution of many different competition and communication systems. The research presented here examines the mechanism of toxin activity and delivery in two such systems. Now we will dissect what we learned in each chapter to elucidate how these systems can be used in nature to shape complex microbial communities.

In the introductory chapter we looked at three systems bacteria use to inhibit the growth of neighboring bacterial cells. While the overall theme of these systems is similar - kill surrounding bacteria to have more space and nutrients for your own kin - they each have dramatically different mechanisms to perform their task. For example colicin-producing cells must lyse to deliver their effectors into the environment, and only target other *E. coli* strains<sup>51</sup>. Therefore, this competition system only seems to be effective under conditions where colicin-producing strains are fighting for resources that are specific to *E. coli*. In contrast, type VI secretion systems (T6SS) deliver across species barriers and require physical contact with neighboring cells to deploy toxic effectors<sup>127</sup>. Contact-dependent growth inhibition (CDI) systems share similarities with colicins and T6S. CDI requires direct physical contact, but the target range is limited to specific species, or even strains within a species of bacteria<sup>112,118</sup>.

CDI and colicin toxin delivery through the cell envelope seems to have more finesse than T6S. It is intriguing that the delivery process of the T6S tube spike into cells is not detrimental to a recipient cell. As long as recipient cells contain cognate immunity proteins for each effector delivered, they show no growth defect or changes in morphology<sup>281</sup>.



Interestingly, the small T6S effectors that get packaged into Hcp or elsewhere on the apparatus are almost exclusively periplasmic acting while Rhs, PAAR, CDI and colicin toxins contain nuclease activity<sup>121,153,164</sup>. This may highlight the differences in the delivery processes amongst these systems. Perhaps the smaller T6 effectors that target the cell wall and membrane only get delivered as far as the periplasm while Rhs and PAAR toxins are capable of crossing the inner membrane to reach their substrate in the cytoplasm. The location of Rhs and PAAR toxins on the tube spike certainly gives them first access to the inner membrane or may even facilitate direct injection into the cytoplasm<sup>139,153</sup>.

In chapter 2, we identified a novel mechanism of CDI toxin entry into *E. coli* cells. We demonstrated the N-terminal domain of CdiA-CT from UPEC536 (CdiA-CT<sup>NT</sup>) can interact with the F pilus subunit, TraA, to bypass the outer membrane upon retraction of the pilus. R17 bacteriophage use a similar mechanism for entry into cells. It is unclear if this mechanism of entry is exploited in nature or if this was a property of early CDI systems before they evolved to more efficiently deliver CdiA-CT. Alternatively, CdiA-CT<sup>NT</sup> could be an evolved phage protein that got fused to a CDI system, but still held onto its phage-like properties including the ability to bind TraA. Regardless, this work demonstrates the adaptability of CDI systems, particularly CdiA-CTs, and possibly illustrates how early CDI systems delivered cargo to the cytoplasm.

Chapter 3 reveals that CDI systems have a narrower target range than previously reported. We identified OmpC and OmpF as the receptor proteins for CdiA<sup>UPEC536</sup>. CDI<sup>UPEC536</sup> exploits the highly variable loops of both proteins, limiting its target range to a sub-population of *E. coli* strains. We sequenced 98 *ompC* alleles from uropathogenic *E. coli* (UPEC) strains isolated from the urinary tract of humans, dogs, or cats, and found that ~50%

of these alleles are resistant to CDI<sup>UPEC536</sup>. Given UPEC536 likely interacts with these UPEC strains in the environment; this raises questions about the use of CDI as a competition system in nature. With this work and work done by Ruhe *et al.* demonstrating CDI<sup>EC93</sup> also has a narrow target range, we propose CDI systems are involved in kin selection. Kin selection is important in microbial communities as it can potentiate homogeneous gene expression and thus give a bacterial population the advantage of behaving like a uniform multicellular organism<sup>235</sup>.

In chapter 4, we identified novel 16S rRNase activity in CDI toxins. CdiA-CT from *E. cloacae* (CdiA-CT<sup>ECL</sup>) was shown to have a similar fold as the toxin domain of colicin E3 (ColE3-CT) and likewise cut 16S rRNA at the same nucleotide position in the A-site of ribosomes. This work highlights the conservation of toxic activities found in colicins and CDI systems. The mechanism of delivery is different for each system; however the toxin domains seem to be conserved and restricted to targeting essential substrates. Colicins have been shown to act as pore-formers, or nucleases towards rRNA, tRNA, and DNA. Likewise, we have identified CdiA-CTs that can perform these same activities, CdiA-CT<sup>ECL</sup> being the first 16S rRNase identified. Furthermore, it seems CdiA-CT<sup>ECL</sup> and ColE3-CT are products of convergent evolution as the primary sequences of the two proteins, active-site residues, and immunity proteins are unique. Taken all together, this suggests targeting central dogma, and in this case the A-site of ribosomes, is the most efficient way to kill cells. Identifying the toxic activities and plasticity of CdiA-CTs will help us understand the potential of CDI systems to deliver a repertoire of diverse cargo. The ability of CdiA-CT/CdiI pairs to easily evolve and become distinct can play a drastic role in microbial competition. For example, if a daughter cell mutates its CdiA-CT/CdiI pair enough to still block parental CdiA-CT but

parental CdiI cannot block the new CdiA-CT activity; this could lead to subsequent proliferation of the mutated population. Events like this have the potential to drive speciation and thus have an impact on the composition of microbes in a given niche.

Chapter 5 tries to elucidate the role of CysK in the tRNase activity exhibited by CdiA-CT<sup>UPEC536</sup>. The crystal structure of CysK/CdiA-CT/CdiI confirms that CdiA-CT binds in the active site of CysK. With the recent proclamation that CDI<sup>UPEC536</sup> is utilized in kin selection, we reasoned CdiA-CT could influence cysteine metabolism amongst kin via self-delivery of CdiA-CT. We were able to demonstrate that CdiA-CT can block CysK activity inside cells, however only when over-expressing the CT. It is unclear if enough CT could be delivered during CDI to have an effect on cysteine metabolism in nature. There could however be certain environmental conditions where delivery of CdiA-CT can influence CysK activity, and therefore CDI would serve as a communication rather than a competition system. In support of this, homologs of CdiA-CT in Gram-positive species, which do not contain CDI systems, do not require CysK for activity. This suggests the role of the tRNase toxin in Gram-positive species is for competition, while in CDI systems it acts as a communication signal. As previously mentioned, communication within kin is an advantage for bacteria, especially in crowded habitats as it allows for nutrient scavenging and biofilm formation<sup>235</sup>.

Contrary to the restricted target range of colicins and CDI systems, chapter 6 covered the more versatile T6SS of *E. cloacae* (ECL). We showed that T6SS-1 deploys at least four different toxic effectors to inhibit the growth of a wide range of Gram-negative bacteria, giving ECL a clear competitive advantage in a laboratory setting. The fact that T6SSs are found in 25% of sequenced Gram-negative bacteria, and especially in enteric strains,

suggests having a T6SS provides an advantage in nature too<sup>135</sup>. Unique to ECL, compared to other T6SS<sup>+</sup> bacteria that have been studied, is the requirement of at least one Rhs protein for T6 function and assembly. Again, this illustrates the importance of studying these systems, as ECL seems to have evolved a different mechanism of toxin delivery. With enough changes this could lead to a distinct delivery system. In fact the T6SS is homologous to T4 contractile bacteriophage. Presumably the T6SS evolved to be able to inject cargo out rather than into a cell<sup>150</sup>. Understanding the mechanism of effector delivery in T6S will help establish the efficiency and role of these systems in nature.

In this thesis we covered three examples of competition systems used by Gram-negative bacteria and highlighted some of the similarities and differences amongst the systems. ECL harbors all three of these competition systems, suggesting they each provide a distinct advantage for ECL in its respective niche. Moreover, colicins, CDI systems, and T6Ss are commonly found in pathogenic species and often in the same strain. T6S has been shown to be a virulence determinant in many species including *Burkholderia pseudomallei*, *Vibrio cholera*, and *Burkholderia mallei*<sup>282-284</sup>. One of the barriers of invading a host is getting past the commensal layer of bacteria present<sup>285</sup>. Perhaps T6S allows pathogens to kill commensal bacteria to more efficiently invade a host. Ultimately, identifying the mechanism of toxin activity and delivery into targets will help elucidate the function of competition systems in nature and their distinct roles in shaping microbial communities.

