Effectors of cell death in bacterial antiphage defense

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

Committee on Virology

have examined a dissertation entitled

Effectors of cell death in bacterial antiphage defense

presented by Brianna Duncan-Lowey

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: Sun Hur

Typed Name: Dr. Sun Hur

Signature:

Typed Name: Dr. Sophie Helaine

Signature:

Typed Name: Dr. Suzanne Walker

Signature: Malcolm White (Nov 1, 2021 09:01 GMT)

Typed Name: Dr. Malcolm White

Date: September 14, 2021

Effectors of cell death in bacterial antiphage defense

A dissertation presented

by

Brianna Duncan-Lowey

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Virology

Harvard University

Cambridge, Massachusetts

September 2021

Dissertation Advisor: Philip Kranzusch Brianna Duncan-Lowey

Effectors of cell death in bacterial antiphage defense

Abstract

Bacteria encode many systems to detect and respond to infection with bacteriophages. Cyclic-oligonucleotide based antiphage signaling systems (CBASS) are widespread antiviral systems encoded in approximately 10% of bacterial genomes. Upon infection of CBASS-encoding bacteria, nucleotide second messengers are synthesized and can diffuse throughout the cell to bind diverse effector proteins. We use structural and biochemical methods to characterize effector proteins that specifically recognize nucleotide second messengers and are activated to induce cell death. Once this abortive infection system is activated, the effector proteins kill the bacterial host before the phage is able to replicate, thereby halting phage infection. We characterize one family of effectors, Cap4 proteins, that use a SAVED domain to specifically recognize nucleotide second messengers and are then activated to kill the cell through the indiscriminate cleavage of double-stranded DNA. These data highlight SAVED domains as widespread ligand-binding domains found in many CBASS effectors and revealed an evolutionary connection between CBASS and CRISPR immunity. We also characterized effector proteins that contain transmembrane domains, revealing that they target the inner membrane to induce cell death after phage infection. These data highlight membrane disruption as a widespread strategy to induce cell death in CBASS immunity. We further characterize one family of transmembrane effectors, Cap15 proteins, which we show use a minimal β-barrel domain to recognize nucleotide second messengers. Together, these studies begin to characterize diverse CBASS effectors that fulfill two requirements: 1. specific recognition of nucleotide second messengers, either by SAVED or β-barrel domains and 2. the induction of

cell death, either through destruction of nucleic acids by Cap4 nucleases or inner membrane disruption by Cap15 transmembrane effectors. These data support an emerging model where CBASS effectors use a modular domain architecture to sense second messengers and induce cell death to halt phage replication.

Table of Contents

Table of Figures

Acknowledgements

I would like to thank my advisory committee, Sun Hur, Sean Whelan, Ann Hochschild, and James DeCaprio for their advice and support. I would also like to thank the Harvard Virology community and Microbiology department, for creating a welcoming environment and so generously offering their time and expertise. I am also grateful to my Virology classmates, who have provided a great environment to develop ideas and have made Boston feel like home.

Thank you to the Kranzusch lab, both past and present, who have provided endless and enthusiastic input and created such a fun environment to do science. I would like to especially thank Nora McNamara-Bordewick, for being such an optimistic and talented scientist, making it a joy to work together.

I would also like to thank my mentor, Philip Kranzusch, for encouraging me to take on challenging projects and always remaining excited about new data every step of the way. I have learned so much these past four years, not only about how to perform experiments, but to identify important questions and effectively communicate results. Learning every aspect of being a scientist in his lab has been an invaluable experience.

I am also grateful to my family for their support. I would especially like to thank my mom, Noreen Lowey, who has learned more about things like the peer review process and beamtime than she ever cared to know, so she could support me throughout this journey.

Finally, I would like to thank my husband, Jeffrey Duncan-Lowey, for being so whole-heartedly supportive and always believing me to be capable of anything. Not only has he been personally supportive, but he has also been an indispensable resource as a knowledgeable bacteriologist. In so many ways, this work would not have been possible without him.

Introduction

Cyclic oligonucleotide-based antiphage signaling system (CBASS) immunity is a widespread form of antiviral defense in bacteria that inhibits phage replication (Cohen et al., 2019). Following infection, activation of CBASS immunity triggers rapid cell death and results in an abortive infection response that limits viral spread within the bacterial population (Cohen et al., 2019; Lowey et al., 2020; Ye et al., 2020). More than 5,000 CBASS operons have been identified in diverse bacteria and archaea, with each system responding to phage through synthesis of a specialized nucleotide second messenger signal (Cohen et al., 2019; Whiteley et al., 2019). CBASS operons are found in genomes of many prokaryotes, demonstrating a broad importance in enabling bacteria to resist phage replication.

CBASS operons function as self-contained antiviral defense systems that typically encode two to four protein components (Figure I.1a) (Cohen et al., 2019; Millman et al., 2020). All CBASS operons contain a cGAS/DncV-like Nucleotidyltransferase (CD-NTase) enzyme that senses phage replication and catalyzes synthesis of a nucleotide second messenger signal to initiate antiviral defense (Figure I.1b) (Whiteley et al., 2019). CD-NTases are named according to clade designations A–H, for example the CBASS operon in *Enterobacter cloacae* encodes the enzyme CD-NTase in clade D (CdnD). Each CBASS operon also encodes a CD-NTase-associated protein (Cap) effector that functions as a receptor to specifically recognize the CD-NTase nucleotide second messenger signal (Lowey et al., 2020; Millman et al., 2020; Severin et al., 2018). Upon nucleotide-binding, CBASS Cap effectors

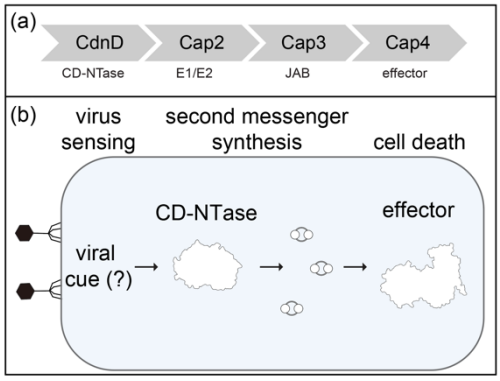


Figure I.1 Overview of cyclic oligonucleotide-based antiphage signaling systems (CBASS)
(a) Representative CBASS operon from the bacterium *Acinetobacter baumannii*. Each CBASS operon encodes a CD-NTase (*e.g.* CdnD) and an effector protein (e.g. Cap4). CBASS operons also frequently encode accessory proteins including Cap2 and Cap3. (b) Schematic overview of CBASS signaling. Upon phage infection, CD-NTases are activated to produce nucleotide second messengers. The nucleotide second messenger signals by directly binding to the effector protein and inducing a cell death response that prevents phage propagation.

become activated and induce cell death through degradation of nucleic acid, depletion of essential cellular metabolites, or direct targeting of the host cell membrane (Duncan-Lowey et al., Chapter 2; Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Severin et al., 2018). Cap protein names are numbered according to domain organization; for example, the effector family containing a restriction endonuclease-like domain and a SAVED domain is called Cap4. Some CBASS operons contain additional accessory *Cap* genes that are hypothesized to regulate CD-NTase function (Figure I.1a) (Cohen et al., 2019; Millman et al., 2020; Whiteley et al., 2019).

Here we summarize recent advances that explain the mechanism of CBASS immune defense. We define general principles that control each step of CBASS signaling including CD-NTase activation, nucleotide second messenger synthesis, and downstream effector function. Surprisingly, structural analysis of CD-NTases and Cap effectors in CBASS immunity has revealed remarkable homology with components of CRISPR antiphage defense and animal cGAS-STING innate immune signaling (Kranzusch et al., 2014; Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019). We highlight features of CBASS immunity that are shared among other signaling systems and describe the evolutionary roots that connect antiviral immunity in bacteria and animal cells.

CD-NTase enzymes

CD-NTase proteins are bacterial and archaeal signaling enzymes that are present in all CBASS operons (Millman et al., 2020; Whiteley et al., 2019). The CD-NTase family is named after the *Vibrio cholerae* enzyme Dinucleotide cyclase in *Vibrio* (DncV) (Davies et al., 2012) and the mammalian cytosolic DNA sensor cyclic GMP–AMP synthase (cGAS) (Sun et al., 2013). The discovery that CBASS CD-NTases are structural and functional homologs of human cGAS and animal cGAS-like receptors (cGLRs) demonstrates a direct evolutionary link between human innate immunity and bacterial antiphage defense (Cohen et al., 2019; Kranzusch et al., 2014; Whiteley et al., 2019).

In response to phage infection, bacterial CD-NTases catalyze synthesis of a nucleotide second messenger signal to initiate downstream antiphage defense. CD-NTase nucleotide second messenger synthesis is known to be induced upon phage infection (Cohen et al., 2019), but the viral cue responsible for CD-NTase activation in CBASS immunity is currently unknown. In human cells, cGAS binds directly to double-stranded DNA mislocalized in the cell cytosol and undergoes a conformational change that results in enzyme activation (Figure I.2c) (Civril et al., 2013; Gao et al., 2013a; Zhou et al., 2018). Other cGLRs including *Drosophila* cGLR1 sense double-stranded RNA, demonstrating that animal cGAS-like enzymes are capable of responding to diverse nucleic acid ligands (Holleufer et al., 2021; Slavik et al., 2021). However, bacterial CD-NTase enzymes do not respond to nucleic acid *in vitro* (Kranzusch et al., 2014; Whiteley et al., 2019), and the lack of spatial compartmentalization in bacteria suggests CBASS immunity does not function through recognition of mislocalized DNA. In contrast to cGAS, many bacterial CD-NTases are robustly active *in vitro* without the addition of an activating ligand (Whiteley et al., 2019). Additionally, structures of *apo* bacterial CD-NTases closely resemble the active conformation observed in cGAS bound to double-stranded DNA (Figure I.2b,c) (Ko et al., 2021; Kranzusch et al., 2014; Whiteley et al., 2019; Zhu et al., 2014). These results are consistent with a model where CBASS immune systems may be held in an

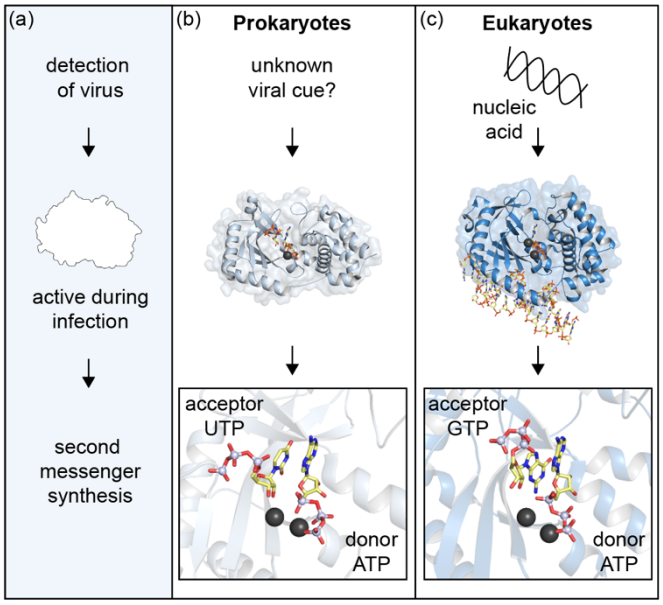


Figure I.2. Bacterial CD-NTases
(a) Overview of CD-NTase function during viral infection. Bacterial CD-NTases sense viral infection and then initiate CBASS defense by producing a nucleotide second messenger. The activating cue sensed during infection is currently unknown. (b) Cartoon representation of the bacterial CD-NTase *Rm*CdnE (top) that synthesizes 3′3′-c-UMP–AMP using a polymerase-like active site (bottom). (c) Cartoon representation of human cGAS (top) and the cGAS active site (bottom), highlighting the structural similarity between animal cGAS-like receptors (cGLRs) and bacterial CD-NTase enzymes.

inactive state by an inhibitory molecule that represses CD-NTase activation in the absence of phage infection (Kranzusch, 2019; Kranzusch et al., 2014; Lowey et al., 2020; Whiteley et al., 2019; Ye et al., 2020; Zhu et al., 2014). Defining the activating cue that initiates CD-NTase signaling in bacteria is a major open question in CBASS immunity.

Bacterial CD-NTases and animal cGLRs share a cage-like architecture formed by an N-terminal nucleotidyltransferase core and a C-terminal α-helix bundle (Figure I.2b,c bottom) (Whiteley et al., 2019). The CD-NTase active site functions as a monomeric unit and catalyzes phosphodiester bond formation through a conserved two-metal dependent mechanism (Kranzusch, 2019). The active site contains distinct nucleotide donor and acceptor pockets with side chains from the enzyme lid that determine nucleotide second messenger product specificity (Govande et al., 2021; Whiteley et al., 2019). *Rhodothermus marinus* CdnE, for example, synthesizes the nucleotide second messenger 3′3′-c-UMP–AMP by selecting an ATP molecule in the donor pocket and a UTP molecule in the acceptor pocket, then forming a phosphodiester bond between the 5′ phosphate of the ATP and the 3′ OH of the UTP. The linear intermediate molecule is flipped within the active site and a second phosphodiester bond is added to release the final cyclized product 3′3′-c-UMP–AMP (Whiteley et al., 2019). cGAS similarly uses a single active site to synthesize the nucleotide second messenger 2′3′-cGAMP (Figure I.2c) (Ablasser et al., 2013; Diner et al., 2013; Gao et al., 2013a; Zhang et al., 2013).

Nucleotide second messenger signals

A key mechanistic aspect of CBASS is the use of antiviral nucleotide second messenger signaling to dramatically amplify initial detection into a robust immune response. Activation of a few CD-NTase or cGAS enzymes following viral infection causes multi-turnover production of hundreds of molecules of nucleotide second messenger product (Figure I.3a) (Gao et al., 2013a; Govande et al., 2021; Whiteley et al., 2019). CD-NTase enzymes in bacteria are extremely diverse at the primary amino-acid level (typically <20% amino acid identity), and many distinct nucleotide second messenger products are used in CBASS immune systems (Davies et al., 2012;

Govande et al., 2021; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019; Ye et al., 2020). The prototypical CD-NTase, *Vibrio cholerae* DncV, synthesizes the cyclic dinucleotide 3′3′-cGAMP composed of two purine bases (Davies et al., 2012; Kranzusch et al., 2014). More recently, bacterial CD-NTases have been characterized that produce pyrimidine-containing signals including 3′3′-c-UMP–AMP and 3′3′-c-UMP–CMP, cyclic trinucleotides including 3′3′3′-c-AMP–AMP–AMP and 3′3′3′-c-AMP–AMP–GMP, and second messengers containing non-canonical 2′–5′ phosphodiester bonds, including 2′3′3′-c-AMP–AMP–AMP (Lowey et al., 2020; Whiteley et al., 2019; Ye et al., 2020) (Figure I.3c). Thus far, 6 distinct CD-NTase products have

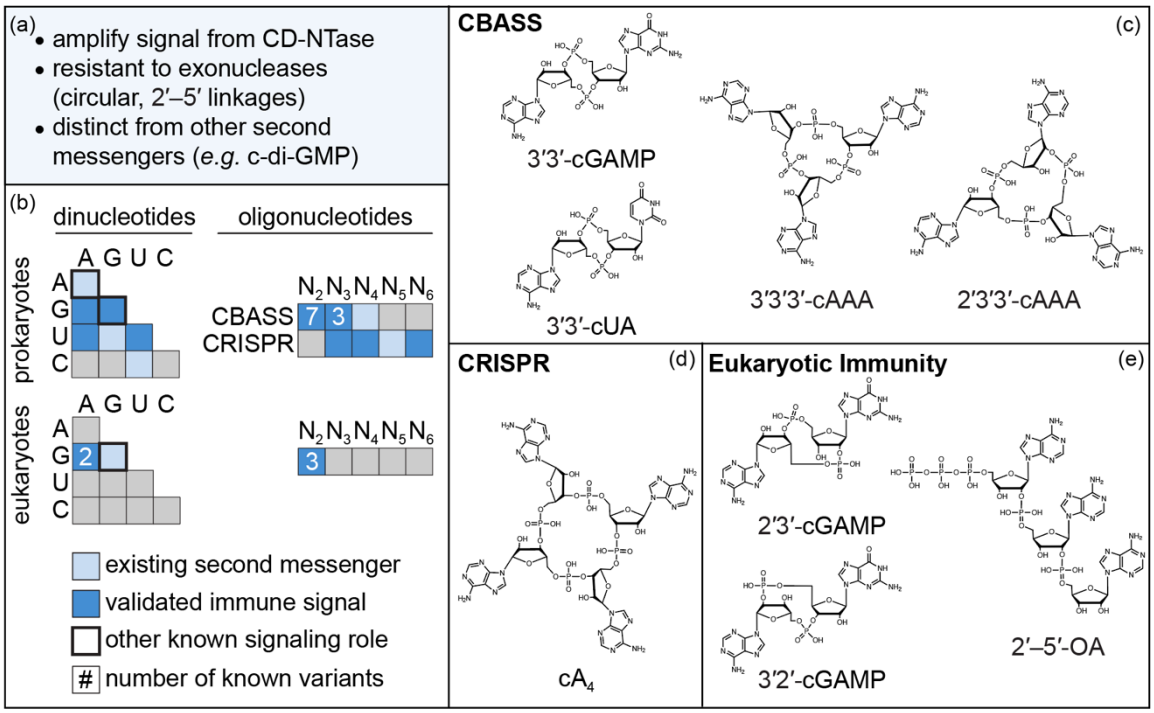


Figure I.3. Nucleotide second messengers
(a) Overview of shared features of nucleotide second messengers in CBASS. (b) Summary of known cyclic oligonucleotide second messengers in prokaryotes (top) and eukaryotes (bottom). (c) Example cyclic oligonucleotides synthesized by bacterial CD-NTases in CBASS operons: 3′3′-cGAMP (*Vc*DncV product), 3′3′-c-UMP–AMP (*Rm*CdnE product), 3′3′3′-c-AMP–AMP–AMP (*Ec*CdnC product), 2′3′3′-c-AMP–AMP–AMP (*Ab*CdnD product). (d) Example cyclic oligonucleotide synthesized by Cas10 in type III CRISPR systems. (e) Nucleotide second messengers known to be involved in metazoan immune signaling: 2′3′-cGAMP (mammalian cGAS and common insect cGLR product), 3′2′-cGAMP (*Drosophila* cGLR1 product), 2′–5′-OA (mammalian OAS1 product).

been verified as functional signals in CBASS immunity, and analysis of CD-NTase specificity suggests up to 180 distinct nucleotide and phosphodiester-linkage combinations may exist in biology (Figure I.3b,c) (Lowey et al., 2020).

In addition to CBASS, antiviral nucleotide second messenger signals also function in type III CRISPR systems and in animal cGLR immune signaling. In type III CRISPR, recognition of phage infection results in activation of the protein Cas10 that synthesizes larger cyclic oligoadenylate rings 3 to 6 nucleotides in size (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Similar to activation of CBASS Cap effectors, type III CRISPR cyclic oligoadenylate signals bind to and activate downstream effector nucleases that inhibit phage replication (Figure I.3b,d) (Kazlauskiene et al., 2017; McMahon et al., 2020; Niewoehner et al., 2017; Rostol et al., 2021).

In animals, the mammalian protein cGAS synthesizes 2′3′-cGAMP to activate the downstream receptor STING and type I interferon responses (Ablasser et al., 2013; Diner et al., 2013; Gao et al., 2013a; Zhang et al., 2013). Recently, 3′2′-cGAMP has been identified as another metazoan immune nucleotide signal synthesized by the dsRNA sensor cGLR1 in *Drosophila* (Figure I.3b,e) (Slavik et al., 2021). Oligoadenylate synthase 1 (OAS1), is an enzyme structurally related to bacterial CD-NTases and animal cGLRs that senses dsRNA and synthesizes a linear nucleotide second messenger 2′–5′-linked oligoadenylate (2′–5′-OA) (Figure I.3b,e) (Kristiansen et al., 2011). However, no linear nucleotide second messengers have been identified in bacterial antiphage defense. While nucleotide second messengers, like 2′3′-cGAMP, seem to be restricted to immune signaling in animals, c-di-GMP is used to regulate cell differentiation in the eukaryote *Dictyostelium* (Chen and Schaap, 2012).

The role of CBASS in controlling cell death and antiphage defense necessitates that CD-NTase nucleotide second messenger products are distinct from signals used during normal bacterial growth. CBASS Cap effectors sensitively respond to 1–10 nM levels of nucleotide second messenger (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020), and cell death responses must therefore be insulated from nucleotide second messenger signals used in normal

housekeeping functions including c-di-GMP and c-di-AMP (Figure I.3a) (Corrigan and Grundling, 2013; Jenal et al., 2017). For example, c-di-GMP is a rare CD-NTase signaling product that has only been observed in bacteria that lack all other c-di-GMP signaling components (Morehouse et al., 2020). It is likely that evolution of atypical nucleotide second messenger signals also facilitates horizontal transfer of CBASS operons between distantly related phyla (Millman et al., 2020). Additionally, rapid divergence of nucleotide signals and evolution of noncanonical 2′–5′ phosphodiester linkages may enable CBASS immune systems to overcome viral nuclease enzymes that degrade second messenger signals (Athukoralage et al., 2020; Eaglesham et al., 2019).

CBASS Cap effectors as nucleotide second messenger receptors

CBASS nucleotide second messengers signal by directly binding to and activating a downstream Cap effector. Each Cap effector has two main functions: (1) specific recognition of the nucleotide second messenger produced by the cognate CD-NTase protein in response to phage infection and (2) induction of cell death to prevent phage propagation (Figure I.4a). Many Cap effectors are two-domain proteins with a discrete “sensor” domain dedicated to nucleotide recognition fused to an enzymatic or toxin-like “effector” domain responsible for executing cell death. The modular architecture of Cap effectors enables diversification of CBASS defense systems including alternative combinations of sensor and effector domain fusions to control cell death (Figure I.4b) (Duncan-Lowey et al., Chapter 2; Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020).

Structures of CBASS Cap effectors explain the mechanism of nucleotide second messenger recognition and reveal homology with components of distantly related bacterial and animal antiviral signaling systems. SAVED domains are a common sensor domain in Cap effectors, representing >1,600 examples and ~30% of all CBASS operons (Figure I.4c) (Burroughs et al., 2015; Lowey et al., 2020). The structure of Cap4 (a restriction endonuclease-

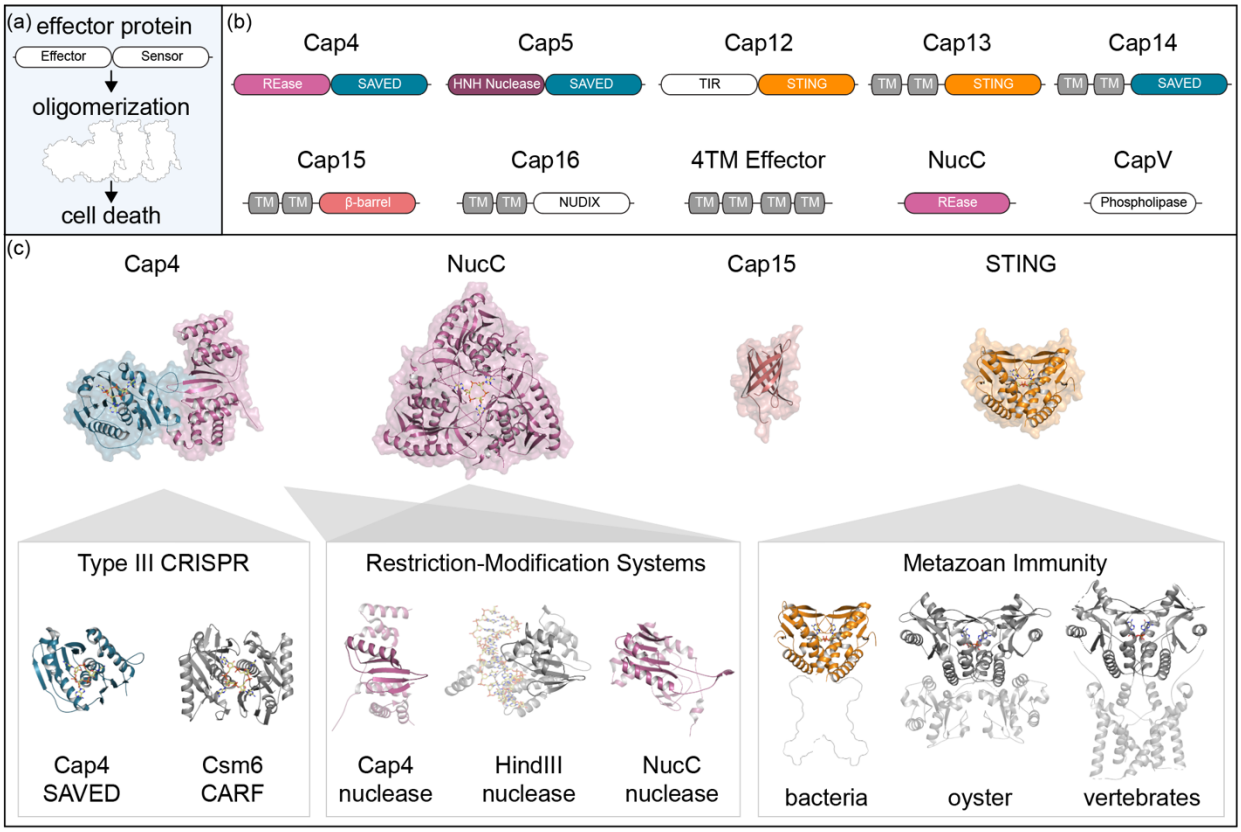


Figure I.4. CBASS Cap effector proteins
(a) Overview of CBASS Cap effector function in inducing cell death. Each effector specifically senses a CD-NTase nucleotide second messenger and is then activated typically through oligomerization. (b) Schematic representation of domain architecture of experimentally characterized CBASS effectors. Common nucleotide sensor domains (*e.g.* SAVED, STING) can be found fused to various effector domains (e.g. nucleases, TIR, TM). (c) Cartoon representation of known CBASS effector structures (top) and comparison to evolutionarily related components in other bacterial defense systems and in metazoan innate immunity (bottom).

SAVED fusion) revealed that SAVED domains form a flat surface with individual pockets to coordinate each base of the nucleotide second messenger signal (Lowey et al., 2020). The Cap4 structure additionally revealed that SAVED is a fusion of two CARF (CRISPR-associated Rossmann Fold) protein subunits which control recognition of cyclic oligonucleotides in type III CRISPR immunity, demonstrating a direct connection between CBASS and CRISPR immunity (Jia et al., 2019; Lowey et al., 2020; Niewoehner and Jinek, 2016). Fusion of CARF proteins into a single-chain SAVED receptor enables recognition of diverse CBASS ligands including cyclic trinucleotides and asymmetric CD-NTase products. CBASS and type III CRISPR immune

systems share additional effector proteins including an REase-type effector named NucC (Lau et al., 2020). Structures of NucC also revealed that some CBASS effectors do not contain a dedicated nucleotide-binding domain and instead recognize the CD-NTase product through loops emerging from the enzymatic or toxin-like domain (Lau et al., 2020). Although rarer in CBASS immunity (~100 operons), Cap12 and Cap13 proteins are notable effectors as they contain a nucleotide sensor domain with direct homology to the animal innate immune protein STING (Morehouse et al., 2020). In both CBASS and animal immunity, STING domains form a homodimeric V-shaped receptor that recognizes cyclic dinucleotide signals in a deep central pocket (Figure I.4c) (Morehouse et al., 2020; Zhang et al., 2013).

Several Cap effectors include additional sensor domains that appear to be unique to CBASS immunity. Recently, structures of Cap15 revealed a minimal eight-stranded beta-barrel domain that is responsible for nucleotide second messenger recognition (Figure I.4c) (Duncan-Lowey et al., Chapter 2). Likewise, bioinformatics analysis has identified an enormous diversity of Cap effectors with many uncharacterized putative nucleotide second messenger sensor domains (Burroughs et al., 2015). Cap effectors like CapV and Cap16 respond to nucleotide second messengers but do not contain any known nucleotide binding domains (Duncan-Lowey et al., Chapter 2; Severin et al., 2018), suggesting that analysis of CBASS components will continue to reveal new modules capable of selective nucleotide second messenger recognition.

CBASS Cap effectors encode a large variety of toxin-like effector domains to execute cell death. The most common strategy to induce cell death in CBASS is host membrane disruption, either through activation of phospholipases that degrade the membrane (Cohen et al., 2019; Severin et al., 2018) or activation of transmembrane-containing effectors that disrupt the bacterial inner membrane (Duncan-Lowey et al., Chapter 2). Many effectors contain a nuclease effector domain that indiscriminately degrades dsDNA upon activation. For example, Cap4 and NucC both contain endonuclease-like domains with homology to bacterial Restriction-Modification system nucleases used to cleave foreign DNA (Figure I.4c) (Watanabe et al., 2009). The CBASS

nucleases Cap4 and NucC, however, lack the structural features that define sequence specificity and instead degrade dsDNA in a sequence-independent manner (Lau et al., 2020; Lowey et al., 2020). Additionally, some Cap effectors contain TIR domains that catalyze rapid degradation of the essential metabolite NAD^+ (Morehouse et al., 2020), and more rare effectors have been implicated in CBASS immunity including those with predicted protease and phosphatase effector domain function (Burroughs et al., 2015; Lowey et al., 2020; Millman et al., 2020).

In most cases, Cap effector activation occurs through sensor domain oligomerization where nucleotide second messenger recognition drives higher-order complex assembly. Electron microscopy analysis demonstrates that Cap4 (REase-SAVED) and Cap12 (TIR-STING) form filaments in the presence of the activating nucleotide second messenger signal (Lowey et al., 2020; Morehouse et al., 2020). Likewise, nucleotide signal recognition induces oligomerization of the CBASS effector NucC and formation of a stable hexameric state capable of promiscuous DNA degradation (Lau et al., 2020). Ligand-induced oligomerization therefore represents a common activation mechanism among structurally diverse CBASS effectors.

Open questions

While recent advances have uncovered many mechanistic details of CBASS immunity, several fundamental questions remain. First, the molecular mechanisms by which phage infection activates CD-NTases and initiates CBASS defense remain unknown. Individual CBASS operons provide defense to phylogenetically diverse phage, indicating that the activating cue sensed during phage replication is broadly conserved and unlikely to be related to a specific viral protein or nucleic acid sequence (Cohen et al., 2019; Lowey et al., 2020; Ye et al., 2020). Similarly, CBASS operons frequently encode accessory *Cap* genes, but no mechanistic role is known for the most common accessory proteins including Cap2 (a predicted E1/E2-like ligase) and Cap3 (a predicted deubiquitinase-family peptidase) (Burroughs et al., 2015; Lowey and Kranzusch, 2020; Millman et al., 2020). Accessory Cap2 and Cap3 proteins are required for efficient defense against some, but not all, phage infections (Cohen et al., 2019), suggesting a role in

modifying the specificity of antiviral defense. Another accessory protein found in some operons, Cap7 (a HORMA-domain containing protein) was demonstrated to directly bind the CBASS operon CD-NTase and contain a potential binding interface for foreign peptide recognition (Ye et al., 2020), suggesting this family of accessory proteins may play a role in phage sensing. Finally, phage encode evasion proteins to circumvent many bacterial defense systems, including anti-CRISPR proteins and inhibitors of Restriction-Modification systems (Jia and Patel, 2021; Samson et al., 2013). Discovery of phage anti-CBASS proteins will provide important tools to explain individual steps of pathway activation and further define the host-pathogen interactions that mediate successful CBASS antiphage defense.

Chapter 1

Bacterial SAVED/CARF-family effectors respond to 3′–5′ and 2′–5′-linked nucleotide signals in CBASS antiviral immunity

Brianna Lowey[1,2], Aaron T. Whiteley[1,2,9], Alexander F.A. Keszei[3], Benjamin R. Morehouse[1,2], Ian Mathews[4,5], Sadie P. Antine[6], Victor J. Cabrera[1], Dmitry Kashin[7], Percy Niemann[7], Mohit Jain[4,5], Frank Schwede[7], John J. Mekalanos[1], Sichen Shao[3], Amy S.Y. Lee[6], Philip J. Kranzusch[1,2,8]

[1] Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

[2] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[3] Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

[4] Department of Pharmacology, University of California, San Diego, La Jolla, CA, USA

[5] Department of Medicine, University of California, San Diego, La Jolla, CA, USA

[6] Department of Biology, Brandeis University, Waltham, MA 02453, USA

[7] Biolog Life Science Institute GmbH & Co. KG, Flughafendamm 9a, 28199 Bremen, Germany

[8] Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA 02115, USA.

[9] Present address: Department of Biochemistry, University of Colorado Boulder, Boulder, CO 80309

The published manuscript may be accessed as follows:

Cell. July 9, 2020. Vol. 182. pp 38–49.

https://doi.org/10.1016/j.cell.2020.05.019

Summary

cGAS/DncV-like Nucleotidyltransferase (CD-NTase) enzymes are immune sensors that synthesize nucleotide second messengers and initiate antiviral responses in both bacterial and animal cells. Here we discover *Enterobacter cloacae* CD-NTase associated protein 4 (Cap4) as a founding member of a diverse family of >2,000 bacterial receptors that respond to CD-NTase signals. Structures of Cap4 reveal a promiscuous DNA endonuclease/DUF4297 domain activated through ligand-induced oligomerization. Oligonucleotide recognition occurs through an appended SAVED domain that is unexpectedly a fusion of two CRISPR-associated Rossman Fold (CARF) subunits co-opted from Type III CRISPR immunity. Like a lock and key, SAVED effectors exquisitely discriminate 2′–5′ and 3′–5′-linked bacterial cyclic oligonucleotide signals and enable specific recognition of at least 180 potential nucleotide second messenger species. Our results reveal SAVED/CARF-family proteins as major nucleotide second messenger receptors in CBASS and CRISPR immune defense and extend the importance of linkage specificity beyond mammalian cGAS-STING signaling.

Introduction

cGAS/DncV-like Nucleotidyltransferases (CD-NTases) are enzymes that synthesize specialized oligonucleotide signals to amplify pathway activation and control downstream effector responses. CD-NTases are conserved in both animal and bacterial signaling systems, and play a key role in innate immunity and phage defense (Ablasser and Chen, 2019; Bernheim and Sorek, 2020; Kranzusch, 2019). In human cells, the CD-NTase cyclic GMP–AMP synthase (cGAS) functions as a sensor for double-stranded DNA aberrantly localized in the cell cytosol during pathogen replication and cancer. Once activated, cGAS produces the nucleotide second messenger 2′–5′, 3′–5′ cyclic GMP–AMP (2′3′-cGAMP) to induce antiviral immunity and interferon signaling (Ablasser et al., 2013; Diner et al., 2013; Gao et al., 2013a; Sun et al., 2013; Zhang et al., 2013). In an analogous system, the *Vibrio cholerae* CD-NTase Dinucleotide cyclase in Vibrio (DncV) synthesizes the nucleotide second messenger 3′–5′, 3′–5′ cGAMP (3′3′-cGAMP) in response to an unknown stimulus during phage infection (Cohen et al., 2019; Davies et al., 2012; Kranzusch et al., 2014). Bacterial 3′3′-cGAMP activates a downstream effector response that results in cell death, limitation of phage replication through abortive infection, and protection of the remaining bacterial population (Cohen et al., 2019; Severin et al., 2018).

Bacterial CD-NTases include >5,600 unique enzymes (Whiteley et al., 2019) that control a diverse array of antiviral immune systems collectively named cyclic-oligonucleotide-based antiphage signaling system (CBASS) immunity (Cohen et al., 2019; Ye et al., 2020). Synthesizing cyclic dinucleotide and cyclic trinucleotide products, bacterial CD-NTases are capable of using all four ribonucleotides as building blocks to dictate signal specificity and enable an increased diversity of CBASS antiviral immune responses (Cohen et al., 2019; Whiteley et al., 2019). For example, in addition to the *V. cholerae* DncV product 3′3′-cGAMP, *Escherichia coli* CdnE synthesizes 3′3′ cyclic UMP–AMP (3′3′-cUA) and *Enterobacter cloacae* CdnD (CD-NTase in clade D) synthesizes the cyclic trinucleotide second messenger 3′3′3′ cyclic AMP–AMP–GMP (3′3′3′-cAAG) (Whiteley et al., 2019). In human cells, the non-canonical 2′–5′ linkage in the cGAS product

2′3′-cGAMP is critical for immune specificity and potent activation of the downstream receptor Stimulator of Interferon Genes (STING) (Ablasser et al., 2013; Diner et al., 2013; Gao et al., 2013b; Zhang et al., 2013), but it is unknown if phosphodiester linkage specificity is an additional determinant of receptor activation in bacterial CBASS antiviral immunity.

The best characterized family of receptors that respond to bacterial CD-NTase nucleotide second messengers are patatin-like phospholipases which upon nucleotide signal binding are activated to degrade membrane phospholipids. The first known example is the *Vibrio cholerae* protein cGAMP-activated phospholipase in *Vibrio* (CapV) that responds to 3′3′-cGAMP and causes membrane rupture and bacterial cell death (Severin et al., 2018). Similarly, the *E. coli* CD-NTase CdnE signals through CapE, a CapV homolog that specifically recognizes 3′3′-cUA, indicating CBASS operons function with high-specificity for a single nucleotide second messenger (Whiteley et al., 2019). However, a majority of bacterial CD-NTase enzymes are encoded in CBASS operons that do not contain proteins with homology to CapV-like receptors (Burroughs et al., 2015; Cohen et al., 2019; Whiteley et al., 2019), suggesting that yet uncharacterized proteins must be responsible for downstream effector functions.

Here we reconstitute *Enterobacter cloacae* CdnD signaling *in vitro* and *in vivo* to discover CD-NTase-associated protein 4 (Cap4) as the founding member of a major family of downstream receptors that specifically respond to nucleotide second messenger signals in CBASS immunity. High-resolution crystal structures of Cap4 in apo and cyclic trinucleotide-bound states combined with negative stain electron microscopy (EM) analysis of a Cap4–3′3′3′-cAAG complex reveal a DNA endonuclease effector domain activated through nucleotide second messenger-induced oligomerization. We demonstrate that a previously uncharacterized protein domain in Cap4 named “SAVED” is responsible for nucleotide second messenger recognition. Remarkably, SAVED is a fusion of two CRISPR-associated Rossman Fold (CARF) domains derived from Type III CRISPR immunity revealing a common ancestry between these two nucleotide second messenger-centric antiviral systems. SAVED domains are widespread in CBASS and Type III

CRISPR immune systems and divergence in the SAVED nucleotide binding pocket enables recognition of an expanded range of CD-NTase products including bacterial second messengers with alternative ring size, nucleobase, and 3′–5′ or 2′–5′ phosphodiester linkages. SAVED domains function as discrete modules fused to a variety of alternative effector proteins and we further show that these receptor–effector fusions are essential for CBASS-mediated protection of bacteria from phage infection. Our results uncover a major family of nucleotide second messenger receptors and explain a mechanism of effector function controlling bacterial viral defense through abortive infection and killing of virus-infected cells.