## References

- 1 Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635-1638, doi:10.1126/science.1110591 (2005).
- 2 Roesch, L. F. *et al.* Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1**, 283-290, doi:10.1038/ismej.2007.53 (2007).
- 3 Bell, T., Newman, J. A., Silverman, B. W., Turner, S. L. & Lilley, A. K. The contribution of species richness and composition to bacterial services. *Nature* **436**, 1157-1160, doi:10.1038/nature03891 (2005).
- 4 Furina, E. K. & Bonartseva, G. A. [The effect of combined and separate inoculation of alfalfa plants with *Azospirillum lipoferum* and *Sinorhizobium meliloti* on denitrification and nitrogen-fixing activities]. *Prikl Biokhim Mikrobiol* **43**, 318-324 (2007).
- 5 Hooper, L. V., Midtvedt, T. & Gordon, J. I. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* **22**, 283-307, doi:10.1146/annurev.nutr.22.011602.092259 (2002).
- 6 Wardwell, L. H., Huttenhower, C. & Garrett, W. S. Current concepts of the intestinal microbiota and the pathogenesis of infection. *Curr Infect Dis Rep* **13**, 28-34, doi:10.1007/s11908-010-0147-7 (2011).
- 7 Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023, doi:10.1038/4441022a (2006).
- 8 Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031, doi:10.1038/nature05414 (2006).
- 9 Manichanh, C. *et al.* Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **55**, 205-211, doi:10.1136/gut.2005.073817 (2006).
- 10 Joossens, M. *et al.* Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* **60**, 631-637, doi:10.1136/gut.2010.223263 (2011).
- 11 Chak, K. F. & James, R. Analysis of the promoters for the two immunity genes present in the ColE3-CA38 plasmid using two new promoter probe vectors. *Nucleic Acids Res* **13**, 2519-2531 (1985).
- 12 Lloubes, R., Baty, D. & Lazdunski, C. The promoters of the genes for colicin production, release and immunity in the ColA plasmid: effects of convergent transcription and Lex A protein. *Nucleic Acids Res* **14**, 2621-2636 (1986).

- 13 Fredericq, P. Sur la p;uralite des recepteurs d'antibiose de *E. coli*. *Paris*, 1189-1194 (1946).
- 14 Gordon, D. M., Riley, M. A. & Pinou, T. Temporal changes in the frequency of colicinogeny in *Escherichia coli* from house mice. *Microbiology* **144** ( Pt 8), 2233-2240 (1998).
- 15 Riley, M. A. & Gordon, D. M. A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col-plasmid lineages. *J Gen Microbiol* **138**, 1345-1352 (1992).
- 16 Gratia, A. & Fredericq, P. Diversite des souches antibiotiques de *Bacterium coli* et etendue variable de leur champ d'action. *C. R. Soc. Biol. (Paris)*, 1032-1033 (1946).
- 17 Dandeu, J. P. & Barbu, E. [Purification of colchicine K]. *C R Acad Sci Hebd Seances Acad Sci D* **265**, 774-776 (1967).
- 18 Herschman, H. R. & Helinski, D. R. Comparative study of the events associated with colicin induction. *J Bacteriol* **94**, 691-699 (1967).
- 19 Herschman, H. R. & Helinski, D. R. Purification and characterization of colicin E2 and colicin E3. *J Biol Chem* **242**, 5360-5368 (1967).
- 20 Konisky, J. & Richards, F. M. Characterization of colicin Ia and colicin Ib. Purification and some physical properties. *J Biol Chem* **245**, 2972-2978 (1970).
- 21 Schwartz, S. A. & Helinski, D. R. Purification and characterization of colicin E1. *J Biol Chem* **246**, 6318-6327 (1971).
- 22 Brunden, K. R., Cramer, W. A. & Cohen, F. S. Purification of a small receptor-binding peptide from the central region of the colicin E1 molecule. *J Biol Chem* **259**, 190-196 (1984).
- 23 Benedetti, H. *et al.* Individual domains of colicins confer specificity in colicin uptake, in pore-properties and in immunity requirement. *J Mol Biol* **217**, 429-439 (1991).
- 24 Braun, V., Patzer, S. I. & Hantke, K. Ton-dependent colicins and microcins: modular design and evolution. *Biochimie* **84**, 365-380 (2002).
- 25 Lazdunski, C. J. *et al.* Colicin import into *Escherichia coli* cells. *J Bacteriol* **180**, 4993-5002 (1998).
- 26 Lazzaroni, J. C., Dubuisson, J. F. & Vianney, A. The Tol proteins of *Escherichia coli* and their involvement in the translocation of group A colicins. *Biochimie* **84**, 391-397 (2002).

- 27 Dankert, J. R., Uratani, Y., Grabau, C., Cramer, W. A. & Hermodson, M. On a domain structure of colicin E1. A COOH-terminal peptide fragment active in membrane depolarization. *J Biol Chem* **257**, 3857-3863 (1982).
- 28 Martinez, M. C., Lazdunski, C. & Pattus, F. Isolation, molecular and functional properties of the C-terminal domain of colicin A. *EMBO J* **2**, 1501-1507 (1983).
- 29 El Ghachi, M. *et al.* Colicin M exerts its bacteriolytic effect via enzymatic degradation of undecaprenyl phosphate-linked peptidoglycan precursors. *J Biol Chem* **281**, 22761-22772, doi:10.1074/jbc.M602834200 (2006).
- 30 Harkness, R. E. & Braun, V. Colicin M inhibits peptidoglycan biosynthesis by interfering with lipid carrier recycling. *J Biol Chem* **264**, 6177-6182 (1989).
- 31 Nagel de Zwaig, R. Mode of action of colicin A. *J Bacteriol* **99**, 913-914 (1969).
- 32 Reeves, P. Mode of action of colicins of types E1, E2, E3, and K. *J Bacteriol* **96**, 1700-1703 (1968).
- 33 De Witt, W. & Helinsky, D. Colicinogenic factor from a noninduced and a mitomycin C-induced *Proteus* strain. *J. Mol. Biol.*, 692-703 (1965).
- 34 Hardy, K., Meynell, G., Dowman, J. & Spratt, B. Two major groups of colicinogenic factors: their evolutionary significance. 217-230 (1973).
- 35 Chak, K. F., Kuo, W. S., Lu, F. M. & James, R. Cloning and characterization of the ColE7 plasmid. *J Gen Microbiol* **137**, 91-100 (1991).
- 36 Pugsley, A. P. Transcription regulation of colicin Ib synthesis. *Mol Gen Genet* **183**, 522-527 (1981).
- 37 Tessman, E. S., Gritzmacher, C. A. & Peterson, P. K. Derepression of colicin E1 synthesis in the constitutive *tif* mutant strain (*spr tif sfi*) and in a *tif sfi* mutant strain of *Escherichia coli* K-12. *J Bacteriol* **135**, 29-38 (1978).
- 38 Little, J. W. & Mount, D. W. The SOS regulatory system of *Escherichia coli*. *Cell* **29**, 11-22 (1982).
- 39 Akutsu, A., Masaki, H. & Ohta, T. Molecular structure and immunity specificity of colicin E6, an evolutionary intermediate between E-group colicins and cloacin DF13. *J Bacteriol* **171**, 6430-6436 (1989).
- 40 Curtis, M. D., James, R. & Coddington, A. An evolutionary relationship between the ColE5-099 and the ColE9-J plasmids revealed by nucleotide sequencing. *J Gen Microbiol* **135**, 2783-2788 (1989).
- 41 Jakes, K. S. & Zinder, N. D. Plasmid ColE3 specifies a lysis protein. *J Bacteriol* **157**, 582-590 (1984).

- 42 Hakkaart, M. J., Veltkamp, E. & Nijkamp, H. J. Protein H encoded by plasmid Clo DF13 involved in lysis of the bacterial host. I. Localisation of the gene and identification and subcellular localisation of the gene H product. *Mol Gen Genet* **183**, 318-325 (1981).
- 43 Hakkaart, M. J., Veltkamp, E. & Nijkamp, H. J. Protein H encoded by plasmid Clo DF13 involved in lysis of the bacterial host. II. Functions and regulation of synthesis of the gene H product. *Mol Gen Genet* **183**, 326-332 (1981).
- 44 Jakes, K., Zinder, N. D. & Boon, T. Purification and properties of colicin E3 immunity protein. *J Biol Chem* **249**, 438-444 (1974).
- 45 Masaki, H. & Ohta, T. Colicin E3 and its immunity genes. *J Mol Biol* **182**, 217-227 (1985).
- 46 Soong, B. W., Hsieh, S. Y. & Chak, K. F. Mapping of transcriptional start sites of the cea and cei genes of the ColE7 operon. *Mol Gen Genet* **243**, 477-481 (1994).
- 47 Chak, K. F. & James, R. Localization and characterization of a gene on the ColE3-CA38 plasmid that confers immunity to colicin E8. *J Gen Microbiol* **130**, 701-710 (1984).
- 48 Cavard, D., Lloubès, R., Morlon, J., Chartier, M. & Lazdunski, C. Lysis protein encoded by plasmid ColA-CA31. Gene sequence and export. *Mol Gen Genet* **199**, 95-100 (1985).
- 49 Sabik, J. F., Suit, J. L. & Luria, S. E. cea-kil operon of the ColE1 plasmid. *J Bacteriol* **153**, 1479-1485 (1983).
- 50 Pilsil, H. & Braun, V. Strong function-related homology between the pore-forming colicins K and 5. *J Bacteriol* **177**, 6973-6977 (1995).
- 51 Cascales, E. *et al.* Colicin biology. *Microbiol Mol Biol Rev* **71**, 158-229, doi:10.1128/MMBR.00036-06 (2007).
- 52 Di Masi, D. R., White, J. C., Schnaitman, C. A. & Bradbeer, C. Transport of vitamin B12 in Escherichia coli: common receptor sites for vitamin B12 and the E colicins on the outer membrane of the cell envelope. *J Bacteriol* **115**, 506-513 (1973).
- 53 Escuyer, V. & Mock, M. DNA sequence analysis of three missense mutations affecting colicin E3 bactericidal activity. *Mol Microbiol* **1**, 82-85 (1987).
- 54 Hill, C. & Holland, I. B. Genetic basis of colicin E susceptibility in Escherichia coli. I. Isolation and properties of refractory mutants and the preliminary mapping of their mutations. *J Bacteriol* **94**, 677-686 (1967).



- 55 Bouveret, E., Rigal, A., Lazdunski, C. & Bénédicti, H. Distinct regions of the colicin A translocation domain are involved in the interaction with TolA and TolB proteins upon import into Escherichia coli. *Mol Microbiol* **27**, 143-157 (1998).
- 56 Cao, Z. & Klebba, P. E. Mechanisms of colicin binding and transport through outer membrane porins. *Biochimie* **84**, 399-412 (2002).
- 57 Kurisu, G. *et al.* The structure of BtuB with bound colicin E3 R-domain implies a translocon. *Nat Struct Biol* **10**, 948-954, doi:10.1038/nsb997 (2003).
- 58 Buchanan, S. K. *et al.* Structure of colicin I receptor bound to the R-domain of colicin Ia: implications for protein import. *EMBO J* **26**, 2594-2604, doi:10.1038/sj.emboj.7601693 (2007).
- 59 Housden, N. G., Loftus, S. R., Moore, G. R., James, R. & Kleanthous, C. Cell entry mechanism of enzymatic bacterial colicins: porin recruitment and the thermodynamics of receptor binding. *Proc Natl Acad Sci U S A* **102**, 13849-13854, doi:10.1073/pnas.0503567102 (2005).
- 60 Duché, D., Frenkian, A., Prima, V. & Llobès, R. Release of immunity protein requires functional endonuclease colicin import machinery. *J Bacteriol* **188**, 8593-8600, doi:10.1128/JB.00941-06 (2006).
- 61 Loftus, S. R. *et al.* Competitive recruitment of the periplasmic translocation portal TolB by a natively disordered domain of colicin E9. *Proc Natl Acad Sci U S A* **103**, 12353-12358, doi:10.1073/pnas.0603433103 (2006).
- 62 Lazzaroni, J. C., Germon, P., Ray, M. C. & Vianney, A. The Tol proteins of Escherichia coli and their involvement in the uptake of biomolecules and outer membrane stability. *FEMS Microbiol Lett* **177**, 191-197 (1999).
- 63 Llobès, R. *et al.* The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Res Microbiol* **152**, 523-529 (2001).
- 64 Benedetti, H., Lazdunski, C. & Llobès, R. Protein import into Escherichia coli: colicins A and E1 interact with a component of their translocation system. *EMBO J* **10**, 1989-1995 (1991).
- 65 Schendel, S. L., Click, E. M., Webster, R. E. & Cramer, W. A. The TolA protein interacts with colicin E1 differently than with other group A colicins. *J Bacteriol* **179**, 3683-3690 (1997).
- 66 Mende, J. & Braun, V. Import-defective colicin B derivatives mutated in the TonB box. *Mol Microbiol* **4**, 1523-1533 (1990).
- 67 Pils, H. *et al.* Domains of colicin M involved in uptake and activity. *Mol Gen Genet* **240**, 103-112 (1993).

- 68 Postle, K. & Kadner, R. J. Touch and go: tying TonB to transport. *Mol Microbiol* **49**, 869-882 (2003).
- 69 Traub, I., Gaisser, S. & Braun, V. Activity domains of the TonB protein. *Mol Microbiol* **8**, 409-423 (1993).
- 70 Ogierman, M. & Braun, V. Interactions between the outer membrane ferric citrate transporter FecA and TonB: studies of the FecA TonB box. *J Bacteriol* **185**, 1870-1885 (2003).
- 71 Cadieux, N. & Kadner, R. J. Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter. *Proc Natl Acad Sci U S A* **96**, 10673-10678 (1999).
- 72 Kleanthous, C. Swimming against the tide: progress and challenges in our understanding of colicin translocation. *Nat Rev Microbiol* **8**, 843-848, doi:10.1038/nrmicro2454 (2010).
- 73 Cavard, D. *et al.* Hydrodynamic properties of colicin A. Existence of a high-affinity lipid-binding site and oligomerization at acid pH. *Eur J Biochem* **172**, 507-512 (1988).
- 74 Elkins, P., Bunker, A., Cramer, W. A. & Stauffacher, C. V. A mechanism for toxin insertion into membranes is suggested by the crystal structure of the channel-forming domain of colicin E1. *Structure* **5**, 443-458 (1997).
- 75 Hilsenbeck, J. L. *et al.* Crystal structure of the cytotoxic bacterial protein colicin B at 2.5 Å resolution. *Mol Microbiol* **51**, 711-720 (2004).
- 76 Vetter, I. R. *et al.* Crystal structure of a colicin N fragment suggests a model for toxicity. *Structure* **6**, 863-874 (1998).
- 77 Wiener, M., Freymann, D., Ghosh, P. & Stroud, R. M. Crystal structure of colicin Ia. *Nature* **385**, 461-464, doi:10.1038/385461a0 (1997).
- 78 Parker, M. W., Pattus, F., Tucker, A. D. & Tsernoglou, D. Structure of the membrane-pore-forming fragment of colicin A. *Nature* **337**, 93-96, doi:10.1038/337093a0 (1989).
- 79 Schein, S. J., Kagan, B. L. & Finkelstein, A. Colicin K acts by forming voltage-dependent channels in phospholipid bilayer membranes. *Nature* **276**, 159-163 (1978).
- 80 Geli, V., Baty, D. & Lazdunski, C. Use of a foreign epitope as a "tag" for the localization of minor proteins within a cell: the case of the immunity protein to colicin A. *Proc Natl Acad Sci U S A* **85**, 689-693 (1988).