Results

Cap4 proteins are DNA endonucleases that respond to CD-NTase nucleotide second messengers

The *E. cloacae* CD-NTase CdnD (*Ec*CdnD) is constitutively active *in vitro* and synthesizes the nucleotide second messenger 3′3′3′-cAAG (Whiteley et al., 2019). *Ec*CdnD is encoded in an operon containing three additional genes of unknown function designated here as *CD-NTase associated protein 2*, *3* and *4* (*Cap2*, *Cap3*, *Cap4*) (Figure 1.1A). To define how bacterial CD-NTase enzymes control downstream signaling we purified each protein from the *Enterobacter cloacae* CdnD02 operon and used a biochemical approach for receptor identification. Incubation of radiolabeled 3′3′3′-cAAG with purified *Ec*Cap4 (Genbank WP_032676399.1) in an electrophoretic mobility shift assay resulted in complete shift and formation of a stable Cap4–3′3′3′-cAAG complex (Figure 1.1B and 1.2A). No interaction was observed between *Ec*Cap4 and the *V. cholerae* DncV product 3′3′-cGAMP, demonstrating that *Ec*Cap4 is a downstream receptor that specifically recognizes the cognate *Ec*CdnD nucleotide second messenger.

To understand the function of Cap4 we screened homologs for suitability in structural analysis and determined a 2.4 Å crystal structure of Cap4 from the bacterium *Moraxella osloensis* (*Mo*Cap4) (Genbank WP_060996052.1) and a 2.6 Å crystal structure of Cap4 from *Acinetobacter baumannii* (AbCap4) (Genbank WP_008942236.1) (Figures 1.1C and S1B, Table 1.1). The structure of Cap4 reveals a two-domain architecture with an N-terminal domain containing a mixed beta-sheet braced on either side with alpha-helical bundles, and a globular C-terminal domain that contains internal two-fold pseudosymmetry (Figures 1.1C and 1.2B). Sequence alignment based on the *Mo*Cap4 structure demonstrates that all Cap4 homologs contain the same domain architecture with *Mo*Cap4 vs. *Ab*Cap4 sharing ~80% identity and *Mo*Cap4 vs. *Ec*Cap4 sharing ~20% identity at the amino-acid level (Figure 1.2B). The Cap4 N-terminal domain is a member of an uncharacterized protein domain classification “domain of unknown function 4297”

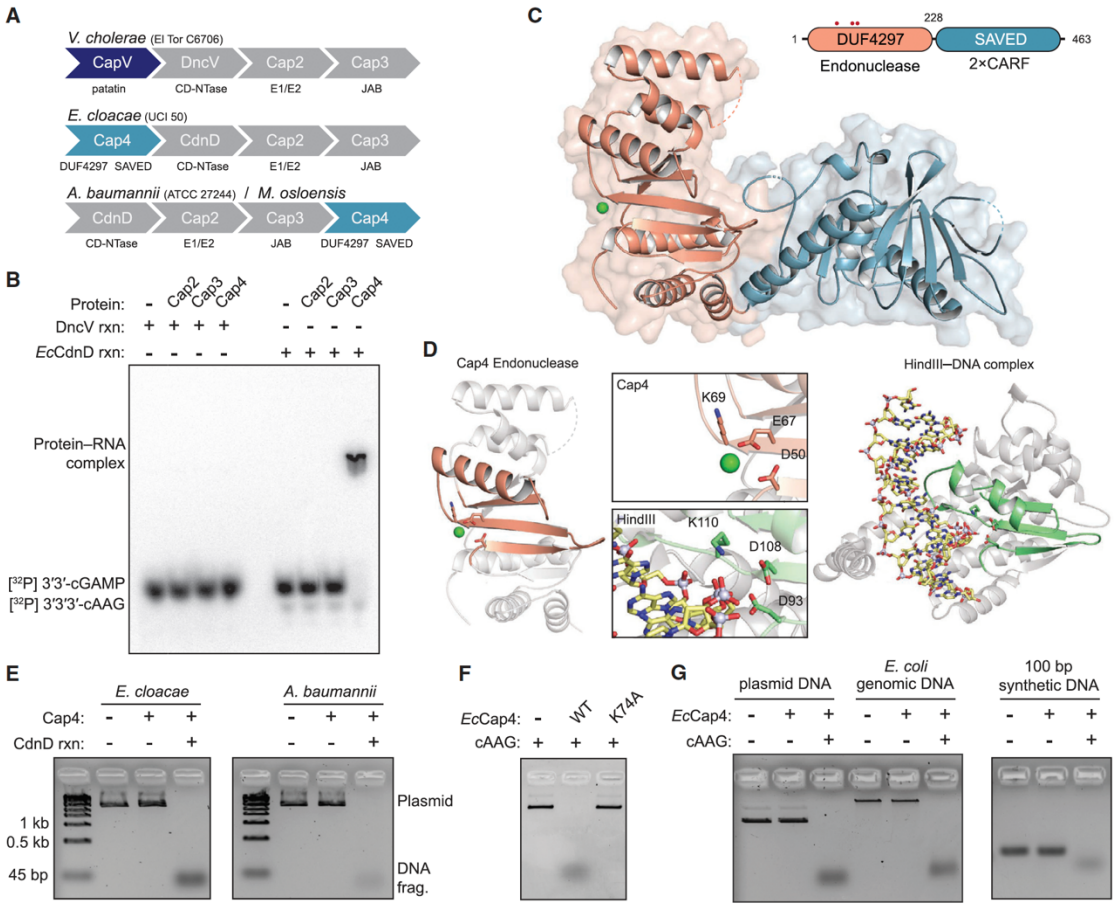


Figure 1.1. Cap4 proteins are endonucleases activated by CD-NTase nucleotide second messengers

(A) Architectures of CBASS operons used for biochemical analysis. Enterobacter cloacae, Acinetobacter baumannii, and Moraxella osloensis systems lack a CapV phospholipase effector homolog and instead encode the protein CD-NTase-associated protein 4 (Cap4). (B) Electrophoretic mobility shift assay measurement of Cap4-3′3′3′-cAAG complex formation. E. cloacae Cap2, Cap3, and Cap4 were incubated with 32P-labeled 3030-cGAMP or 3′3′3′-cAAG, and bound complexes were resolved by nondenaturing polyacrylamide gel electrophoresis. (C) Cartoon schematic and crystal structure of apo MoCap4. MoCap4 contains an N-terminal DUF4297 domain (orange) with homology to restriction endonucleases and a C-terminal SAVED domain (blue) with homology to CARF domains found in type III CRISPR immunity. Red circles indicate locations of active-site residues. (D) Structural comparison of the Cap4 DUF4297 domain and the restriction endonuclease HindIII, showing structural homology and shared catalytic residues. The core endonuclease domains in Cap4 and HindIII are highlighted in orange and green, respectively. (E and F) Agarose gel analysis of plasmid DNA degradation by Cap4. (E) Cap4 proteins degrade target DNA only in the presence of activating nucleotide second messenger synthesized by the neighboring CD-NTase CdnD within the CBASS operon. (F) Activity is dependent on the conserved Cap4 catalytic active site, with no cleavage observed with the EcCap4 mutant K74A. (G) Agarose gel analysis of Cap4 DNA cleavage promiscuity. In the presence of activating 3′3′3′-cAAG, EcCap4 is capable of degrading all sources of dsDNA, including plasmid DNA, E. coli genomic DNA, and synthetic 100-bp DNA. Biochemical data are representative of at least 3 independent experiments.

Table 1.1. Crystallographic statistics

	***Ab*Cap4 Apo**	***Ab*Cap4–2′3′3′-cAAA**	**AbCap4–3′3′3′-cAAA**	***Ab*Cap4–2′3′3′-cAAA** (SeMet)
Data Collection				
Resolution (Å)[a]	39.60–2.60 (2.64–2.60)	49.21–2.10 (2.14–2.10)	38.80–2.40 (2.44–2.40)	49.46–2.45 (2.49–2.45)
Wavelength (Å)	0.97918	0.97918	0.97918	0.97918
Space group	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1
Unit cell: a, b, c (Å)	99.72 111.38 173.30	106.58, 111.46, 164.83	106.50 111.18 163.75	106.37, 112.23, 163.35
Unit cell: α, β, γ (°)	90.00 103.06 90.00	90.00, 100.21, 90.00	90.00 100.36 90.00	90.00, 100.03, 90.00
Molecules per ASU	6	6	6	6
Total reflections	335131	1567459	519258	3953693
Unique reflections	100775	217277	145531	138853
Completeness (%)[a]	88.7 (48.8)	98.3 (95.5)	99.0 (86.2)	99.9 (99.9)
Multiplicity[a]	3.3 (2.2)	7.2 (7.1)	3.6 (3.1)	28.5 (27.8)
$I/\sigma I$[a]	12.9 (2.3)	12.4 (1.6)	8.4 (1.0)	14.1 (1.3)
CC(1/2)[b] (%)[a]	99.7 (80.8)	99.8 (55.0)	99.6 (42.8)	99.9 (66.4)
Rpim[c] (%)[a]	3.8 (29.9)	4.2 (81.1)	5.8 (88.1)	3.4 (63.3)
Sites				36
Refinement				
Resolution (Å)	39.60–2.60	49.21–2.10	38.62–2.40	
Free reflections	1993	2000	2000	
R-factor / R-free	18.7/23.3	18.7 / 21.8	23.5 / 20.5	
Bond distance (RMS Å)	0.007	0.002	0.005	
Bond angles (RMS °)	0.903	0.561	1.03	
Structure/Stereochemistry				
No. atoms: protein	20800	21078	20894	
No. atoms: ligand	30 (SO_4)	370 (2′3′3′ cAAA)	426 (3′3′3′ cAAA, SO_4)	
No. atoms: solvent	358	1515	459	
Average B-factor: protein	54.70	54.78	66.61	
Average B-factor: ligand	62.19	80.84	68.89	
Average B-factor: water	42.73	50.25	53.03	
Ramachandran plot: favored	96.13%	97.84%	96.61%	
Ramachandran plot: allowed	3.76%	2.16%	3.31%	
Ramachandran plot: outliers	0.12%	0%	0.08%	
Rotamer outliers	0.0%	0.77%	0.0%	
MolProbity[d] score	1.58	1.16	1.40	
Protein Data Bank ID	6WAM	6VM6	6WAN	

[a] Highest resolution shell values in parenthesis
[b] (Karplus and Diederichs, 2012)
[c] (Weiss, 2001)
[d] (Chen et al., 2010)

Table 1.1. Crystallographic statistics, continued

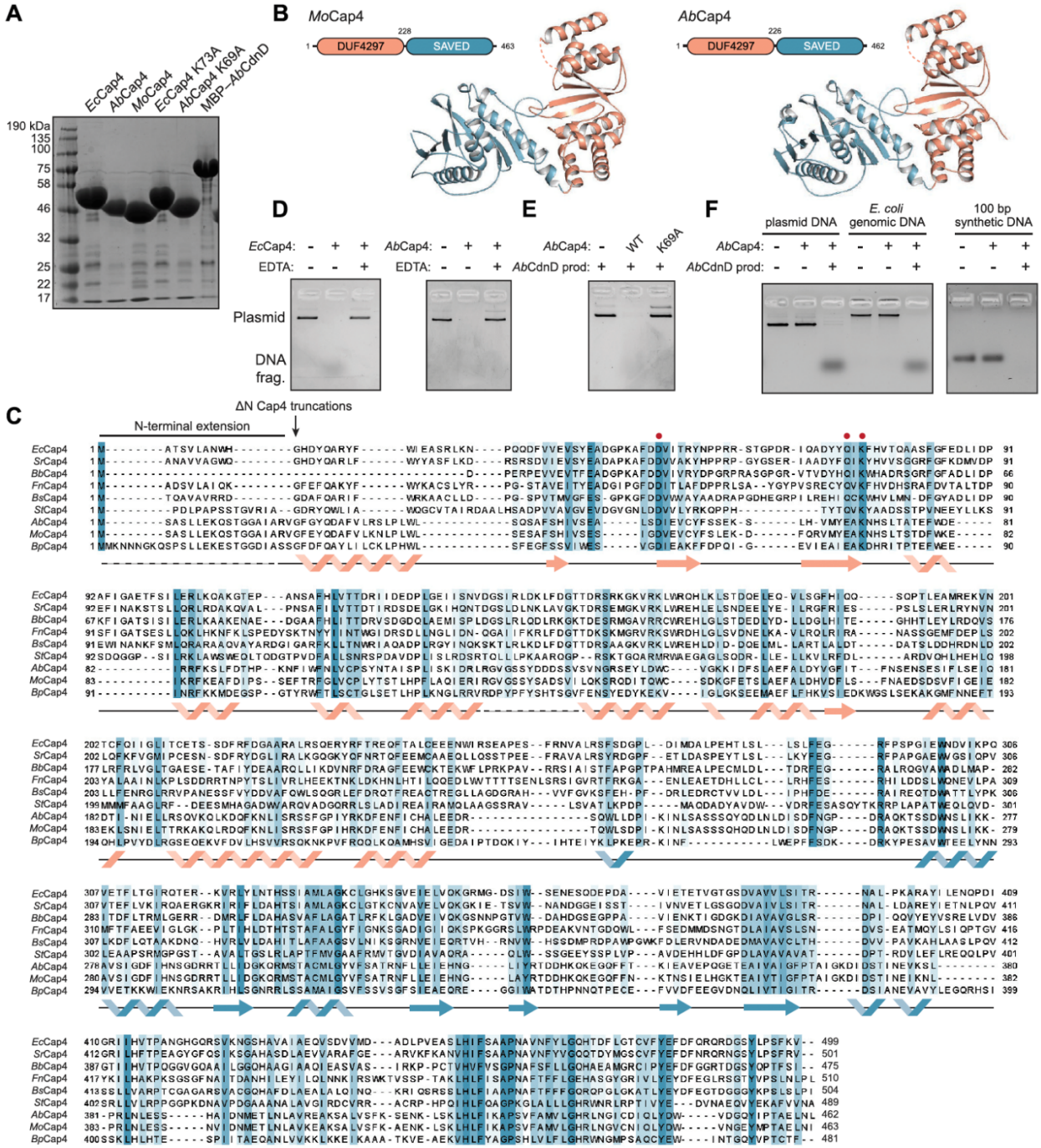
	***Mo*Cap4 Apo**	***Mo*Cap4** (SeMet)
Data Collection		
Resolution (Å)[a]	49.36–2.34 (2.43–2.34)	49.30–2.80 (2.95–2.80)
Wavelength (Å)	0.99998	0.97949
Space group	P 3_2 2 1	P 3_2 2 1
Unit cell: a, b, c (Å)	157.59, 157.59, 63.32	157.82, 157.82, 63.16
Unit cell: α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Molecules per ASU	1	1
Total reflections	423237	2521929
Unique reflections	37860	22581
Completeness (%)[a]	99.2 (92.4)	99.8 (98.8)
Multiplicity[a]	11.2 (10.6)	111.7 (108.1)
$I/\sigma I$[a]	15.3 (0.9)	19.8 (1.7)
CC(1/2)[b] (%)[a]	99.9 (41.3)	100 (69.5)
Rpim[c] (%)[a]	3.2 (89.5)	3.2 (60.8)
Sites		5
Refinement		
Resolution (Å)	49.36–2.35	
Free reflections	2004	
R-factor / R-free	20.7 / 23.4	
Bond distance (RMS Å)	0.006	
Bond angles (RMS °)	0.710	
Structure/Stereochemistry		
No. atoms: protein	3385	
No. atoms: ligand	1 (Mg)	
No. atoms: solvent	67	
Average B-factor: protein	83.51	
Average B-factor: ligand	99.78	
Average B-factor: water	65.29	
Ramachandran plot: favored	97.07%	
Ramachandran plot: allowed	2.93%	
Ramachandran plot: outliers	0.0%	
Rotamer outliers	0.0%	
MolProbity[d] score	1.46	
Protein Data Bank ID	6VM5	

[a] Highest resolution shell values in parenthesis
[b] (Karplus and Diederichs, 2012)
[c] (Weiss, 2001)
[d] (Chen et al., 2010)

(DUF4297) that is widespread in bacteria. Comparative analysis using the *Mo*Cap4 structure reveals that the DUF4297 domain shares structural homology with Type II restriction endonucleases including the enzymes AgeI and HindIII (DALI Z-scores 3.7 and 2.8, respectively) (Figure 1.1D). Sequence alignment of Cap4 and Type II restriction enzymes confirms conservation of all putative active-site residues required for metal coordination and suggests that Cap4 proteins are functional nuclease enzymes (Figure 1.2B) (Burroughs et al., 2015; Watanabe et al., 2009). We therefore tested the *Ec*Cap4 and *Ab*Cap4 proteins for the ability to cleave DNA substrates *in vitro*. In the presence of activating nucleotide second messenger, Cap4 proteins catalyze complete degradation of plasmid DNA to small <45 bp DNA fragments (Figure 1.1E). Cap4 alone exhibits no nuclease activity, indicating recognition of a nucleotide second messenger is a strict requirement of nuclease activation. Cap4 activity is divalent cation metal-dependent (Figure 1.2D) and mutation of a conserved lysine residue in the predicted active site ablates nuclease activity (Figures 1.1F and 1.2E), further confirming that Cap4 proteins use a Type II restriction enzyme-like reaction mechanism to degrade DNA substrates.

Type II restriction endonucleases contain loops or helices that extend from the nuclease domain to form DNA major and minor groove contacts and control target-sequence cleavage specificity (Tamulaitiene et al., 2017; Watanabe et al., 2009). The Cap4 N-terminal endonuclease domain is minimized with nearly all extensions absent suggesting relaxation of DNA targeting specificity (Figure 1.1D). Consistent with these structural observations, *Ec*Cap4 degrades plasmid DNA, *E. coli* genomic DNA, and synthetic dsDNA *in vitro* with no apparent target sequence or DNA modification specificity (Figures 1.1G and 1.2F). In each case, Cap4 endonuclease activity remains strictly dependent on the presence of activating nucleotide second messenger. Together, these results demonstrate Cap4 proteins are sequence non-specific dsDNA endonucleases that are controlled through nucleotide second messenger-gated enzyme activation.

Figure 1.2. Biochemical reconstitution of Cap4 endonuclease activity
(A) SDS-PAGE and Coomassie stain analysis of purified recombinant Cap4 and CdnD proteins. (B) Cartoon schematics and crystal structures of apo MoCap4 and AbCap4 enzymes. (C) Alignment of divergent Cap4 homologs. Endonuclease active site residues are indicated with a red circle, and the known secondary structure of the Cap4 endonuclease domain (orange) and SAVED sensor domain (blue) determined from the AbCap4 crystal structure is displayed below. (D) EcCap4 and AbCap4 endonuclease activity is inhibited by the divalent metal chelator EDTA. A requirement of divalent cation metals for catalysis agrees with structural analysis of the Cap4 active site and homology with Type II restriction endonucleases. (E) Like EcCap4, AbCap4 mutation of the putative catalytic residue K69A ablates endonuclease activity. (F) Agarose gel analysis of AbCap4 endonuclease cleavage activity. Like EcCap4, AbCap4 is a promiscuous DNA endonuclease that degrades plasmid DNA, E. coli genomic DNA, and synthetic DNA but requires recognition of the activating nucleotide second messenger synthesized by the cognate CdnD CD-NTase. Biochemical data are representative of at least 3 independent experiments.


A
EcCap4
AbCap4
MoCap4
EcCap4 K73A
AbCap4 K69A
MBP–AbCdnD
190 kDa
135
100
75
58
46
32
25
22
17
B
MoCap4
1
DUF4297
228
SAVED
463
AbCap4
1
DUF4297
226
SAVED
462
D
EcCap4: - + +
EDTA: - - +
Plasmid
DNA frag.
E
AbCap4: - + +
EDTA: - - +
AbCap4: - WT K69A
AbCdnD prod: + + +
F
plasmid DNA
E. coli genomic DNA
100 bp synthetic DNA
AbCap4: - + + - + + - + +
AbCdnD prod: - - + - - + - - +
C
ΔN Cap4 truncations
N-terminal extension
EcCap4
SrCap4
BbCap4
FnCap4
BsCap4
StCap4
AbCap4
MoCap4
BpCap4

Cap4 proteins respond to specific 3′–5′ and 2′–5′-linked nucleotide second messengers

To determine the specificity of CD-NTase-Cap4 signaling we next compared cross-activation between the *E. cloacae* and *A. baumannii* CdnD operons. Although each CdnD activates the robust endonuclease activity of its cognate Cap4 protein *in vitro*, heterologous reactions with one CdnD or Cap4 component derived from each operon fail to reconstitute signaling and do not result in DNA degradation (Figures 1.3A and 1.4A). Selective Cap4 activation reveals that *Ab*CdnD must therefore synthesize a nucleotide second messenger distinct from the *Ec*CdnD product 3′3′3′-cAAG. We tested nucleotide combinations and observed that ATP is necessary and alone sufficient to allow *Ab*CdnD to synthesize the activating nucleotide second messenger signal (Figure 1.4B). However, all known canonically-linked cyclic di-, tri-, and tetra-AMP RNA products failed to reconstitute activation of *Ab*Cap4 DNA degradation activity (Figure 1.3B). We therefore hypothesized that like human cGAS-STING, *Ab*CdnD-Cap4 signaling may require a noncanonical 2′–5′-linked second messenger. To assess whether the *Ab*CdnD nucleotide product contains a non-canonical linkage, we analyzed sensitivity of eukaryotic and prokaryotic CD-NTase reaction products to enzymatic digestion. The enzyme nuclease P1 specifically hydrolyzes 3′–5′-linked phosphodiester bonds and is unable to cleave non-canonical 2′–5′ phosphodiester bonds (Ablasser et al., 2013; Diner et al., 2013; Whiteley et al., 2019). Unlike the *V. cholerae* DncV and *Ec*CdnD products 3′3′-cGAMP and 3′3′3′-cAAG, both the mammalian cGAS 2′3′-cGAMP and the *Ab*CdnD reaction products contain phosphates resistant to nuclease P1 digestion confirming the presence of a 2′–5′-linked bond (Figure 1.3C).

We next used ion-exchange and C18 column purification to isolate the *Ab*CdnD nucleotide second messenger directly from enzymatic reactions for further characterization. *Ab*CdnD synthesizes two closely related products that co-elute with nearly all purification steps (Figure 1.3D) but can be separated on a C18 reverse-phase column (Figure 1.3E). Biochemical analysis demonstrates the major (~65%) and minor (~35%) *Ab*CdnD products are each cyclic oligoadenylate species containing P1-resistant 2′–5′ phosphodiester linkages (Figure 1.4C).

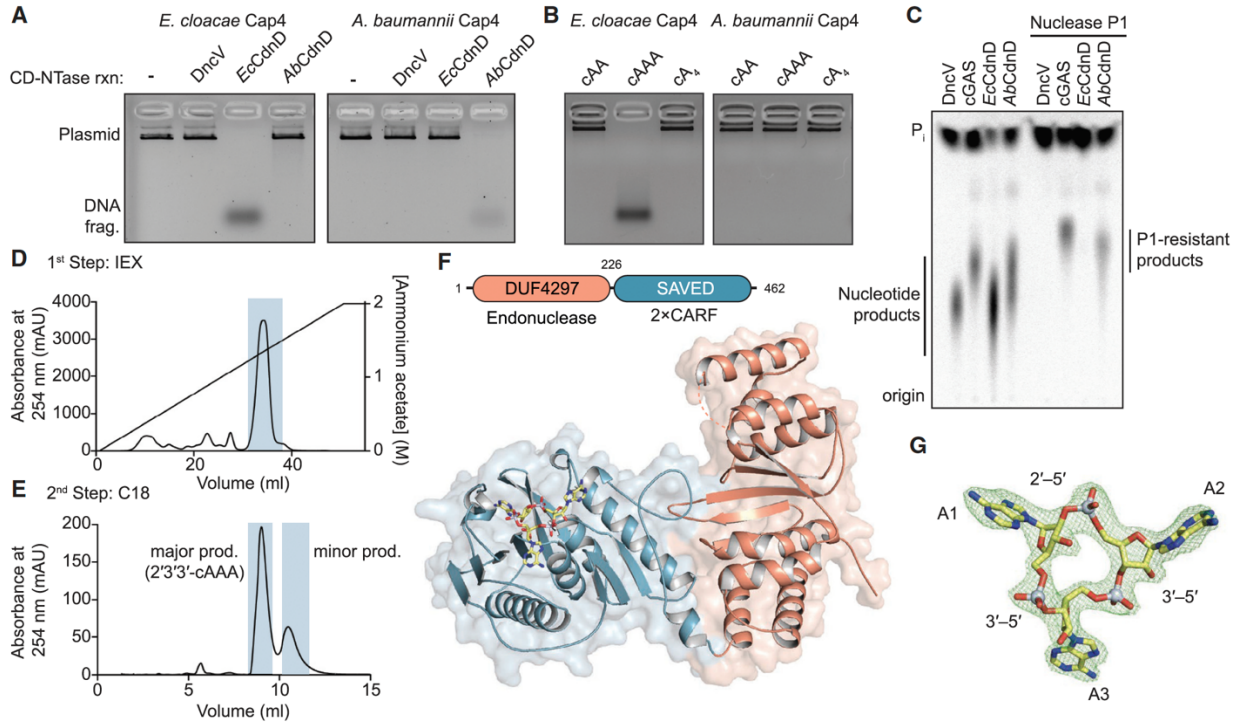


Figure 1.3. Cap4 proteins respond to specific 3′–5′- and 2′–5′-linked nucleotide second messengers

(A) Analysis of Cap4 activation specificity using nucleotide second messengers synthesized with CD-NTase enzymes (DncV reaction, 3′3′-cGAMP; EcCap4 reaction, 3′3′3′-cAAG). EcCap4 and AbCap4 activation requires the specific nucleotide second messenger synthesized by the cognate CD-NTase, with no crossactivation occurring between distantly related operons. (B) Activation of EcCap4 and AbCap4 DNA degradation activity with synthetic nucleotide second messengers. Synthetic 3′3′3′-cAAA is able to activate the EcCap4 enzyme normally responsive to 3′3′3′-cAAG, but no tested canonically 3′–5′-linked cyclic oligonucleotide species is able to activate AbCap4. (C) P1 nuclease digestion and thin-layer chromatography analysis of CD-NTase products. Nuclease P1 specifically degrades 3′–5′ bonds, leaving non-canonical 2′–5′ bonds intact and resistant to phosphatase treatment. The human enzyme cGAS product 2′3′-cGAMP and the bacterial AbCdnD product contain nuclease P1-resistant 2′–5′ linkages. (D and E) Large-scale synthesis and purification of the AbCdnD nucleotide second messenger products. AbCdnD synthesizes two closely related cyclic oligoadenylate products that co-migrate as (D) a single peak on anion-exchange (IEX) chromatography but can be separated into (E) major and minor product species with C18 chromatography. (F) Cartoon schematic and crystal structure of AbCap4 bound to the AbCdnD major product 2′3′3′-cAAA. The cyclic trinucleotide 2′3′3′-cAAA binds within the Cap4 SAVED domain (blue), indicating that this domain is responsible for CD-NTase signal recognition. (G) Simulated-annealing Fo Fc omit map (contoured at 3 s) of the ligand density, demonstrating unambiguous assignment of the major AbCdnD nucleotide second messenger as 2′3′3′-cAAA. Biochemical data are representative of at least 3 independent experiments.

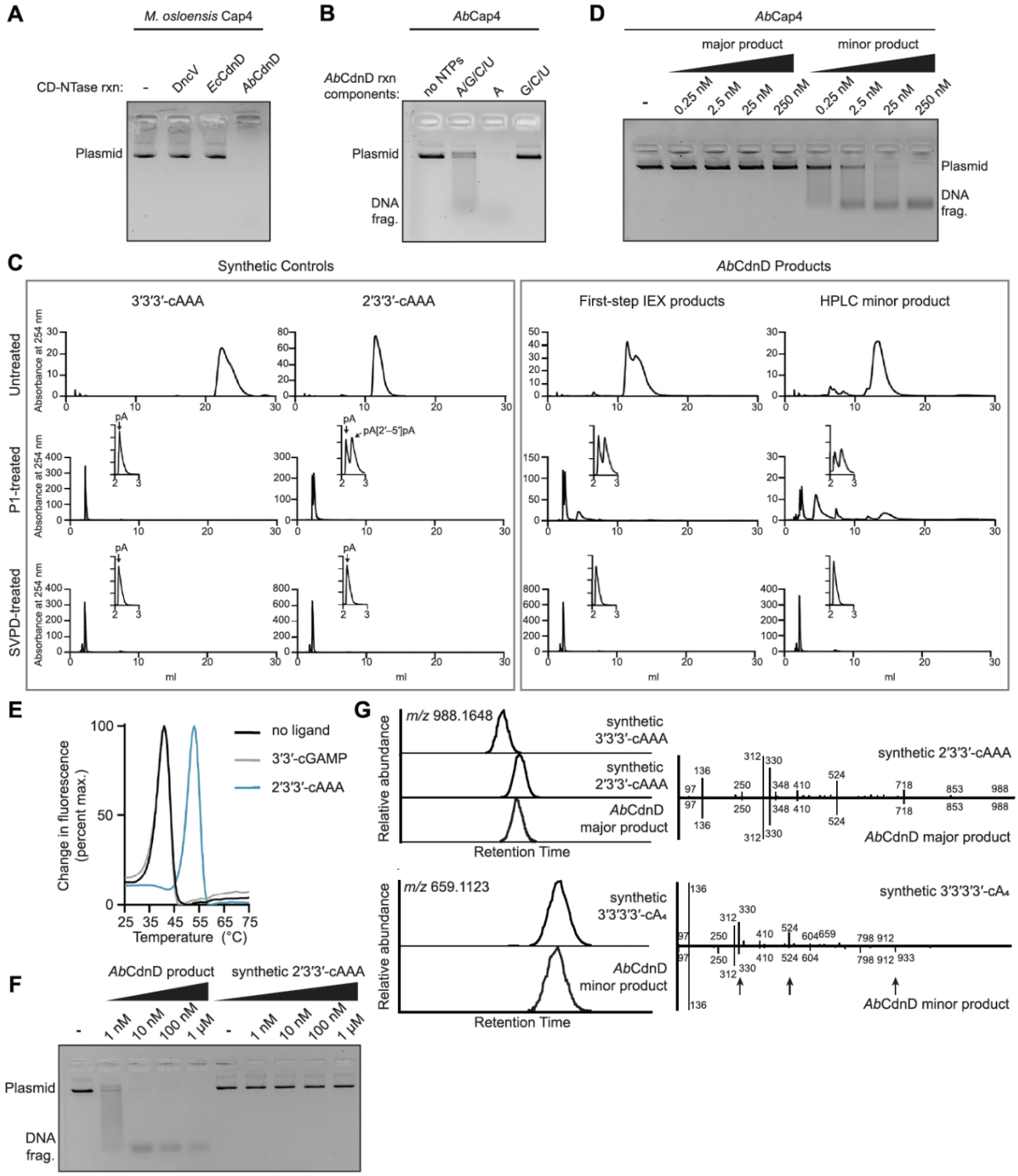
Mass spectrometry analysis demonstrates that the major product is a cyclic trinucleotide and the minor product peak is a mixture of cAAA and cA_4 species (Figure 1.4G). Interestingly, the minor product most potently activates *Ab*Cap4, with full enzyme activation and DNA degradation requiring low nanomolar concentrations of nucleotide second messenger (Figure 1.4D). Although the major product does not activate *Ab*Cap4 DNA degradation activity, this *Ab*CdnD product is still capable of binding and stabilizing the *Ab*Cap4 enzyme and we were able to determine a 2.2 Å co-crystal structure of the *Ab*Cap4–nucleotide second messenger complex (Figures 1.3F and 1.4E). Clear ligand density was observed for a cyclic trinucleotide bound within the *Ab*Cap4 globular C-terminal domain allowing unambiguous assignment of this *Ab*CdnD product as the cyclic trinucleotide 2′–5′, 3′–5′, 3′–5′ c-AMP–AMP–AMP (2′3′3′-cAAA) and direct structural confirmation of the ability of a bacterial CD-NTase to synthesize noncanonical 2′–5′ linked RNAs (Figures 1.3G and 1.4F). To further confirm these findings, we verified using mass spectrometry fragmentation analysis that the *Ab*CdnD major product matches a chemically synthesized 2′3′3′-cAAA standard (Figure 1.4G). 2′–5′ phosphodiester linkages are rare in biology and their role in nucleotide signaling was previously suggested to be a unique adaptation evolved within eukaryotic innate immune signaling (Danilchanka and Mekalanos, 2013; Kranzusch et al., 2015; Margolis et al., 2017). These results demonstrate that 2′–5′-linked products are also involved in bacterial antiviral signaling and that Cap4 nucleases function as selective sensors that can use linkage-specificity to adapt to distinct CD-NTase nucleotide second messenger signals.

Cap4 contains a C-terminal CARF-family domain that controls ligand specificity

The structure of the *Ab*Cap4–2′3′3′-cAAA complex reveals that the Cap4 C-terminal domain is responsible for nucleotide second messenger recognition. Previously, the Cap4 C-terminal SAVED domain had been identified bioinformatically as enriched in CD-NTase-containing operons (Burroughs et al., 2015; Cohen et al., 2019). Surprisingly, analysis of the *Ab*Cap4 SAVED

Figure 1.4. Purification and characterization of the *Ab*CdnD nucleotide second messenger product

(A) DNA cleavage analysis of *Mo*Cap4 endonuclease activity. The *Ab*CdnD nucleotide second messenger product is able to activate both *Ab*Cap4 and the more closely related *Mo*Cap4 (~80% identity) but not the distantly related *Ec*Cap4 enzyme (~20% identity). (B) Analysis of *Ab*CdnD nucleotide specificity. *Ab*Cap4 endonuclease activity was used to detect the presence of activating nucleotide second messenger signal and determine which nucleotides are required for *Ab*CdnD product formation. ATP is alone necessary and sufficient for AbCdnD-dependent activation demonstrating that the nucleotide second messenger is an oligoadenylate species. (C) Enzymatic degradation and HPLC analysis of synthetic control cyclic oligoadenylate species and *Ab*CdnD products. Synthetic and *Ab*CdnD product nucleotide second messengers were analyzed following cleavage with Nuclease P1 (P1, cleaves 3′–5′ phosphodiester bonds) and snake venom phosphodiesterase (SVPD, cleaves both 3′–5′ and 2′–5′ phosphodiester bonds) by HPLC and C18 chromatography. Degradation of the AbCdnD major product matches that of synthetic 2′3′3′-cAAA further confirming the structural assignment of this nucleotide second messenger species. The *Ab*CdnD minor product that potently activates AbCap4 activity is partially P1-resistant confirming the presence of a 2′–5′ phosphodiester bond in this closely related species. (D) Assessment of DNA degradation after activation of AbCap4 with either the major or minor *Ab*CdnD products from Figure 1.3E. (E) Thermal denaturation assay to measure nucleotide second messenger-dependent stabilization of *Ab*Cap4. *Ab*Cap4 is stabilized in the presence of the *Ab*CdnD major product 2′3′3′-cAAA but not in presence of the unrelated DncV product 3′3′-cGAMP demonstrating specific recognition of 2′3′3′-cAAA. (F) *Ab*Cap4 DNA degradation assay comparing candidate activating ligands. The AbCdnD product mixture, but not synthetic 2′3′3′-cAAA, is capable of activating AbCap4 further confirming that the minor 2′–5′-linked product is primarily responsible for AbCap4 activation. (G) LC-MS/MS analysis of *Ab*CdnD major (top) and minor (bottom) enzymatic products. Parent mass extracted ion trace (left) and tandem mass spectra comparison (right) of the major *Ab*CdnD product confirms 2′3′3′-cAAA. Distinct patterning in tandem mass spectra (black arrows) in the minor product relative to 3′3′3′3′-cA_4 indicates a cA_4 molecule of distinct phosphodiester linkages. Ions detected: m/z 988.1648 [M+H] + (cAAA); m/z 659.1123 [M+2H] 2+ (cA_4). Biochemical data are representative of at least 3 independent experiments.


A
M. osloensis Cap4
CD-NTase rxn:
-
DncV
EcCdnD
AbCdnD
Plasmid
B
AbCap4
AbCdnD rxn components:
no NTPs
A/G/C/U
A
G/C/U
Plasmid
DNA frag.
D
AbCap4
major product
minor product
-
0.25 nM
2.5 nM
25 nM
250 nM
0.25 nM
2.5 nM
25 nM
250 nM
Plasmid
DNA frag.
C
Synthetic Controls
AbCdnD Products
3′3′3′-cAAA
2′3′3′-cAAA
First-step IEX products
HPLC minor product
Untreated
P1-treated
SVPD-treated
Absorbance at 254 nm
pA
pA[2′–5′]pA
ml
E
Change in fluorescence (percent max.)
Temperature (°C)
no ligand
3′3′-cGAMP
2′3′3′-cAAA
F
AbCdnD product
synthetic 2′3′3′-cAAA
-
1 nM
10 nM
100 nM
1 μM
-
1 nM
10 nM
100 nM
1 μM
Plasmid
DNA frag.
G
m/z 988.1648
synthetic 3′3′3′-cAAA
synthetic 2′3′3′-cAAA
AbCdnD major product
Relative abundance
Retention Time
synthetic 2′3′3′-cAAA
AbCdnD major product
m/z 659.1123
synthetic 3′3′3′3′-cA4
AbCdnD minor product
synthetic 3′3′3′3′-cA4
AbCdnD minor product

domain structure reveals clear structural homology with CRISPR-associated Rossman fold (CARF) proteins including Csm6 (DALI Z-score: 4.2), with each half of the pseudo-symmetric SAVED domain containing homology to a single CARF domain (Figures 1.5A, B). The CARF domain of Csm6 from different CRISPR systems binds cA_4 or cA_6 cyclic oligoadenylate signals made by Cas10 following target recognition in Type III CRISPR systems (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Canonical CARF-domain proteins like Csm6 homodimerize through CARF–CARF interactions to form a two-fold symmetric binding surface for nucleotide second messenger recognition (Jia et al., 2019; Niewoehner and Jinek, 2016). In contrast, the Cap4 SAVED domain is comprised of two individual CARF-like subunits joined with an ~25 amino-acid internal linker between beta-strands 4 and 5. Using a thermal shift assay to measure *Ec*Cap4 and *Ab*Cap4 nucleotide second messenger complex formation, we confirmed that the SAVED domain is alone sufficient to specifically recognize the activating signal (Figure 1.6A).

Comparison of the *Ab*Cap4 SAVED–2′3′3′-cAAA structure with previous CARF domain structures bound to RNA ligands reveals a mixture of shared and divergent features required for nucleotide second messenger recognition. The Csm6 homodimer has two-fold symmetry, with each monomer recognizing two nucleobases of the cA_4 signal (Jia et al., 2019; Molina et al., 2019). Unlike the larger cA_4 or cA_6 signals in Type III CRISPR immunity, the majority of CD-NTase enzymes synthesize asymmetric nucleotide signals that lack internal two-fold symmetry (Whiteley et al., 2019). Pseudo-symmetry of the *Ab*Cap4 SAVED domain permits the existence of three unique pockets and specific recognition of each base of 2′3′3′-cAAA (Figures 1.5C, D, and 1.6B). The “A1 pocket” in *Ab*Cap4 is formed through a highly conserved aromatic W449 position that stacks against the nucleobase and Y454 that hydrogen bonds with the adenine N7 nitrogen (Figures 1.5D, E). Nucleobase A2 is recognized through base-stacking interactions within a pocket formed by K299 and R301 (Figure 1.6B).

The A1 pocket is conserved in both Cap4 and Csm6 homologs demonstrating that nucleobase interactions at this site have been maintained throughout divergence of SAVED

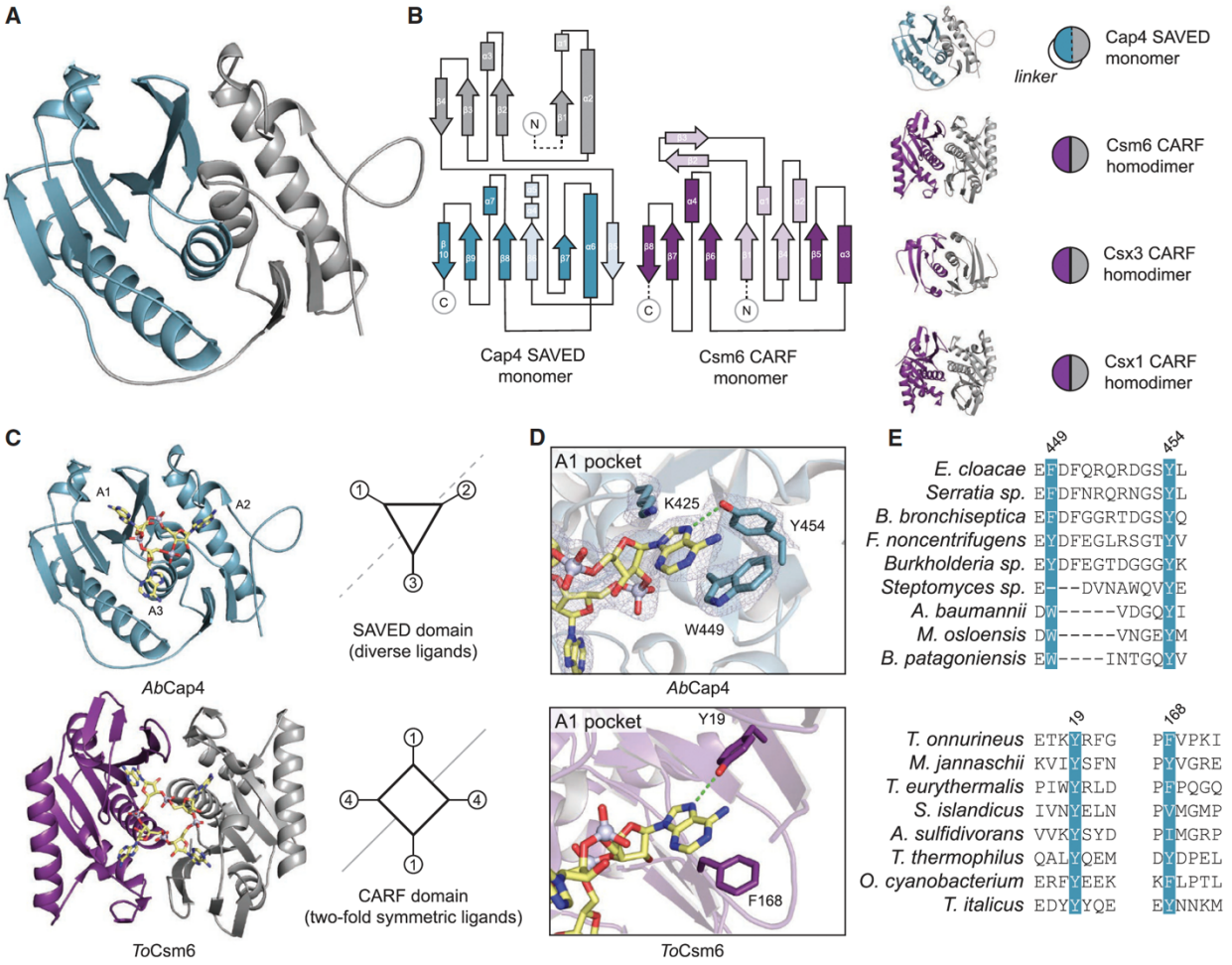


Figure 1.5. Cap4 ligand specificity is controlled by the SAVED domain, a divergent CARF family domain

(A) Structure of the AbCap4 SAVED domain, colored in blue and gray to highlight internal two-fold pseudosymmetry. Each half of the SAVED domain monomer has homology to a CARF subunit. Within the SAVED domain, these two CARF subunits are fused into a single chain with an ~25-amino acid linker. (B) Topology diagram of AbCap4 and comparison with the T. onnurineus Csm6 CARF domain from type III CRISPR immunity. Each half of the AbCap4 SAVED domain shares the common core CARF domain topology, including a b strand that leads into a bracing outer helix (b1-to-a2 and b7-to-a6) and a second b strand and central helix (b2-to-a3 and b8-to-a7) at the normal CARF dimerization interface. All CARF family proteins contain these conserved structural features, including Csm6 (PDB: 606V), Csx3 (PDB: 3WZI), and Csx1 (PDB: 2I71). (C) Structural comparison of the AbCap4 SAVED-2′3′3′-cAAA and ToCsm6 CARF-cA4 complexes. SAVED and CARF domains recognize nucleotide ligands with a similar binding surface. The single-domain architecture of the SAVED domain breaks the restriction of two-fold symmetry (dashed line) and enables recognition of diverse CD-NTase nucleotide second messengers. In contrast, the homo-dimeric architecture of canonical CARF proteins necessitates ligands with two-fold symmetry (solid line). (D and E) Structure of AbCap4 adenosine 1 (A1) pocket interactions and conservation of key residues shared in the SAVED and CARF domains. (E) Strict conservation of the A1 pocket with an aromatic residue and a tyrosine residue further supports emergence of SAVED domains through duplication and fusion of an ancient CARF family protein.

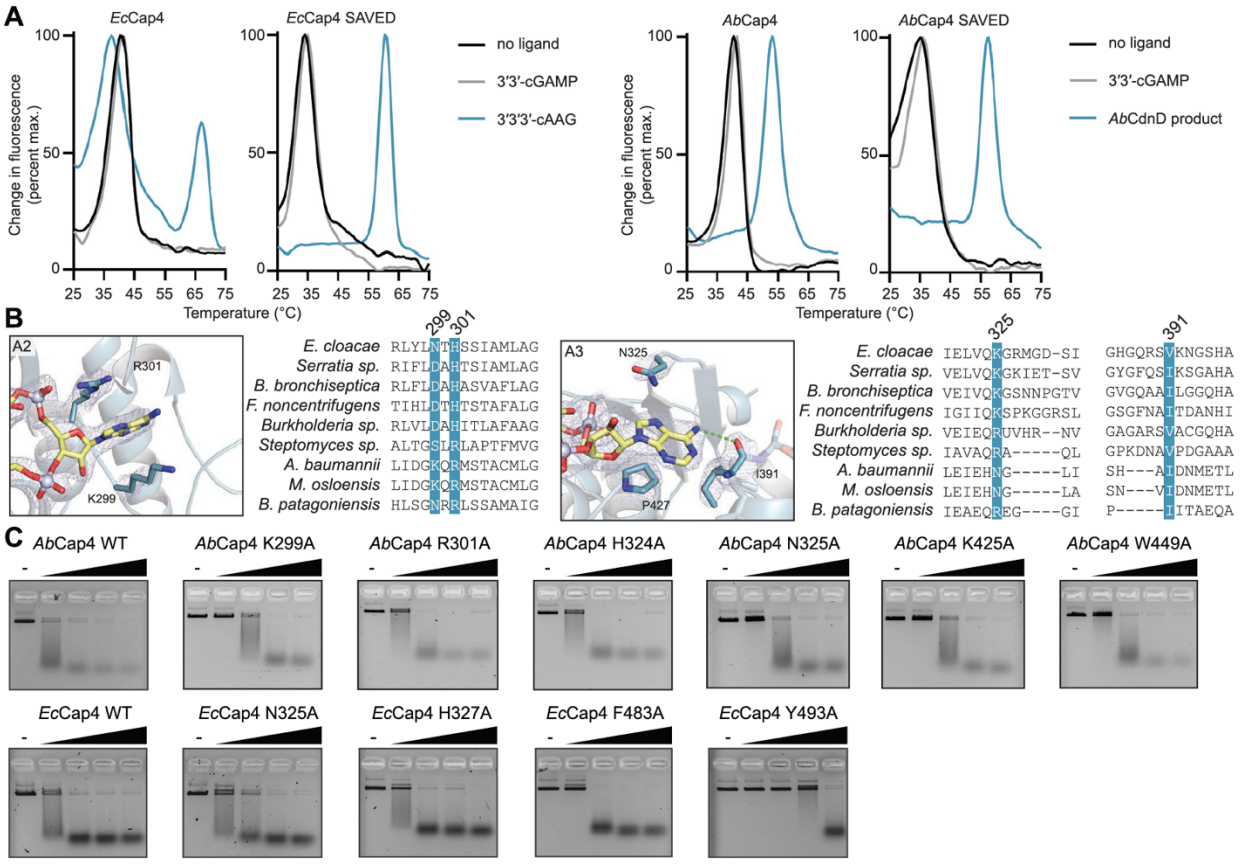


Figure 1.6. Biochemical and mutagenesis analysis of Cap4-ligand interactions
(A) Thermal denaturation assay measuring nucleotide second messenger-dependent stabilization of EcCap4 and AbCap4. Activating nucleotide second messengers bind and stabilize both full-length Cap4 and isolated SAVED-domain only constructs, confirming the SAVED domain is the ligand-binding domain. EcCap4 SAVED domain (258–end), AbCap4 SAVED domain (235–end). (B) Zoom-in cutaways showing A2 and A3 nucleotide pocket interactions in the AbCap4–2′3′3′-cAAA complex structure, and alignments of key pocket residues in various Cap4 homologs. (C) DNA degradation assays with AbCap4 and EcCap4 proteins with point mutations within the SAVED domain verifies the importance of individual contacts for nucleotide second messenger recognition. Cap4 mutants were activated with increasing concentrations of activating ligand (1 nM, 10 nM, 100 nM, 1 mM). Biochemical data are representative of at least 3 independent experiments.

proteins from a CARF-protein ancestor (Figure 1.5E). Mutagenesis confirms the importance of the A1 pocket in both *Ab*Cap4 and *Ec*Cap4, with Cap4 proteins containing mutations within this pocket requiring >10-fold more ligand to induce similar levels of activation (Figure 1.6C). In contrast, the contacts to nucleobases A2 and A3 occur in pockets that do not exist in homodimeric CARF proteins. We determined an additional 2.4 Å structure of the *Ab*Cap4–3′3′3′-cAAA complex

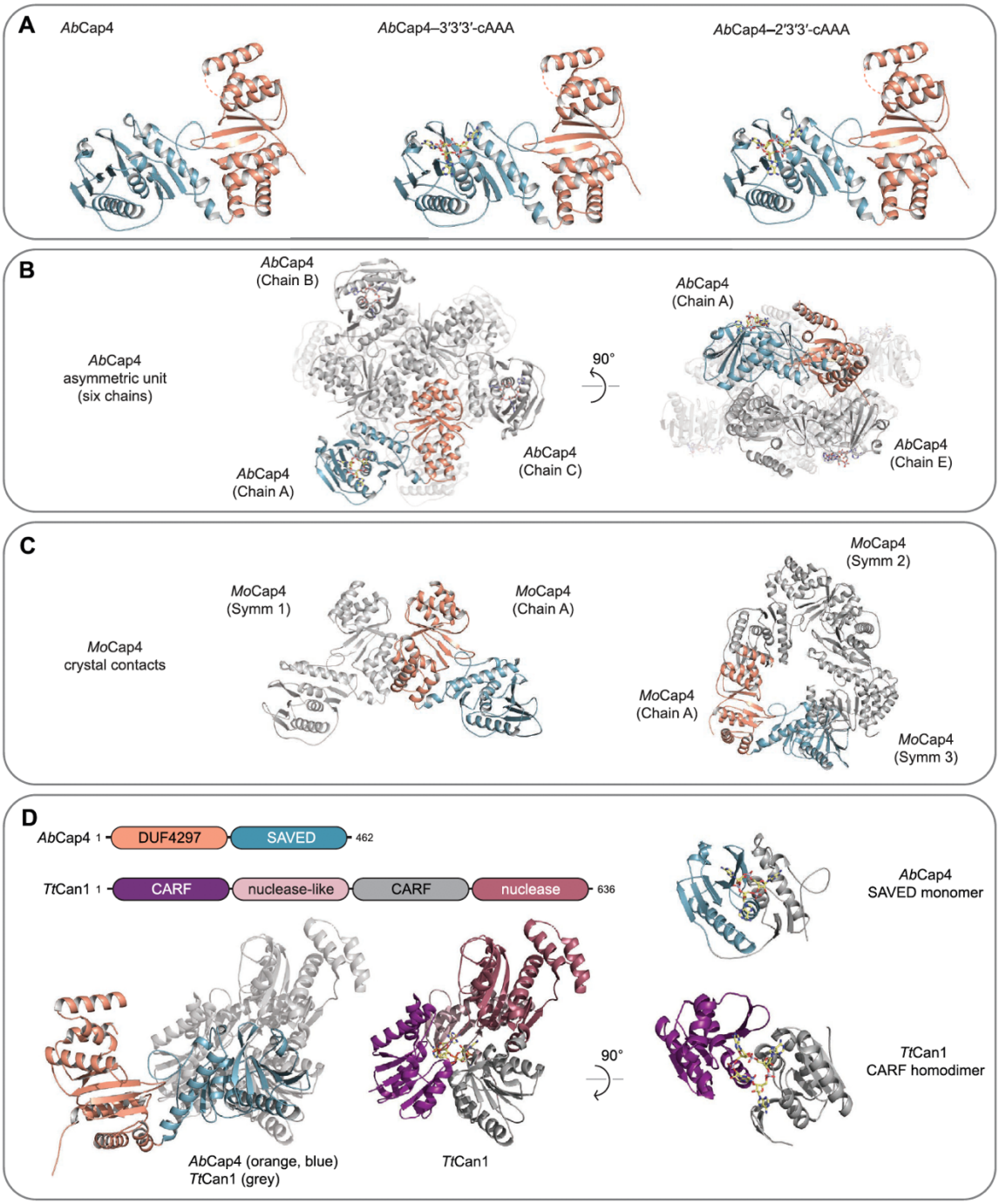


Figure 1.7. **Analysis of Cap4 ligand-bound and apo structures**
(A) Comparison of apo *Ab*Cap4, *Ab*Cap4–3′3′3′-cAAA, and *Ab*Cap4–2′3′3′-cAAA crystal structures. Each *Ab*Cap4 structure exhibits the same overall conformation supporting that the *Ab*Cap4-ligand bound structures represent an inactive-state. (B) Schematic of packing observed of each of the six chains in the *Ab*Cap4 asymmetric unit. (C) Packing of *Mo*Cap4 monomers with symmetry-related chains. (D) Structural comparison of Cap4 and Can1 from the *Thermus thermophilus* Type III CRISPR immune system (PDB 6SCE). Cap4 and Can1 each contain two CARF domains in a single polypeptide, but the proteins have distinct domain organizations and represent independent evolutionary events.

to further guide analysis of Cap4-ligand interactions and phosphodiester linkage specificity (Figure 1.7A). Comparison of the *Ab*Cap4–2′3′3′-cAAA and –3′3′3′-cAAA complexes demonstrates that a flat, parallel orientation of the nucleotide ribose at the A2 position along the SAVED domain surface allows the neighboring gap between the A1 and A2 pockets to accommodate either a 3′–5′ or 2′–5′ linkage. In contrast, the perpendicular orientation of the A3 ribose constrains the nucleotide backbone to permit only a 3′–5′ linkage and explains how architecture within the SAVED domain binding pocket can dictate signal specificity. Diversification of the ligand binding interface through genetic fusion of two CARF-like subunits into a single-chain SAVED domain was likely a key evolutionary intermediate enabling specific recognition of diverse CD-NTase nucleotide second messengers that lack two-fold symmetry. Together these data demonstrate that SAVED domains are divergent members of the CARF protein family and that the single-chain architecture of the Cap4 SAVED domain allows recognition of diverse asymmetric nucleotide recognition signals.

Cap4 proteins are activated through ligand-dependent oligomerization

To define the mechanism of Cap4 ligand-dependent activation, we next analyzed *Ec*Cap4 and *Ab*Cap4 proteins using negative stain EM. In the presence of activating nucleotide ligand, the monomeric Cap4 proteins oligomerize and form higher-order complexes (Figures 1.8A and 1.9A). The activated Cap4 complexes primarily adopt a dimeric state with two Cap4 proteins stacked against each other in an SS-shaped configuration. We also observed Cap4 particles that correspond to higher-order oligomeric complexes (Figure 1.9B), and we confirmed with SEC-MALS analysis that activated Cap4 in solution can oligomerize into multiple higher-order oligomeric states (Figures 1.8B and 1.9C). We classified the oligomerization status of ~20,000 particles from each condition to quantify the frequency of ligand-induced oligomerization. In agreement with Cap4–ligand interactions functioning as a requirement for high-order complex

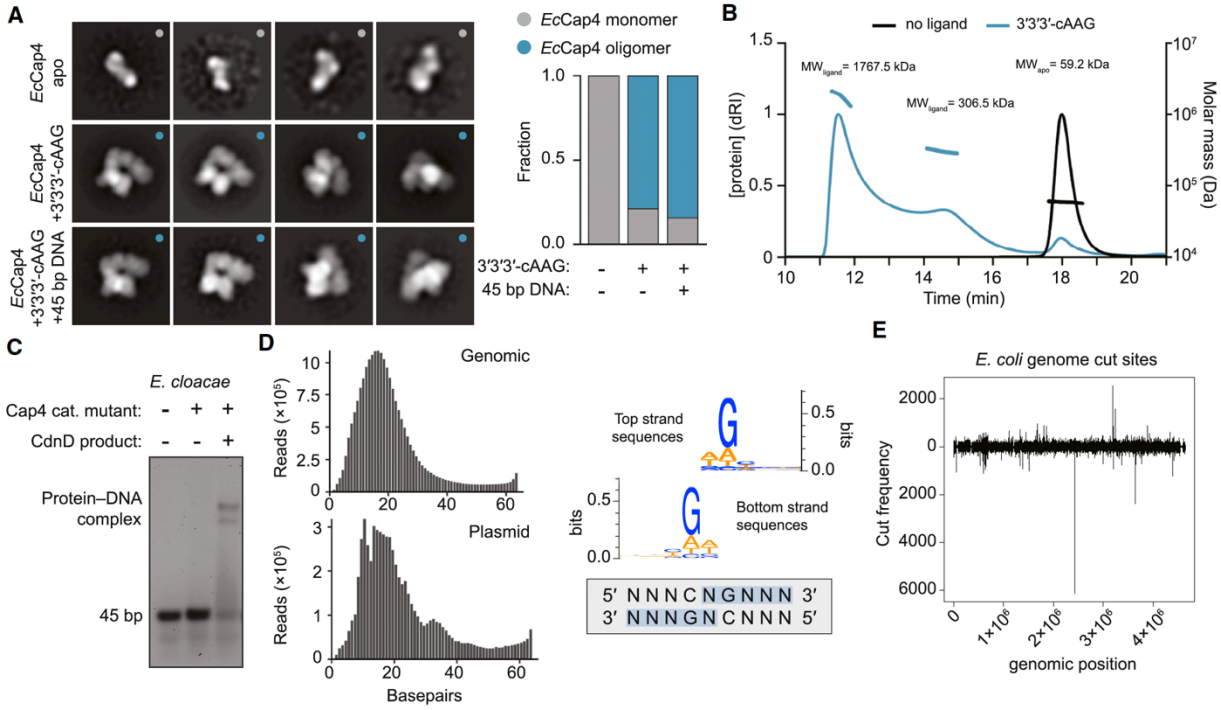
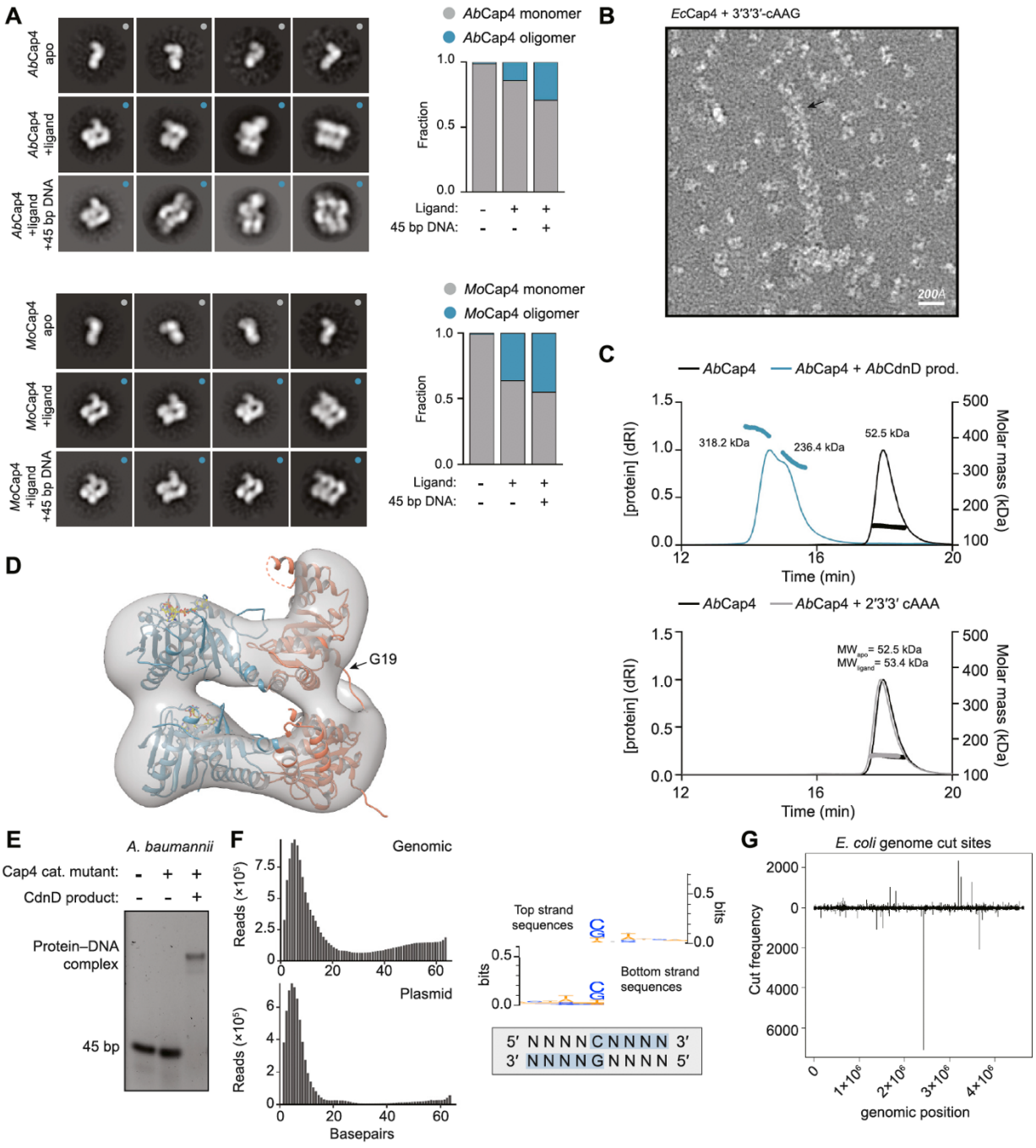


Figure 1.8. Cap4 proteins are activated through ligand-dependent oligomerization
(A) Negative-stain EM class averages of EcCap4 (K74A) with or without activating 3′3′3′-cAAG and 45 bp DNA. Particle classification and quantification for each condition demonstrates that Cap4 oligomerization only occurs in the presence of activating nucleotide second messenger. (B) Size-exclusion chromatography-multiangle light scattering (SEC-MALS) analysis of EcCap4 in the presence or absence of activating ligand 303030-cAAG. Absolute molecular masses of apo and ligand-bound complexes confirms EcCap4 oligomerization in the presence of activating nucleotide second messenger (EcCap4 expected molecular weight [MW], 56.1 kDa). (C) Electrophoretic mobility shift assay assessing DNA binding of EcCap4 (K74A). Cap4 was incubated with 45-bp DNA and activating ligand 3′3′3′-cAAG as indicated. Stable EcCap4-DNA complex formation only occurs in the presence of activating nucleotide second messenger. (D) Deep sequencing of EcCap4 DNA fragments and analysis of EcCap4 target specificity and distribution of DNA fragment sizes following E. coli genomic DNA or plasmid DNA degradation by EcCap4 (left). Cut site mapping demonstrates a minimal 50 CNG cut site preference (right). (E) Mapping of EcCap4 cleaved DNA fragments across the E. coli genomic DNA confirms the relaxed targeting specificity of EcCap4 endonuclease activity. y axis positive and negative values indicate strand-specific cutting. Data are representative of at least 3 independent experiments.

formation, <1% of the Cap4 particles occupied the dimeric or oligomeric state in the absence of activating nucleotide signal (Figure 1.8A). The same higher-order oligomerization is observed in the presence of 45 bp target DNA, suggesting that nucleotide second messenger binding alone allows assembly of Cap4 into a fully active oligomeric state.

Figure 1.9. Oligomerization of Cap4 proteins upon nucleotide second messenger recognition

(A) Negative stain EM class averages of *Ab*Cap4 (K69A) and *Mo*Cap4 (K69A) apo proteins, with activating ligand (purified using ion exchange chromatography from *Ab*CdnD reactions), or with ligand and 45 bp DNA. Cap4 proteins were incubated with components as indicated, then analyzed by negative stain EM (left). Fractions of averaged classes of each oligomeric state were calculated (right). (B) Micrograph image of a larger filament (highlighted with arrow) formed by *Ec*Cap4 in the presence of activating ligand 3′3′3′-cAAG. Higher-order oligomerization supports larger complexes observed in SEC-MALS experiments (Figure 1.8B). (C) SEC-MALS analysis of *Ab*Cap4 in the presence or absence of either purified *Ab*CdnD product or synthetic 2′3′3′-cAAA. Absolute molecular masses of apo and ligand-bound complexes were quantified by multiangle light scattering (*Ab*Cap4 expected MW: 52.1 kDa). (D) 3D reconstruction of the activated *Mo*Cap4 dimeric complex from negative stain EM data with high-resolution *Ab*Cap4–2′3′3′-cAAA structures docked within the density. The location of the Cap4 N-terminal extension and truncation point used for mutagenesis (G19) is denoted with an arrow and supports a possible role for this extension in further controlling endonuclease domain DNA binding and activation. See Figure 1.11. (E) Electrophoretic mobility shift assay assessing DNA binding of *Ab*Cap4 (K69A). Cap4 was incubated with 45 bp DNA and activating ligand as indicated. Like *Ec*Cap4, stable *Ab*Cap4–DNA complex formation only occurs in the presence of activating nucleotide second messenger. (F,G) Analysis of DNA fragments remaining after complete digestion with *Ab*Cap4. Distribution of fragments sizes (left), cut site consensus sequences (center), and location of cuts within *E. coli* genomic DNA showing relaxation of cut site sequence preference.


A
AbCap4 apo
AbCap4 +ligand
AbCap4 +ligand +45 bp DNA
AbCap4 monomer
AbCap4 oligomer
Fraction
Ligand:
45 bp DNA:
MoCap4 apo
MoCap4 +ligand
MoCap4 +ligand +45 bp DNA
MoCap4 monomer
MoCap4 oligomer
B
EcCap4 + 3′3′3′-cAAG
200Å
C
AbCap4
AbCap4 + AbCdnD prod.
318.2 kDa
236.4 kDa
52.5 kDa
[protein] (dRI)
Molar mass (kDa)
Time (min)
AbCap4 + 2′3′3′ cAAA
MW apo = 52.5 kDa
MW ligand = 53.4 kDa
D
G19
E
A. baumannii
Cap4 cat. mutant:
CdnD product:
Protein–DNA complex
45 bp
F
Genomic
Plasmid
Reads (×10⁵)
Basepairs
Top strand sequences
Bottom strand sequences
bits
5′ NNNNCNNNN 3′
3′ NNNNGNNNN 5′
G
E. coli genome cut sites
Cut frequency
genomic position

3D reconstruction at ~15 Å of the activated dimeric *Mo*Cap4 complex allowed docking of the high-resolution *Ab*Cap4–2′3′3′-cAAA crystal structure and further analysis of the mechanism of activation (Figure 1.9D). In the docked assembly, two SAVED domains stack against each other and result in alignment of the Cap4 endonuclease active-sites. Docking of Cap4 within the 3D reconstructions suggests that no major domain rearrangement is required for activation (Figure 1.9B). In agreement, all four of the crystal structures of *Ab*Cap4 and *Mo*Cap4 share a fixed, rigid orientation of the Cap4 N-terminal endonuclease and C-terminal SAVED domains in spite of distinct packing and crystal forms (Figure 1.7B, C). These results support a model where recognition of the activating nucleotide second messenger likely induces a local conformational change in the SAVED domain that drives oligomerization and endonuclease activation. Oligomerization and positioning of two adjacent Cap4 endonuclease domains creates a singular surface to engage target DNA (Figure 1.9D). Using an electrophoretic mobility shift assay, we observed that Cap4 alone is unable to bind DNA and that ligand-recognition is a pre-requisite for target DNA interaction (Figures 1.8C and 1.9E). An ~10–20 amino-acid extension is conserved at the N-terminus of each Cap4 effector but is not visible in any of our *Ab*Cap4 or *Mo*Cap4 crystal structures (Figure 1.2B, C). Interestingly, biochemical analysis of *Ec*Cap4 or *Ab*Cap4 proteins with N-terminal deletions demonstrates the N-terminal extension is dispensable for nucleotide second messenger recognition and Cap4 oligomerization but is required for DNA-binding and endonuclease domain activation (Figure 1.10). Together, these results suggest a two-step model of Cap4 activation where nucleotide second messenger recognition in the SAVED domain induces Cap4 oligomerization and subsequent DNA binding and target degradation.

To further define the result of ligand-induced Cap4 activation we next developed a deep-sequencing approach to map the cleavage specificity and fragment distribution or Cap4 nuclease activity. Sequencing of DNA fragments remaining following degradation of genomic or plasmid DNA with *Ec*Cap4 reveals a consistent fragment length of ~17 base pairs and further suggests that DNA degradation occurs through a defined oligomerization of multiple nuclease active sites

(Figure 1.8D). In agreement with the promiscuous nuclease activity observed in the biochemical assays (Figure 1.1), analysis of the *Ec*Cap4 cut site preference from mapped DNA fragments demonstrates a highly degenerate recognition sequence distinct from the strict sequence-preference characteristic of Type II restriction enzymes. *Ec*Cap4 exhibits preference for a minimal recognition sequence 5′ CNG while *Ab*Cap4 is less specific with targeting of a 5′ C or G and a final average fragment size of only ~6 bp (Figures 1.8D and 1.9F). Mapping of the observed Cap4 cut sites in *E. coli* genomic DNA demonstrates complete degradation across nearly all regions of the genome (Figure 1.8E and 1.9G). Together these data support a model of Cap4 regulation where ligand-induced oligomerization activates the endonuclease domain and results in promiscuous cleavage of DNA to minimal fragments.

SAVED domain-containing proteins are a major form of viral defense in diverse bacteria

SAVED domain-containing proteins occur in 29.8% of sequenced CD-NTase-containing operons and comprise one of the most prevalent effector modules in CBASS phage immunity (Figure 1.11A) (Burroughs et al., 2015; Cohen et al., 2019; Whiteley et al., 2019). In addition to the endonuclease–SAVED architecture in Cap4 proteins, CD-NTase operons encode SAVED domains fused to additional protein partners including putative HNH nucleases, caspase-like proteases, calcineurin-like phosphatases, TIR NAD^+ hydrolase enzymes, and transmembrane-containing segments (Figure 1.11B). Nucleotide second messenger-induced oligomerization of SAVED domains is likely a general strategy to activate diverse effector functions and coordinate abortive infection systems to limit phage replication.

Analysis of the bacterial CD-NTase phylogenetic tree reveals that SAVED domain-containing proteins are encoded near vastly divergent CD-NTase enzymes from across clades B, C, D, G and H (Figure 1.11A) (Whiteley et al., 2019). The broad distribution of SAVED domain-containing effectors suggests this domain is capable of recognizing diverse classes of nucleotide second messengers. To test this hypothesis, we conducted a genetic screen to identify SAVED-

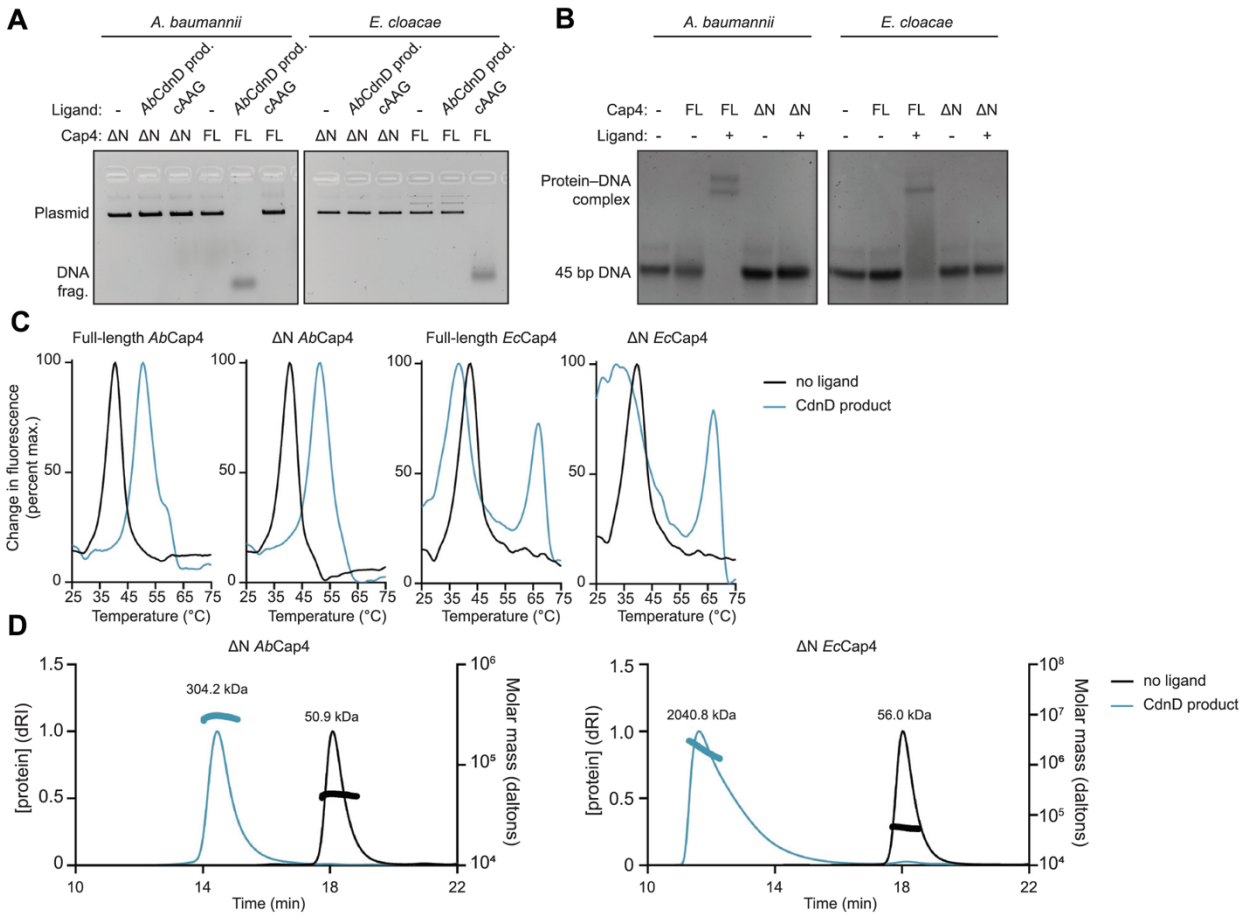


Figure 1.10. Characterization of Cap4 N-terminus in DNA binding, ligand binding, and oligomerization

(A) Agarose gel analysis of Cap4 activity and plasmid DNA degradation. Full-length Cap4 proteins degrade target DNA in the presence of activating nucleotide second messenger, but N-terminal truncations of *Ab*Cap4 (G19–end) and *Ec*Cap4 (G11–end) do not. (B) Electrophoretic mobility shift assay assessing DNA binding of N-terminal truncations of *Ab*Cap4 (G19–end) and *Ec*Cap4 (G11–end). The conserved N-terminal extension not visible in the *Ab*Cap4 or *Mo*Cap4 crystal structures is required for DNA recognition. (C) Thermal denaturation assay measuring nucleotide second messenger-dependent stabilization of N-terminal truncations of *Ec*Cap4 and *Ab*Cap4. Activating nucleotide second messengers bind and stabilize truncated Cap4 proteins, indicating the N-terminus is not required for ligand-binding. (D) SEC-MALS analysis of N-terminal truncations of *Ec*Cap4 and *Ab*Cap4 in the presence or absence of activating ligand. Quantification of absolute molecular masses of *Ec*Cap4– and *Ab*Cap4–ligand complexes confirm the endonuclease domain N-terminal extension does not impact ligand-binding (ΔN *Ec*Cap4 expected MW: 55.1 kDa; ΔN *Ab*Cap4 expected MW: 50.4 kDa).