- 81 Goldman, K., Suit, J. L. & Kayalar, C. Identification of the plasmid-encoded immunity protein for colicin E1 in the inner membrane of *Escherichia coli*. *FEBS Lett* **190**, 319-323 (1985).
- 82 Geli, V., Baty, D., Pattus, F. & Lazdunski, C. Topology and function of the integral membrane protein conferring immunity to colicin A. *Mol Microbiol* **3**, 679-687 (1989).
- 83 Song, H. Y., Cohen, F. S. & Cramer, W. A. Membrane topography of ColE1 gene products: the hydrophobic anchor of the colicin E1 channel is a helical hairpin. *J Bacteriol* **173**, 2927-2934 (1991).
- 84 Song, H. Y. & Cramer, W. A. Membrane topography of ColE1 gene products: the immunity protein. *J Bacteriol* **173**, 2935-2943 (1991).
- 85 Geli, V. & Lazdunski, C. An alpha-helical hydrophobic hairpin as a specific determinant in protein-protein interaction occurring in *Escherichia coli* colicin A and B immunity systems. *J Bacteriol* **174**, 6432-6437 (1992).
- 86 Zhang, Y. L. & Cramer, W. A. Intramembrane helix-helix interactions as the basis of inhibition of the colicin E1 ion channel by its immunity protein. *J Biol Chem* **268**, 10176-10184 (1993).
- 87 Pils, H. & Braun, V. Evidence that the immunity protein inactivates colicin 5 immediately prior to the formation of the transmembrane channel. *J Bacteriol* **177**, 6966-6972 (1995).
- 88 Espeset, D., Duché, D., Baty, D. & Géli, V. The channel domain of colicin A is inhibited by its immunity protein through direct interaction in the *Escherichia coli* inner membrane. *EMBO J* **15**, 2356-2364 (1996).
- 89 Chauleau, M., Mora, L., Serba, J. & de Zamaroczy, M. FtsH-dependent processing of RNase colicins D and E3 means that only the cytotoxic domains are imported into the cytoplasm. *J Biol Chem* **286**, 29397-29407, doi:10.1074/jbc.M111.242354 (2011).
- 90 de Zamaroczy, M., Mora, L., Lecuyer, A., Géli, V. & Buckingham, R. H. Cleavage of colicin D is necessary for cell killing and requires the inner membrane peptidase LepB. *Mol Cell* **8**, 159-168 (2001).
- 91 Schaller, K. & Nomura, M. Colicin E2 is DNA endonuclease. *Proc Natl Acad Sci U S A* **73**, 3989-3993 (1976).
- 92 Cooper, P. C. & James, R. Two new E colicins, E8 and E9, produced by a strain of *Escherichia coli*. *J Gen Microbiol* **130**, 209-215 (1984).
- 93 Kleanthous, C. *et al.* Structural and mechanistic basis of immunity toward endonuclease colicins. *Nat Struct Biol* **6**, 243-252, doi:10.1038/6683 (1999).

- 94 Ko, T. P., Liao, C. C., Ku, W. Y., Chak, K. F. & Yuan, H. S. The crystal structure of the DNase domain of colicin E7 in complex with its inhibitor Im7 protein. *Structure* **7**, 91-102 (1999).
- 95 Gorbalenya, A. E. Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci* **3**, 1117-1120, doi:10.1002/pro.5560030716 (1994).
- 96 Pommer, A. J. *et al.* Mechanism and cleavage specificity of the H-N-H endonuclease colicin E9. *J Mol Biol* **314**, 735-749, doi:10.1006/jmbi.2001.5189 (2001).
- 97 Shub, D. A., Goodrich-Blair, H. & Eddy, S. R. Amino acid sequence motif of group I intron endonucleases is conserved in open reading frames of group II introns. *Trends Biochem Sci* **19**, 402-404 (1994).
- 98 Kleanthous, C. & Walker, D. Immunity proteins: enzyme inhibitors that avoid the active site. *Trends Biochem Sci* **26**, 624-631 (2001).
- 99 Bowman, C. M. Inactivation of ribosomes by colicin E3 in vitro: Requirement for 50 S ribosomal subunits. *FEBS Lett* **22**, 73-75 (1972).
- 100 Bowman, C. M., Dahlberg, J. E., Ikemura, T., Konisky, J. & Nomura, M. Specific inactivation of 16S ribosomal RNA induced by colicin E3 in vivo. *Proc Natl Acad Sci U S A* **68**, 964-968 (1971).
- 101 Ogawa, T. *et al.* A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science* **283**, 2097-2100 (1999).
- 102 Tomita, K., Ogawa, T., Uozumi, T., Watanabe, K. & Masaki, H. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc Natl Acad Sci U S A* **97**, 8278-8283, doi:10.1073/pnas.140213797 (2000).
- 103 Carr, S., Walker, D., James, R., Kleanthous, C. & Hemmings, A. M. Inhibition of a ribosome-inactivating ribonuclease: the crystal structure of the cytotoxic domain of colicin E3 in complex with its immunity protein. *Structure* **8**, 949-960 (2000).
- 104 Walker, D., Lancaster, L., James, R. & Kleanthous, C. Identification of the catalytic motif of the microbial ribosome inactivating cytotoxin colicin E3. *Protein Sci* **13**, 1603-1611, doi:10.1110/ps.04658504 (2004).
- 105 Ng, C. L. *et al.* Structural basis for 16S ribosomal RNA cleavage by the cytotoxic domain of colicin E3. *Nat Struct Mol Biol* **17**, 1241-1246, doi:10.1038/nsmb.1896 (2010).
- 106 Ben-Zeev, E., Zarivach, R., Shoham, M., Yonath, A. & Eisenstein, M. Prediction of the structure of the complex between the 30S ribosomal subunit and colicin E3 via

- weighted-geometric docking. *J Biomol Struct Dyn* **20**, 669-676, doi:10.1080/07391102.2003.10506883 (2003).
- 107 Baan, R. A., Frijmann, M., van Knippenberg, P. H. & Bosch, L. Consequences of a specific cleavage in situ of 16-S ribosomal RNA for polypeptide chain elongation. *Eur J Biochem* **87**, 137-142 (1978).
- 108 Lancaster, L. E., Savelsbergh, A., Kleanthous, C., Wintermeyer, W. & Rodnina, M. V. Colicin E3 cleavage of 16S rRNA impairs decoding and accelerates tRNA translocation on Escherichia coli ribosomes. *Mol Microbiol* **69**, 390-401, doi:10.1111/j.1365-2958.2008.06283.x (2008).
- 109 Graille, M., Mora, L., Buckingham, R. H., van Tilbeurgh, H. & de Zamaroczy, M. Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity protein. *EMBO J* **23**, 1474-1482, doi:10.1038/sj.emboj.7600162 (2004).
- 110 Lin, Y. L., Elias, Y. & Huang, R. H. Structural and mutational studies of the catalytic domain of colicin E5: a tRNA-specific ribonuclease. *Biochemistry* **44**, 10494-10500, doi:10.1021/bi050749s (2005).
- 111 Luna-Chávez, C., Lin, Y. L. & Huang, R. H. Molecular basis of inhibition of the ribonuclease activity in colicin E5 by its cognate immunity protein. *J Mol Biol* **358**, 571-579, doi:10.1016/j.jmb.2006.02.014 (2006).
- 112 Aoki, S. K. *et al.* Contact-dependent inhibition of growth in Escherichia coli. *Science* **309**, 1245-1248, doi:10.1126/science.1115109 (2005).
- 113 Clantin, B. *et al.* The crystal structure of filamentous hemagglutinin secretion domain and its implications for the two-partner secretion pathway. *Proc Natl Acad Sci U S A* **101**, 6194-6199, doi:10.1073/pnas.0400291101 (2004).
- 114 Mazar, J. & Cotter, P. A. Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Mol Microbiol* **62**, 641-654, doi:10.1111/j.1365-2958.2006.05392.x (2006).
- 115 Aoki, S. K. *et al.* A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**, 439-442, doi:nature09490 [pii] 10.1038/nature09490 (2010).
- 116 Webb, J. S. *et al.* Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. *PLoS One* **8**, e57609, doi:10.1371/journal.pone.0057609 (2013).
- 117 Beck, C. M. *et al.* CdiA from Enterobacter cloacae delivers a toxic ribosomal RNase into target bacteria. *Structure* **22**, 707-718, doi:10.1016/j.str.2014.02.012 (2014).