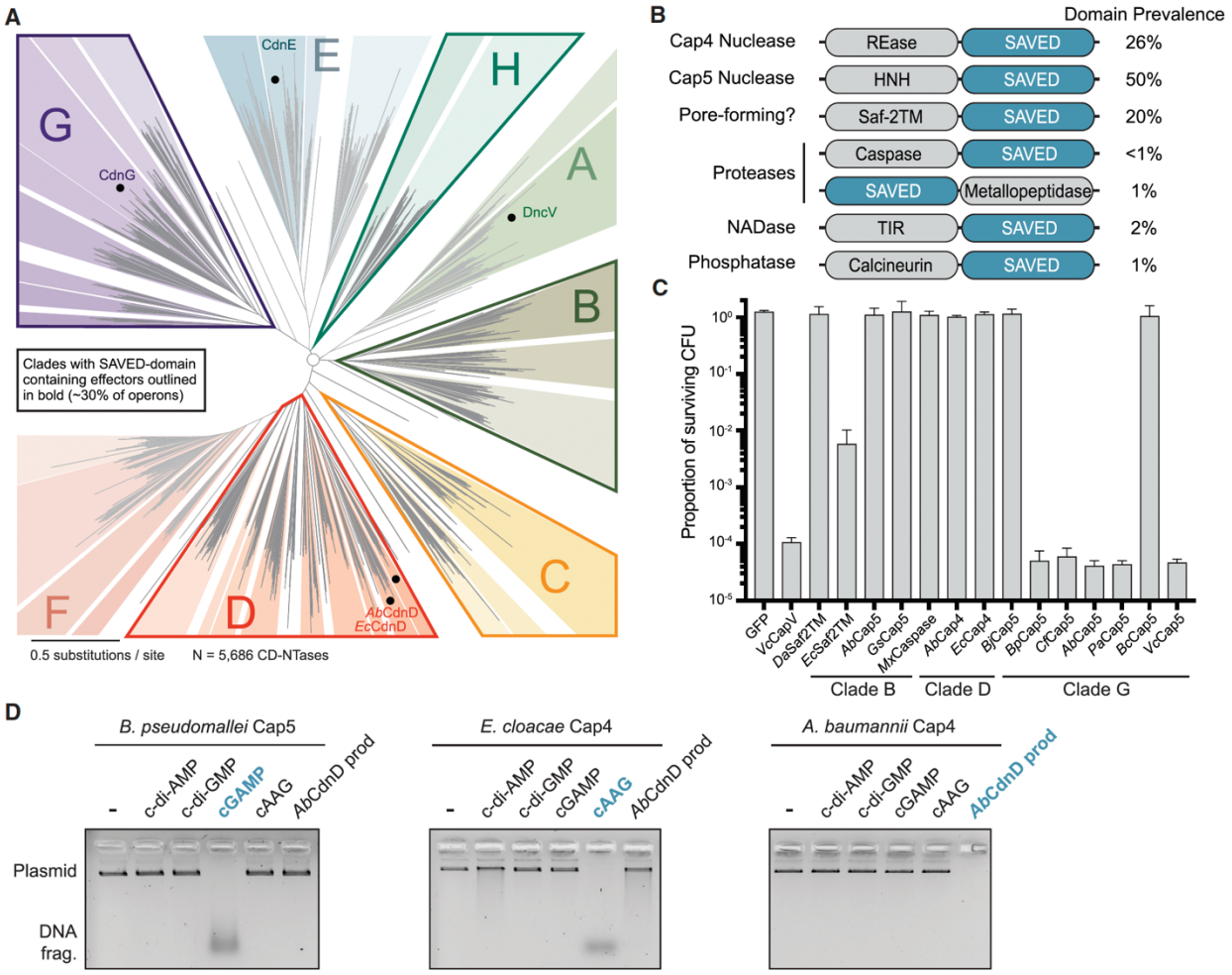


Figure 1.11. SAVED domain-containing proteins are a major form of viral defense in diverse bacteria.

(A) Phylogenetic tree showing distribution of SAVED-containing effectors, in bacterial CD-NTase operons. Bacterial CD-NTase clades are displayed as an unrooted tree (Whiteley et al., 2019), with clades containing SAVED effectors (totaling ~30% of all bacterial CD-NTase operons) outlined in bold (clades B, C, D, G, H). Dots indicate locations of CD-NTases of interest. (B) Domain organization of diverse SAVED-containing effectors (Burroughs et al., 2015). Prevalence of each domain architecture in sequenced bacterial genomes is listed as a percent of all SAVED-containing effectors (REase, restriction endonuclease). (C) A genetic screen to identify SAVED-effectors responsive to 3′3′-cGAMP. SAVED-effectors were co-expressed with the 3′3′-cGAMP-producing CD-NTase DncV, and a spot-dilution assay with quantification of recovered colony forming units (CFU) used to assess SAVED activation and cytotoxicity. SAVED-domain effectors from CD-NTase clades B and G respond to 3′3′-cGAMP demonstrating that SAVED domains are capable of responding to both cyclic dinucleotide and cyclic trinucleotide second messengers. See methods for CBASS effector species designations. (D) *In vitro* reconstitution of nucleotide second messenger specificity for divergent Cap4 proteins and the HNH endonuclease–SAVED fusion *B. pseudomallei CD-NTase associated protein 5* (Cap5). *Bp*Cap5 stimulated DNA cleavage activity demonstrates that SAVED-domain dependent nucleotide second messenger sensing is capable of activating structurally distinct enzymatic domains.

containing proteins that specifically respond to the cyclic dinucleotide 3′3′-cGAMP. Co-expression analysis in *E. coli* with the promiscuously active CD-NTase *V. cholerae* DncV identified multiple SAVED domain-containing proteins in CD-NTase clades B and G that respond to 3′3′-cGAMP and induce cell death (Figure 1.11C). These effectors contain an HNH–SAVED architecture structurally distinct from the Cap4 DUF4297 endonuclease-SAVED architecture, and we therefore designated genes with this architecture as *CD-NTase-associated protein 5* (*Cap5*). Purified *Burkholderia pseudomallei* Cap5 (Genbank WP_004556385.1) is active in the presence of 3′3′-cGAMP and catalyzes robust dsDNA degradation *in vitro* (Figure 1.11D). Activation of each tested SAVED domain-containing protein requires recognition of a specific nucleotide second messenger further demonstrating the specificity of CD-NTase signaling pathways (Figure 1.11D). Notably, the 3′3′-cGAMP-responsive SAVED-domain effector *Bp*Cap5 failed to respond to the common bacterial second messengers 3′3′ cyclic di-GMP and 3′3′ cyclic di-AMP. Exquisite nucleotide ligand specificity therefore insulates CBASS antiviral immune systems from other host signals and limits inappropriate activation.

CD-NTase operons have been demonstrated to protect bacterial populations during phage infection (Cohen et al., 2019; Ye et al., 2020). Interestingly, SAVED domain-containing proteins are also present within Type III CRISPR operons, suggesting the shared mechanism of SAVED- and CARF-dependent activation allows exchange of CBASS and CRISPR components and further diversification of antiviral immune defenses (Figure 1.12A). To verify that CD-NTase operons with SAVED domain-containing effectors also restrict phage replication, we used the *E. cloacae* operon containing *CdnD, Cap2, Cap3,* and *Cap4* and an *E. coli* operon containing *CdnG* (Genbank WP_000064266.1), *Cap2*, *Cap3*, and *Cap5* for expression and analysis of phage resistance (Figure 1.12B). Introduction of plasmids encoding either of these operons in *E. coli* mediated a >3-log decrease in efficiency of plaque formation for phage T2. Expression of the CdnG operon resulted in an ~1-log decrease in phage T5 plaque formation while the CdnD operon had no effect. No effect was observed on replication of phage T7, further supporting an emerging

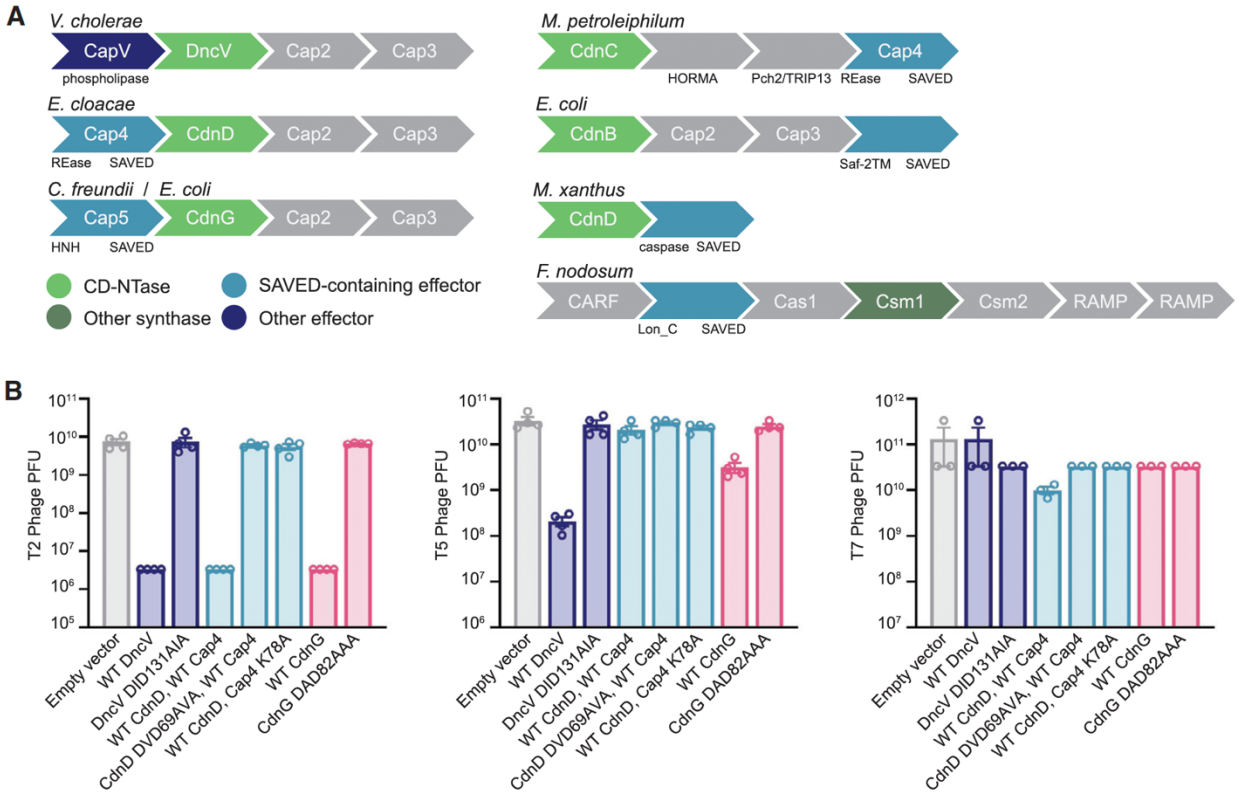


Figure 1.12. SAVED effectors are essential for CBASS antiviral defense
(A) Example CBASS and CRISPR immunity operons encoding SAVED domain-containing effector proteins. *Vibrio cholerae* (WP_001901330), *Enterobacter cloacae* (WP_032676400), *Citrobacter freundii* (WP_032942206), *Methylibium petroleiphilum* (WP_011829962), *Escherichia coli* (WP_000058223), *Myxococcus xanthus* (YP_635404), *Fervidobacterium nodosum* (WP_011994539). (B) CBASS operons encoding SAVED-domain effectors protect *E. coli* from phage replication. *E. coli* containing either an empty vector or vector encoding a CBASS operon were infected with phage as indicated and efficiency of plating was quantified by plaque forming units. CBASS operons enable phage-specific immunity.

model where CBASS operons enable protection from specific subsets of phage (Figure 1.12B) (Cohen et al., 2019; Ye et al., 2020). Restriction of phage replication is strictly dependent on the catalytic activity of the CD-NTase and the nuclease active-site of Cap4 (Figure 1.12B). Together, these data demonstrate discovery of a broad strategy in bacteria that couples CD-NTase signaling to activation of downstream SAVED-containing proteins to protect against viral replication.

Discussion

Here we identify *E. cloacae* Cap4 as the founding member of a widespread family of antiviral effector proteins in bacteria. We show that Cap4 proteins are nucleotide second messenger sensors that function as downstream receptors and effector DNA endonucleases in CBASS immune systems. CBASS immunity relies on bacterial CD-NTase enzymes that catalyze synthesis of specialized oligonucleotide second messengers including cyclic di- and tri-nucleotide products (Cohen et al., 2019; Whiteley et al., 2019). Although the mechanism of CBASS activation upon phage infection remains unclear, many bacterial CD-NTase enzymes are constitutively active *in vitro* and synthesize nucleotide second messengers once highly purified (Whiteley et al., 2019). Similar to sensors that function as guards in plant innate immunity pathways (Dodds and Rathjen, 2010), many CD-NTases may therefore sense phage replication indirectly by detecting abnormal changes in metabolite concentrations or loss of bacteria homeostasis (Kranzusch, 2019). Following phage recognition and CD-NTase activation, Cap4 recognizes the resulting oligonucleotide signals and uses a two-domain protein architecture to couple ligand-binding with induction of DNA endonuclease activity. Ligand-binding occurs within a C-terminal SAVED-domain and mediates activation of an appended N-terminal endonuclease/DUF4297 domain with homology to Type II restriction endonucleases. Cap4-mediated DNA cleavage is promiscuous and restricts phage replication in a bacterial population. SAVED-domain containing proteins like Cap4 are found in ~30% of CBASS operons and provide a clear mechanism for how of these systems mediate broad resistance to viral infection.

Structures of *A. baummannii* and *M. osloensis* Cap4 explain a molecular mechanism for ligand-recognition and effector function activation. Following nucleotide second messenger recognition in the SAVED domain, Cap4 oligomerizes in a conformation that stacks two nuclease domains in close proximity (Figures 1.5, 1.7D, and 1.8). A nucleotide second messenger-dependent transition therefore maintains the Cap4 nuclease effector domain in an inactive monomeric state until detection of the correct CD-NTase signal enables higher-order complex

assembly and target DNA degradation. The *Ab*Cap4 SAVED domain is alone sufficient to recognize the activating cyclic trinucleotide signal (Figure 1.6A) indicating that the fused effector domain does not participate in ligand specificity or the initial step of receptor activation. Additionally, we characterize an HNH endonuclease–SAVED effector Cap5 from *B. pseudomallei* and demonstrate that CD-NTase nucleotide second messenger recognition is capable of activating structurally divergent effector domains (Figure 1.11D). These results support a sensor–effector model of immunity where fusion of a SAVED domain is sufficient to enable nucleotide second messenger-gated control of diverse partner effector proteins. SAVED domains are found fused to a wide variety of additional effector domains including protease, phosphodiesterase, and potentially pore-forming TM-proteins (Burroughs et al., 2015; Cohen et al., 2019) supporting that the generalizability of SAVED-dependent activation allows diverse domains to be co-opted as modules for viral defense.

We show that the SAVED protein domain formed through fusion of two CARF-like subunits into a single-chain nucleotide second messenger sensor (Figure 1.5). These results explain how SAVED domains recognize diverse asymmetric ligands and reveal an unexpected shared evolutionary history between CBASS and CRISPR antiviral systems. Nearly all characterized Type III CRISPR systems function through homodimeric CARF proteins that recognize a nucleotide second messenger with two-fold symmetry (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Fusion of two CARF-like subunits into a SAVED domain breaks the requirement for two-fold symmetrical ligands and explains a key evolutionary event that enabled CD-NTase–SAVED systems to signal with diverse asymmetric ligands. A similar evolutionary event of fusion of two CARF domains into a single protein was recently observed with an accessory nuclease in a *Thermus thermophilus* Type III CRISPR immune system (McMahon et al., 2020) (Figure 1.7D), further highlighting how genetic fusion of multiple CARF domains can provide an evolutionary advantage in antiviral immune systems. Interestingly, a structurally unrelated 3′3′3′-cAAA-activated DNA endonuclease named NucC has been identified in an *E. coli* CdnC CD-NTase

system (Lau et al., 2020). *E. coli* NucC is a homo-trimeric receptor with a three-fold symmetric ligand-binding pocket revealing that SAVED and NucC effectors are likely alternative evolutionary paths that enabled nucleotide second messenger diversification within CBASS immunity. CD-NTase enzymes synthesize nucleotide second messengers using all four nucleobases as building blocks (Whiteley et al., 2019), and we now show that CD-NTase product second messengers are further specified with incorporation of both 2′–5′ and 3′–5′ phosphodiester linkages. These findings dramatically expand the known nucleotide second messenger signaling space and indicate that CD-NTase enzymes are theoretically capable of synthesizing at least 180 distinct oligonucleotide variants. Combined with cA_4 and cA_6 species in Type III CRISPR immunity and linear 2′–5′-linked chains produced by mammalian oligoadenylate synthase enzymes (Hornung et al., 2014), these results reveal a vast diversity of natural nucleotide second messenger signals in viral defense.

The enormous diversity of CD-NTase nucleotide second messenger signals likely enables CBASS antiviral systems to specialize and adapt to phage resistance. Mammalian poxviruses encode a 2′3′-cGAMP-specific nuclease that degrades the cGAS product and prevents STING-dependent immune responses (Eaglesham et al., 2019). Likewise, a phage protein was recently identified that degrades cA_4 to evade Type III CRISPR immunity (Athukoralage et al., 2020). Viral evolution of nucleases that specifically degrade second messenger signals likely creates evolutionary pressure to force diversification of antiviral signaling systems. The ability of CD-NTase–SAVED systems to morph between cyclic dinucleotide and cyclic trinucleotide signals may represent larger “leaps” that allow escape and temporary relief from an evolutionary arms race (Daugherty and Malik, 2012). Importantly, our discovery of a role for 2′–5′ linkages in the *A. baumannii* CD-NTase signaling system demonstrates that incorporation of noncanonical linkages is likely another mechanism that evolved in bacteria to subvert viral resistance. Along with the previous identification of diverse cGAS-like enzymes in bacteria and effector proteins with sequence homology to mammalian STING (Cohen et al., 2019; Whiteley et al., 2019), the discovery of bacterial signals with 2′–5′ linkages demonstrates that all fundamental components

of mammalian cGAS-STING signaling appear to be functionally shared within ancestrally-related bacterial CBASS immune systems.

Author Contributions

Brianna Lowey: Conceived the project, designed and conducted all biochemical and crystallography experiments and CD-NTase nucleotide product purification, and wrote the manuscript.

Aaron T. Whiteley: Designed and conducted phage challenge and bacterial cytotoxicity experiments. **Alexander F.A. Keszei**: Conducted negative stain EM experiments. **Benjamin R. Morehouse**: Assisted with CD-NTase nucleotide product purification. **Ian Mathews**: Conducted with LC-MS/MS experiments. **Sadie P. Antine**: Conducted DNA fragment deep-sequencing experiments and cleavage site analysis. **Victor J. Cabrera**: Assisted with phage challenge and bacterial cytotoxicity experiments. **Dmitry Kashin**: Performed synthetic nucleotide second messenger synthesis and characterization experiments. **Percy Niemann**: Performed synthetic nucleotide second messenger synthesis and characterization experiments. **Mohit Jain**: Assisted with LC-MS/MS experiments. **Frank Schwede**: Designed synthetic nucleotide second messenger synthesis and characterization experiments. **John J. Mekalanos**: Assisted with phage challenge and bacterial cytotoxicity experiments. **Sichen Shao**: Conducted negative stain EM experiments. **Amy S.Y. Lee**: Conducted DNA fragment deep-sequencing experiments and cleavage site analysis.

Philip J. Kranzusch: Conceived the study, designed experiments, assisted with structure determination, and wrote the manuscript.

Specific Acknowledgements

The authors are grateful to Kevin Corbett (U.C. San Diego) and members of the Kranzusch lab for helpful comments and discussion, to Kelly Arnett and Harvard University's Center for Macromolecular Interactions for advice regarding SEC-MALS experiments, and the Molecular Electron Microscopy Suite at Harvard Medical School.

METHOD DETAILS

Protein expression and purification

Recombinant CD-NTase and Cap4 proteins were purified from *E. coli* as previously described (Whiteley et al., 2019; Zhou et al., 2018). Briefly, *Ec*Cap4, Cap4 homologs, and CD-NTase enzymes were cloned from synthetic DNA (Integrated DNA Technologies) into a custom pET vector encoding a 6×His-SUMO2-tagged fusion protein and transformed into BL21-RIL *E. coli* (Agilent) containing the rare tRNA plasmid pRARE2. Single transformants were inoculated into a 20 ml MDG media starter culture, grown overnight at 37°C, and used to seed 1 L M9ZB media cultures grown at 37°C, 230 RPM until an OD_{600} of ~2.5. Cultures were induced by chilling flasks on ice for 20 min, supplementing cultures with to 0.5 mM IPTG, and then incubating at 16°C, 230 RPM for ~15 h. Bacteria were pelleted and lysed by sonication in 1× Lysis Buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 30 mM imidazole, 10% glycerol and 1 mM DTT) and recombinant protein was purified by gravity chromatography and binding to Ni-NTA resin (Qiagen). Resin was washed with 1× Lysis Buffer supplemented to 1 M NaCl and then eluted with 1× Lysis Buffer supplemented to 300 mM imidazole. Purified protein was supplemented with recombinant human SENP2 protease (D364–L589, M497A) to remove the SUMO2 tag, and dialyzed overnight against 20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM DTT. Proteins were concentrated using a 30K-cutoff concentrator (Millipore) and further purified by size-exclusion chromatography on a 16/600 Superdex 200 column in 20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM TCEP. Final purified fractions were analyzed by SDS-PAGE and Coomassie staining, concentrated to >30 mg ml^{-1}, flash-frozen in liquid nitrogen, and stored as 35 µl aliquots at −80°C.

Nucleotide product electrophoretic mobility shift assay

Radiolabeled nucleotide products were enzymatically synthesized with purified *V. cholerae* DncV or *Ec*CdnD in reactions containing 25 µM ATP, 25 µM GTP, trace α^{32}P-GTP in 1× Reaction Buffer (5 mM $MgCl_2$, 50 mM KCl, 1 mM TCEP and either 50 mM Tris-HCl pH 7.5 (DncV) or 50 mM

CAPSO pH 9.4 (*Ec*CdnD)) in a final reaction volume of 20 µl. Synthesis reactions were incubated ~15 h at 37°C and then inactivated by treating with 1 µl of 5 units μl^{-1} Calf Intestinal Phosphatase (New England Biolabs) and incubating at 37°C for 1 h. Radiolabeled products were then diluted 1:20 into electrophoretic mobility shift reactions containing either no protein, ~1 µM 6×His-SUMO2-*Ec*Cap2, 6×His-SUMO2-*Ec*Cap3, or 6×His-SUMO2-*Ec*Cap4 and 50 mM KCl, 50 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, and 1 mM TCEP. Reactions were incubated for 20 min at 25°C, then separated on an 8 cm 10% nondenaturing polyacrylamide gel run at 100 V for 45 min. The gel was dried and then exposed to a phosphor-screen and imaged with a Typhoon Trio Variable Mode Imager (GE Healthcare).

Protein crystallization and structure determination

Cap4 proteins were crystallized at 18°C using the hanging drop method. Concentrated protein stocks were thawed on ice and diluted in buffer (25 mM HEPES-KOH pH 7.5, 1 mM TCEP). Drops were set by mixing protein stock and reservoir solution in 2 µl drops over a 350 µl reservoir in Easy-Xtal 15-well trays (Qiagen). Each protein was crystallized as follows: 1) Apo *Mo*Cap4: Native or selenomethionine *Mo*Cap4 protein was diluted to 10 mg ml^{-1}, supplemented with 10.5 mM $MgCl_2$, mixed at a 1:1 ratio of protein:reservoir solution (110 mM Tris-HCl pH 8.5, 21% PEG-400), crystals were allowed to grow for 5 days and then were cryoprotected using reservoir solution supplemented with 10% ethylene glycol. 2) *Ab*Cap4–2′3′3′-cAAA: Native *Ab*Cap4 protein was diluted to 10 mg ml^{-1} and pre-incubated with 10.5 mM $MgCl_2$ and ~200 µM purified *Ab*CdnD product, mixed at a 1:1 ratio of protein:reservoir solution (10% PEG-3350, 200 mM lithium sulfate, 100 mM imidazole pH 8.0), crystals were allowed to grow for 25 days and then were cryoprotected using reservoir solution supplemented with 25% ethylene glycol; Selenomethionine *Ab*Cap4 was diluted to 10 mg ml^{-1} and pre-incubated with 10.5 mM $MgCl_2$ and ~200 µM purified *Ab*CdnD product, mixed at a 1:1 ratio of protein:reservoir solution (0.2 M lithium sulfate, 0.1 M Tris-HCl pH 8.5, 25% PEG-400), crystals were allowed to grow for 13 days and then were cryoprotected using

reservoir solution supplemented with 10% PEG-400. 3) Apo *Ab*Cap4: Native *Ab*Cap4 protein was diluted to 10 mg ml^{-1} and mixed at a 1:1 ratio of protein:reservoir solution (0.1 M Tris pH 8.5, 0.2 M $LiSO_4$, 16% PEG-3350), crystals were allowed to grow for 6 days and then were cryoprotected using reservoir solution supplemented with 30% ethylene glycol. 4) *Ab*Cap4–3′3′3′-cAAA: Native *Ab*Cap4 protein was diluted to 10 mg ml^{-1} and pre-incubated with 500 μM 3′3′3′-cAAA (Biolog), mixed at a 1:1 ratio of protein:reservoir solution (0.1 M Tris pH 8.5, 0.2 M $LiSO_4$, 16% PEG-3350), crystals were allowed to grow for 6 days and then were cryoprotected using reservoir solution supplemented with 25% ethylene glycol.

X-ray crystallography data were processed with XDS and AIMLESS (Kabsch, 2010) using the SSRL autoxds script (A. Gonzalez, Stanford SSRL). Experimental phase information for *Mo*Cap4 and *Ab*Cap4 was determined using data collected from selenomethionine-substituted crystals. In total, 5 and 36 sites respectively were identified with HySS in PHENIX (Adams et al., 2010) and an initial map was produced using SOLVE/RESOLVE (Terwilliger, 1999). Model building was performed using Coot (Emsley and Cowtan, 2004), before refinement in PHENIX. Statistics were analyzed as described in Table 1 (Chen et al., 2010; Karplus and Diederichs, 2012; Weiss, 2001).

DNA degradation assays

DNA degradation assays were performed by incubating 50 nM Cap4 protein with 10 ng μl^{-1} pGEM9z plasmid (Promega), *E. coli* genomic DNA, or 100 bp synthetic DNA as specified in a 25 μl reaction for 20 min at 37°C with 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, and 1 mM TCEP. Degradation reactions were supplemented with activating nucleotide second messenger product at 50 nM synthetic or purified product (*Ab*CdnD product), or unpurified cognate CD-NTase reaction (see below). Reactions were stopped by addition of 6× loading buffer (60 mM EDTA, 20 mM Tris-HCl pH 7.5, 30% glycerol, 0.1% SDS), and then 10 μl was separated on a 1.5% or 2% TAE agarose gel. Gels were run at 100 V for 20 min, then stained by rocking at room-temperature

in buffer with 10 μg ml^{-1} ethidium bromide for 30 min. Gels were de-stained in water for ~10 min and then imaged with a ChemiDoc MP Imaging System.

*Ab*CdnD product purification

Large scale *Ab*CdnD product synthesis was carried out in 250 ml reactions with 500 nM purified *Ab*CdnD and 250 μM ATP in 1× reaction buffer (50 mM KCl, 20 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM TCEP, 40 mM Tris pH 9.0), incubated overnight at 37°C with gentle shaking. The reaction was then treated with 20 μl of 5 units μl^{-1} Calf Intestinal Phosphatase (New England Biolabs) at 37°C for 3 h, then 0.2 μm filtered. The reaction was diluted to 12.5 mM salt with water, then loaded on to a HiTrap Q column (GE Healthcare) at 5 ml min^{-1} with a peristaltic pump and eluted on an AKTA FPLC (GE Healthcare) with a linear gradient of ammonium acetate from 0 to 2 M over 50 min. The *Ab*CdnD product elutes as a single major peak at approximately 1.5 M ammonium acetate. Fractions containing the major product were identified by absorbance at 254 nm, pooled and dried by speedvac, and then re-suspended in approximately 1 ml of water. Initial ion exchange-purified product was desalted by gel filtration on a Superdex 30 Increase 10/300 GL column (GE Healthcare) run at 0.5 ml min^{-1} with water as running buffer.

Second-step purification was carried out using a C18 column (Agilent Zorbax Bonus-RP 4.6×150 mm, 3.5-micron). The column was heated to 50°C and run at 1 ml min^{-1} with a mobile phase of 50 mM NaH_2PO_4 (pH 6.8 with NaOH) supplemented with 3% acetonitrile. The sample was injected with ~10 μg per run to ensure separation of major and minor products. Fractions from several runs were pooled and dried by speedvac overnight, then desalted using C18 cartridges (Thermo Scientific Hypersep C18 100 mg columns). Speedvac-dried fractions were re-suspended in nuclease-free water with 2% acetic acid. The C18 cartridge was pre-washed with 500 μl MeOH, then 2× with 500 μl water with 2% acetic acid before loading the sample over the cartridge 3×. The cartridge was washed 2× with 500 μl water with 2% acetic acid, then eluted with 500 μl MeOH. Sample was dried by speedvac for approximately 2 h, then re-suspended in

nuclease-free water. Concentrations were estimated using an extinction coefficient of ε = 40,500 l mole^{-1} cm^{-1}.

CD-NTase reactions and thin layer chromatography

CD-NTase reactions were performed with 5 μM recombinant enzyme in 20 μl reactions with 25 μM NTPs and trace α^{32}P-ATP (mcGAS: ATP, GTP; DncV: ATP, GTP; *Ec*CdnD: ATP, GTP; *Ab*CdnD: ATP) overnight at 37°C with 50 mM KCl, 10 mM $MgCl_2$, 1 mM TCEP, and 50 mM CAPSO pH 9.4 (CdnD) or 50 mM Tris-HCl pH 7.5 (cGAS, DncV). cGAS reactions were additionally supplemented with 2 μM 45-bp DNA, and *Ab*CdnD reactions were additionally supplemented with 1 mM $MnCl_2$. Reactions were terminated with addition of 1 μl of 5 units μl^{-1} Calf Intestinal Phosphatase (New England Biolabs) and incubating at 37°C for 1 h. For P1 degradation assays, 10 μl of each reaction was treated with 1 μl P1 nuclease (Sigma Product N8630) for 1 h at 37°C. Reactions were analyzed by thin-layer chromatography by diluting each reaction 1:10 in 100 mM sodium acetate pH 5.2 and then spotting 1 μl of each reaction on a PEI-cellulose plate (Millipore) developed in ~1 cm of 1.5 M KH_2PO_4 pH 3.8 until buffer was 1 cm from the top of plate. Plates were dried then expose to a phosphor-screen and imaged with a Typhoon Trio Variable Mode Imager (GE Healthcare).

***Ab*CdnD nucleotide product degradation and fragment analysis**

Synthetic controls (3′3′3′-cAAA and 2′3′3′-cAAA) or purified *Ab*CdnD product were diluted to 25 μM in 100 μl reactions and supplemented with 10× P1 buffer (final 1× concentration: 30 mM NaOAc pH 5.3, 5 mM $ZnSO_4$, 50 mM NaCl) or snake venom phosphodiesterase buffer (final 1× concentration: 50 mM Tris pH 9, 10 mM MgCl2, 50 mM NaCl). Snake venom phosphodiesterase (0.8 mU) (Sigma, Phosphodiesterase I from *Crotalus adamanteus* venom, Product P3243) or Nuclease P1 (80 mU) (Sigma Product N8630) was added and reactions were incubated for 1 h or 5 h at 37°C, respectively. Reactions were diluted 1:2 in nuclease-free water, then filtered

through a 10-kDa filter (Millipore) by centrifugation. The degradation products were analyzed using a C18 column (Agilent Zorbax Bonus-RP 4.6×150 mm, 3.5-micron) on HPLC held at 50°C and run at 1 ml min^{-1} with 50 mM NaH_2PO_4 (pH 6.8 with NaOH) supplemented with 3% acetonitrile.

Chemical synthesis of cyclic (adenosine-(2′–5′)-monophosphate-adenosine-(3′–5′)-monophosphate-adenosine-(3′–5′)-monophosphate) (c[A(2′,5′)pA(3′,5′)pA(3′,5′)p] / 2′–5′, 3′–5′, 3′–5′ c-AMP–AMP–AMP / 2′3′3′-cAAA), sodium salt

All reagents and solvents for chemical operations were of analytical grade or the best grade available from commercial suppliers. Solvents for chromatographic operations were specified as analytical grade, HPLC grade, or gradient HPLC grade. YMC*Gel SIL (6 nm, S-75 µm) was used for preparative flash chromatography and TLC was performed with Merck 60 F254 silica gel plates. All chromatographic operations were performed at ambient temperature. Evaporation of solvents was accomplished by rotary evaporation *in-vacuo* either with membrane pump vacuum or oil pump high vacuum with water bath temperatures not exceeding 30–33°C.

UV-spectra were recorded on a JASCO V-650 spectrometer in phosphate buffered aqueous solution (pH 7). Mass spectra were generated with a Bruker Esquire LC 6000 spectrometer in the electrospray ionization mass spectrometry (ESI-MS) mode with 50% water / 49.5% methanol / 0.5% NH_3 (pH 9–10) as matrix. Nuclear magnetic resonance (NMR) spectra were recorded with a 400 MHz Bruker Avance III HD and chemical shifts are expressed in parts per million (ppm). Chemical shifts were referenced to the DMSO solvent signal, 2.50 ppm for ^{1}H. 85% phosphoric acid was used as external standard for ^{31}P NMR spectra with 0 ppm. All ^{31}P NMR spectra were recorded with proton decoupling. VWR / Hitachi: LaChromElite L-2130 Pump; VWR / Hitachi: LaChromElite L-2420 UV/Vis detector**;** VWR / Hitachi: LaChromElite organizer, Agilent Technologies: OpenLAB Control Panel A.02.01.

3 mmol of cyanoethyl phosphoramidite 5′-DMTr-2′-TBDMS-3′-CEP-N^6-Bz-adenosine (ChemGenes, Wilmington, MA, USA, Cat. No. ANP-5671) were used as starting material for the synthesis of the protected dimeric linear precursor 5′-OH-2′-TBDMS-N^6-Bz-adenosine-(3′→5′)-cyanoethyl-phosphate-2′-TBDMS-3′-H-phosphonate-N^6-Bz-adenosine with a standard oligonucleotide coupling protocol, originally developed for cyclic dinucleotides (Gaffney et al., 2010). After preparative flash chromatography purification on silica gel with chloroform / methanol (1:1), the linear dimer was evaporated to dryness. 7.5 mmol (2.5 eq.) 5-ethylthio-tetrazole as coupling reagent were added and the resulting mixture was evaporated 4 times from 40 ml absolute acetonitrile. The last evaporation was stopped at ~15 ml total volume, 0.5 g molecular sieves 3 Å were added and the mixture was stirred at room temperature for 30 min. 6 mmol (2 eq.) cyanoethyl phosphoramidite 5′-DMTr-2′-CEP-3′-TBDMS-N6-Bz-adenosine (ChemGenes, Wilmington, MA, USA, Cat. No. ANP-5681) were added and the protected trimeric linear precursor 5′-OH-3′-TBDMS-N^6-Bz-adenosine-(2′→5′)-cyanoethyl-phosphate-2′-TBDMS-N^6-Bz-adenosine-(3′→5′)-cyanoethyl-phosphate-2′-TBDMS-3′-H-phosphonate-N^6-Bz-adenosine was prepared as previously described (Gaffney et al., 2010). After preparative flash chromatography purification on silica gel with chloroform / methanol (1:1), the linear trimer was evaporated to dryness. The final cyclization step and the release of protection groups was performed according to the standard protocol in Gaffney *et al*., leading to the raw product 2′3′3′-cAAA after evaporation of solvents.

100 ml water was added and the resulting suspension was placed in an ultrasonic bath at room temperature for 15 min, followed by 3 extraction cycles with 50 ml chloroform each. The combined organic phases were extracted with 50 ml water and the combined product-containing aqueous phase was filtered with a 0.45 µm regenerated cellulose (RC) filter and partially concentrated under reduced pressure to remove traces of chloroform. The complex product solution was diluted with water to 1000 ml and applied to a Q Sepharose Fast Flow anion exchange column (40–165 µm; 380 × 50 mm) Cl^--form (Sigma), previously regenerated with 2 M

sodium chloride and washed with water. The column was washed with water (1000 ml), followed by a gradient of 0–600 mM triethylammonium bicarbonate buffer (TEAB, pH 7, 7500 ml) in water, 500 ml 600 mM TEAB and 1500 ml 1 M TEAB (detection wavelength 254 nm). The title compound eluted with ~400 mM TEAB. Product-containing fractions were carefully concentrated to a final volume of approximately 20 ml with a rotary evaporator equipped with a drop catcher *in-vacuo*. Subsequent purification of 2′3′3′-cAAA were accomplished by repeated preparative reversed phase medium pressure liquid chromatography (MPLC). The product solution was applied to a LiChroprep RP-18 column (15–25 µm; 450 x 50 mm) (Merck) previously equilibrated with 100 mM triethylammonium formate (TEAF, pH 6.8) in water. Elution was performed with 100 mM TEAF, 20 mM TEAF, followed by a step-gradient of 1%, 2%, 3%, 4% and 10% 2-propanol, 20 mM TEAF (pH 6.8) in water. Final purification of product containing fractions was accomplished by a LiChroprep RP-18 column (15–25 µm; 430 x 25 mm) (Merck) with a similar step-gradient of 2-propanol, 20 mM TEAF as organic modifier. For desalting, 2′3′3′-cAAA fractions of sufficient purity were applied to an in-line set-up of two LiChroprep RP-18 columns (15–25 µm; 450 × 50 mm each) (Merck), previously equilibrated with water. The columns were washed with water to remove excess TEAF buffer. Afterwards, 2% 2-propanol in water was used to elute the desalted 2′3′3′-cAAA. To generate the sodium salt form of 2′3′3′-cAAA, pooled product-containing fractions were partially concentrated under reduced pressure and subsequently applied to a Toyopearl SP-650M cation exchange column (65 µm; 90 x 35 mm) Na^+-form (Sigma), previously regenerated with 2 M sodium chloride and washed with water. For elution the column was washed with water until no UV-absorbance was detectable at 254 nm anymore. After filtration and careful evaporation under reduced pressure, 845.6 µmol 2′3′3′-cAAA, sodium salt was isolated with a purity of 99.02% by HPLC (theoretical yield: 28.2%).

Formula (free acid): $C_{30}H_{36}N_{15}O_{18}P_3$ (MW 987,63 g/mol)

UV-Vis (water pH 7.0): λ_{max} 259 nm; ε 40500.

ESI-MS pos. mode: m/z 988 $(M+H)^+$, m/z 1010 $(M+Na)^+$, m/z 1032 $(M-H+2Na)^+$.

ESI-MS neg. mode: m/z 986 $(M-H)^-$, m/z 1008 $(M-2H+Na)^-$, m/z 1030 $(M-3H+2Na)^-$.

^{1}H NMR (400 MHz, Deuterium Oxide) δ 8.51 (s, 1H), 8.47 (s, 1H), 8.24 (s, 1H), 8.22 (s, 1H), 8.21 (s, 1H), 8.18 (s, 1H), 6.34 (d, *J* = 7.8 Hz, 1H), 6.20 (d, *J* = 7.0 Hz, 1H), 6.17 (d, *J* = 4.2 Hz, 1H), 5.25 - 5.18 (m, 1H), 4.89 – 4.71 (m, 7H), 4.58 - 4.54 (m, 2H), 4.37 – 4.20 (m, 6H) ppm.

^{31}P NMR (162 MHz, Deuterium Oxide) δ 0.25 (s, 1P), 0.02 (s, 1P), -0.29(s, 1P) ppm.

Analytical HPLC: (Kromasil 100-10, RP-8 (10 µm; 250 x 4 mm)) 3% acetonitrile, 50 mM sodium dihydrogen phosphate buffer, pH 6.8; 1.5 mL/min; UV 259 nm; t_{RET} 5.84 min.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

To characterize the enzymatic products of *Ab*CdnD, LC-MS/MS analysis was performed as previously described (Lau et al., 2020; Ye et al., 2020). LC-MS/MS was performed on a Thermo Vanquish UHPLC coupled to a Thermo QExactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific). Chromatography utilized a Sequant ZIC-pHILIC polymeric column (100 mm × 2.1 mm, 5 µm) (EMD Millipore) maintained at 45°C and a flow rate of 0.4 ml min^{-1}. *Ab*CdnD products and cyclic oligonucleotide standards were separated by injecting 2 µl of sample and eluting on the following linear gradient: (A) 20 mM ammonium bicarbonate in water, pH 9.6, and (B) acetonitrile; 90% B for 0.25 min, a linear gradient to 55% B at 4 min, and sustained until 6 min. Column was re-equilibrated for 2.5 min at 90% B. Detection of cyclic oligonucleotides was performed in positive ionization mode using an heated electrospray ionization (HESI) source with the following parameters: spray voltage of 3.5 kV; sheath gas, auxiliary gas, and sweep gas flow rates of 40, 20, and 2 AU, respectively; capillary and auxiliary gas heater temperature of 275 and 350°C, respectively. Profile MS1 spectra were acquired under the following parameters: mass resolution of 35,000, AGC volume of 1×10^6 maximum IT of 75 ms and a scan range from 450 to 1400 m/z to account for z = 1 and z = 2 ions of cyclic tri- and tetra- adenosine nucleotides. Data-dependent tandem mass spectra were acquired using CID of the following settings: mass resolution of 17,500, AGC volume of 1×10^5, maximum IT of 50 ms; a loop count of 5, isolation

window of 1.5 m/z; normalized collision energy of 25 eV; dynamic exclusion was not used. Data reported are for the most common ion for each indicated cyclic oligonucleotide.

Thermal shift assay

Proteins were mixed with 3× Sypro dye and 100 μM purified nucleotide ligand (3′3′3′-cAAG or IEX-purified *Ab*CnD product) in 20 mM HEPES-KOH pH 7.5 and 75 mM KCl, with a final protein concentration of 10 μM. Using a BioRad CFX96 thermocycler, samples were brought from 25 to 95°C, reading fluorescence in HEX channel every 0.5°C. The derivative of each curve over time was calculated using BioRad CFX Manager, then normalized as a percent maximum change in fluorescence for each sample.

Negative stain electron microscopy

0.05–0.1 mM purified Cap4 proteins were mixed with equimolar amounts of cyclic trinucleotide and DNA as indicated. For negative staining, all samples were adjusted to 0.0025–0.005 mg ml^{-1} in 20 mM HEPES-KOH pH 7.5, 150 mM KCl, 1 mM TCEP without or with 1 μM activating ligand. Negative stain grids were prepared by applying the sample (3 μl) to a glow-discharged (30 s, 30 mA) 400-mesh Cu grid (Electron Microscopy Sciences) coated with an ~15 nm layer of continuous carbon (Safematic CCU-010). After 30 s, the grid was blotted from the side, immediately stained with 1.5% uranyl formate, and blotted again. The staining procedure was repeated two times, with a 30 s incubation with uranyl formate before the final blotting step. The grid was allowed to dry for at least 15 min before imaging. Samples were imaged with a Tecnai T12 (FEI) transmission electron microscope operated at 120 keV and equipped with a Gatan 4K × 4K CCD camera. Images were collected at a nominal magnification of 67,000× and pixel size of 1.68 Å with defocus values of ~1.0–2.0 μm. Between 80 and 333 micrograph images were collected for each of the 9 datasets. Image processing was done in RELION-3.0 (Zivanov et al., 2018). After CTF estimation with GCTF (Zhang, 2016), particle picking was carried out with gautomatch (Kai Zhang,

https://www.mrc-lmb.cam.ac.uk/kzhang/) or LoG-based autopicking (Zivanov et al., 2018) and manually inspected. Particles were extracted with a box size of 156 pixels and subjected to reference-free 2D classification. Particles in the best-resolved classes were selected and subjected to an additional one to two rounds of 2D classification. In addition, for each dataset, random subsets of 10,000 particles were subjected to 2D classification to compare the distributions of particles in different image groups. For datasets with more than 10,000 particles, the standard error of mean of the distributions between four independent subsets of 10,000 particles ranged from 0.1 to 1.1%. Particles retained after 2D classification were used to generate a de novo initial model using the stochastic gradient descent algorithm in RELION, followed by 3D refinement and classification. Overall resolutions of EM maps of dimer complexes were ~15 Å. Rigid body docking was performed using Chimera (Pettersen et al., 2004).

SEC-MALS

Cap4 and Cap4–nucleotide second messenger complex samples were prepared by diluting in SEC-MALS running buffer (150 mM KCl, 20 mM HEPES-KOH pH 7.5 and 1 mM TCEP) to 2 mg ml^{-1} and incubating on ice for 5 min with 100 μM nucleotide second messenger and then separated on an SRT SEC-300 column (Sepax). Protein concentration was calculated using refractive index on a Wyatt Optilab T-rex Refractive Index Detector assuming dn dc^{-1} of 0.185 and a molar mass was calculated using a Wyatt Dawn Heleos II Multi-Angle Light Scattering detector and ASTRA software.

DNA EMSA assay

Catalytically inactive Cap4 proteins (*Ab*Cap4 K69A, *Ec*Cap4 K74A) were mixed on ice at a concentration of 10 μM with 1 μM 45 bp DNA and 50 μM 3′3′3′-cAAG (*Ec*Cap4) or IEX-purified *Ab*CdnD product (*Ab*Cap4) in a final reaction volume of 20 μl containing 20 mM HEPES-KOH pH 7.5, 50 mM KCl, 5 mM $MgCl_2$ and 1 mM TCEP. Reactions were incubated on ice for 5 min and

then supplemented with to a final concentration of 5% glycerol (v/v). Samples were separated on a 2% native agarose TB gel containing 100 mM Tris and 45 mM boric acid by running at 250 V for 45 min at 4°C. The gel was stained by soaking in buffer with 10 μg ml^{-1} ethidium bromide solution for 20 min and then imaged with ChemiDoc MP Imaging System.

DNA fragment sequencing

Cap4 proteins were incubated with 400 ng plasmid or genomic DNA in 80 μl reactions for 2 h (500 nM protein, 100 nM activating ligand, 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 1 mM TCEP, 5 mM $MgCl_2$), then DNA fragments were purified by phenol-chloroform extraction. Sequencing libraries were made from single stranded DNA fragments using an Accel-NGS 1S Plus DNA Library kit (Swift Biosciences) as previously described (Lau et al., 2020) and sequenced on a NextSeq500. For bioinformatics analyses, given that Cap4 digestion results in small fragments, we took advantage of that fact that reads containing the 3′ adapter sequence have the full fragment sequence. Reads containing the 3′ adapter were selected and adapter trimmed using Cutadapt (Martin, 2011). The 8 nucleotide low complexity sequence was trimmed by Cutadapt, and reads were mapped to the pGEM9z(-) plasmid (Promega) or the *E. coli* K12 genomic sequence (Genbank U00096.3) using Bowtie2 (Langmead and Salzberg, 2012). The location of the 5′ end of each read was used as the cut site and extracted from the output sam files. The 10 nucleotide sequence upstream and downstream of the cut sites were compiled and used to identify the consensus cut site using WebLogo3 (Crooks et al., 2004). The lengths of the mapped reads were extracted using custom python scripts and histogram plots were made using ggplot2 (Wickham, 2016).

DncV co-expression with diverse CBASS effectors

A plasmid expressing *dncV* (pAW1371-pBAD33-*dncV*) and a second plasmid expressing a predicted CBASS effector were electroporated into competent *Escherichia coli* BL21-DE3

(Invitrogen). Bacteria were recovered for 1 h shaking at 37°C in Super Optimal Broth with Catabolite repression (SOC), then plated onto selective LB Agar with glucose to repress *dncV* expression (LB: 10 g L^{-1} tryptone, 5 10 g L^{-1} yeast extract, 5 10 g L^{-1} NaCl, 20 µg ml^{-1} chloramphenicol, 100 µg ml^{-1} carbenicillin, 0.2% w/v glucose). Single colonies were inoculated into selective LB Agar plus glucose medium and cultured at 37°C shaking for ~16 h. 5 µl spots of 10-fold serial dilutions in LB were pipetted onto selective high-salt LB (20 µg ml^{-1} chloramphenicol, 100 µg ml^{-1} carbenicillin, NaCl concentration adjusted to 2% w/v) under inducing conditions (0.2% arabinose, 5 µM IPTG) or repressive conditions (0.2% glucose). Data are measured as CFU per ml and the ratio of inducing to repressive conditions and are the mean of three independent experiments. The predicted effector plasmids are pETSUMO2 expressing either: green fluorescent protein (*gfp*, negative control); *Vibrio dncV* native effector *capV* (WP_001133548.1, positive control); *Desulfotomaculum alkaliphilium* CD-NTase005 effector Saf2TM-SAVED (WP_031517735.1); *Escherichia coli* CD-NTase010 effector Saf2TM-SAVED (WP_001057904.1); *Acinetobacter baumannii* CD-NTase011 effector *Ab*Cap5 (WP_031984940.1); *Geobacillus sp.* CD-NTase012 effector *Gs*Cap5 (WP_013400843.1); *Myxococcus xanthus* CD-NTase022 effector caspase-SAVED (WP_020479061.1); *Acinetobacter baumannii* CD-NTase037 effector a.k.a. *Ab*CdnD effector *Ab*Cap4 (WP_008942236.1); *Enterobacter cloacae* CD-NTase038 effector a.k.a. *Ec*CdnD effector *Ec*Cap4 (WP_032676399.1); *Bradyrhizobium japonicum* CD-NTase039 effector *Bj*Cap5 (WP_011082906.1); *Burkholderia pseudomallei* CD-NTase041 a.k.a. *Bp*CdnG effector *Bp*Cap5 (WP_004556385.1); *Citrobacter freundii* CD-NTase042 effector highly similar to *Ec*CdnG effector *Cf*Cap5 (ETX65525.1); *Acinetobacter baumannii* CD-NTase043 effector *Ab*Cap5 (WP_000539314.1); *Pseudomonas aeruginosa* CD-NTase044 effector *Pa*Cap5 (WP_023082129.1); *Bacillus coagulans* CD-NTase046 effector *Bc*Cap5 (WP_013858316.1); *Vibrio cholerae* CD-NTase053 effector *Vc*Cap5 (WP_000259919.1).

Phage resistance assays

Electrocompetent *E. coli* MG1655 was electroporated with individual medium-copy plasmids encoding entire CBASS operons under their native promoters and analyzed similarly to previously described methods (Cohen et al., 2019; Doron et al., 2018). *E. coli* were recovered and plated on selective LB (carbenicillin 100 μg ml^{-1}). Single colonies were inoculated into selective MMCG medium (1× M9 Minimal Salts, 0.4% glucose, 0.02% $MgSO_4$, 0.001% $CaCl_2$, plus 100 μg ml^{-1} carbenicillin) and cultivated at 37°C shaking for ~20 h. Cultures were diluted 1:100 into selective MMCG medium and cultivated at 37°C shaking for 4 h to harvest mid-log cultures. Phage resistance was measured by a modified double-agar overlay technique. Bacteria were immobilized in soft-agar overlays by thoroughly mixing 400 μl of mid-log culture with 3.5 ml molten MMCG top agar (MMCG medium plus 0.35% agar, 5 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1 mM $MnCl_2$). Bacteria combined with top agar was immediately poured onto a 100 × 15 mm petri dish containing 20 ml solidified MMCG Agar (1.6%) and allowed to cool for 10 min at room temperature. A high-titer T2 phage lysate (Coli Genetic Stock Center CGSC12141), T5 lysate (CGSC12144), or T7 lysate (CGSC12146) prepared from MG1655 *E. coli* was 10-fold serially diluted into SM Buffer (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-HCl pH 7.5). 3 μl spots of dilutions were pipetted onto the solidified double-agar overlays. Spots were allowed to dry for 20 min at room temperature. Plates were incubated at 37°C for ~16 h and plaques were quantified to compare efficiency of plating. Data are the mean of at least three independent experiments. Plasmids expressing CBASS operons were pLOCO2-*dncV* operon (Genome AE003852.1, 178097-185003), pLOCO2-Ec-*cdnD02* operon (Locus JCKK01000002, 2261880-2268858), and pLOCO2-Ec-*cdnG* operon a.k.a. CD-NTase042 operon (Locus JSNY01000125, 10563-17743).

Chapter 2

Effector-mediated membrane disruption controls cell death in CBASS antiphage defense

Brianna Duncan-Lowey[1,2], Nora K. McNamara-Bordewick[2], Nitzan Tal[3], Rotem Sorek[3], Philip J. Kranzusch[1,2,4,5]

[1] Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

[2] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[3] Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

[4] Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA 02115, USA

[5] Lead Contact

Summary

Cyclic oligonucleotide-based antiphage signaling systems (CBASS) are antiviral defense operons that protect bacteria from phage replication. Here we discover a widespread class of CBASS transmembrane (TM) effector proteins that respond to antiviral nucleotide signals and limit phage propagation through direct membrane disruption. Crystal structures of the *Yersinia* TM effector Cap15 reveal a compact 8-stranded β-barrel scaffold that forms a novel cyclic dinucleotide receptor domain that oligomerizes upon activation. We demonstrate that activated Cap15 relocalizes throughout the cell and specifically induces rupture of the inner membrane. Screening for active effectors, we identify the function of distinct families of CBASS TM effectors and demonstrate that cell death via disruption of inner membrane integrity is a common mechanism of defense. Our results reveal the function of the most prominent class of effector protein in CBASS immunity and define disruption of the inner membrane as a widespread strategy of abortive infection in bacterial phage defense.

Introduction

In bacteria, cyclic oligonucleotide-based antiphage signaling systems (CBASS) are a major form of antiviral defense that limits replication of diverse phages (Cohen et al., 2019). Each CBASS operon contains a cGAS/DncV-like nucleotidyltransferase (CD-NTase) enzyme that functions to sense phage replication, and a partner CD-NTase-associated protein (Cap) effector that directly impairs host cell function to prevent viral spread (Lowey et al., 2020; Whiteley et al., 2019; Ye et al., 2020). Following recognition of phage infection, the CD-NTase synthesizes a nucleotide second messenger, which then directly binds the Cap effector and results in activation and induction of antiviral defense (Cohen et al., 2019; Lau et al., 2020; Lowey et al., 2020; Whiteley et al., 2019).

A critical step in CBASS immunity is the ability of the Cap effector to respond to a specific CD-NTase nucleotide second messenger signal (Govande et al., 2021; Lowey et al., 2020). Bacterial CD-NTase enzymes produce a diverse array of nucleotide signals with distinct base and phosphodiester linkage specificity including cyclic dinucleotide (*e.g.* 3′3′ cyclic GMP–AMP, 3′3′-cGAMP) and cyclic trinucleotide (*e.g.* 3′3′ cyclic AMP–AMP–GMP, 3′3′3′-cAAG) products (Lowey et al., 2020; Whiteley et al., 2019). Recent structures of CBASS effectors in complex with nucleotide second messengers explain how correct signal recognition controls protein function (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020). The structure of *A. baumannii* Cap4 demonstrates that a protein domain named SAVED is responsible for nucleotide recognition and subsequent protein oligomerization to control effector activation (Lowey et al., 2020). Interestingly, the Cap4 SAVED domain is a structural homolog of CRISPR-associated Rossman fold (CARF)-family proteins, revealing structural and functional overlap between components of CBASS and CRISPR immune systems (Jia et al., 2019; Lowey et al., 2020; Niewoehner and Jinek, 2016). Likewise, discovery of a bacterial domain homologous to the human cyclic dinucleotide binding protein Stimulator of Interferon Genes (STING) within CBASS effectors Cap12 and Cap13 demonstrates an evolutionary link between CBASS nucleotide-second messenger binding and

components of animal innate immunity (Morehouse et al., 2020). These CBASS effectors containing STING and SAVED domains specifically bind nucleotide second messengers, then oligomerize when activated. The majority of CBASS operons, however, contain effectors with no characterized nucleotide-binding domain, suggesting that major forms of nucleotide second messenger recognition remain to be discovered.

Following nucleotide second messenger recognition, Cap effectors induce cell death to kill the host bacteria and block phage replication through a form of antiviral defense termed abortive infection. Nearly all characterized abortive infection defense systems function through effectors with enzymatic domains that degrade or modify target host or phage proteins (Lopatina et al., 2020). Several CBASS enzymatic effectors downstream of CD-NTases have been mechanistically characterized, including phospholipases (CapV), DNA endonucleases (Cap4, Cap5, NucC), and NADases (Cap12) (Cohen et al., 2019; Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Severin et al., 2018). The nuclease effectors Cap4, Cap5, and NucC indiscriminately degrade dsDNA upon activation (Lau et al., 2020; Lowey et al., 2020), while the phospholipase effector CapV cleaves membrane phospholipids leading to cell death (Cohen et al., 2019; Severin et al., 2018). However, the majority of CBASS operons do not contain enzymatic effectors and instead encode uncharacterized proteins with predicted transmembrane (TM) segments (Burroughs et al., 2015; Millman et al., 2020). While it has been hypothesized that these effectors lead to membrane disruption, it is unknown how TM effectors function to restrict phage replication.

Here, we discover that CBASS TM effectors are potent antiphage defense proteins that destroy bacterial inner membrane integrity and induce host cell death. Through development of a screen to identify active TM effectors, we define a new family of CBASS proteins named Cap15 that respond to cyclic dinucleotide signals and block phage replication. High-resolution crystal structures of the Cap15 nucleotide binding domain reveal a compact β-barrel scaffold with a central pocket for ligand recognition, and oligomerization interfaces essential for phage defense.

Mechanistically, we show that Cap15 oligomerization and activation disrupts bacterial cell integrity and leads to specific disruption of the inner membrane. We further demonstrate that inner membrane disruption is a widespread mechanism shared between diverse families of CBASS TM effectors that are encoded in both gram-negative and gram-positive bacteria. Our results demonstrate the function of a major class of CBASS effectors controlling antiviral defense and define the host membrane as a common target of bacterial abortive infection defense systems.

Results

Discovery of CBASS transmembrane effectors cause cell death

Building upon previous bioinformatic analysis (Burroughs et al., 2015; Cohen et al., 2019; Millman et al., 2020; Whiteley et al., 2019), we analyzed putative effector proteins in CBASS operons and identified a diverse set of 2,234 proteins containing transmembrane segments. TM effectors are encoded in >40% of CBASS operons, representing a dominant form of effector that occurs more frequently than previously characterized phosphodiesterase, nuclease, or NADase proteins with enzymatic function (Figure 2.1A) (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Severin et al., 2018). The most common TM effector contains two transmembrane segments fused to an uncharacterized β-strand-rich C-terminal domain, and we named this effector CD-NTase associated protein 15 (Cap15) (Figures 2.1B,C). To determine if Cap15 is capable of responding to CBASS cyclic dinucleotide signaling and inducing cell toxicity in bacteria, we leveraged the known ability of the CD-NTase *Vibrio cholerae* DncV (*Vc*DncV; NCBI ref. WP_001901330.1) to be auto-activated as a promiscuous 3′3′-cGAMP synthase during over-expression in *E. coli* (Whiteley et al., 2019; Zhou et al., 2018). Expression of *Escherichia albertii* Cap15 (*Ea*Cap15, NCBI ref. WP_206748793.1) in the presence of active 3′3′-cGAMP signaling prevents culture growth (Figure 2.1D) and induces cell death as confirmed by flow cytometry (Figure 2.2D). A *Vc*DncV D131A, D133A substitution that disrupts CD-NTase catalytic function rescues cell viability, demonstrating that *Ea*Cap15 only induces cell death in the presence of active 3′3′-cGAMP nucleotide second messenger signaling (Figure 2.1D).

We next screened a panel of 23 CBASS TM effectors and identified a wide diversity of proteins capable of inducing cell toxicity specifically in response to 3′3′-cGAMP (Figure 2.1E). Toxic TM effector proteins belong to several families including Cap14 proteins (TM-SAVED, pfam PF18145) and Cap16 proteins (TM-NUDIX, pfam PF18167) (Figure 2.1E, Figure 2.2A). TM effectors are encoded throughout the CD-NTase family tree (Whiteley et al., 2019), suggesting

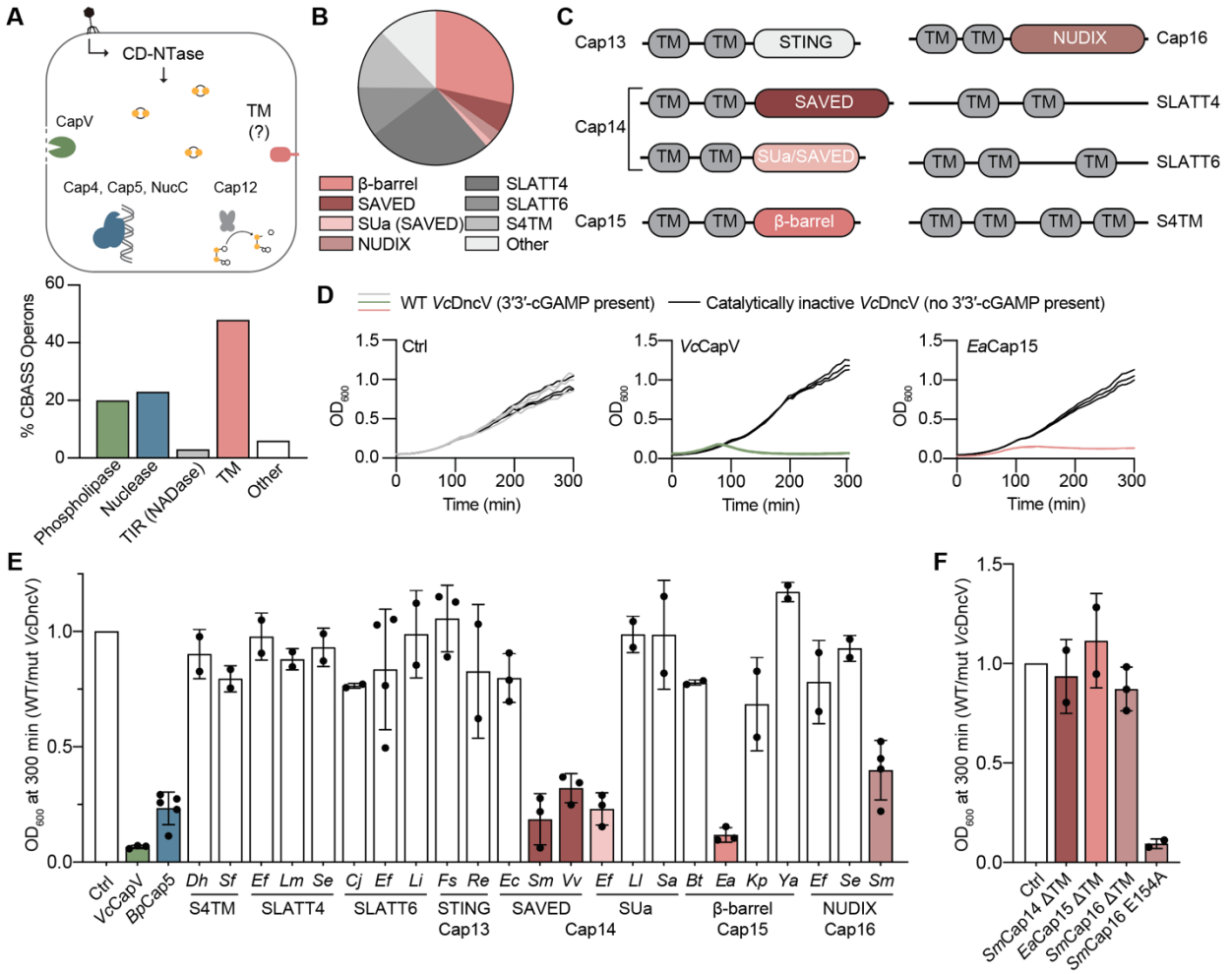
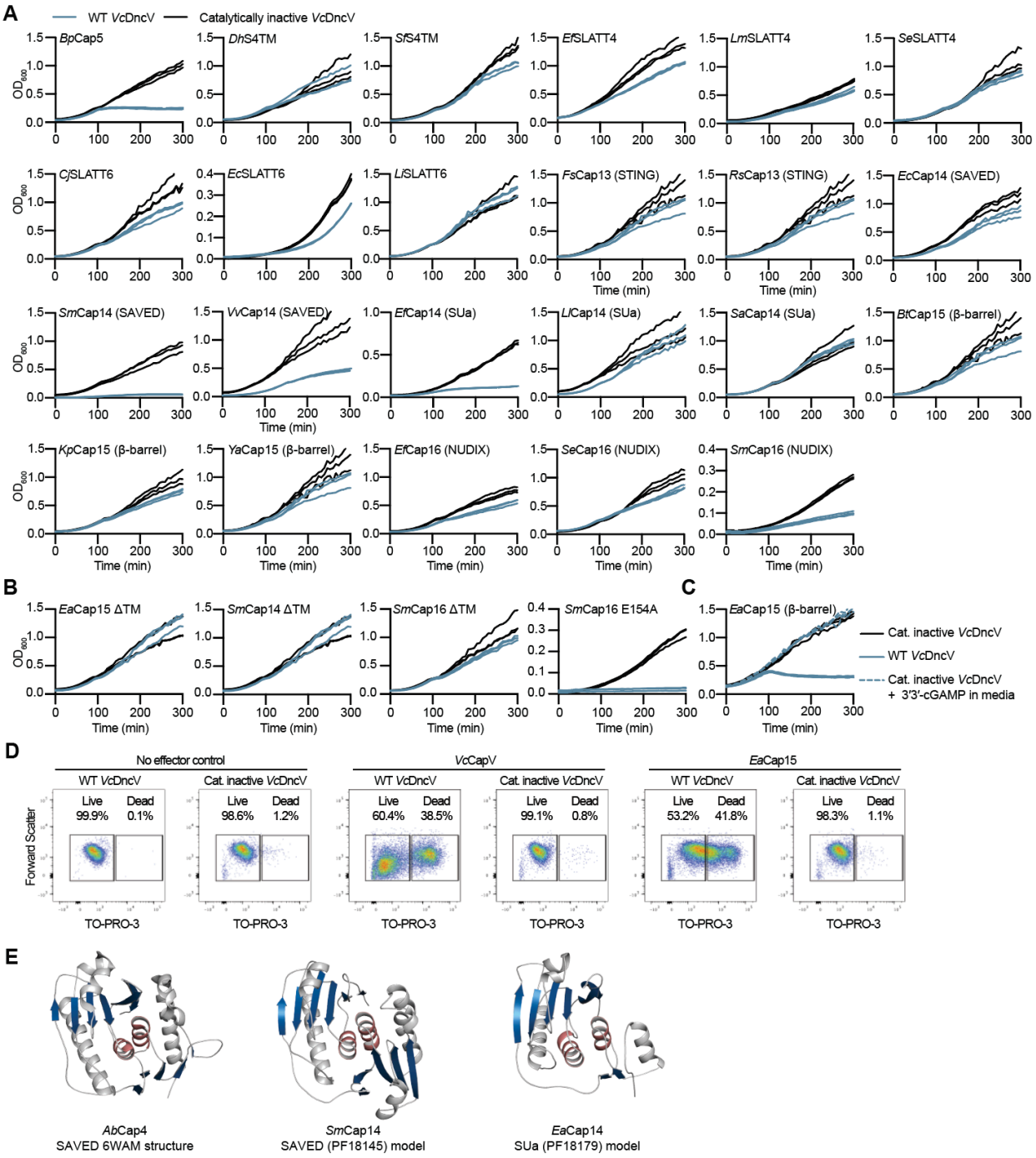


Figure 2.1. Discovery of CBASS transmembrane effectors that control cell death
(A) Schematic showing CBASS effector types and abundance in sequenced genomes. Transmembrane (TM) effectors are encoded in >40% of CBASS operons. (B) Quantification of individual families of TM effectors in CBASS operons. (C) Domain organization of common CBASS TM effector families. Structure prediction analysis demonstrates that SUa (pfam PF18179) is a variant of the SAVED nucleotide-binding domain See also Figure 2.2. (D) Growth curves of *E. coli* expressing a CBASS effector protein and the 3′3′-cGAMP synthase *Vibrio cholerae* DncV (*Vc*DncV). *Vc*DncV is known to catalyze active 3′3′-cGAMP signaling upon overexpression in *E. coli* (Whiteley et al., 2019; Zhou et al., 2018). Cells expressing WT *Vc*DncV are indicated in color and catalytically inactive *Vc*DncV in gray. 3′3′-cGAMP-responsive effectors induce cell death upon co-expression with active *Vc*DncV. (E) Quantification of growth curves and cell death induced by 3′3′-cGAMP-responsive effector proteins. Data are displayed as the difference in OD_{600} of cultures expressing WT vs. catalytically inactive *Vc*DncV 300 min after induction, relative to a control culture with no effector expression. Many families of TM-containing effectors cause cell death in response to 3′3′-cGAMP production. (F) Quantification of cell death induced by full-length or truncated (ΔTM) TM effectors where the TM domain was replaced by SUMO to ensure solubility. Cap16 contains a putative NUDIX hydrolase domain and was additionally tested with a NUDIX inactivating E154A mutation. TM-containing effectors require the appended transmembrane segments to induce cell death. Toxicity data are representative of at least 2 independent experiments.

Figure 2.2. Diverse transmembrane effectors cause cell death upon activation with 3′3′-cGAMP

(A) Growth curves of *E. coli* expressing WT (blue) or catalytically inactive (gray) 3′3′-cGAMP synthase *Vc*DncV and an effector protein, as summarized in Figure 2.1E. (B) Growth curves of *E. coli* expressing WT (blue) or catalytically inactive (gray) 3′3′-cGAMP synthase *Vc*DncV and an effector protein either truncated to remove the transmembrane domain (ΔTM) or with an inactivating mutation to the putative NUDIX active site (E154A), as summarized in Figure 2.1F. Toxicity data are representative of at least 2 independent experiments. (C) Growth curves of *E. coli* expressing WT (blue) or catalytically inactive (gray) 3′3′-cGAMP synthase *Vc*DncV and *Ea*Cap15, with or without 3′3′-cGAMP added to the media. (D) Flow cytometry of *E. coli* expressing *Vc*DncV and effector proteins as in (A). Bacteria were stained with live-dead stain TO-PRO-3. (E) Comparison of SAVED domains in Cap4 and Cap14 effectors. Crystal structure of SAVED domain from Cap4 (PDB 6WAM) (Lowey et al., 2020) and models of SAVED domains from *Sm*Cap14 (with Pfam domain annotation PF18145) and *Ea*Cap14 (with Pfam domain annotation PF18179), which both encode SAVED domains. Conserved central helices are highlighted in pink and bracing sheets in blue.


A
WT VcDncV
Catalytically inactive VcDncV
BpCap5
DhS4TM
SfS4TM
EfSLATT4
LmSLATT4
SeSLATT4
CjSLATT6
EcSLATT6
LiSLATT6
FsCap13 (STING)
RsCap13 (STING)
EcCap14 (SAVED)
SmCap14 (SAVED)
VvCap14 (SAVED)
EfCap14 (SUa)
LlCap14 (SUa)
SaCap14 (SUa)
BtCap15 (β-barrel)
KpCap15 (β-barrel)
YaCap15 (β-barrel)
EfCap16 (NUDIX)
SeCap16 (NUDIX)
SmCap16 (NUDIX)
OD600
Time (min)
B
EaCap15 ΔTM
SmCap14 ΔTM
SmCap16 ΔTM
SmCap16 E154A
C
EaCap15 (β-barrel)
Cat. inactive VcDncV
WT VcDncV
Cat. inactive VcDncV
+ 3′3′-cGAMP in media
D
No effector control
VcCapV
EaCap15
WT VcDncV
Cat. inactive VcDncV
Live
Dead
99.9%
0.1%
98.6%
1.2%
60.4%
38.5%
99.1%
0.8%
53.2%
41.8%
98.3%
1.1%
Forward Scatter
TO-PRO-3
E
AbCap4
SAVED 6WAM structure
SmCap14
SAVED (PF18145) model
EaCap14
SUa (PF18179) model

that the effectors not responsive to 3′3′-cGAMP in our screen likely induce cell death but respond to other CD-NTase products including pyrimidine-containing cyclic dinucleotide and cyclic trinucleotide signals (Lowey et al., 2020; Whiteley et al., 2019). In each case, cell toxicity is strictly dependent on the presence of transmembrane segments revealing that membrane association is essential for TM effector function (Figure 2.1F, Figure 2.2B). The only TM effector with a putative enzymatic domain is Cap16, which contains a NUDIX hydrolase domain (Figures 2.1B,C). Mutation of the putative Cap16 active site (E154A) does not impact cell toxicity (Figure 2.1F, Figure 2.2B), further confirming that TM effector function is dependent on transmembrane segments and not catalytic activity. These data define Cap15 and diverse TM effectors in CBASS immunity that respond to antiviral nucleotide second messenger signaling and induce potent cell toxicity.

CBASS Cap15 transmembrane effectors protect bacteria from phage infection and encode a minimal β-barrel domain

To determine the role of TM effectors in antiviral defense, we expressed complete Cap15-containing CBASS operons in *E. coli* and challenged the bacteria with a panel of 10 diverse phages. CBASS operons containing Cap15 effectors potently restrict viral plaque formation (Figure 2.3A,B). The *Escherichia* CBASS (*Ea*CdnB [NCBI ref. WP_000995828.1] and *Ea*Cap15 [NCBI ref. WP_206748793.1]), which we identified as a 3′3′-cGAMP signaling system (Figure 2.1D), protects *E. coli* against infection with T2, T5, and T6. Interestingly, a related *Yersinia* CBASS operon which encodes homologous CD-NTase and Cap15 components (*Ya*CdnE [NCBI ref. WP_145567548.1] and *Ya*Cap15 [NCBI ref. WP_145567547.1]) protects *E. coli* against all tested *Myoviridae* (T2, T4, T6), *Siphoviridae* (T5, SECphi4, SECphi6, SECphi18, SECphi27), and the ssDNA phage SECphi17, but not the *Podoviridae* phage T7, demonstrating a variable range of phage defense between CBASS operons encoding the same class of effector (Figure 2.3A).

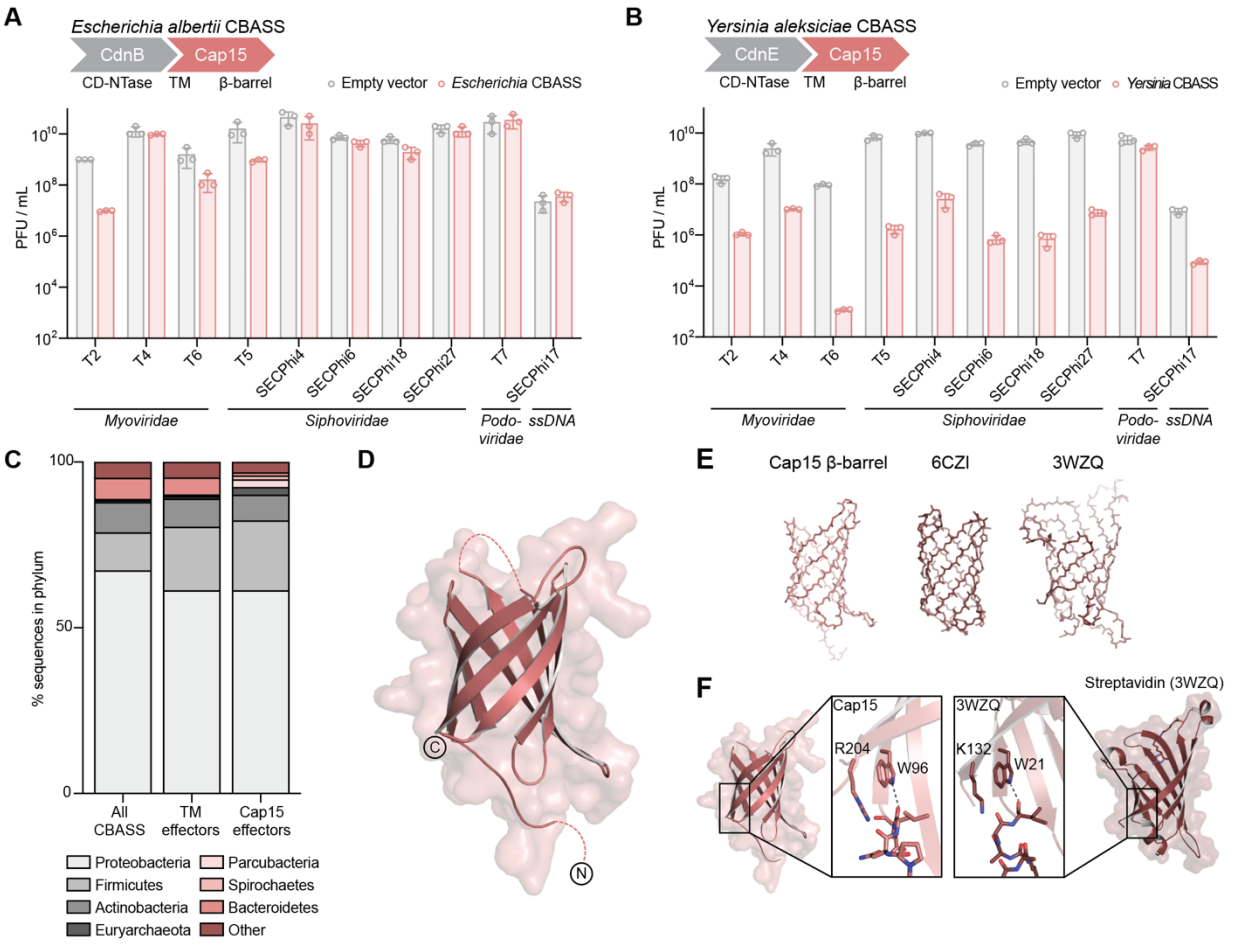


Figure 2.3. Cap15 protects bacteria from phage infection and encodes a minimal β-barrel nucleotide-binding domain

(A, B) Phage challenge of *E. coli* expressing *Escherichia* or *Yersinia* CBASS operons. CBASS operons containing Cap15 β-barrel TM effectors protect against infection with diverse phages. Cartoon schematic shows organization of the CBASS operons encoding CD-NTases from clade B (CdnB) or E (CdnE) and the Cap15 β-barrel-containing TM effector. (C) Phylogenetic analysis of the distribution of Cap15 β-barrel effectors and other CBASS TM effectors across bacterial phyla demonstrates widespread occurrence in both gram-negative and gram-positive bacteria. (D) Crystal structure of the *Yersinia aleksiciae* Cap15 (*Ya*Cap15) β-barrel nucleotide-binding domain. The *Ya*Cap15 structure is minimized and formed by eight β-strands. (E) Comparison of Cap15 to an *in silico* designed small molecule sensor (6CZI) (Dou et al., 2018) and the biotin-binding protein streptavidin (3WZQ) (Kawato et al., 2015). (F) The Cap15 β-barrel domain contains a common tryptophan-corner capping motif at the bottom of the β-barrel similar to streptavidin proteins. A capping helix common to many β-barrels is replaced by a capping β-strand in Cap15. Phage defense data are representative of at least 3 independent experiments.

We analyzed CBASS Cap15-containing operons and observed that they are widely distributed in both gram-negative and gram-positive bacteria (Figure 2.3C). CBASS Cap15-containing operons protect against a wide diversity of phages, including both dsDNA and ssDNA phage, suggesting the viral signal that activates these CBASS operons is broadly conserved. In each case, disruption of the CD-NTase active site or deletion of the Cap15 transmembrane region abolished all phage protection, demonstrating that Cap15-mediated antiviral defense is strictly dependent upon functional CD-NTase signaling and interaction with the bacterial membrane (Figure 2.4B,C).