- 118 Ruhe, Z. C., Wallace, A. B., Low, D. A. & Hayes, C. S. Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *MBio* **4**, doi:mBio.00480-13 [pii] 10.1128/mBio.00480-13 (2013).
- 119 Ruhe, Z. C., Nguyen, J. Y., Beck, C. M., Low, D. A. & Hayes, C. S. The proton-motive force is required for translocation of CDI toxins across the inner membrane of target bacteria. *Mol Microbiol* **94**, 466-481, doi:10.1111/mmi.12779 (2014).
- 120 Nikolakakis, K. *et al.* The toxin/immunity network of *Burkholderia pseudomallei* contact-dependent growth inhibition (CDI) systems. *Mol Microbiol* **84**, 516-529, doi:10.1111/j.1365-2958.2012.08039.x (2012).
- 121 Poole, S. J. *et al.* Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet* **7**, e1002217, doi:10.1371/journal.pgen.1002217 (2011).
- 122 Morse, R. P. *et al.* Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proc Natl Acad Sci U S A* **109**, 21480-21485, doi:10.1073/pnas.1216238110 (2012).
- 123 Diner, E. J., Beck, C. M., Webb, J. S., Low, D. A. & Hayes, C. S. Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). *Genes Dev* **26**, 515-525, doi:gad.182345.111 [pii] 10.1101/gad.182345.111 (2012).
- 124 Campanini, B. *et al.* Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein. *Biochim Biophys Acta*, doi:10.1016/j.bbapap.2015.02.013 (2015).
- 125 Aoki, S. K., Webb, J. S., Braaten, B. A. & Low, D. A. Contact-dependent growth inhibition causes reversible metabolic downregulation in *Escherichia coli*. *J Bacteriol* **191**, 1777-1786, doi:10.1128/JB.01437-08 (2009).
- 126 Aoki, S. K. *et al.* 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, doi:10.1111/j.1365-2958.2008.06404.x (2008).
- 127 Ho, B. T., Dong, T. G. & Mekalanos, J. J. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* **15**, 9-21, doi:10.1016/j.chom.2013.11.008 (2014).
- 128 Russell, A. B., Peterson, S. B. & Mougous, J. D. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* **12**, 137-148, doi:10.1038/nrmicro3185 (2014).
- 129 Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343-347, doi:10.1038/nature10244 (2011).

- 130 Russell, A. B. *et al.* A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. *Cell Host Microbe* **11**, 538-549, doi:10.1016/j.chom.2012.04.007 (2012).
- 131 Pukatzki, S. *et al.* Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc Natl Acad Sci U S A* **103**, 1528-1533, doi:10.1073/pnas.0510322103 (2006).
- 132 Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D. & Mekalanos, J. J. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104**, 15508-15513, doi:10.1073/pnas.0706532104 (2007).
- 133
- 134 Boyer, F., Fichant, G., Berthod, J., Vandembrouck, Y. & Attree, I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* **10**, 104, doi:10.1186/1471-2164-10-104 (2009).
- 135 Schwarz, S., Hood, R. D. & Mougous, J. D. What is type VI secretion doing in all those bugs? *Trends Microbiol* **18**, 531-537, doi:10.1016/j.tim.2010.09.001 (2010).
- 136 Zheng, J. & Leung, K. Y. Dissection of a type VI secretion system in *Edwardsiella tarda*. *Mol Microbiol* **66**, 1192-1206, doi:10.1111/j.1365-2958.2007.05993.x (2007).
- 137 Zheng, J., Ho, B. & Mekalanos, J. J. Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. *PLoS ONE* **6**, e23876, doi:10.1371/journal.pone.0023876PONE-D-11-08816 [pii] (2011).
- 138 Lin, J. S., Ma, L. S. & Lai, E. M. Systematic dissection of the agrobacterium type VI secretion system reveals machinery and secreted components for subcomplex formation. *PLoS One* **8**, e67647, doi:10.1371/journal.pone.0067647 (2013).
- 139 Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483**, 182-186, doi:10.1038/nature10846 (2012).
- 140 Leiman, P. G. *et al.* Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A* **106**, 4154-4159, doi:10.1073/pnas.0813360106 (2009).
- 141 Mesyanzhinov, V. V. Bacteriophage T4: structure, assembly, and initiation infection studied in three dimensions. *Adv Virus Res* **63**, 287-352, doi:10.1016/S0065-3527(04)63005-3 (2004).
- 142 Aschtgen, M. S., Gavioli, M., Dessen, A., Lloubès, R. & Cascales, E. The SciZ protein anchors the enteroaggregative *Escherichia coli* Type VI secretion system to

- the cell wall. *Mol Microbiol* **75**, 886-899, doi:10.1111/j.1365-2958.2009.07028.x (2010).
- 143 Aschtgen, M. S., Bernard, C. S., De Bentzmann, S., Llobès, R. & Cascales, E. SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative *Escherichia coli*. *J Bacteriol* **190**, 7523-7531, doi:10.1128/JB.00945-08 (2008).
- 144 Felisberto-Rodrigues, C. *et al.* Towards a structural comprehension of bacterial type VI secretion systems: characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar. *PLoS Pathog* **7**, e1002386, doi:10.1371/journal.ppat.1002386 (2011).
- 145 Ma, L. S., Lin, J. S. & Lai, E. M. An IcmF family protein, ImpLM, is an integral inner membrane protein interacting with ImpKL, and its walker a motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*. *J Bacteriol* **191**, 4316-4329, doi:10.1128/JB.00029-09 (2009).
- 146 Aschtgen, M. S., Zoued, A., Llobès, R., Journet, L. & Cascales, E. The C-tail anchored TssL subunit, an essential protein of the enteroaggregative *Escherichia coli* Sci-1 Type VI secretion system, is inserted by YidC. *Microbiologyopen* **1**, 71-82, doi:10.1002/mbo3.9 (2012).
- 147 Ma, L. S., Narberhaus, F. & Lai, E. M. IcmF family protein TssM exhibits ATPase activity and energizes type VI secretion. *J Biol Chem* **287**, 15610-15621, doi:10.1074/jbc.M111.301630 (2012).
- 148 Pell, L. G., Kanelis, V., Donaldson, L. W., Howell, P. L. & Davidson, A. R. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc Natl Acad Sci U S A* **106**, 4160-4165, doi:10.1073/pnas.0900044106 (2009).
- 149 Mougous, J. D. *et al.* A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**, 1526-1530, doi:10.1126/science.1128393 (2006).
- 150 Brunet, Y. R., Hénin, J., Celia, H. & Cascales, E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. *EMBO Rep* **15**, 315-321, doi:10.1002/embr.201337936 (2014).
- 151 Zoued, A. *et al.* Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta* **1843**, 1664-1673, doi:10.1016/j.bbamcr.2014.03.018 (2014).
- 152 Leiman, P. G. & Shneider, M. M. Contractile tail machines of bacteriophages. *Adv Exp Med Biol* **726**, 93-114, doi:10.1007/978-1-4614-0980-9\_5 (2012).
- 153 Shneider, M. M. *et al.* PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* **500**, 350-353, doi:10.1038/nature12453 (2013).



- 154 Brunet, Y. R., Espinosa, L., Harchouni, S., Mignot, T. & Cascales, E. Imaging type VI secretion-mediated bacterial killing. *Cell Rep* **3**, 36-41, doi:10.1016/j.celrep.2012.11.027 (2013).
- 155 Kapitein, N. *et al.* ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion. *Mol Microbiol* **87**, 1013-1028, doi:10.1111/mmi.12147 (2013).
- 156 Zoued, A. *et al.* TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. *J Biol Chem* **288**, 27031-27041, doi:10.1074/jbc.M113.499772 (2013).
- 157 Bönemann, G., Pietrosiuk, A., Diemand, A., Zentgraf, H. & Mogk, A. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J* **28**, 315-325, doi:10.1038/emboj.2008.269 (2009).
- 158 Pukatzki, S., McAuley, S. B. & Miyata, S. T. The type VI secretion system: translocation of effectors and effector-domains. *Curr Opin Microbiol* **12**, 11-17, doi:10.1016/j.mib.2008.11.010 (2009).
- 159 Brooks, T. M., Unterweger, D., Bachmann, V., Kostiuk, B. & Pukatzki, S. Lytic activity of the *Vibrio cholerae* type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB. *J Biol Chem* **288**, 7618-7625, doi:10.1074/jbc.M112.436725 (2013).
- 160 Chou, S. *et al.* Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. *Cell Rep* **1**, 656-664, doi:10.1016/j.celrep.2012.05.016 (2012).
- 161 LeRoux, M. *et al.* Quantitative single-cell characterization of bacterial interactions reveals type VI secretion is a double-edged sword. *Proc Natl Acad Sci U S A* **109**, 19804-19809, doi:10.1073/pnas.1213963109 (2012).
- 162 Srikannathasan, V. *et al.* Structural basis for type VI secreted peptidoglycan DL-endopeptidase function, specificity and neutralization in *Serratia marcescens*. *Acta Crystallogr D Biol Crystallogr* **69**, 2468-2482, doi:10.1107/S0907444913022725 (2013).
- 163 Whitney, J. C. *et al.* Identification, structure, and function of a novel type VI secretion peptidoglycan glycoside hydrolase effector-immunity pair. *J Biol Chem* **288**, 26616-26624, doi:10.1074/jbc.M113.488320 (2013).
- 164 Silverman, J. M. *et al.* Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. *Mol Cell* **51**, 584-593, doi:10.1016/j.molcel.2013.07.025 (2013).
- 165 Russell, A. B. *et al.* Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* **496**, 508-512, doi:10.1038/nature12074 (2013).