Each Cap15 protein is comprised of N-terminal transmembrane helices fused to an uncharacterized C-terminal domain (Figure 2.1C). To define the molecular basis of Cap15 effector function, we determined a 1.9 Å crystal structure of the *Ya*Cap15 C-terminal domain (Figure 2.3D, Table 2.1). The *Ya*Cap15 structure reveals eight β-strands (β2–β9) that wrap and form a compact β-barrel (Figure 2.3D). Strikingly, the closest structural homologs to Cap15 are not natural proteins but instead are computationally derived β-barrel proteins designed to bind fluorescent small molecules (Figure 2.3E) (Dou et al., 2018). Similar to the rationally designed proteins, in Cap15 a set of 9 glycine residues distributed throughout the 8 β-strands are positioned to increase strand curvature, reduce side chain density within the barrel center, and enable folding of the minimized β-barrel structure (Figure 2.4D,E). Cap15 is further stabilized by a conserved "tryptophan corner" interaction previously observed in other β-barrel structures (Dou et al., 2018). In *Ya*Cap15 this interaction occurs where W96 on the N-terminal strand β2 stacks against R204 on the C-terminal strand β9 and forms a hydrogen bond with the peptide carboxyl on strand β1 below the β-barrel (Figure 2.3F). Together, these data define the TM effector Cap15 as a widespread β-barrel domain-containing protein that enables CBASS anti-phage defense.

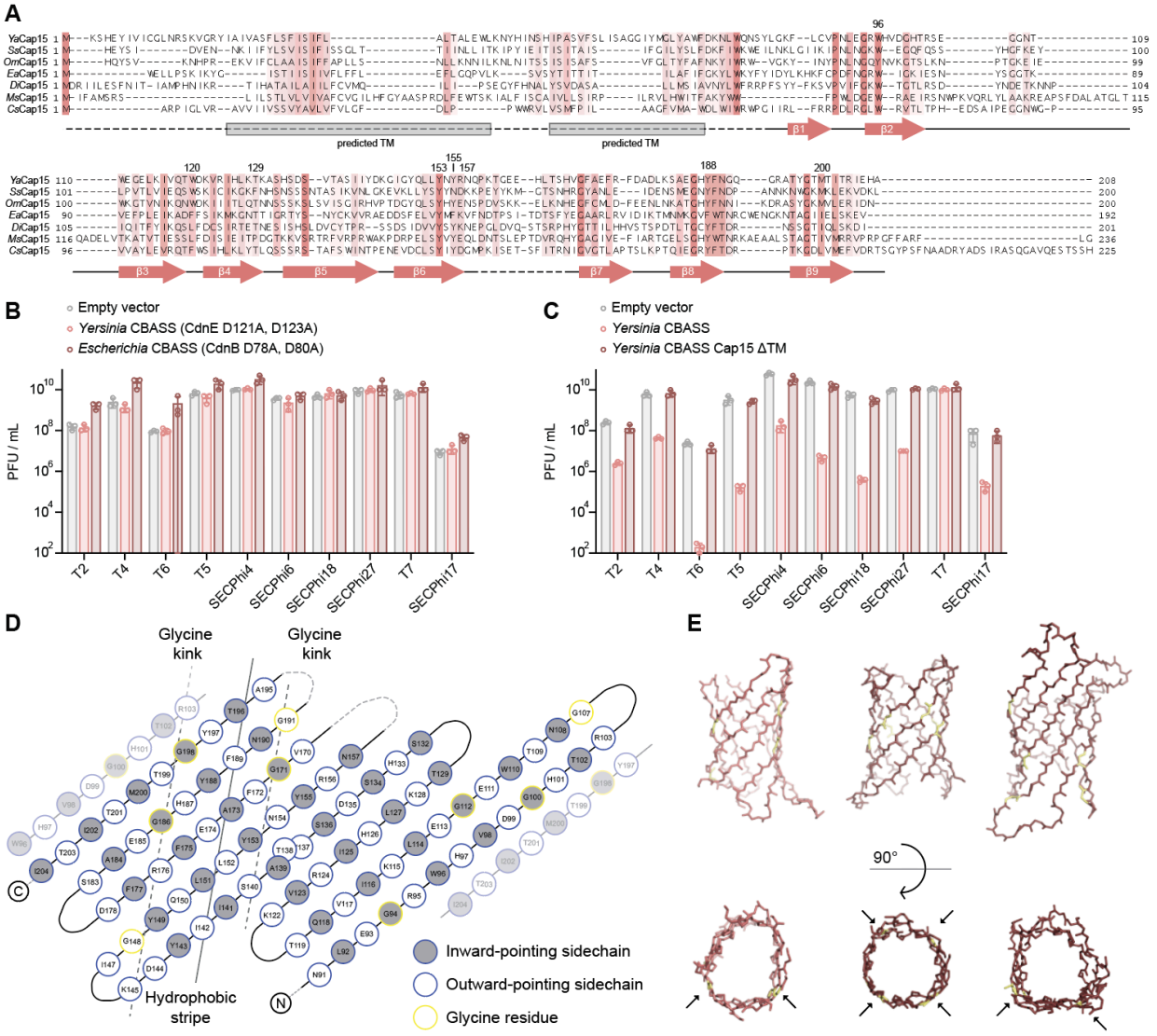


Figure 2.4. Cap15 contains a minimal β-barrel domain.
(A) Alignment of divergent Cap15 homologs annotated with the determined secondary structure of *Ya*Cap15. *Ya*Cap15 recognizes UMP-containing cyclic dinucleotides; *Ea*Cap15 recognizes 3′3′-cGAMP. (B) Phage challenge of *E. coli* expressing the *Escherichia* CBASS operon with catalytically inactive CdnB (D78A, D80A) or the *Yersinia* CBASS operon with catalytically inactive CdnE (D121A, D123A). Phage defense requires functional CD-NTase activity (empty vector controls shown in Figure 2.3B are included again for reference). (C) Phage challenge of *E. coli* expressing the *Yersinia* CBASS operon with Cap15 ΔTM construct. Phage defense requires Cap15 transmembrane domains. (D) Schematic representation of Cap15 β-barrel showing alternative inward- and outward-pointing amino acid side chains, highlighting glycine residues (yellow) that replace “inward-pointing” side chains, decreasing side chain density in the center of the β-barrel. Glycine kinks are shown as dashed lines. A stripe of hydrophobic residues that point outward from the barrel is highlighted with a solid line. (E) Cartoon representation of the peptide backbone of Cap15, a rationally designed β-barrel (6CZI), and streptavidin (3WZQ) with glycine residues within strands highlighted in yellow and the “corners” enabled by the glycine kinks highlighted by arrows.

Table 2.1. Crystallographic statistics

	*Ya*Cap15 (SeMet Phasing)	*Ya*Cap15 (SeMet Refinement)	*Ya*Cap15
Data Collection			
Resolution (Å)[a]	37.09–2.30 (2.38–2.30)	37.01–1.90 (1.94–1.90)	44.05–2.60 (2.72–2.60)
Wavelength (Å)	0.97918	0.97918	0.97918
Space group	C 1 2 1	C 1 2 1	P 6_1
Unit cell: a, b, c (Å)	93.51, 31.20, 50.04	93.29, 31.04, 50.02	50.87, 50.87, 155.02
Unit cell: α, β, γ (°)	90.0, 100.33, 90.0	90.0, 100.22, 90.0	90.0, 90.0, 120.0
Molecules per ASU	1	1	2
Total reflections	487276	77567	111881
Unique reflections	6554	11344	7040
Completeness (%)[a]	99.8 (98.1)	99.8 (98.9)	100.0 (100.0)
Multiplicity[a]	74.3 (71.4)	6.9 (6.3)	15.9 (15.2)
$I/\sigma I$[a]	25.8 (6.1)	13.3 (1.3)	19.3 (1.3)
CC(1/2)[b] (%)[a]	99.9 (43.2)	99.9 (56.2)	100.0 (51.2)
Rpim[c] (%)[a]	2.7 (75.0)	2.9 (64.3)	1.8 (67.4)
Sites	1		
Refinement			
Resolution (Å)		37.0–1.90	44.05–2.60
Free reflections		1134	679
R-factor / R-free		21.5 / 25.4	29.7 / 31.3
Bond distance (RMS Å)		0.006	0.002
Bond angles (RMS °)		0.752	0.491
Structure/Stereochemistry			
No. atoms: protein		869	1687
No. atoms: solvent		24	0
Average B-factor: protein		54.89	108.24
Average B-factor: water		54.08	N/A
Ramachandran plot: favored		99.02%	95.41%
Ramachandran plot: allowed		0.98%	4.59%
Ramachandran plot: outliers		0.00%	0.00%
Rotamer outliers		1.11%	4.55%
MolProbity[d] score		1.12	2.20
Protein Data Bank ID		7N34	7N35

[a] Highest resolution shell values in parenthesis
[b] (Karplus and Diederichs, 2012)
[c] (Weiss, 2001)
[d] (Chen et al., 2010)

Cap15 β-barrel nucleotide-binding domain specifically binds uracil-containing cyclic dinucleotides

The Cap15 β-barrel domain creates a solvent-exposed pocket at the top of the protein that is in the same location as the small-molecule binding site of rationally designed β-barrels (Figure 2.5A) (Dou et al., 2018). Additionally, Cap15 exhibits more distantly related structural homology to streptavidin proteins that form a similar pocket at the top of a β-barrel domain to create a binding site for biotin (Figure 2.6A). We therefore hypothesized that the function of the Cap15 β-barrel domain is to recognize the CD-NTase product nucleotide second messenger. To test this hypothesis, we purified the *Yersinia aleksiciae* CD-NTase (*Ya*CdnE) encoded adjacent to *Ya*Cap15 and determined the nucleotide second messenger product for this CBASS operon using ^{32}P-labeled nucleotides and thin-layer chromatography. *Ya*CdnE synthesizes a mixture of di-pyrimidine and purine–pyrimidine cyclic dinucleotide products with a strong preference for incorporation of UTP (Figure 2.5B). HPLC (Figure 2.5C, Figure 2.6B) and mass spectrometry analysis of the purified reaction products demonstrate that the most abundant products synthesized by *Ya*CdnE in the presence of ATP, GTP, CTP, and UTP are 3′3′-c-di-UMP (*m/z* = 613.058) and 3′3′-cUMP–AMP (*m/z* = 636.085).

We next analyzed the ability of the *Ya*Cap15 C-terminal β-barrel domain to bind radiolabeled *Ya*CdnE cyclic dinucleotide products and observed specific recognition of 3′3′-c-di-UMP and 3′3′-cUMP–AMP (Figure 2.5D,E, Figure 2.6C,D). *Ya*Cap15 binds 3′3′-c-di-UMP and 3′3′-cUMP–AMP with ~100 nM affinity and exhibits no ability to interact with the control cyclic dinucleotide 3′3′-c-di-AMP (Figure 2.5D,E, Figure 2.6C,D). Additionally, we observed that Cap15 binding to any UMP-containing cyclic dinucleotide results in a dramatic increase in the thermostability of the complex, further supporting a direct role for the Cap15 C-terminal β-barrel domain in sensing the CD-NTase cyclic dinucleotide antiviral signal (Figure 2.5F, Figure 2.6E).

Analysis of the Cap15 solvent exposed pocket at the top of the β-barrel reveals highly conserved residues surrounding the putative nucleotide-binding site (Figure 2.5G,H, Figure 2.4A).

Figure 2.5. Cap15 β-barrel nucleotide-binding domain specifically binds UMP-containing cyclic dinucleotides

(A) Cutaway slice through the center of Cap15 (left) and the rationally designed small molecule sensor 6CZI (right). Protein surface is represented in mesh, and the small molecule bound to 6CZI is shown in sticks. A solid, hydrophobic core stabilizes the bottom of each β-barrel creating a solvent-exposed pocket at the top lined with hydrophilic amino acids for ligand recognition. (B) Analysis of *Ya*CdnE nucleotide second messenger synthesis. *Ya*CdnE was incubated with α-^{32}P-NTPs, and reactions were phosphatase treated and separated by thin layer chromatography. (C) HPLC analysis of *Ya*CdnE products after separation by ion exchange chromatography. *Ya*CdnE products separated as two major peaks from ion exchange and were analyzed separately by HPLC (top) compared to cyclic dinucleotide standards (bottom). *Ya*CdnE synthesizes 3′3′-c-di-UMP, 3′3′-cUMP–AMP, and 3′3′-cUMP–CMP as major products. See also Figure 2.6B (D) Electrophoretic mobility shift assay measurement of *Ya*Cap15–3′3′-cUMP–AMP complex formation. A titration of *Ya*Cap15 (5 nM – 5 μM) was incubated with 10 nM ^{32}P-labeled 3′3′-cUMP–AMP, and bound complexes were resolved by nondenaturing polyacrylamide gel electrophoresis. (E) Quantification of *Ya*Cap15–cyclic dinucleotide complex formation as shown in (E). See also Figure 2.6C. (F) Thermal denaturation assay to quantify stabilization of the *Ya*Cap15 β-barrel upon binding to UMP-containing cyclic dinucleotides. Cyclic dinucleotide recognition leads an ~25°C increase in *Ya*Cap15 β-barrel melting temperature demonstrating significant stabilization. (G) Cartoon highlighting conserved residues lining the Cap15 β-barrel solvent-exposed nucleotide-binding pocket. (H) Conservation of select residues lining the Cap15 nucleotide-binding pocket. *Ya*Cap15 binds UMP-containing cyclic dinucleotides; *Ea*Cap15 is activated by 3′3′-cGAMP. (I) Quantification of 3′3′-c-di-UMP binding assays using *Ya*Cap15 proteins with point mutations within the nucleotide-binding pocket verifies the importance of individual contacts for nucleotide second messenger recognition. *Ya*Cap15 mutants were incubated at 100 nM or 1 μM concentration with 3′3′-c-di-UMP and complex formation was quantified as in (F). See also Figure 2.6F. (J) Phage challenge of *E. coli* expressing *Yersinia* CBASS with mutations to the Cap15 nucleotide-binding pocket (empty vector controls shown in Figure 2.3B are included again for reference). Mutations that inhibit nucleotide-binding abolish protection from phage infection. Biochemical and phage defense data are representative of at least 3 independent experiments.

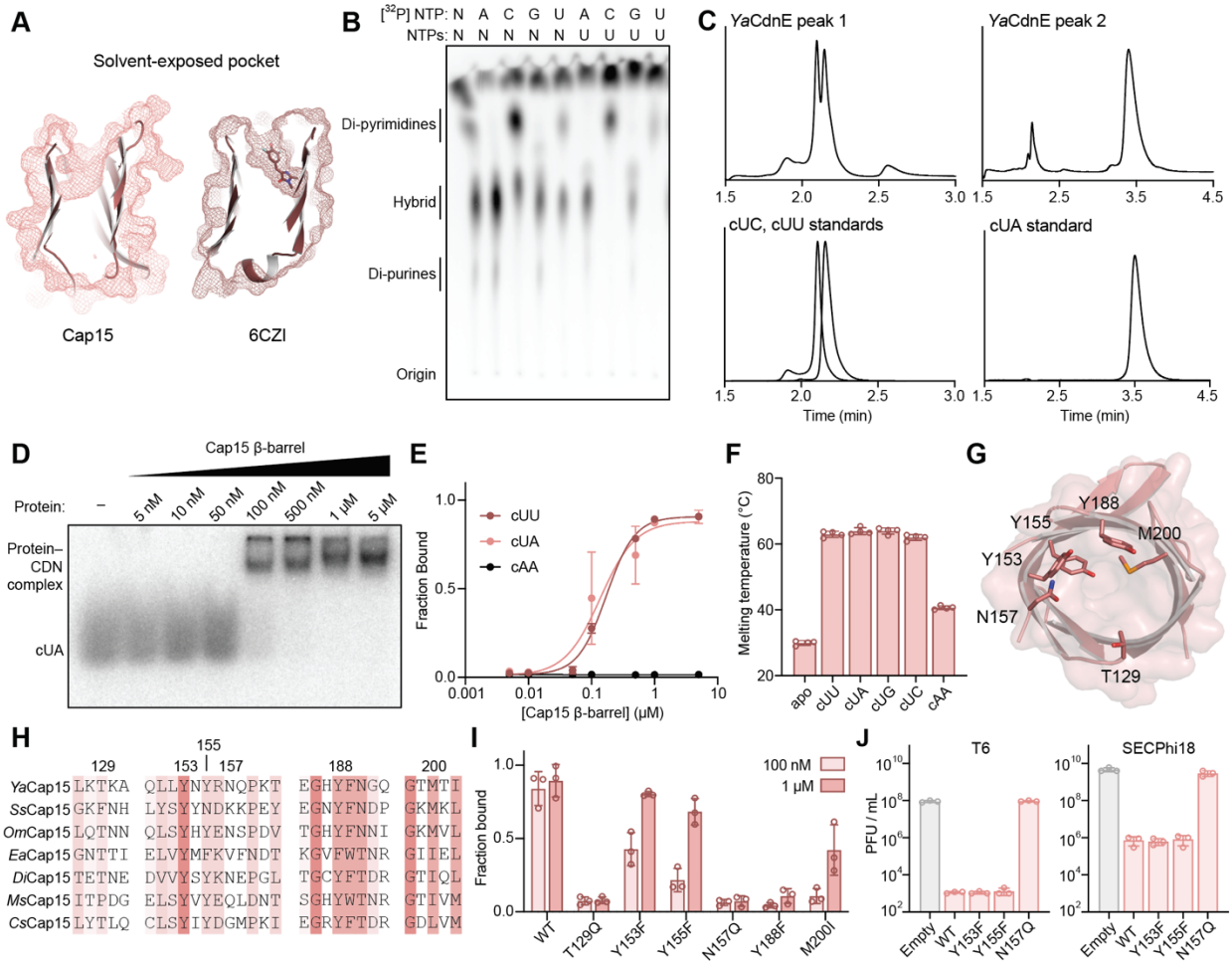
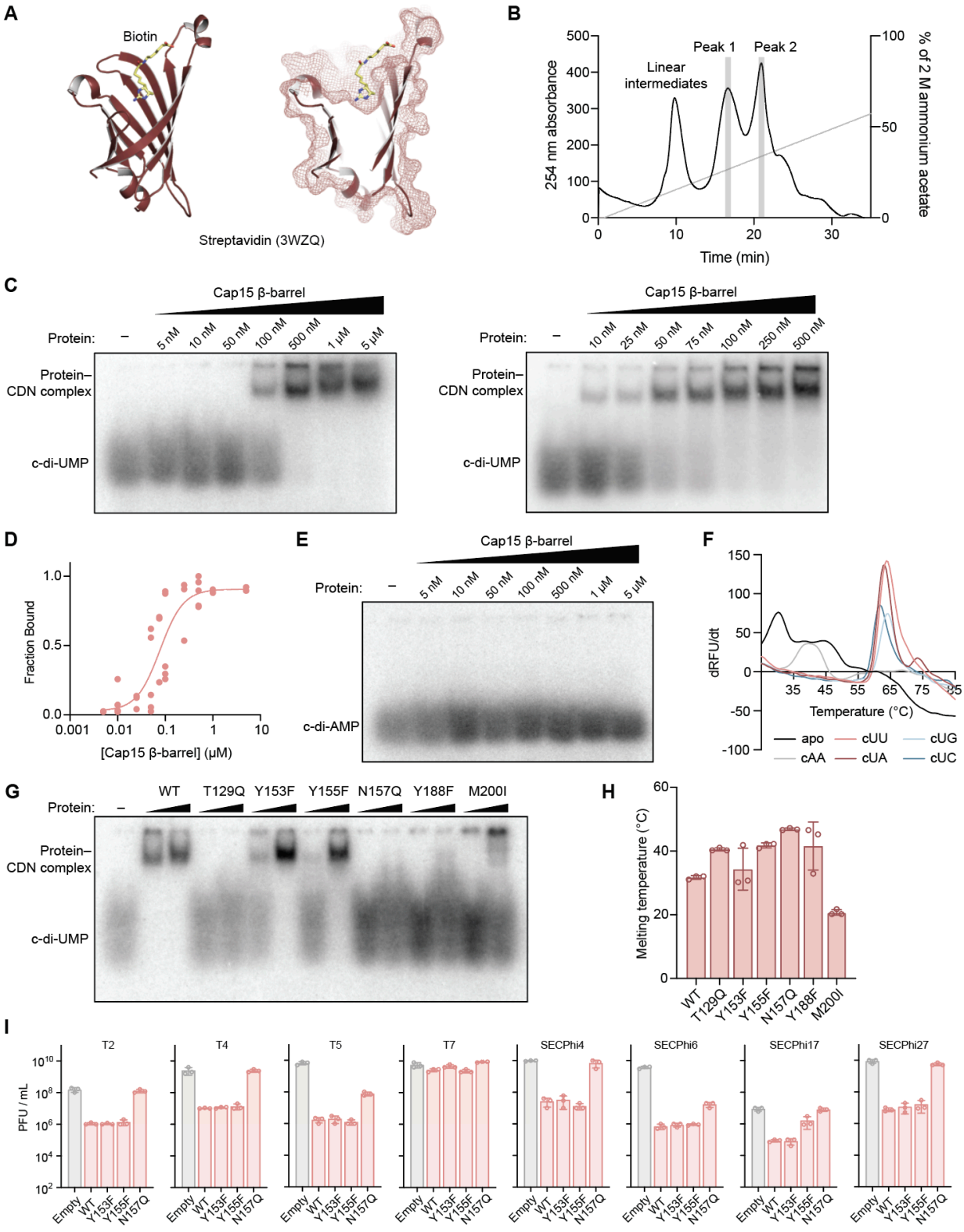

A
Solvent-exposed pocket
Cap15
6CZI
B
[32P] NTP: N A C G U A C G U
NTPs: N N N N N U U U U
Di-pyrimidines
Hybrid
Di-purines
Origin
C
YaCdnE peak 1
YaCdnE peak 2
cUC, cUU standards
cUA standard
Time (min)
Time (min)
D
Cap15 β-barrel
Protein: – 5 nM 10 nM 50 nM 100 nM 500 nM 1 μM 5 μM
Protein–CDN complex
cUA
E
cUU
cUA
cAA
Fraction Bound
[Cap15 β-barrel] (μM)
F
Melting temperature (°C)
apo cUU cUA cUG cUC cAA
G
Y188
Y155
M200
Y153
N157
T129
H
155
129 153 157 188 200
YaCap15 LKTKA QLLYNYRNQPKT EGHYFNGQ GTMTI
SsCap15 GKFNH LYSYYNDKKPEY EGNYFNDP GKMKL
OmCap15 LQTNN QLSYHYENSPDV TGHYFNNI GKMVL
EaCap15 GNTTI ELVYMFKVFNDT KGVFWTNR GIIEL
DiCap15 TETNE DVVYSYKNEPGL TGCYFTDR GTIQL
MsCap15 ITPDG ELSYVYEQLDNT SGHYWTNR GTIVM
CsCap15 LYTLQ CLSYIYDGMPKI EGRYFTDR GDLVM
I
100 nM
1 μM
Fraction bound
WT T129Q Y153F Y155F N157Q Y188F M200I
J
T6
SECPhi18
PFU / mL
Empty WT Y153F Y155F N157Q

Figure 2.6. Cap15 β-barrel binds UMP-containing cyclic dinucleotides.
(A) Streptavidin (3WZQ) bound to biotin. Streptavidin binds biotin within a shallow pocket at the top of the β-barrel. (B) Ion exchange chromatography separation of *Ya*CdnE products. Peak 1 and peak 2 are compared to cyclic dinucleotide standards in Figure 2.5C. (C) Electrophoretic mobility shift assay (EMSA) measuring Cap15 β-barrel complex formation with ^{32}P-labeled UMP-containing cyclic dinucleotides, as quantified in Figure 2.5E and Figure 2.6D. (D) Quantification of Cap15–cyclic c-di-UMP complex formation as shown in (C). (E) Electrophoretic mobility shift assay (EMSA) measuring Cap15 β-barrel complex formation with ^{32}P-labeled 3′3′-c-di-AMP, as quantified in Figure 2.5E. (F) Thermal denaturation assay to quantify stabilization of the Cap15 β-barrel upon binding to UMP-containing cyclic di-nucleotides, as summarized in Figure 3G. (G) Electrophoretic mobility shift assay (EMSA) measuring Cap15 β-barrel to ^{32}P-labeled 3′3′-c-di-UMP, as quantified in Figure 3I. Biochemical data are representative of at least 2 independent experiments. (H) Thermal denaturation assay to quantify the stability of Cap15 β-barrel mutants. (I) Phage challenge of *E. coli* expressing the *Yersinia* CBASS operon with *Ya*Cap15 mutations in the putative nucleotide-binding pocket (empty vector controls shown in Figure 2.3B are included again for reference). Phage defense data are representative of at least 3 independent experiments.


A
Biotin
Streptavidin (3WZQ)
B
Peak 1
Peak 2
Linear intermediates
254 nm absorbance
% of 2 M ammonium acetate
Time (min)
C
Cap15 β-barrel
Protein:
Protein–CDN complex
c-di-UMP
D
Fraction Bound
[Cap15 β-barrel] (μM)
E
Cap15 β-barrel
c-di-AMP
F
dRFU/dt
Temperature (°C)
apo
cUU
cUG
cAA
cUA
cUC
G
WT
T129Q
Y153F
Y155F
N157Q
Y188F
M200I
H
Melting temperature (°C)
I
T2
T4
T5
T7
SECPhi4
SECPhi6
SECPhi17
SECPhi27
PFU / mL
Empty

To confirm the role of this Cap15 pocket in mediating cyclic dinucleotide recognition, we next mutated each conserved residue to eliminate potential hydrogen-bonding interactions (Y153F, Y155F, Y188F, M200I) or insert bulkier side chains to limit access to the binding site (T129Q, N157Q). Substitutions to the conserved *Ya*Cap15 pocket residues each reduce 3′3′-c-di-UMP binding or abrogate all ligand recognition *in vitro* (Figure 2.5I, Figure 2.6F). We confirmed that all mutations, with the exception of M200I, did not negatively impact protein stability (Figure 2.6G). We next determined whether mutations that disrupt Cap15 ligand-binding prevent CBASS anti-phage defense. We observed that a *Ya*Cap15 N157Q mutation predicted to occlude the nucleotide-binding pocket results in complete loss of phage protection while mutations that only mildly decrease ligand-binding (Y153F, Y155F) do not affect phage defense (Figure 2.5I,J, Figure 2.6H). Together, these results demonstrate the discovery of a minimal β-barrel as a new class of cyclic dinucleotide binding domain and reveal the mechanistic basis of Cap15 cyclic dinucleotide recognition in CBASS defense.

Cap15 effector function requires protein oligomerization

Following nucleotide second messenger recognition, soluble CBASS effectors including Cap4, Cap12, and NucC are known to activate by oligomerizing into higher-order protein complexes (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020). To determine if protein oligomerization also has a role in activation of CBASS TM effectors, we next examined packing within the *Ya*Cap15 crystal for potential oligomeric interfaces. In the *Ya*Cap15 crystal, strand β1 extends away from the folded domain and forms a tight anti-parallel β-sheet interaction with a partnering strand β1 on the neighboring *Ya*Cap15 protomer (Figure 2.7A). This interface (interface 1) between neighboring protomers is further stabilized by a stacking interaction between the W120 side-chain of each protein (Figure 2.7A). We identified an additional oligomerization

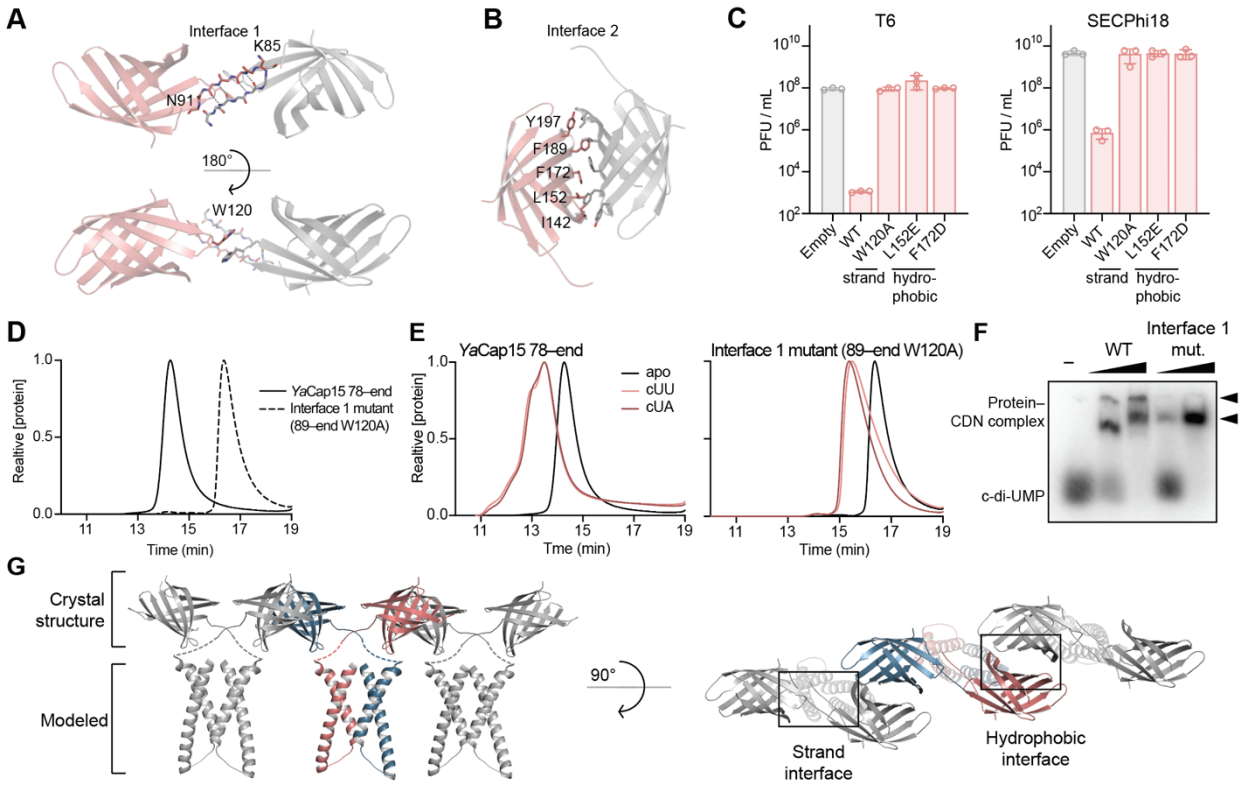
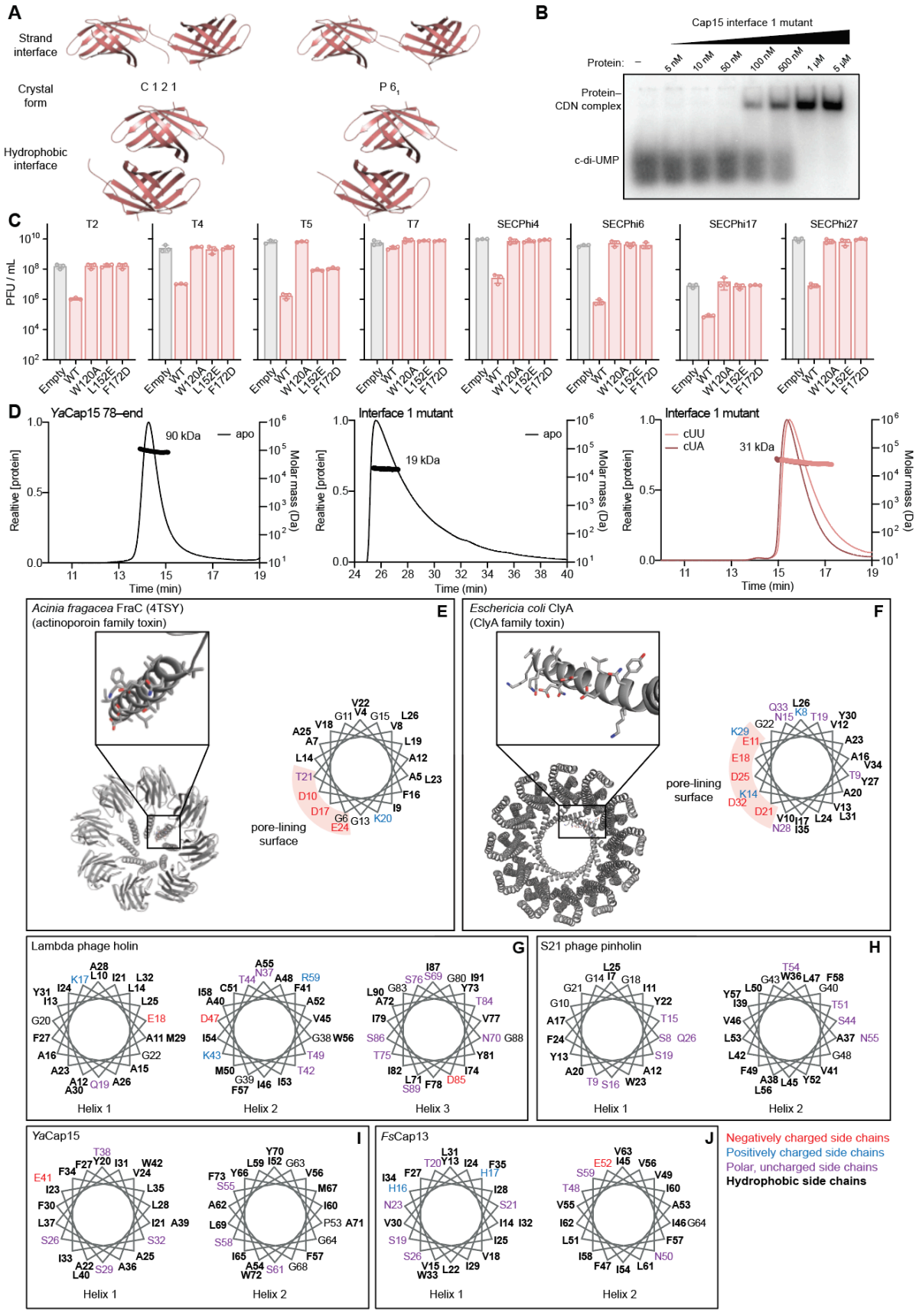


Figure 2.7. Cap15 oligomerizes upon ligand binding
(A,B) Analysis of packing with symmetry mates in the Cap15 crystal structure defines a β-strand oligomerization interface (Interface 1) and a hydrophobic packing interface (Interface 2).(C) Phage challenge of *E. coli* expressing *Yersinia* CBASS with mutations to the oligomerization interfaces (empty vector controls shown in Figure 2.3B are included again for reference). (D) Size exclusion chromatography demonstrating oligomerization of WT *Ya*Cap15, which is disrupted by mutations to *Ya*Cap15 Interface 1 (89–end, W120A). (E) Size exclusion chromatography demonstrating that recognition of activating cyclic dinucleotides results in *Ya*Cap15 higher-order complex formation (left). Mutations to *Ya*Cap15 Interface 1 prevent full higher-order complex formation but do not prevent cyclic dinucleotide recognition (right). (F) Electrophoretic mobility shift assay measurement of WT or "interface 1" mutant Cap15 binding to 3′3′-c-di-UMP. 100 nM or 1 μM of indicated protein was incubated with 10 nM ^{32}P-labeled 3′3′-c-di-UMP, and bound complexes were resolved by nondenaturing polyacrylamide gel electrophoresis. Arrows indicate the two complexes formed by Cap15 after ligand binding. The "interface 1" mutant disrupts formation of the higher-order *Ya*Cap15 complex (top arrow). (G) Structural model of Cap15 oligomerization. The experimentally determined structure of the Cap15 β-barrel is shown with modeled transmembrane domains. Both interface 1 and interface 2 are compatible with clustering of Cap15 along a two-dimensional membrane surface.

Figure 2.8. Analysis of Cap15 oligomerization interfaces and transmembrane domains.
(A) Cartoon schematic showing conserved packing interfaces between Cap15 β-barrels in two unique crystal forms. The β-strand interface (top) and hydrophobic interface (bottom) are the only interfaces in common among the two uniquely packed crystal forms. (B) Electrophoretic mobility shift assay (EMSA) measuring interface 1 mutant Cap15 β-barrel complex formation with ^{32}P-labeled 3′3′-c-di-UMP. (C) Phage challenge of *E. coli* expressing the *Yersinia* CBASS operon with *Ya*Cap15 mutations in the oligomerization interfaces (empty vector controls shown in Figure 2.3B are included again for reference). (D) SEC-MALS quantification of YaCap15 β-barrels shown in Figure 2.7D,E (E–H) Analysis of transmembrane domains of soluble proteins including (E) FraC and (F) ClyA proteins that form pores upon activation, and proteins that accumulate in raft-like assemblies in the membrane and then form lesions upon activation including phage (G) holin and (H) pinholin proteins. (I,J) Analysis of the transmembrane domains of the CBASS TM effectors Cap15 and Cap13. CBASS effector TMs do not contain a clear hydrophilic face, inconsistent with known requirements for pore formation. Biochemistry data are representative of at least 2 independent experiments and phage defense data are representative of independent experiments.


A
Strand interface
Crystal form
C 1 2 1
P 6$_1$
Hydrophobic interface
B
Cap15 interface 1 mutant
Protein:
–
5 nM
10 nM
50 nM
100 nM
500 nM
1 µM
5 µM
Protein–CDN complex
c-di-UMP
C
T2
T4
T5
T7
SECPhi4
SECPhi6
SECPhi17
SECPhi27
PFU / mL
Empty
WT
W120A
L152E
F172D
D
YaCap15 78–end
90 kDa
apo
Interface 1 mutant
19 kDa
31 kDa
cUU
cUA
Realtive [protein]
Molar mass (Da)
Time (min)
E
Acinia fragacea FraC (4TSY)
(actinoporoin family toxin)
pore-lining surface
F
Escherichia coli ClyA
(ClyA family toxin)
pore-lining surface
G
Lambda phage holin
Helix 1
Helix 2
Helix 3
H
S21 phage pinholin
Helix 1
Helix 2
I
YaCap15
Helix 1
Helix 2
J
FsCap13
Helix 1
Helix 2
Negatively charged side chains
Positively charged side chains
Polar, uncharged side chains
Hydrophobic side chains

interface (interface 2) between adjacent *Ya*Cap15 protomers mediated by a stripe of hydrophobic residues I142, L152, F172, F189, and Y197 facing out of the barrel (Figure 2.7B, Figure 2.3D). We determined a second 2.6 Å structure of *Ya*Cap15 and verified preservation of identical interface 1 and interface 2 contacts within this distinct crystal form (Figure 2.8A, Table 2.1). To determine the importance of the observed contacts, we designed mutations that block oligomerization and tested the impact *in vivo* during phage challenge. Mutations that disrupt interface 1 (W120A) or interface 2 (L152E, F172D) completely abolish phage defense (Figure 2.7C, Figure 2.8C), demonstrating that the *Ya*Cap15 oligomerization interfaces are essential for effector function.

We next measured *Ya*Cap15 oligomerization in solution using size exclusion chromatography with multi-angle ligand scattering (SEC-MALS). Wildtype *Ya*Cap15 migrates as

We next measured *Ya*Cap15 oligomerization in solution using size exclusion chromatography with multi-angle ligand scattering (SEC-MALS). Wildtype *Ya*Cap15 (15 kDa) migrates as an oligomeric complex with a molecular weight of 91 kDa, consistent with a hexameric assembly (Figure 2.7D, Figure 2.8D). Interestingly, in the presence of 3′3′-c-di-UMP or 3′3′-cUMP–AMP, the migration profile of *Ya*Cap15 dramatically shifts to a mixture of higher-order species indicating that CD-NTase signal recognition triggers further protein oligomerization (Figure 2.7E). *Ya*Cap15 forms a heterogenous population of species after binding to activating cyclic dinucleotides, and it was therefore it not possible to define a discrete molecular weight for the higher-order complex.

To determine if *Ya*Cap15 oligomerization is required for nucleotide second messenger recognition, we purified a *Ya*Cap15 β-barrel domain variant where oligomerization interface 1 is disrupted by a W120A mutation and deletion of strand β1 (*Ya*Cap15 89–end W120A, 14 kDa). In contrast to wildtype *Ya*Cap15, *Ya*Cap15 89–end W120A migrates as a monomer with a calculated molecule weight of 19 kDa (Figure 2.7D, Figure 2.8D). In the presence of 3′3′-c-di-UMP or 3′3′-cUMP–AMP, *Ya*Cap15 89–end W120A undergoes a shift to form a dimer (calculated mass of 31 kDa) but fails to form the fully oligomerized assemblies observed with the wildtype protein (Figure

2.7E, Figure 2.8D). Using radiolabeled cyclic dinucleotides, we confirmed that *Ya*Cap15 89–end W120A retains the ability to bind 3′3′-c-di-UMP, demonstrating that full oligomerization is not required for cyclic dinucleotide recognition (Figure 2.7E, Figure 2.8B). Similar to previous analysis of Cap12 (Morehouse et al., 2020), we further observed that the *Ya*Cap15 oligomerization-disrupting mutations prevent formation of larger complexes that migrate higher in the gel well during electrophoretic mobility shift assay experiments (Figure 2.7F, Figure 2.8B).

Using the experimentally defined *Ya*Cap15 oligomerization interfaces, we constructed a model of full-length *Ya*Cap15 to predict orientation of the N-terminal transmembrane helices (Figure 2.7G). Notably, the experimentally defined *Ya*Cap15 oligomerization interfaces are compatible with oligomerization along the two-dimensional surface of the bacterial cell membrane. These data support a mechanism where Cap15 exists as a monomer or lower-order oligomer, and subsequently oligomerizes within the membrane after binding to the activating nucleotide second messenger. Together, these results confirm the importance of *Ya*Cap15 oligomerization interfaces *in vivo* and demonstrate that effector oligomerization is feature conserved between divergent CBASS systems with cytosolic or membrane-bound effector proteins.

CBASS TM effectors cause membrane disruption and cell death

Given the requirement of both the Cap15 transmembrane domains and Cap15 protein oligomerization to induce cell death (Figure 2.1F; Figure 2.7D), we hypothesized that TM effectors may inhibit phage replication through direct targeting of the cell membrane after activation. To understand how cell death is induced by Cap15 activation, we examined *E. coli* expressing activated *Ea*Cap15 using phase contrast and fluorescence microscopy. Following induction of *Vc*DncV 3′3′-cGAMP signaling, *E. coli* expressing *Ea*Cap15 exhibit a dramatic shrinking of the inner membrane that results in separation from the cell wall and expansion of the periplasmic space (Figure 2.9A). Activation of *Ea*Cap15 induced no observable defect in the outer membrane (Figure 2.10A). Bacteria expressing *Ea*Cap15 with catalytically inactive *Vc*DncV, or *Vc*DncV

Figure 2.9. Cap15 activation causes membrane disruption and cell death
(A) Phase contrast image of *E. coli* expressing *Ea*Cap15 effector with WT or catalytically inactive 3′3′-cGAMP synthase *Vc*DncV. Activation of *Ea*Cap15 by 3′3′-cGAMP leads to inner membrane disruption. (B) Phase contrast image of *E. coli* expressing GFP control, *Ea*Cap15, or *Vc*CapV effector with WT or catalytically inactive 3′3′-cGAMP synthase *Vc*DncV. Activation of the Cap15 TM effector or CapV phospholipase effector each induce a similar membrane defect (arrows). (C) Phase contrast and fluorescence microscopy of *E. coli* expressing WT or mutant *Vc*DncV and *Ea*Cap15 or an empty vector control. A periplasmic mCherry marker was co-expressed to allow visualization of expansion of the periplasm (arrows) during Cap15 activation. (D) *E. coli* expressing GFP or GFP-labeled *Ea*Cap15 with WT or catalytically inactive *Vc*DncV. In the inactive state, *Ea*Cap15-GFP localizes to the poles of the bacterial cell. 3′3′-cGAMP signaling induces Cap15 activation and redistribution throughout the cell inner membrane. (E) Time-lapse images of *E. coli* in (C). *Ea*Cap15-GFP redistributes throughout the bacterial cell upon activation and moves with the inner membrane upon collapse (arrows). (F) Time-lapse images of *E. coli* expressing WT *Vc*DncV and *Ea*Cap15 in the presence of the live-dead stain TO-PRO-3 and periplasmic mCherry. After inner membrane collapse, the cells return to phase dark, are permeable to TO-PRO-3, and leak mCherry throughout the cell, confirming inner membrane rupture and cell death. (G) Schematic representation of membrane morphology observed upon Cap15 activation, with membranes shown in black, peptidoglycan shown in gray, and periplasm shown in red. (H) Transmission electron microscopy of *E. coli* dying after *Ea*Cap15 activation, highlighting collapse of the inner membrane from the cell wall. Light microscopy images are representative of at least three independent experiments.

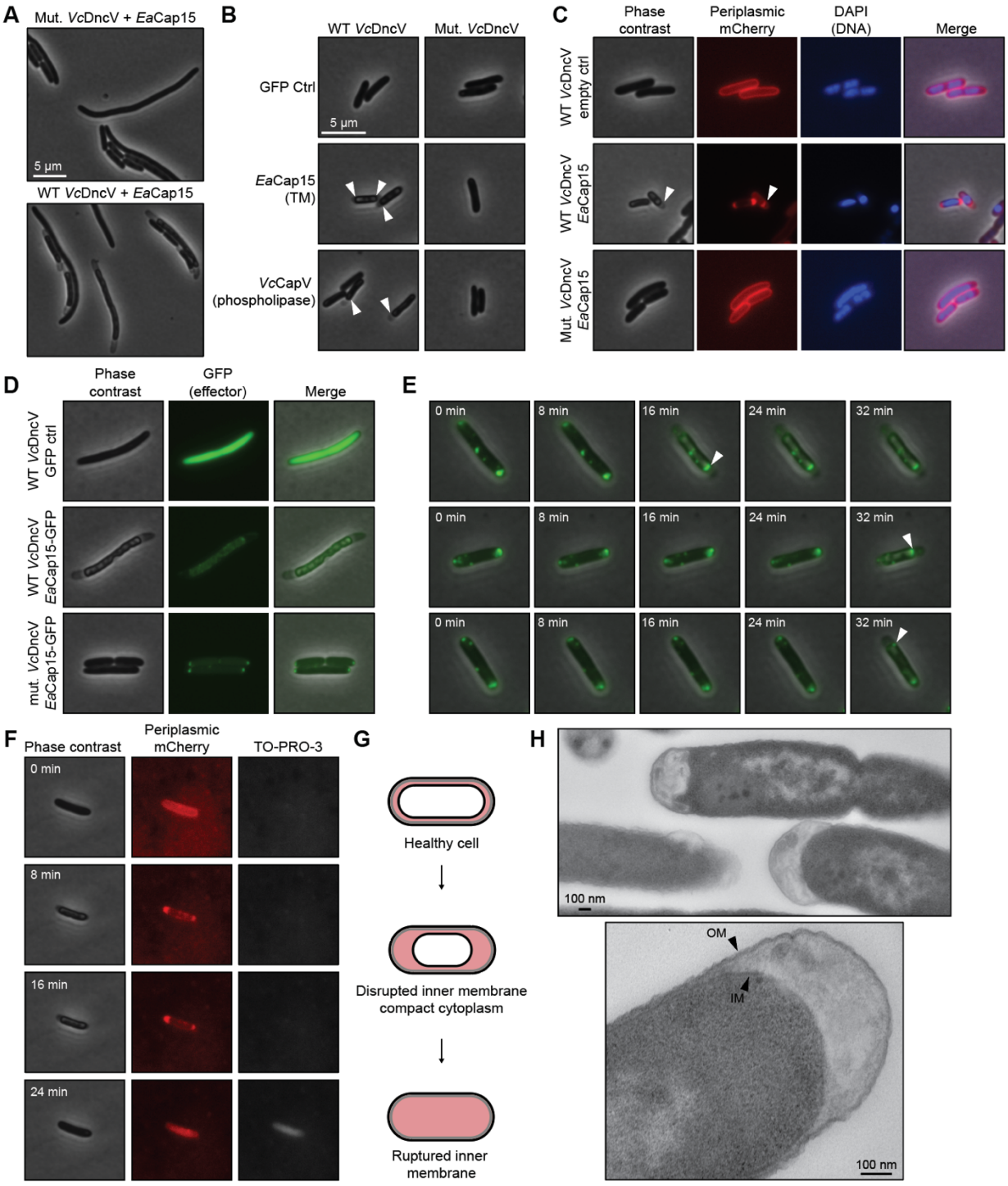
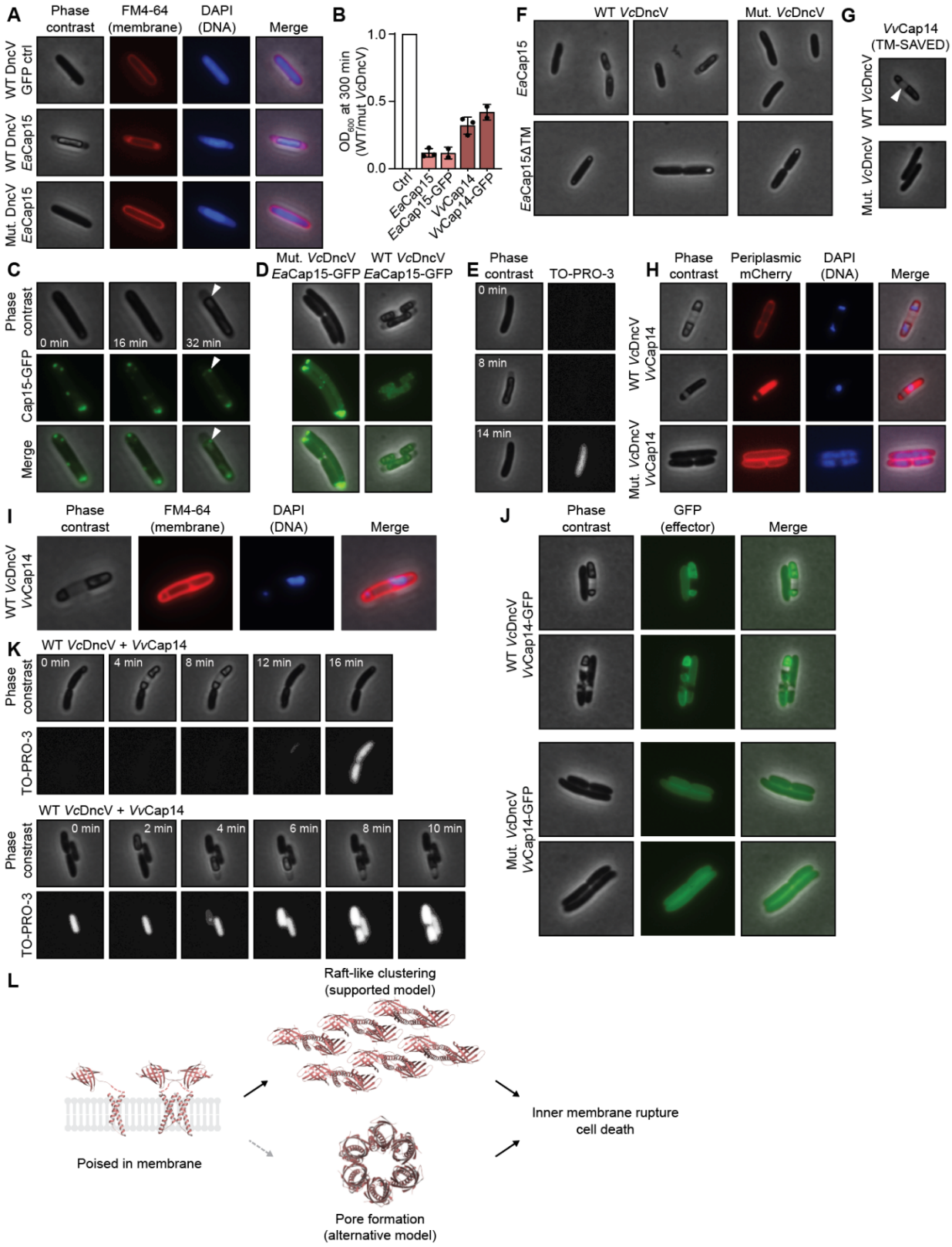

A
Mut. VcDncV + EaCap15
5 µm
WT VcDncV + EaCap15
B
WT VcDncV
Mut. VcDncV
GFP Ctrl
5 µm
EaCap15 (TM)
VcCapV (phospholipase)
C
Phase contrast
Periplasmic mCherry
DAPI (DNA)
Merge
WT VcDncV empty ctrl
WT VcDncV EaCap15
Mut. VcDncV EaCap15
D
Phase contrast
GFP (effector)
Merge
WT VcDncV GFP ctrl
WT VcDncV EaCap15-GFP
mut. VcDncV EaCap15-GFP
E
0 min
8 min
16 min
24 min
32 min
F
Phase contrast
Periplasmic mCherry
TO-PRO-3
0 min
8 min
16 min
24 min
G
Healthy cell
Disrupted inner membrane compact cytoplasm
Ruptured inner membrane
H
100 nm
OM
IM
100 nm

Figure 2.10. Cap14 and Cap15 similarly cause membrane disruption and cell death upon activation.
(A) *E. coli* stained with membrane (FM4-64) and DNA (DAPI) dyes. *Ea*Cap15 activation results in inner membrane collapse, compression of the cytoplasm, and expansion of the periplasmic space. (B) Quantification of cell death induced by GFP-labeled *Ea*Cap15 (C-terminal label) and *Vv*Cap14 (N-terminal label) (untagged *Ea*Cap15 and *Vv*Cap14 from Figure 2.1E included for reference). GFP-fusion does not disrupt effector function. (C) Individual channel images of *E. coli* expressing an *Ea*Cap15–GFP fusion with WT *Vc*DncV, as in Figure 2.9E. (D) *E. coli* expressing GFP or GFP-labeled *Ea*Cap15 with WT or catalytically inactive *Vc*DncV; additional images for Figure 2.9D. (E) Time-lapse images of *E. coli* expressing WT *Vc*DncV and *Ea*Cap15 in the presence of the live-dead stain TO-PRO-3. (F) Expression of WT *Vc*DncV and either full-length *Ea*Cap15 or *Ea*Cap15 ΔTM, which lacks transmembrane domains, showing the transmembrane segments are required for membrane disruption. (G) Analysis of *E. coli* expressing the TM effector *Vv*Cap14 (TM-SAVED). Cap14 causes inner membrane disruption (arrow) upon activation of 3′3′-cGAMP signaling. (H) *E. coli* expressing WT or mutant *Vc*DncV, *Vv*Cap14, and periplasmic mCherry. Images show bacteria before (top row) and after (middle row) Cap14-mediated inner membrane rupture. (I) *E. coli* as in (G) stained with membrane (FM4-64) and DNA (DAPI) dyes. (J) *E. coli* expressing GFP or a GFP–*Vv*Cap14 fusion with WT or catalytically inactive *Vc*DncV. *Vv*Cap14 is evenly distributed in the absence of 3′3′-cGAMP signaling, and then becomes enriched on the cell periphery upon activation. (K) Time-lapse images of *E. coli* expressing WT *Vc*DncV and *Vv*Cap14 in the presence of the live-dead stain TO-PRO-3. Following inner membrane collapse, the cells return to phase dark and are permeable to TO-PRO-3, demonstrating rupture of the inner membrane. (L) Model of Cap15 activation. Cap15 begins poised in the inner membrane and oligomerizes upon binding to activating nucleotide second messenger. Analysis of the TM domains and oligomerization of the β-barrel support a model in which Cap15 forms raft-like clusters and disrupts the inner membrane, as opposed to an alternate model where Cap15 forms well-defined pores. Microscopy data are representative of at least 3 independent experiments.


A
Phase contrast
FM4-64 (membrane)
DAPI (DNA)
Merge
WT DncV GFP ctrl
WT DncV EaCap15
Mut. DncV EaCap15
B
OD600 at 300 min (WT/mut VcDncV)
1.0
0.5
0.0
Ctrl
EaCap15
EaCap15-GFP
VvCap14
VvCap14-GFP
F
WT VcDncV
Mut. VcDncV
EaCap15
EaCap15ΔTM
G
VvCap14 (TM-SAVED)
WT VcDncV
Mut. VcDncV
C
Phase contrast
Cap15-GFP
Merge
0 min
16 min
32 min
D
Mut. VcDncV EaCap15-GFP
WT VcDncV EaCap15-GFP
E
Phase contrast
TO-PRO-3
0 min
8 min
14 min
H
Phase contrast
Periplasmic mCherry
DAPI (DNA)
Merge
WT VcDncV VvCap14
Mut. VcDncV VvCap14
I
Phase contrast
FM4-64 (membrane)
DAPI (DNA)
Merge
WT VcDncV VvCap14
J
Phase contrast
GFP (effector)
Merge
WT VcDncV VvCap14-GFP
Mut. VcDncV VvCap14-GFP
K
WT VcDncV + VvCap14
Phase constrast
TO-PRO-3
0 min
4 min
8 min
12 min
16 min
WT VcDncV + VvCap14
0 min
2 min
4 min
6 min
8 min
10 min
L
Raft-like clustering (supported model)
Poised in membrane
Pore formation (alternative model)
Inner membrane rupture cell death

alone, exhibit normal cellular morphology, demonstrating that membrane disruption requires both nucleotide second messenger synthesis and effector activation (Figure 2.9C). Using a periplasmic-localized mCherry marker (Uehara et al., 2010), we further confirmed that *Ea*Cap15 activation results in clear expansion of the periplasmic space as the disrupted inner membrane pulls away from the cell wall (Figure 2.9C). Moreover, we observed that activation of the CBASS phospholipase effector *Vibrio cholerae* CapV (*Vc*CapV), which is known to enzymatically degrade the inner membrane (Severin et al., 2018), results in a similar phenotype, supporting a mechanism of membrane collapse due to direct targeting of inner membrane integrity (Figure 2.9B).

To better understand the consequence of Cap15 oligomerization and activation, we analyzed Cap15 transmembrane domain sequences in comparison to the transmembrane domains of proteins known to induce membrane disruption. Many classes of membrane-disrupting proteins (*e.g.* actinoporin, ClyA) are soluble until they are activated, when exposed transmembrane segments then imbed within the membrane and form a pore (Dal Peraro and van der Goot, 2016; Mueller et al., 2009; Tanaka et al., 2015). In each case, the transmembrane regions of these proteins contain a clear hydrophilic face that lines the central channel of the final pore structure (Figure 2.8E,F). In contrast, other membrane-targeting proteins (*e.g.* phage holin, pinholin) begin poised in the membrane and then cluster into raft-like assemblies to create membrane lesions (Figure 2.8G,H) (Cahill and Young, 2019; Pang et al., 2013; Young et al., 2000). The transmembrane regions of *Ea*Cap15 and CBASS effectors lack hydrophilic faces and more closely match the hydrophobic, uncharged helices that define membrane disrupting proteins that form raft-like assemblies (Figure 2.8I,J). Further supporting potential mechanistic similarity, activation of phage holin proteins results in a similar phenotype in which collapse of the bacterial inner membrane leads to expansion of the periplasm (Dewey et al., 2010).

To further interrogate this model, we used time-lapse imaging to track Cap15 localization and cell fate throughout each stage of TM effector activation. We labeled *Ea*Cap15 with GFP at the C-terminus and confirmed that the fusion protein induces cell death similar to wildtype Cap15

when co-expressed with active *Vc*DncV (Figure 2.10B). Consistent with the model of Cap15 beginning poised in the membrane prior to activation, we observed that in the resting state, *Ea*Cap15 localizes to the membrane at the poles of the *E. coli* cell (Figure 2.9D, Figure 2.10C). Upon induction of *Vc*DncV expression and onset of 3′3′-cGAMP signaling, *Ea*Cap15 redistributes as puncta throughout the cell membrane, and co-localizes with the inner membrane as it collapses away from the cell wall (Figure 2.9D,E, Figure 2.10C,D). Several minutes after shrinking of the inner membrane, bacteria return to a phase-dark state and become permeable to the live-dead stain TO-PRO-3, indicating cell death (Figure 2.9F,G, Figure 2.10E). At this stage, periplasmic mCherry signal leaks throughout the bacterium, confirming rupture of the inner membrane (Figure 2.9F,G). We further confirmed the *Ea*Cap15 transmembrane domains are required to induce the disruption of the membrane (Figure 2.10F). Finally, we used transmission electron microscopy to examine the membranes of cells dying after Cap15 activation and confirmed that the inner membrane specifically collapses away from the cell wall (Figure 2.9H).

To expand upon these observations with Cap15, we assessed the structurally distinct CBASS TM effector Cap14 (TM-SAVED). Similar to *Ea*Cap15, activation of *Vibrio vulnificus* Cap14 (*Vv*Cap14; NCBI ref. WP_017790126.1) specifically induces inner membrane disruption (Figure 2.10G,H). Interestingly, cells expressing each TM effector exhibit distinct morphologies with *Ea*Cap15 inducing collapse associated with pulling the inner membrane away from the poles of the cell and *Vv*Cap14 separating the cytoplasm into two segments from the center (Figure 2.10G). After several minutes, *E. coli* expressing active *Vv*Cap14 similarly undergo inner membrane rupture and cell death, with the release of periplasmic mCherry into the bacterium and the cells becoming permeable to TO-PRO-3 dye (Figure 2.10H,K).

Together, these data demonstrate that CBASS TM effectors limit phage replication by disrupting the bacterial cell membrane and initiating a cell death cascade. The data support a model in which CBASS nucleotide second messenger signaling triggers membrane-bound TM

effectors to oligomerize into raft-like assemblies that disrupt the inner membrane and induce cell death (Figure 2.10L).

Discussion

Our results define bacterial cell membrane disruption as a widespread mechanism controlling abortive infection in CBASS antiphage defense. Following phage infection, CBASS immunity begins with activation of a CD-NTase enzyme to synthesize an antiviral nucleotide second messenger signal. We demonstrate that TM effector proteins directly sense nucleotide second messenger signaling and subsequently initiate a cell death cascade that results in bacterial membrane disruption and restriction of phage propagation. TM effectors are the most prevalent form of CD-NTase-associated effector proteins (Burroughs et al., 2015; Millman et al., 2020), revealing direct membrane targeting as a critical feature of CBASS antiviral defense.

Analysis of the CBASS protein Cap15 explains a mechanism for how TM effectors disrupt membrane integrity and terminate phage replication. Cap15 activation specifically triggers collapse of the bacterial inner membrane and subsequent cytoplasmic condensation. Disruption of the inner membrane causes shrinking away from the outer membrane and cell wall, and eventually results in complete membrane rupture and cell death (Figure 2.9). Microscopy images of bacteria succumbing to Cap15-induced death reveal a phenotype similar to *E. coli* cells dying from genetic disruption of lipid homeostasis (Sutterlin et al., 2016). While Cap15 is the most common TM effector, our results identify several additional families of CBASS TM effectors including Cap14 (TM-SAVED) and Cap16 (TM-NUDIX), that respond to antiviral nucleotide second messenger signaling and similarly induce cell death (Figure 2.1E). We show that Cap14 also specifically triggers inner membrane rupture (Figure 2.10G–K) demonstrating that direct targeting of membrane integrity is a shared feature among diverse CBASS operons. Interestingly, Cap14- and Cap15-mediated death exhibit altered dynamics and collapsed membrane morphologies suggesting that while the overall mechanism of inner membrane disruption is conserved, CBASS TM effectors may use distinct molecular mechanisms to induce cell death.

Our structural characterization of Cap15 additionally explains a mechanism for how TM effectors sense and respond to antiviral nucleotide second messenger signals. Crystal structures

of *Ya*Cap15 reveal a compact β-barrel domain with a hydrophobic lower core that enables formation of a central ligand-binding pocket for specific nucleotide recognition (Figures 2.3 and 2.5). The minimized architecture of *Ya*Cap15 mirrors the structural principles used for rational design of small molecule sensors (Dou et al., 2018), revealing a remarkable convergence between synthetic protein design and natural evolution of a nucleotide-sensing domain. The Cap15 β-barrel domain occurs in ~25% of the TM effectors in CBASS operons representing discovery of >600 new protein sensors capable of sensitively responding to diverse cyclic oligonucleotide signals. These novel β-barrel nucleotide-binding domains expand the list of structurally defined nucleotide second messenger receptors in bacterial and metazoan antiviral immunity including CARF/SAVED-family proteins and NucC receptors in bacteria (Jia et al., 2019; Lau et al., 2020; Lowey et al., 2020; Niewoehner and Jinek, 2016), STING-family receptors in bacteria and animal cells (Morehouse et al., 2020; Ouyang et al., 2012; Yin et al., 2012), and RECON-like receptors in animals (McFarland et al., 2017). We further show that, like soluble CBASS effectors, Cap15 is activated through protein oligomerization. The organization of Cap15 as a discrete nucleotide-binding domain fused to cell death-inducing transmembrane segments further supports that CBASS effectors use a modular architecture to couple nucleotide second messenger recognition with diverse downstream effector functions (Lowey et al., 2020). The modular organization of CBASS effectors facilitates rapid diversification of new effector functions to combat phage replication and likely enabled CD-NTase and STING acquisition during evolution of metazoan cGAS-STING signaling from CBASS precursor components (Morehouse et al., 2020; Whiteley et al., 2019).

Given membrane-targeting effectors are predominant in CBASS defense operons, inner membrane disruption likely provides a key evolutionary advantage in phage defense. Sequence analysis of CBASS operons supports frequent horizontal transfer and exchange between divergent bacterial species (Burroughs et al., 2015; Millman et al., 2020). CBASS effector proteins mediating cell death must therefore target ubiquitous host components to remain functional when

shared between gram-negative and gram-positive bacteria. In line with this prediction, our results support that the vast majority of CBASS effectors target the conserved inner membrane (TM effectors and phospholipases) or double-stranded DNA (DNA endonucleases). Targeting the host inner membrane to control abortive infection may be particularly advantageous as successful phage replication requires maintenance of normal host membrane physiology until precisely timed activation of lytic machinery facilitates virion release (Cahill and Young, 2019). Interestingly, phages in the order *Caudovirales*, including all dsDNA phages demonstrated here to be susceptible to CBASS TM effectors, use perturbations of the membrane to trigger the inner membrane proteins holin or pinholin to initiate cell lysis and virion release (Cahill and Young, 2019). The ability of CBASS TM effectors to prematurely disrupt membrane integrity likely short-circuits this process and enhances antiviral defense. Phages have evolved elaborate modification systems to protect viral DNA from recognition and cleavage (Stern and Sorek, 2011), and it will be interesting to uncover if mechanisms exist that allow escape from direct host membrane disruption.

Limitations of the Study

Here we report the structure of the CBASS Cap15 ligand-binding domain and present mutational analysis of a putative ligand-binding pocket. We also show that ligand-binding induces oligomerization of Cap15, consistent with the activation mechanism observed for other CBASS effectors (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020). Further structural studies of full-length Cap15 in an activated state will be required to define how ligand-binding induces oligomerization and how CBASS TM effector clustering results in membrane disruption. An additional limitation of our study is that the CBASS TM effector experiments only analyze death in the gram-negative host *E. coli*. Cap15 and CBASS TM effectors are broadly distributed among bacteria phyla, and it will be intriguing to compare how the mechanism of cell death and antiphage defense may differ in gram-positive hosts.

METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Escherichia coli strains and phages

BL21-RIL *E. coli* (Agilent) and BL21(DE3) *E. coli* (NEB) were transformed and plated on MDG plates (0.5% glucose, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM MgSO4, 0.25% aspartic acid, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals), then colonies were used to inoculate overnight MDG cultures. Overnight MDG cultures were used to inoculate M9ZB cultures (0.5% glycerol, 1% Cas-amino Acids, 47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 2 mM $MgSO_4$, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals), which were grown in each experiment as described below.

For phage challenge studies, *E. coli* strain MG1655 (ATCC 47076) was grown in MMB (LB media supplemented with 0.1 mM $MnCl_2$, 5 mM $MgCl_2$, with or without 0.5% agar) at room temperature. Whenever applicable, media were supplemented with ampicillin (100 µg mL^{-1}), to ensure the maintenance of plasmids. Infection was performed in MMB media at room temperature. Phages used in this study are listed in Table 2.3.

METHOD DETAILS

Protein expression and purification

Recombinant TM effector proteins, Cap15 effector mutants, and CD-NTase enzymes were purified as previously described (Lowey et al., 2020; Zhou et al., 2018). Briefly, CD-NTases and effectors were cloned into an N-terminal 6×His-SUMO2-tagged pET vector and transformed into BL21-RIL *E. coli* (Agilent). Large scale cultures (2–4 liters) were grown for ~5 h at 37°C, then induced with IPTG overnight at 16°C. Bacterial pellets were resuspended and sonicated in lysis buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 30 mM imidazole, 10% glycerol and 1 mM DTT) and purified using Ni-NTA resin (Qiagen). Ni-NTA resin was washed with lysis buffer

supplemented to 1 M NaCl and eluted with lysis buffer supplemented to 300 mM imidazole. The Ni-NTA elution fraction was dialyzed into 20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM DTT overnight while removing the SUMO2 tag with recombinant human SENP2 protease (D364–L589, M497A). Proteins were concentrated using a 10K-cutoff concentrator (Millipore) and purified by size exclusion chromatography on a 16/600 Superdex 75 column. Proteins were concentrated to >30 mg mL^{-1}, flash frozen with liquid nitrogen, and stored at −80°C.

Effector-induced killing assay

Plasmids expressing wildtype or catalytically inactive *Vc*DncV (D131A, D133A) (pBAD33) and an effector protein (custom pET, see Table 2.2) were transformed into competent *Escherichia coli* BL21(DE3) (NEB) and plated onto plates with MDG media (0.5% glucose, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM MgSO4, 0.25% aspartic acid, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals). After overnight incubation, 5 mL MDG starter cultures were inoculated with 3 colonies each and grown overnight at 37°C with 230 RPM shaking. Cultures were diluted 1:50 into 5 mL M9ZB cultures (0.5% glycerol, 1% Cas-amino Acids, 47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 2 mM $MgSO_4$, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals) and grown for 3 h at 37°C with 230 RPM shaking. Cultures were then diluted 1:5 into fresh M9ZB media with a final concentration of 0.2% arabinose and 5 μM IPTG to induce protein expression; 200 μL of culture were added to 96-well plate and OD_{600} was recorded in technical triplicate over 300 min in a Synergy H1 plate reader (BioTek), while shaking at 37°C. Wells containing media alone were used for OD_{600} background subtraction. Each biological replicate was measured in technical triplicate. OD_{600} at 300 min (WT/mut *Vc*DncV) was calculated as the OD_{600} of cultures expressing WT *Vc*DncV and effector divided by the OD_{600} of cultures expressing catalytically inactive *Vc*DncV and effector. This value was then divided by same ratio calculated for cultures expressing WT or mut. *Vc*DncV

and a GFP control in place of effector. This calculation is represented by one data point in Figure 2.1E; all cultures used within one calculation were grown concurrently.

Quantification of cell death by flow cytometry

E. coli expressing *Vc*DncV and effector proteins were grown as induced as described above in effector-induced killing assays. After cultures were grown with inducer for 2 h, bacteria were diluted into 1× PBS with an approximate final dilution of 1:500 of cultures that had an OD_{600} of 1. The cell-impermeable nucleic acid dye, TO-PRO-3 iodide (ThermoFisher) was added to a final concentration of 500 nM. Cells were analyzed using an LSR-II Analyzer (BD Biosciences) with excitation with a 633 nm laser and 660/20 emission filter. Cells expressing WT *Vc*DncV and no effector were used to define gating for live cells.

Bioinformatics Analyses

Previously-identified CBASS effectors (Millman et al., 2020) were assigned to clusters with at least 20% identity over 60% of the sequence using MMseqs2 (Gabler et al., 2020; Steinegger and Soding, 2017; Zimmermann et al., 2018). Clusters were further grouped into families using Pfam to identify at least two sequences within the cluster with a Pfam assignment (Sonnhammer et al., 1997).

SAVED domains within Cap14 were modeled using AlphaFold2.0 (Jumper et al., 2021). The AlphaFold prediction for the domain previously annotated as SUa (Pfam PF18179) allowed for the identification of this family of proteins as containing SAVED domains.

Transmembrane domains were identified within experimentally determined structures where available (Mueller et al., 2009; Tanaka et al., 2015), and were otherwise predicted using TMHMM (Krogh et al., 2001). Where modeled transmembrane domains are displayed, they were modeled using AlphaFold2.0 (Jumper et al., 2021).

Crystallization and structure determination

Cap15 proteins were crystallized at 18°C using the hanging drop method. Concentrated protein stocks were thawed on ice and diluted in buffer (25 mM HEPES-KOH pH 7.5, 1 mM TCEP). Drops were set by mixing protein stock and reservoir solution in 2 µL drops over a 350 µL reservoir in 15-well Easy-Xtal trays (NeXtal). Each protein was crystallized as follows: *Ya*Cap15 W78–end C 1 2 1 crystal form, selenomethionine *Ya*Cpa15 W78–end was diluted to 10 mg mL^{-1} and mixed at a 1:1 ratio of protein:reservoir solution (40% PEG-200, 0.1 M ADA pH 5.8); *Ya*Cap15 W78–end P 6_1 crystal form, *Ya*Cap15 W78–end was diluted to 10 mg mL^{-1} and mixed at a 1:1 ratio of protein:reservoir solution (40% PEG-200, 0.1 M ADA pH 5.3). In each case, crystals were harvested directly from the mother liquor without additional cryoprotectant, and flash frozen in liquid nitrogen.

X-ray data were collected at the Northeastern Collaborative Access Team beamlines 24-ID-C (P30 GM124165), and used a Pilatus detector (S10RR029205), an Eiger detector (S10OD021527) and the Argonne National Laboratory Advanced Photon Source (DE-AC02-06CH11357). X-ray crystallography data were processed with XDS and AIMLESS (Kabsch, 2010) using the SSRL autoxds script (A. Gonzalez, Stanford SSRL). Experimental phase information for *Ya*Cap15 was determined using data collected from selenomethionine-substituted C 1 2 1 crystals. One site was identified with HySS and an initial map was produced using SOLVE/RESOLVE in Phenix (Liebschner et al., 2019). Model building was performed using Coot (Emsley and Cowtan, 2004), then refined in Phenix. Statistics were analyzed as described in Table 2.1 (Chen et al., 2010; Karplus and Diederichs, 2012; Weiss, 2001).

CD-NTase product thin layer chromatography

Purified *Ya*CdnE was mixed at a final concentration of 1 µM with 25 µM NTPs and trace α-^{32}P-NTP in 10 µL reactions with 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, and 50 mM CAPSO pH 9.0. These reactions were incubated for 2 h at 37°C and then terminated by treatment with 1 µL of 5

units μL^{-1} Calf Intestinal Phosphatase (New England Biolabs) for 1 h at 37°C. 1 µL of each reaction was spotted on a PEI-cellulose thin layer chromatography (TLC) plate (Millipore) and developed in a 1.5 M KH_2PO_4 pH 3.8 buffer until the solvent front was ~1 cm from the top of plate. TLC plates were dried and the exposed to a phosphor screen followed by imaging with an Amersham Typhoon IP (Cytiva).

*Ya*CdnE product purification and analysis by HPLC

Large scale *Ya*CdnE product synthesis was carried out in 20 mL reactions with 1 µM purified *Ya*CdnE and 250 µM ATP, CTP, GTP, and UTP in 1× reaction buffer (50 mM KCl, 20 mM $MgCl_2$, 40 mM CAPSO pH 9.0, 1 mM TCEP), then incubated overnight at 37°C with gentle shaking. The reaction was then treated with 10 µL of 5 units μL^{-1} Calf Intestinal Phosphatase (New England Biolabs) at 37°C overnight before filtering through a 0.2 µm filter. The reaction was diluted to 12.5 mM KCl with water, then loaded onto a 5 mL HiTrap Q column (GE Healthcare) at 1 mL min^{-1} with a peristaltic pump and eluted on an AKTA FPLC (GE Healthcare) with a linear gradient of ammonium acetate from 0–2 M over 60 min. Fractions from the two major peaks were identified by absorbance at 254 nm.

Comparison of major peak products to cyclic dinucleotide standards (3′3′-c-di-UMP [Biolog C 256], 3′3′-cUMP–CMP [Biolog C 375], 3′3′-cUMP–AMP [Biolog C 357]) was carried out using a C18 column (Agilent Zorbax Bonus-RP 4.6×150 mm, 3.5-micron). The column was heated to 40°C and run at 1 mL min^{-1} with a mobile phase of 50 mM NaH_2PO_4 (pH 6.8 with NaOH) supplemented with 3% acetonitrile.

Mass spectrometry

*Ya*CdnE reactions containing 25 µM ATP, GTP, CTP, and UTP, 10 mM $MgCl_2$, 50 mM KCl, 50 mM CAPSO pH 9, 1 mM DTT, and 10 µM enzyme with a final volume of 200 µL. Reactions were incubated overnight and treated with 1 µL of 5 units μL^{-1} Calf Intestinal Phosphatase (New

England Biolabs) for 1 h at 37°C. Sample analysis was carried out by MS-Omics as follows. Samples were diluted 1:3 in 10 % ultra-pure water and 90% acetonitrile containing 10 mM ammonium acetate at pH 9 then filtered through a Costar Spin-X centrifuge tube filter 0.22 µm Nylon membrane. The analysis was carried out using a Thermo Scientific Vanquish LC coupled to Thermo Q Exactive HF MS. An electrospray ionization interface was used as ionization source. Analysis was performed in positive ionization mode. Peak integration was performing using Tracefinder 4.0 and the extracted data was used to create a calibration curve to quantify the samples The UPLC was modified from previously-reported protocols (Hsiao et al., 2018).

Electrophoretic mobility shift assay

Radiolabeled 3′3′-c-di-UMP, 3′3′-cUMP–AMP, and 3′3′-c-di-AMP were synthesized by combining 25 µM NTPs and trace α^{32}P-NTP with purified enzymes *Ya*CdnE (3′3′-c-di-UMP: UTP, α^{32}P-UTP), *Ec*CdnE (3′3′-cUMP–AMP: UTP, ATP, α^{32}P-UTP), and *V. cholerae* DncV (3′3′-c-di-AMP: ATP, α^{32}P-ATP) at a final protein concentration of 1 µM in a 20 µL reaction with 5 mM $MgCl_2$, 50 mM KCl, 50 mM Tris-HCl pH 9.0, 1 mM DTT. These reactions were incubated at 37°C for ~24 h and then terminated by treatment with 1 µL of 5 units μL^{-1} Calf Intestinal Phosphatase (New England Biolabs) for 1 h at 37°C followed by boiling at 95°C for 5 min. The radiolabeled cyclic dinucleotide reaction products were then added at a concentration of 10 nM to reactions containing the Cap15 nucleotide-binding domain or a no protein control in a buffer of 50 mM KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl pH 7.5, and 1 mM TCEP. In wildtype *Ya*Cap15 nucleotide binding assays, binding was measured at protein concentrations ranging from 1 nM to 1 µM and in mutant *Ya*Cap15 nucleotide binding assays binding was measured at concentrations of 100 nM and 1 µM. Reactions were incubated for 10 min at 25°C then separated on a 6% nondenaturing polyacrylamide gel (7.2 cm, 29:1 acrylamide-bisacrylamide, 0.5× TBE buffer) run at 100 V for 20 min in 0.5× TBE buffer. The gel was immediately dried at 80°C for 1 h and then exposed to a phosphor screen followed by imaging with an Amersham Typhoon IP (Cytiva).