- 166 Aoki, S. K., Poole, S. J., Hayes, C. S. & Low, D. A. Toxin on a stick: modular CDI toxin delivery systems play roles in bacterial competition. *Virulence* **2**, 356-359, doi:16463 [pii] (2011).
- 167 Ruhe, Z. C., Low, D. A. & Hayes, C. S. Bacterial contact-dependent growth inhibition. *Trends Microbiol* **21**, 230-237, doi:S0966-842X(13)00023-1 [pii] 10.1016/j.tim.2013.02.003 (2013).
- 168 Aoki, S. K. *et al.* 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 (2008).
- 169 Webb, J. S. *et al.* Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. *PLoS ONE* **8**, e57609 (2013).
- 170 Aoki, S. K., Webb, J. S., Braaten, B. A. & Low, D. A. Contact-dependent growth inhibition causes reversible metabolic downregulation in *Escherichia coli*. *J. Bacteriol.* **191**, 1777-1786 (2009).
- 171 Morse, R. P. *et al.* Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proc Natl Acad Sci U S A* **109**, 21480-21485, doi:1216238110 [pii] 10.1073/pnas.1216238110 (2012).
- 172 Zhang, D., Iyer, L. M. & Aravind, L. 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, doi:gkr036 [pii] 10.1093/nar/gkr036 (2011).
- 173 Anderson, M. S., Garcia, E. C. & Cotter, P. A. The *Burkholderia* *bcpAIOB* genes define unique classes of two-partner secretion and contact dependent growth inhibition systems. *PLoS Genet* **8**, e1002877, doi:10.1371/journal.pgen.1002877 PGENETICS-D-12-00079 [pii] (2012).
- 174 Dautin, N., Barnard, T. J., Anderson, D. E. & Bernstein, H. D. Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J* **26**, 1942-1952, doi:7601638 [pii] 10.1038/sj.emboj.7601638 (2007).
- 175 Zarivach, R. *et al.* Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. *Nature* **453**, 124-127, doi:nature06832 [pii] 10.1038/nature06832 (2008).
- 176 Poole, S. J. *et al.* Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet.* **7**, e1002217, doi:10.1371/journal.pgen.1002217 PGENETICS-D-11-00234 [pii] (2011).

- 177 Beckwith, J. R. & Signer, E. R. 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 (1966).
- 178 Casadaban, M. J. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J Mol Biol* **104**, 541-555 (1976).
- 179 Garza-Sánchez, F., Janssen, B. D. & Hayes, C. S. Prolyl-tRNA(Pro) in the A-site of SecM-arrested ribosomes inhibits the recruitment of transfer-messenger RNA. *J Biol Chem* **281**, 34258-34268, doi:M608052200 [pii] 10.1074/jbc.M608052200 (2006).
- 180 Bolivar, F. *et al.* Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95-113 (1977).
- 181 Chang, A. C. & Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141-1156 (1978).
- 182 Hayes, C. S. & Sauer, R. T. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol. Cell* **12**, 903-911, doi:S109727650300385X [pii] (2003).
- 183 Johnson, R. C. & Reznikoff, W. S. Copy number control of Tn5 transposition. *Genetics* **107**, 9-18 (1984).
- 184 Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006 0008 (2006).
- 185 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640-6645, doi:10.1073/pnas.120163297120163297 [pii] (2000).
- 186 Datta, S., Costantino, N. & Court, D. L. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**, 109-115, doi:10.1016/j.gene.2006.04.018 (2006).
- 187 Aiyar, A. & Leis, J. Modification of the megaprimer method of PCR mutagenesis: improved amplification of the final product. *Biotechniques* **14**, 366-369 (1993).
- 188 Beck, C. M. *et al.* CdiA from *Enterobacter cloacae* delivers a toxic ribosomal RNase into target bacteria. *Structure* **22**, 707-718, doi:S0969-2126(14)00052-5 [pii] 10.1016/j.str.2014.02.012 (2014).
- 189 Cascales, E. *et al.* Colicin biology. *Microbiol. Mol. Biol. Rev.* **71**, 158-229, doi:71/1/158 [pii] 10.1128/MMBR.00036-06 (2007).

- 190 Guyer, M. S., Reed, R. R., Steitz, J. A. & Low, K. B. Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring Harb Symp Quant Biol* **45 Pt 1**, 135-140 (1981).
- 191 Achtman, M., Kennedy, N. & Skurray, R. Cell-cell interactions in conjugating *Escherichia coli*: role of *traT* protein in surface exclusion. *Proc Natl Acad Sci U S A* **74**, 5104-5108 (1977).
- 192 Manning, P. A., Beutin, L. & Achtman, M. Outer membrane of *Escherichia coli*: properties of the F sex factor *traT* protein which is involved in surface exclusion. *J Bacteriol* **142**, 285-294 (1980).
- 193 Manchak, J., Anthony, K. G. & Frost, L. S. Mutational analysis of F-pilin reveals domains for pilus assembly, phage infection and DNA transfer. *Mol Microbiol* **43**, 195-205, doi:2731 [pii] (2002).
- 194 Daehnel, K., Harris, R., Maddera, L. & Silverman, P. Fluorescence assays for F-pili and their application. *Microbiology* **151**, 3541-3548, doi:151/11/3541 [pii] 10.1099/mic.0.28159-0 (2005).
- 195 Lin, A. *et al.* Inhibition of bacterial conjugation by phage M13 and its protein g3p: quantitative analysis and model. *PLoS One* **6**, e19991, doi:10.1371/journal.pone.0019991 PONE-D-11-00061 [pii] (2011).
- 196 Clarke, M., Maddera, L., Harris, R. L. & Silverman, P. M. F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci U S A* **105**, 17978-17981, doi:0806786105 [pii] 10.1073/pnas.0806786105 (2008).
- 197 Maneewannakul, S., Maneewannakul, K. & Ippen-Ihler, K. Characterization, localization, and sequence of F transfer region products: the pilus assembly gene product TraW and a new product, TrbI. *J Bacteriol* **174**, 5567-5574 (1992).
- 198 Novotny, C. P. & Fives-Taylor, P. Retraction of F pili. *J Bacteriol* **117**, 1306-1311 (1974).
- 199 Moore, D., Sowa, B. A. & Ippen-Ihler, K. Location of an F-pilin pool in the inner membrane. *J Bacteriol* **146**, 251-259 (1981).
- 200 Chandran, V. *et al.* Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**, 1011-1015, doi:nature08588 [pii] 10.1038/nature08588 (2009).
- 201 Waksman, G. & Fronzes, R. Molecular architecture of bacterial type IV secretion systems. *Trends Biochem Sci* **35**, 691-698, doi:S0968-0004(10)00114-3 [pii] 10.1016/j.tibs.2010.06.002 (2010).
- 202 Silverman, P. M. & Clarke, M. B. New insights into F-pilus structure, dynamics, and function. *Integr Biol (Camb)* **2**, 25-31, doi:10.1039/b917761b (2010).

- 203 Burke, J. M., Novotny, C. P. & Fives-Taylor, P. Defective F pili and other characteristics of *Flac* and Hfr *Escherichia coli* mutants resistant to bacteriophage R17. *J Bacteriol* **140**, 525-531 (1979).
- 204 Willetts, N. S., Moore, P. M. & Paranchych, W. Variant pili produced by mutants of the *Flac* plasmid. *J Gen Microbiol* **117**, 455-464 (1980).
- 205 Riechmann, L. & Holliger, P. The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell* **90**, 351-360, doi:S0092-8674(00)80342-6 [pii] (1997).
- 206 Roberts, J. W. & Steitz, J. E. The reconstitution of infective bacteriophage R17. *Proc Natl Acad Sci U S A* **58**, 1416-1421 (1967).
- 207 Steitz, J. A. Identification of the A protein as a structural component of bacteriophage R17. *J Mol Biol* **33**, 923-936, doi:0022-2836(68)90328-8 [pii] (1968).
- 208 Dent, K. C. *et al.* The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. *Structure* **21**, 1225-1234, doi:S0969-2126(13)00194-9 [pii] 10.1016/j.str.2013.05.012 (2013).
- 209 Krahn, P. M., O'Callaghan, R. J. & Paranchych, W. Stages in phage R17 infection. VI. Injection of A protein and RNA into the host cell. *Virology* **47**, 628-637 (1972).
- 210 Paranchych, W., Ainsworth, S. K., Dick, A. J. & Krahn, P. M. Stages in phage R17 infection. V. Phage eclipse and the role of F pili. *Virology* **45**, 615-628 (1971).
- 211 Paranchych, W., Krahn, P. M. & Bradley, R. D. Stages in phage R17 infection. *Virology* **41**, 465-473, doi:0042-6822(70)90168-6 [pii] (1970).
- 212 Oriol, P. J. The thermal and alkaline degradation of MS2 bacteriophage. *Arch Biochem Biophys* **132**, 8-15 (1969).
- 213 Osborn, M., Weiner, A. M. & Weber, K. Large scale purification of A-protein from bacteriophage R17. *Eur J Biochem* **17**, 63-67 (1970).
- 214 Lang, S., Gruber, C. J., Raffl, S., Reisner, A. & Zechner, E. L. Common Requirement for the Relaxosome of Plasmid R1 in Multiple Activities of the Conjugative Type IV Secretion System. *J Bacteriol* **196**, 2108-2121, doi:JB.00045-13 [pii] 10.1128/JB.00045-13 (2014).
- 215 Lang, S. *et al.* An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. *Mol Microbiol* **82**, 1071-1085, doi:10.1111/j.1365-2958.2011.07872.x (2011).
- 216 Schoulaker, R. & Engelberg-Kulka, H. *Escherichia coli* mutant temperature sensitive for group I RNA bacteriophages. *J Virol* **25**, 433-435 (1978).

- 217 Willetts, N. & Achtman, M. Genetic analysis of transfer by the *Escherichia coli* sex factor F, using P1 transductional complementation. *J Bacteriol* **110**, 843-851 (1972).
- 218 Rumnieks, J. & Tars, K. Diversity of pili-specific bacteriophages: genome sequence of IncM plasmid-dependent RNA phage M. *BMC Microbiol* **12**, 277, doi:1471-2180-12-277 [pii] 10.1186/1471-2180-12-277 (2012).
- 219 Friedman, S. D., Genthner, F. J., Gentry, J., Sobsey, M. D. & Vinje, J. Gene mapping and phylogenetic analysis of the complete genome from 30 single-stranded RNA male-specific coliphages (family Leviviridae). *J Virol* **83**, 11233-11243, doi:JVI.01308-09 [pii] 10.1128/JVI.01308-09 (2009).
- 220 Kajava, A. V. *et al.* Beta-helix model for the filamentous haemagglutinin adhesin of *Bordetella pertussis* and related bacterial secretory proteins. *Mol Microbiol* **42**, 279-292 (2001).
- 221 Kim, S. *et al.* Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961-964, doi:10.1126/science.1143993 (2007).
- 222 Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262-265, doi:10.1126/science.1078973 (2003).
- 223 Smith, D. L. *et al.* Short-tailed stx phages exploit the conserved YaeT protein to disseminate Shiga toxin genes among enterobacteria. *J Bacteriol* **189**, 7223-7233, doi:10.1128/JB.00824-07 (2007).
- 224 Beck, C. M., Diner, E. J., Kim, J. J., Low, D. A. & Hayes, C. S. The F pilus mediates a novel pathway of CDI toxin import. *Mol Microbiol* **93**, 276-290, doi:10.1111/mmi.12658 (2014).
- 225 Cherepanov, P. P. & Wackernagel, W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14 (1995).
- 226 Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006.0008, doi:10.1038/msb4100050 (2006).
- 227 Pratt, L. A., Hsing, W., Gibson, K. E. & Silhavy, T. J. From acids to osmZ: multiple factors influence synthesis of the OmpF and OmpC porins in *Escherichia coli*. *Mol Microbiol* **20**, 911-917 (1996).
- 228 Yamashita, E., Zhalnina, M. V., Zakharov, S. D., Sharma, O. & Cramer, W. A. Crystal structures of the OmpF porin: function in a colicin translocon. *EMBO J* **27**, 2171-2180, doi:10.1038/emboj.2008.137 (2008).

- 229 Baslé, A., Rummel, G., Storici, P., Rosenbusch, J. P. & Schirmer, T. Crystal structure of osmoporin OmpC from *E. coli* at 2.0 Å. *J Mol Biol* **362**, 933-942, doi:10.1016/j.jmb.2006.08.002 (2006).
- 230 Petersen, L., Bollback, J. P., Dimmic, M., Hubisz, M. & Nielsen, R. Genes under positive selection in *Escherichia coli*. *Genome Res* **17**, 1336-1343, doi:10.1101/gr.6254707 (2007).
- 231 Lou, H. *et al.* Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant *E. coli*. *PLoS One* **6**, e25825, doi:10.1371/journal.pone.0025825 (2011).
- 232 Nikaido, H. Porins and specific diffusion channels in bacterial outer membranes. *J Biol Chem* **269**, 3905-3908 (1994).
- 233 Gehring, K. B. & Nikaido, H. Existence and purification of porin heterotrimers of *Escherichia coli* K12 OmpC, OmpF, and PhoE proteins. *J Biol Chem* **264**, 2810-2815 (1989).
- 234 Chen, S. L. *et al.* Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci U S A* **103**, 5977-5982, doi:10.1073/pnas.0600938103 (2006).
- 235 Stewart, P. S. & Franklin, M. J. Physiological heterogeneity in biofilms. *Nat Rev Microbiol* **6**, 199-210, doi:10.1038/nrmicro1838 (2008).
- 236 West, S. A. & Gardner, A. Altruism, spite, and greenbeards. *Science* **327**, 1341-1344, doi:10.1126/science.1178332 (2010).
- 237 Strassmann, J. E., Gilbert, O. M. & Queller, D. C. Kin discrimination and cooperation in microbes. *Annu Rev Microbiol* **65**, 349-367, doi:10.1146/annurev.micro.112408.134109 (2011).
- 238 Cascales, E. *et al.* Colicin biology. *Microbiol Mol Biol Rev* **71**, 158-229, doi:10.1128/MMBR.00036-06 [pii] 10.1128/MMBR.00036-06 (2007).
- 239 Senior, B. W. & Holland, I. B. Effect of colicin E3 upon the 30S ribosomal subunit of *Escherichia coli*. *Proc Natl Acad Sci U S A* **68**, 959-963 (1971).
- 240 Chao, L. & Levin, B. R. Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proc Natl Acad Sci U S A* **78**, 6324-6328 (1981).
- 241 Czaran, T. L., Hoekstra, R. F. & Pagie, L. Chemical warfare between microbes promotes biodiversity. *Proc Natl Acad Sci U S A* **99**, 786-790, doi:10.1073/pnas.012399899012399899 [pii] (2002).

- 242 Hood, R. D. *et al.* A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* **7**, 25-37, doi:S1931-3128(09)00417-X [pii] 10.1016/j.chom.2009.12.007 (2010).
- 243 MacIntyre, D. L., Miyata, S. T., Kitaoka, M. & Pukatzki, S. The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A* **107**, 19520-19524, doi:1012931107 [pii] 10.1073/pnas.1012931107 (2010).
- 244 Silverman, J. M., Brunet, Y. R., Cascales, E. & Mougous, J. D. Structure and regulation of the type VI secretion system. *Annu Rev Microbiol* **66**, 453-472, doi:10.1146/annurev-micro-121809-151619 (2012).
- 245 Holberger, L. E., Garza-Sanchez, F., Lamoureux, J., Low, D. A. & Hayes, C. S. A novel family of toxin/antitoxin proteins in *Bacillus* species. *FEBS Lett* **586**, 132-136, doi:S0014-5793(11)00892-1 [pii] 10.1016/j.febslet.2011.12.020 (2012).
- 246 Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M. & Aravind, L. Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct* **7**, 18, doi:1745-6150-7-18 [pii] 10.1186/1745-6150-7-18 (2012).
- 247 Thomason, L. *et al.* Recombineering: genetic engineering in bacteria using homologous recombination. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al Chapter 1*, Unit 1 16 (2007).
- 248 Garza-Sanchez, F., Janssen, B. D. & Hayes, C. S. Prolyl-tRNA(Pro) in the A-site of SecM-arrested ribosomes inhibits the recruitment of transfer-messenger RNA. *J Biol Chem* **281**, 34258-34268, doi:M608052200 [pii] 10.1074/jbc.M608052200 (2006).
- 249 Hayes, C. S. & Sauer, R. T. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol Cell* **12**, 903-911 (2003).
- 250 Chaverroche, M. K., Ghigo, J. M. & d'Enfert, C. A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res* **28**, E97 (2000).
- 251 Hayes, C. S., Bose, B. & Sauer, R. T. Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *J Biol Chem* **277**, 33825-33832, doi:10.1074/jbc.M205405200 (2002).
- 252 Koskiniemi, S. *et al.* Rhs proteins from diverse bacteria mediate intercellular competition. *Proc Natl Acad Sci U S A* **110**, 7032-7037, doi:1300627110 [pii] 10.1073/pnas.1300627110 (2013).
- 253 Perez, A. *et al.* Cloning, nucleotide sequencing, and analysis of the AcrAB-TolC efflux pump of *Enterobacter cloacae* and determination of its involvement in



- antibiotic resistance in a clinical isolate. *Antimicrob Agents Chemother* **51**, 3247-3253, doi:AAC.00072-07 [pii] 10.1128/AAC.00072-07 (2007).
- 254 Aiyar, A., Xiang, Y. & Leis, J. Site-directed mutagenesis using overlap extension PCR. *Methods Mol Biol* **57**, 177-191, doi:10.1385/0-89603-332-5:177 (1996).
- 255 Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105-124, doi:S0022-2836(83)71012-0 [pii] 10.1006/jmbi.1993.1012 (1993).
- 256 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
- 257 Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601, doi:S0907444909012098 [pii] 10.1107/S0907444909012098 (2009).
- 258 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:S0907444909052925 [pii] 10.1107/S0907444909052925 (2010).
- 259 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:S0907444904019158 [pii] 10.1107/S0907444904019158 (2004).
- 260
- 261 Diner, E. J. & Hayes, C. S. Recombineering reveals a diverse collection of ribosomal proteins L4 and L22 that confer resistance to macrolide antibiotics. *J Mol Biol* **386**, 300-315, doi:S0022-2836(08)01600-8 [pii] 10.1016/j.jmb.2008.12.064 (2009).
- 262 Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549, doi:gkq366 [pii] 10.1093/nar/gkq366 (2010).
- 263 Soelaiman, S., Jakes, K., Wu, N., Li, C. & Shoham, M. Crystal structure of colicin E3: implications for cell entry and ribosome inactivation. *Mol Cell* **8**, 1053-1062, doi:S1097-2765(01)00396-3 [pii] (2001).
- 264 Desveaux, D., Marechal, A. & Brisson, N. Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci* **10**, 95-102, doi:S1360-1385(04)00279-1 [pii] 10.1016/j.tplants.2004.12.008 (2005).
- 265 Rojas, C. M., Ham, J. H., Deng, W. L., Doyle, J. J. & Collmer, A. HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Erwinia*

- chrysanthemi* EC16 on *Nicotiana clevelandii* seedlings. *Proc Natl Acad Sci U S A* **99**, 13142-13147, doi:10.1073/pnas.202358699202358699 [pii] (2002).
- 266 Basturea, G. N., Rudd, K. E. & Deutscher, M. P. Identification and characterization of RsmE, the founding member of a new RNA base methyltransferase family. *RNA* **12**, 426-434, doi:rna.2283106 [pii] 10.1261/rna.2283106 (2006).
- 267 Richardson, R. M., Pares, X. & Cuchillo, C. M. Chemical modification by pyridoxal 5'-phosphate and cyclohexane-1,2-dione indicates that Lys-7 and Arg-10 are involved in the p2 phosphate-binding subsite of bovine pancreatic ribonuclease A. *Biochem J* **267**, 593-599 (1990).
- 268 Gite, S., Reddy, G. & Shankar, V. Active-site characterization of S1 nuclease. I. Affinity purification and influence of amino-group modification. *Biochem J* **285** ( Pt 2), 489-494 (1992).
- 269 Pingoud, A. & Jeltsch, A. Structure and function of type II restriction endonucleases. *Nucleic Acids Res* **29**, 3705-3727 (2001).
- 270 Riley, M. A. Positive selection for colicin diversity in bacteria. *Mol Biol Evol* **10**, 1048-1059 (1993).
- 271 Tan, Y. & Riley, M. A. Positive selection and recombination: major molecular mechanisms in colicin diversification. *Trends Ecol Evol* **12**, 348-351, doi:S0169-5347(97)01127-0 [pii] (1997).
- 272 Paruchuri, D. K., Seifert, H. S., Ajioka, R. S., Karlsson, K. A. & So, M. Identification and characterization of a *Neisseria gonorrhoeae* gene encoding a glycolipid-binding adhesin. *Proc Natl Acad Sci U S A* **87**, 333-337 (1990).
- 273 Sekowska, A., Kung, H. F. & Danchin, A. Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. *J Mol Microbiol Biotechnol* **2**, 145-177 (2000).
- 274 Levy, S. & Danchin, A. Phylogeny of metabolic pathways: O-acetylserine sulphhydrylase A is homologous to the tryptophan synthase beta subunit. *Mol Microbiol* **2**, 777-783 (1988).
- 275 Kredich, N. M. The molecular basis for positive regulation of *cys* promoters in *Salmonella typhimurium* and *Escherichia coli*. *Mol Microbiol* **6**, 2747-2753 (1992).
- 276 Wang, T. & Leyh, T. S. Three-stage assembly of the cysteine synthase complex from *Escherichia coli*. *J Biol Chem* **287**, 4360-4367, doi:10.1074/jbc.M111.288423 (2012).
- 277 Rodriguez-Hernandez, A. *et al.* Structural and mechanistic basis for enhanced translational efficiency by 2-thiouridine at the tRNA anticodon wobble position. *J Mol Biol* **425**, 3888-3906, doi:10.1016/j.jmb.2013.05.018 (2013).

- 278 Pingoud, A., Fuxreiter, M., Pingoud, V. & Wende, W. Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci* **62**, 685-707, doi:10.1007/s00018-004-4513-1 (2005).
- 279 Diniz, J. A. & Coulthurst, S. J. Intra-species Competition in *Serratia marcescens* is Mediated by Type VI Secretion Rhs Effectors and a Conserved Effector-Associated Accessory Protein. *J Bacteriol*, doi:10.1128/JB.00199-15 (2015).
- 280 Busby, J. N., Panjikar, S., Landsberg, M. J., Hurst, M. R. & Lott, J. S. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature* **501**, 547-550, doi:10.1038/nature12465 (2013).
- 281 Basler, M., Ho, B. T. & Mekalanos, J. J. Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152**, 884-894, doi:10.1016/j.cell.2013.01.042 (2013).
- 282 Burtnick, M. N. *et al.* The cluster 1 type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*. *Infect Immun* **79**, 1512-1525, doi:10.1128/IAI.01218-10 (2011).
- 283 Miyata, S. T., Kitaoka, M., Brooks, T. M., McAuley, S. B. & Pukatzki, S. *Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*. *Infect Immun* **79**, 2941-2949, doi:10.1128/IAI.01266-10 (2011).
- 284 Schell, M. A. *et al.* Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol Microbiol* **64**, 1466-1485, doi:10.1111/j.1365-2958.2007.05734.x (2007).
- 285 Packey, C. D. & Sartor, R. B. Commensal bacteria, traditional and opportunistic pathogens, dysbiosis and bacterial killing in inflammatory bowel diseases. *Curr Opin Infect Dis* **22**, 292-301, doi:10.1097/QCO.0b013e32832a8a5d (2009).