Signal intensity was calculated using ImageQuant TL 8.2 software (Cytiva). Background signal intensity, taken from the no protein control lane, was subtracted from each signal intensity reading. Percent ligand bound was calculated by dividing the signal intensity of the bound protein-ligand complex by the total signal intensity of both the unbound ligand and the protein-ligand complex. Where two bands are present in the bound complex, the intensity of both bands was combined to calculate the bound fraction.

Thermal shift assay

Purified *Ya*Cap15 W78–end β-barrel, WT or containing point mutations as indicated, was added at a final concentration of 10 μM with 100 μM synthetic nucleotide ligands (Biolog) and 2× SYPRO Orange Dye in a buffer of 75 mM KCl, 20 mM HEPES-KOH pH 7.5. Samples were slowly heated from 25 to 95°C over the course of 2 h using a qPCR CFX96 thermocycler (Biorad) and fluorescence in the SYBR channel was measured every 0.5°C. The first derivative of each fluorescence curve was calculated, and the melting temperature was identified as the peak of each derivative curve.

SEC-MALS

*Ya*Cap15 W78–end samples were prepared by diluting in SEC-MALS running buffer (150 mM KCl, 20 mM HEPES-KOH pH 7.5, 1 mM TCEP) to 1.25 mg mL^{-1} and incubating on ice for 5 min, then separated on an SRT SEC-150 column or SRT SEC-300 column (Sepax). Protein concentration was calculated using refractive index on a Wyatt Optilab T-rex Refractive Index Detector assuming dn dc^{-1} of 0.185 and a molar mass was calculated using a Wyatt Dawn Heleos II Multi-Angle Light Scattering detector and ASTRA software.

Plasmid and strain construction

Yersinia aleksiciae (genome accession NZ_CP011975.1) and *Escherichia albertii* (genome accession MOD1-EC1698) CBASS operons used in this study were synthesized by Genscript Corp. and cloned into the pBAD plasmid (Thermofisher, cat. #43001), as previously described (Bernheim et al., 2021). *Escherichia albertii* and *Yersinia aleksiciae* CBASS operon mutants were constructed by Genscript Corp or cloned using PCR fragments and Gibson assembly. Plasmids were transformed into *E. coli* strain MG1655.

Plaque assays

Phages were propagated by picking a single phage plaque into a liquid culture of *E. coli* MG1655 grown at 37°C to OD_{600} of 0.3 in MMB medium until culture collapse. The culture was then centrifuged for 10 min at 4,000× RPM and the supernatant was filtered through a 0.2 μm filter to remove remaining bacteria and bacterial debris. Lysate titer was determined using the small drop plaque assay method as previously described (Mazzocco et al., 2009).

Plaque assays were performed as previously described (Mazzocco et al., 2009). Bacteria (*E. coli* MG1655 with CBASS-expressing cells) or negative control (*E. coli* MG1655 with a pBAD-GFP vector) were grown overnight at 37°C. Then 300 μL of the bacterial culture was mixed with 30 mL melted MMB agar (LB media supplemented with 0.1 mM $MnCl_2$, 5 mM $MgCl_2$, 0.5% agar, and 0.2% arabinose) and left to dry for 1 h at room temperature. CBASS operons containing point mutations as indicated or Cap15 transmembrane regions were removed (*Ya*Cap15 W78–end). 10-fold serial dilutions in MMB were performed for each of the tested phages and 10 μL drops were spotted onto the bacterial layer. Plates were incubated overnight at room-temperature and plaque forming units (PFUs) were determined by counting the derived plaques after overnight incubation.

Microscopy

Plasmids expressing wildtype or catalytically inactive *Vc*DncV (D131A, D133A) (pBAD33) and an effector protein (custom pET) were transformed into competent *E. coli* BL21(DE3) (NEB) and plated onto plates with MDG media (0.5% glucose, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM MgSO4, 0.25% aspartic acid, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol, and trace metals), then incubated overnight at 37°C. The next day, 5 mL MDG starter cultures were inoculated with 3 colonies each, then incubated overnight at 37°C with 230 RPM shaking. The next morning, overnight cultures were resuspended in 5× volume M9ZB (0.5% glycerol, 1% Cas-amino Acids, 47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 2 mM $MgSO_4$, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals) and grown for 4 h at 37°C with 230 RPM shaking. Protein expression was then induced by resuspending cultures in 5× volume of inducing media (M9ZB, 0.2% arabinose, 5 µM IPTG), and cultures were grown for 1.5 h at 37°C with 230 RPM shaking. Where indicated, the bacteria were additionally transformed with a plasmid expressing periplasmic mCherry as previously described (Uehara et al., 2010), and subjected to additional selection with 50 µg mL^{-1} spectinomycin.

Where indicated, 100 µL bacteria were stained with dyes by adding 1 µL of 100× stocks: DAPI (ThermoFisher D1306) 0.2 mg mL^{-1}, FM4-64 (ThermoFisher T13320) 0.15 mg mL^{-1}, TO-PRO-3 iodide (ThermoFisher T3605) 50 µM. Bacteria were then plated on agarose pads (1.2% agarose, M9ZB, 0.4% arabinose, 10 µM IPTG, with additional FM4-64 15 µg mL^{-1} or 500 nM TO-PRO-3 iodide where indicated) and immediately covered with coverslip, as previously described (Wang and Montero Llopis, 2016). Slides were imaged between 2 and 4 h after initial induction; during time-lapse imaging, slides were maintained at 30°C. Images were acquired using a Nikon Ti microscope equipped with Plan Apo 100×/1.40 Oil Ph3 DM objective and Andor Zyla 4.2 Plus sCMOS camera and Nikon Elements 4.30 acquisition software. Adobe Photoshop was used for brightness and contrast adjustments.

Transmission electron microscopy

E. coli expressing WT *Vc*DncV and *Ea*Cap15 were grown and induced as described for phase contrast microscopy. After 2.5 h, cells were pelleted and fixed in 2.5% paraformaldehyde, 5.0% glutaraldehyde, 0.06% picric acid in 0.2 M cacodylate buffer, followed by osmication and uranyl acetate staining, dehydration in graded alcohols and embedded in Taab 812 Resin (Marivac Ltd., Nova Scotia, Canada). Blocks/samples:80nm sections werw cut with the Leica Ultracut S microtome, picked up on formvar-carbon coated slot Cu grids, stained with 0.2% Lead Citrate, and viewed and imaged under the Philips Tecnai BioTwin Spirit or JEOL 1200x Electron Microscope.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details for each experiment can be found in the figure legends and outlined in the corresponding methods details section. Data are plotted with error bars representing the standard deviation (SD).

Table 2.2. CBASS transmembrane effector protein accession numbers

Effector	Family	Pfam	Genus species	Accession
*Vc*CapV	patatin	PF01734	*Vibrio cholerae*	WP_001133548.1
*Bp*Cap5	HNH-SAVED	PF01844/PF18145	*Berkholderia pseudomallei*	WP_004556385.1
*Dh*S4TM	S4TM	PF18159	*Desulfotomaculum hydrothermale*	WP_031517912.1
*Sf*S4TM	S4TM	PF18159	*Shigella flexneri*	EIQ80516.1
*Ef*SLATT4	SLATT4	PF18186	*Enterococcus faecium*	WP_002320756.1
*Lm*SLATT4	SLATT4	PF18186	*Listeria monocytogenes*	WP_009929204.1
*Se*SLATT4	SLATT4	PF18186	*Salmonella enterica*	WP_000369761.1
*Cj*SLATT6	SLATT6	PF18169	*Campylobacter jejuni*	WP_044779458.1
*Ef*SLATT6	SLATT6	PF18169	*Enterococcus faecium*	WP_002320756.1
*Li*SLATT6	SLATT6	PF18169	*Leptospira interrogans*	WP_000159261.1
*Fs*Cap13	TM-STING	N/A	*Flavobacteriaceae sp.*	P0DUD7.1
*Re*Cap13	TM-STING	N/A	*Roseivirga ehrenhergii*	A0A150XSR0.2
*Ec*Cap14	TM-SAVED	PF18303/PF18145	*Escherichia coli*	EST64817.1
*Sm*Cap14	TM-SAVED	PF18303/PF18145	*Serratia marcescens*	WP_015376197.1
*Vv*Cap14	TM-SAVED	PF18303/PF18145	*Vibrio vulnificus*	WP_017790126.1
*Ef*Cap14	TM-SUa	PF18179	*Enterococcus faecium*	WP_002302464.1
*Ll*Cap14	TM-SUa	PF18179	*Lactococcus lactis*	WP_033899737.1
*Sa*Cap14	TM-SUa	PF18179	*Streptococcus agalactiae*	WP_016480073.1
*Bt*Cap15	TM-β-barrel	PF18153	*Bacillus thermoamylovorans*	WP_041847731.1
*Ea*Cap15	TM-β-barrel	PF18153	*Escherichia albertii*	WP_206748793.1
*Kp*Cap15	TM-β-barrel	PF18153	*Klebsiella pneumoniae*	WP_016154438.1
*Ya*Cap15	TM-β-barrel	PF18153	*Yersinia aleksiciae*	WP_145567547.1
*Ef*Cap16	TM-NUDIX	PF18167	*Enterococcus faecalis*	WP_142956467.1
*Se*Cap16	TM-NUDIX	PF18167	*Staphylococcus epidermidis*	WP_020363756.1
*Sm*Cap16	TM-NUDIX	PF18167	*Streptococcus mutans*	WP_012997811.1

Table 2.3. Phage accession numbers

Phage	Source	Identifier	Accession
Phage SECPhi17	Doron *et al.* 2018	N/A	LT960607.1
Phage SECPhi18	Doron *et al.* 2018	N/A	LT960609.1
Phage SECPhi27	Doron *et al.* 2018	N/A	LT961732.1
Phage SECPhi6	Millman *et al.* 2020	N/A	CADCZA010000001.1
Phage SECPhi4	Millman *et al.* 2020	N/A	MT331608
Phage T2	German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)	DSM 16352	LC34380.1
Phage T4	U. Qimron	N/A	AF158101.6
Phage T5	U. Qimron	N/A	AY543070.1
Phage T6	German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)	DSM 4622	MH550421.1
Phage T7	U. Qimron	N/A	NC_001604.1

Author Contributions

Brianna Duncan-Lowey: Designed experiments, performed crystallography, biochemical, bacterial toxicity, and microscopy experiments, wrote the manuscript and prepared figures.

Nora K. McNamara-Bordewick: Performed crystallography and biochemical experiments.

Nitzan Tal: Performed phage challenge assays.

Rotem Sorek: Performed phage challenge assays.

Philip J. Kranzusch: Designed experiments and wrote the manuscript.

All authors contributed to editing the manuscript and support the conclusions.

Specific Acknowledgements

The authors are grateful to Tobias Herrmann, Thomas Bernhardt, Stephen Harrison, Amy S.Y. Lee, and members of the Kranzusch lab for helpful comments and discussion, Harvard University's Center for Macromolecular Interactions, Paula Montero Llopis and the Microscopy Resources on the North Quad (MicRoN) core at Harvard Medical School, and Morten Danielsen and Daniel Malhiero for assistance with mass spectrometry. Electron microscopy imaging was performed in the HMS Electron Microscopy Facility. The work was funded by the Pew Biomedical Scholars Program (P.J.K.), Burroughs Wellcome Fund PATH award (P.J.K.), Mark Foundation for Cancer Research (P.J.K.), Mathers Foundation (P.J.K.), and Parker Institute for Cancer Immunotherapy (P.J.K.), the European Research Council (grant ERC-CoG 681203 to R.S.), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine (R.S.), the Minerva Foundation (R.S.) and the Knell Family Center for Microbiology (R.S.). B.D.-L. is supported as a Herchel Smith Graduate Research Fellow.

Chapter 3

Molecular basis of CD-NTase nucleotide selection in CBASS anti-phage defense

Apurva A. Govande[1,2], Brianna Duncan-Lowey[1,2], James B. Eaglesham[1,2], Aaron T. Whiteley[3], Philip J. Kranzusch[1,2,4]

[1] Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

[2] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[3] Department of Biochemistry, University of Colorado Boulder, Boulder, Colorado, 80309, USA

[4] Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA 02115, USA

The published manuscript and an additional supplementary table may be accessed as follows:

Cell Reports. June 1, 2021. Vol. 35, 109206.

https://doi.org/10.1016/j.celrep.2021.109206

Summary

cGAS/DncV-like Nucleotidyltransferase (CD-NTase) enzymes are signaling proteins that initiate antiviral immunity in animal cells and CBASS phage defense in bacteria. Upon phage recognition, bacterial CD-NTases catalyze synthesis of cyclic oligonucleotide signals that activate downstream effectors and execute cell death. How CD-NTases control nucleotide selection to specifically induce defense remains poorly defined. Here we combine structural and nucleotide analog interference mapping approaches to identify molecular rules controlling CD-NTase specificity. Structures of the cyclic trinucleotide synthase *E. cloacae* CdnD reveal coordinating nucleotide interactions and a possible role for inverted nucleobase positioning during product synthesis. We demonstrate that correct nucleotide selection in the CD-NTase donor pocket results in formation of a thermostable protein–nucleotide complex and extend our analysis to establish specific patterns governing selectivity for each of the major bacterial CD-NTase Clades A–H. Our results explain CD-NTase specificity and enable predictions of nucleotide second messenger signals within diverse antiviral systems.

Introduction

cGAS/DncV-like Nucleotidyltransferase (CD-NTase) enzymes are signaling proteins in bacteria and animals that sense viral infection and synthesize nucleotide second messengers to initiate antiviral defense (Kranzusch, 2019; Whiteley et al., 2019). In bacteria, CD-NTases function as part of defense operons named cyclic-oligonucleotide-based anti-phage signaling systems (CBASS) that provide bacteriophage resistance (Cohen et al., 2019). Following phage infection, bacterial CD-NTase enzymes sense a yet undefined viral cue and initiate nucleotide second messenger synthesis. CBASS operons additionally encode CD-NTase-associated protein effectors that respond to the nucleotide second messenger signal and induce rapid cell death to limit phage propagation (Cohen et al., 2019; Lowey et al., 2020; Ye et al., 2020). Recent evidence demonstrates that the human cGAS-STING immune pathway responsible for sensing pathogen- and tumor-derived cytosolic DNA evolved from components conserved within bacterial CBASS operons (Morehouse et al., 2020; Whiteley et al., 2019). Broad conservation of CD-NTase signaling across both the bacterial and animal kingdoms demonstrates the importance of understanding the molecular rules controlling antiviral nucleotide second messenger synthesis.

Bacterial CD-NTases synthesize a remarkable diversity of signals including cyclic dinucleotide and cyclic trinucleotide species. CD-NTase enzymes use all four standard nucleobases and are capable of synthesizing both 3′–5′ and 2′–5′ linked products resulting in a predicted diversity of >180 nucleotide second messenger species (Lowey et al., 2020; Whiteley et al., 2019). Examples include *Vibrio cholerae* DncV which synthesizes the cyclic dinucleotide 3′–5′ / 3′–5′ cyclic GMP–AMP (3′3′-cGAMP) and *Enterobacter cloacae* CdnD which synthesizes the cyclic trinucleotide 3′3′3′-cyclic AMP–AMP–GMP (3′3′3′-cAAG) (Davies et al., 2012; Whiteley et al., 2019). CBASS effectors recognize nucleotide second messengers with exquisite specificity, necessitating that CD-NTase enzymes must catalyze rapid synthesis of the correct nucleotide product in order to initiate antiviral defense (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019). Crystal structures of bacterial CD-NTases define enzyme

features that direct nucleotide second messenger formation (Kranzusch et al., 2014; Whiteley et al., 2019; Ye et al., 2020; Zhu et al., 2014), but current understanding is insufficient to predict product diversity across distantly related CD-NTases.

Here we combine biochemical and structural biology approaches to characterize CD-NTase nucleotide second messenger formation and define the determinants of nucleotide selectivity. We determine a crystal structure of the cyclic trinucleotide synthase *E. cloacae* CdnD that enables direct comparison with the cyclic dinucleotide synthase *V. cholerae* DncV. Using a panel of modified nucleotide analogs, we define how product formation is controlled through nucleotide selection, alteration of reaction order, and formation of distinct reaction intermediates prior to cyclization. We extend these observations to structures of CD-NTase enzymes from previously uncharacterized clades and establish clade-specific patterns of selectivity that control nucleotide selection. Together, our results explain the molecular rules controlling formation of antiviral nucleotide second messengers and create a foundation to predict CD-NTase product specificity.

Results and Discussion

Structure of the 3′3′3′-cAAG synthase *Ec*CdnD in complex with ATP

To define the mechanism of CD-NTase product specificity, we determined crystal structures of the 3′3′3′-cAAG synthase *Enterobacter cloacae* CdnD (*Ec*CdnD) in complex with two molecules of ATP (1.45 Å) (Table 3.1). The *Ec*CdnD structure exhibits a canonical CD-NTase fold forming a cage-like architecture with a recessed central pocket for nucleotide coordination and a lid with hydrophilic sidechains lining the top of the active-site (Figure 3.1A) (Whiteley et al., 2019). Previous structures of CD-NTase–nucleotide complexes define “donor” and “acceptor” nucleotide pockets critical for substrate coordination and the sequential steps of product cyclization (Civril et al., 2013; Gao et al., 2013a; Kranzusch, 2019; Kranzusch et al., 2014; Whiteley et al., 2019; Zhu et al., 2014). *Ec*CdnD active site residues D69, D71, and D121 coordinate two magnesium ions and the triphosphate of one molecule of ATP positioned to act as the phosphate donor during initial bond formation (Figure 3.1B, left).The *Ec*CdnD donor ATP adopts a similar orientation to the donor nucleotide in structures of the 3′3′-cGAMP synthase *Vibrio cholerae* DncV (*Vc*DncV) (Figure 3.1C, Figure 3.2A, inset) (Kato et al., 2015; Kranzusch et al., 2014; Zhu et al., 2014). In contrast, the ATP within the acceptor pocket of *Ec*CdnD is held in an unexpected inverted orientation that is rotated 120° relative to the *Vc*DncV acceptor nucleotide (Figure 3.1B, C).

To further understand the positioning of the nucleotides in complex with *Ec*CdnD, we screened related Clade D CD-NTases and determined a crystal structure of *Salmonella enterica* CdnD (2.60 Å, ~80% identical to *Ec*CdnD, 0.48 Å Cα RMSD) in complex with two molecules of GTP (Table 3.1, Figure 3.2B). The *Se*CdnD–GTP structure reveals nearly identical nucleotide positioning with an inverted acceptor nucleotide (Figure 3.2B) suggesting that the *Ec*CdnD–ATP nucleotide orientation is not artificially induced by specific crystallization conditions. However, the inverted orientation of the acceptor nucleotide results in a 3′ OH position incompatible for nucleophilic attack, indicating the trapped ATP conformation is likely not directly relevant for 3′3′3′-cAAG synthesis (Figure 3.1B, C). The inverted acceptor nucleotide is further coordinated by an

Table 3.1. Crystallographic statistics

	Enterobacter cloacae **CdnD**	*Enterobacter cloacae* **CdnD** (SeMet)	*Salmonella enterica* **CdnD**
Data Collection			
Resolution (Å)[a]	1.45	1.498	2.59
Wavelength (Å)	0.97918	0.97949	0.97918
Space group	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1
Unit cell: a, b, c (Å)	66.07 101.37 68.36	66.31 102.36 68.55	74.47 69.78 88.70
Unit cell: α, β, γ (°)	90.00 117.61 90.00	90.00 118.04 90.00	90.00 110.35 90.00
Molecules per ASU	2	2	2
Total reflections	987705	5739620	93109
Unique reflections	140248	128307	26345
Completeness (%)[a]	99.5 (97.8)	99.2 (92.9)	99.3 (96.4)
Multiplicity[a]	7.0 (6.8)	44.7 (31.1)	3.5 (3.5)
$I/\sigma I$[a]	13.8 (1.4)	13.6 (0.9)	3.5 (1.1)
CC(1/2)[b] (%)[a]	99.9 (61.9)	99.9 (37.8)	98.5 (38.1)
Rpim[c] (%)[a]	2.5 (55.8)	4.2 (91.9)	13.5 (96.5)
Sites		12	
Refinement			
Resolution (Å)	38.87 – 1.45		38.49 – 2.59
Free reflections	2046		1983
R-factor / R-free	15.2/16.5		22.9/26.5
Bond distance (RMS Å)	0.008		0.003
Bond angles (RMS °)	1.04		0.66
Structure/Stereochemistry			
No. atoms: protein	5615		5581
No. atoms: ligand	156		134
No. atoms: water	806		91
Average B-factor: protein	26.6		54.6
Average B-factor: ligand	19.8		50.9
Average B-factor: water	36.4		46.1
Ramachandran plot: favored	97.96%		96.62%
Ramachandran plot: allowed	2.04%		3.67%
Ramachandran plot: outliers	0.00%		0.00%
Rotamer outliers	0.33%		1.15%
MolProbity[d] score	1.00		1.62
Protein Data Bank ID	**7LJL**		**7LJM**

[a] Highest resolution shell values in parentheses
[b] (Karplus and Diederichs, 2012)
[c] (Weiss, 2001)
[d] (Chen et al., 2010)

Table 3.1. Crystallographic Statistics, continued.

	***Bacteroides fragilis* CdnB**	***Bacteroides fragilis* CdnB** (SeMet)	***Bradyrhizobium diazoefficiens* CdnG**	***Bradyrhizobium diazoefficiens* CdnG** (SeMet)
Data Collection				
Resolution (Å)[a]	1.76	3.095	1.598	3.685
Wavelength (Å)	0.97918	0.97910	0.97920	0.97920
Space group	P 2_1 2_1 2_1	P 3_1 2 1	P 1 2_1 1	P 6_2
Unit cell: a, b, c (Å)	63.37 69.61 92.40	69.71 69.71 182.71	112.00 68.69 112.43	128.99 128.99 59.56
Unit cell: α, β, γ (°)	90.00 90.00 90.00	90.00 90.00 120.00	90.00 108.40 90.00	90.00 90.00 120.00
Molecules per ASU	2	2	2	2
Total reflections	273132	189210	725390	498887
Unique reflections	41249	9979	212025	6251
Completeness (%)[a]	99.9 (99.1)	99.0 (100.0)	99.0 (90.7)	99.6 (98.7)
Multiplicity[a]	6.6 (4.5)	19.0 (18.1)	3.4 (3.2)	79.8 (75.7)
$I/\sigma I$[a]	10.6 (1.2)	9.5 (1.9)	9.1 (2.7)	9.0 (1.4)
CC(1/2)[b] (%)[a]	99.7 (36.4)	99.5 (52.1)	99.5 (76.1)	100.00 (75.0)
Rpim[c] (%)[a]	4.1 (83.6)	9.2 (107.5)	4.5 (32.6)	8.8 (270.3)
Sites		6		6
Refinement				
Resolution (Å)	46.862 – 1.760		37.94 – 1.59	
Free reflections	2000		2018	
R-factor / R-free	17.8/19.5		19.7/21.7	
Bond distance (RMS Å)	0.003		0.006	
Bond angles (RMS °)	0.79		0.86	
Structure/Stereochemistry				
No. atoms: protein	2356		10660	
No. atoms: ligand	55		132	
No. atoms: water	394		2038	
Average B-factor: protein	31.0		22.3	
Average B-factor: ligand	49.0		13.4	
Average B-factor: water	41.3		31.8	
Ramachandran plot: favored	98.26%		99.10%	
Ramachandran plot: allowed	1.74%		0.90%	
Ramachandran plot: outliers	0.00%		0.00%	
Rotamer outliers	0.39%		0.54%	
MolProbity[d] score	1.06		1.37	
Protein Data Bank ID	**7LJO**		**7LJN**	

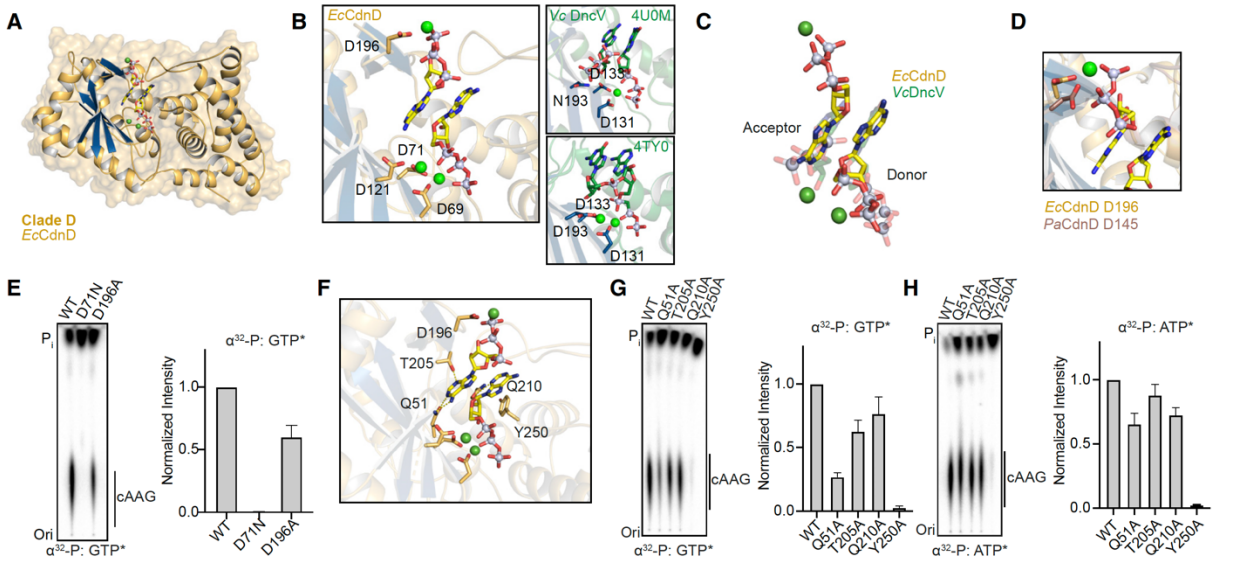
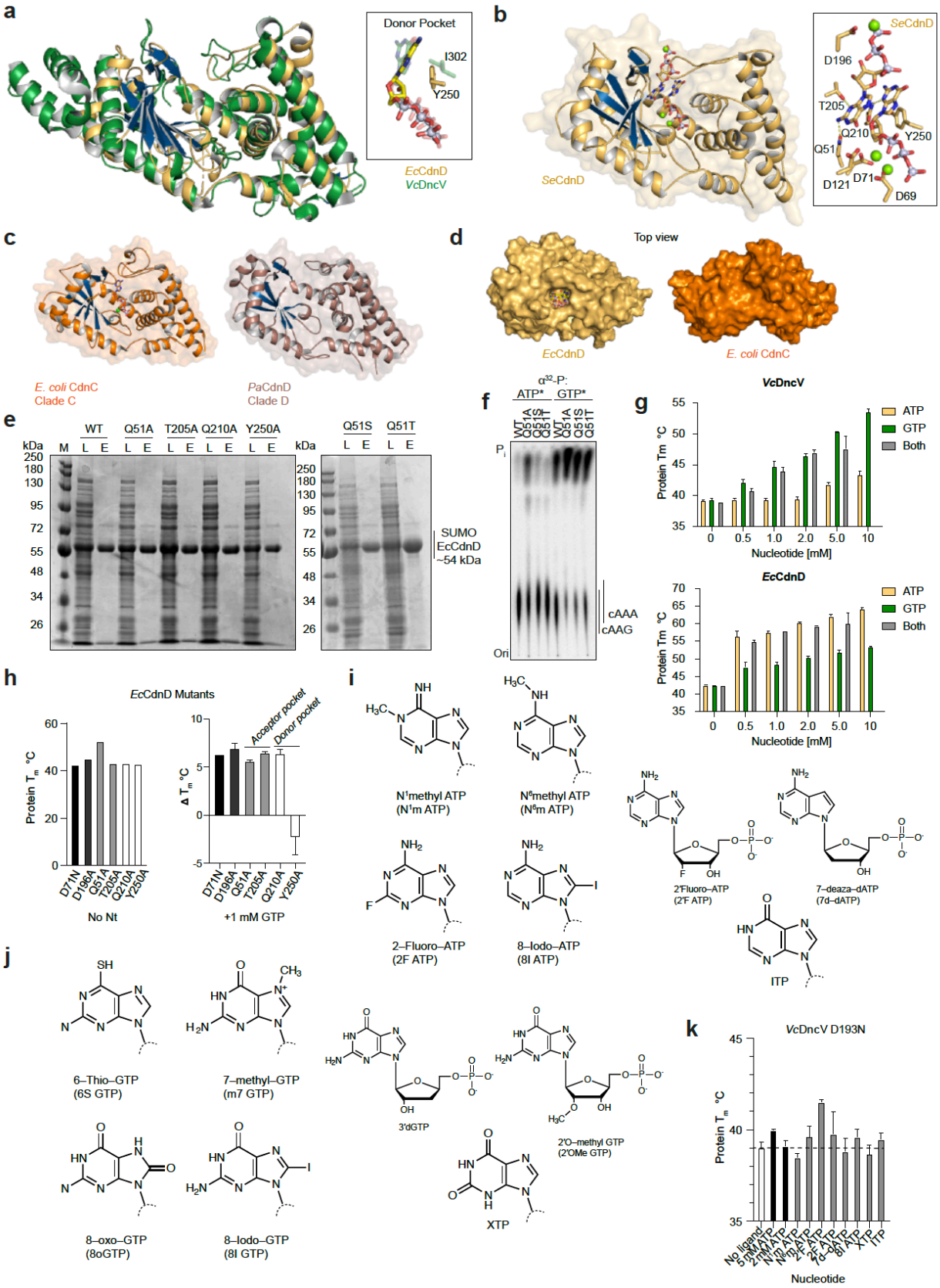


Figure 3.1. Crystal structure of the 3′3′3′-cAAG synthase *Ec*CdnD in a pre-reactive state
(A) Crystal structure of the Clade D CD-NTase from *Enterobacter cloacae* (*Ec*CdnD) in complex with two molecules of ATP. *Ec*CdnD shares all conserved structural features of the CD-NTase family of enzymes. (**B)** Left: Zoom-in cutaway of the *Ec*CdnD active site highlighting catalytic triad D69, D71, and D121 coordinating two magnesium ions and the triphosphate of one molecule of ATP. The second molecule of ATP is bound in an inverted orientation and coordinated by a third magnesium ion and D196. Top right, active site of *Vc*DncV with catalytic aspartates D131 and D133 coordinating a single magnesium ion and ATP and GTP in an orientation conducive to bond formation. A *Vc*DncV D193N mutation stalls the reaction in a pre-reactive state (PDB 4U0M). Bottom, active site of *Vc*DncV in complex with the non-hydrolyzable intermediate ppcpA[3′–5′]pG. Adenosine and guanosine bases are reoriented and rotated 180° relative to the pre-reactive state (PDB 4TY0). **(C)** Overlay of ATP nucleotides from the *Ec*CdnD crystal structure, in yellow, and pre-reactive ATP and GTP from *Vc*DncV crystal structure, in green (PDB 4U0M). Phosphate donor nucleotides are oriented similarly while the phosphate acceptor nucleotide in *Ec*CdnD is inverted relative to the *Vc*DncV acceptor nucleotide. **(D)** Overlay highlighting *Ec*CdnD D196 and coordination of a third metal ion compared to the crystal structure of a related Clade D CD-NTase from *Pseudomonas aeruginosa*, *Pa*CdnD, with an aspartate in a similar position (PDB 6P82). D196 is conserved in 57% of Clade D but is not conserved in other clades. **(E)** Thin-layer chromatography analysis of *Ec*CdnD and quantification of active site residue D71 and third metal ion coordinating residue D196. D71N mutation abrogates synthesis of 3′3′3′-cAAG, while D196A mutation does not prevent product formation but may show a slight defect. **(F)** Detail of nucleotide coordinating residues in the *Ec*CdnD crystal structure. **(G)** Thin-layer chromatography analysis and quantification of nucleotide coordinating residues. Q51A mutation affects the incorporation of GTP. Y250A mutation completely abrogates synthesis of 3′3′3′-cAAG. **(H)** Thin-layer chromatography analysis and quantification of nucleotide coordinating residues. Q51A mutation shows little defect in ATP incorporation and synthesizes minor product 3′3′3′-cAAA. Y250A mutation completely abrogates synthesis. All TLC data are representative of three independent experiments and error bars denote the standard deviation.

Figure 3.2. *Ec*CdnD and *Vc*DncV structural comparison and complex formation with nucleotide substrates.
(A) Structural overlay and comparison of *Ec*CdnD and *Vc*DncV. *Ec*CdnD shares the pol-β like palm domain and core CD-NTase fold. Inset, detail of phosphate donor nucleotide from EcCdnD and *Vc*DncV. Y250 stacks the donor nucleotide in the EcCdnD structure, while I302 plays a similar role in *Vc*DncV. **(B)** Crystal structure of GTP bound *Se*CdnD, inset shows GTP nucleotides and conserved coordinating residues. **(C)** Crystal structures of cyclic trinucleotide synthases E. coli CdnC (PDB 6P80) and *Pa*CdnD (PDB 6P82). **(D)** Left, top view surface representation of *Ec*CdnD highlighting the inverted acceptor nucleotide extruding through a surface exposed channel, right, surface representation of *E. coli* CdnC showing the enclosed active site. **(E)** Coomassie stained gel of *Ec*CdnD mutants showing expression level and purity. L, clarified lysate. E, elution of final Ni-NTA purified protein with a 6 His-SUMO2 tag. **(F)** Thin-layer chromatography analysis of nucleotide coordinating residue 51. Q51A mutation affects the incorporation of GTP, Q51S and Q51T mutations similarly reduce the incorporation of GTP. **(G)** Plot of Tm for *Vc*DncV D193N and *Ec*CdnD D71N with titrated ATP, GTP, or both between 0–10 mM total nucleotide concentration. Bars are the mean of two independent experiments with error bars indicating standard deviation. **(H)** *Ec*CdnD mutant Tm without nucleotides or ΔT_m of EcCdnD mutants bound to 1 mM GTP. $\Delta T_m = T_m(1\ mM\ GTP) - T_m(Apo)$. Data plotted are the mean of two independent experiments with error bars indicating standard deviation. **(I)** Chemical structures of modified ATP nucleotides and ITP. **(J)** Chemical structures of modified GTP nucleotides and XTP. **(K)** Plot of T_m for VcDncV and modified ATP nucleotides at higher concentrations. 5 mM: N^1m ATP, N^6m ATP, 2′F ATP, XTP, ITP; 2 mM: 7d–dATP, 8I ATP, 2F ATP. Data plotted are the mean of two independent experiments with error bars indicating standard deviation.


a
Donor Pocket
I302
Y250
EcCdnD
VcDncV
b
SeCdnD
D196
T205
Q210
Y250
Q51
D121
D71
D69
c
E. coli CdnC
Clade C
PaCdnD
Clade D
d
Top view
EcCdnD
E. coli CdnC
e
WT
Q51A
T205A Q210A
Y250A
Q51S
Q51T
kDa
M
L E
SUMO
EcCdnD
~54 kDa
f
α32-P:
ATP*
GTP*
WT
Q51A
Q51S
Q51T
Pi
Ori
cAAA
cAAG
g
VcDncV
Protein Tm °C
ATP
GTP
Both
Nucleotide [mM]
EcCdnD
h
EcCdnD Mutants
Protein Tm °C
ΔTm °C
Acceptor pocket
Donor pocket
D71N
D196A
Q51A
T205A
Q210A
Y250A
No Nt
+1 mM GTP
i
N1methyl ATP
(N1m ATP)
N6methyl ATP
(N6m ATP)
2–Fluoro–ATP
(2F ATP)
8–Iodo–ATP
(8I ATP)
2'Fluoro–ATP
(2'F ATP)
7–deaza–dATP
(7d–dATP)
ITP
j
6–Thio–GTP
(6S GTP)
7–methyl–GTP
(m7 GTP)
8–oxo–GTP
(8oGTP)
8–Iodo–GTP
(8I GTP)
3'dGTP
2'O–methyl GTP
(2'OMe GTP)
XTP
k
VcDncV D193N
Protein Tm °C
No ligand
5 mM ATP
2 mM ATP
1 mM ATP
N1m ATP
N6m ATP
2F ATP
2'F ATP
7d–dATP
8I ATP
XTP
ITP
Nucleotide

auxiliary metal-binding site created between D196 and the nucleotide triphosphate (Figure 3.1D). *Ec*CdnD and *Se*CdnD are most closely related to previous structures *Pseudomonas aeruginosa* CdnD (*Pa*CdnD, 3.1 Å Cα RMSD) and *E. coli* CdnC (2.7 Å Cα RMSD) that each synthesize 3′3′3′-cAAA (Ye et al., 2020). Residue D196 is conserved in 57% of Clade D CD-NTases and is present in the structure of *Pa*CdnD (Figure 3.1D, Figure 3.2C). However, residue D196 is not present in *E. coli* CdnC or Clade C CD-NTase enzymes and *E. coli* CdnC lacks a channel present in *Ec*CdnD that allows the inverted ATP to extend from the active site (Figure 3.2D). We mutated *Ec*CdnD residue D196 to alanine and observed only a slight defect in synthesis of 3′3′3′-cAAG, suggesting the metal coordinated by D196 is not required for catalysis (Figure 3.1E). In contrast, a D71N mutation of the main active site results in complete abrogation of *Ec*CdnD enzymatic activity (Figure 3.1E).

We next analyzed nucleotide contacts in the active site of *Ec*CdnD (Figure 3.1F). In the donor pocket, Y250 stacks with the bound ATP nucleotide and stabilizes the base, while Q210 hydrogen bonds with the ATP N3 position. Mutation of Y250 leads to a complete disruption of 3′3′3′-cAAG synthesis that is not the result of reduced enzyme stability (Figure 3.1G, Figure 3.2E), while a Q210A mutation has little effect on product formation. The inverted ATP in the *Ec*CdnD acceptor pocket is coordinated by hydrogen bonding interactions between T205 and the N3 position of ATP, and Q51 and the N1 and N6 positions of ATP (Figure 3.1F). Interestingly, *Ec*CdnD Q51 is positioned to make Watson-Crick edge contacts that discriminate between ATP and GTP. We observed that an *Ec*CdnD Q51A mutant enzyme exhibits a defect in incorporation of GTP resulting in increased synthesis of a minor reaction product 3′3′3′-cAAA (Figure 3.1H). Similarly, a Q51S (present in *E. coli* CdnC) or Q51T (present in *Pa*CdnD) substitution results in a shift toward 3′3′3′-cAAA production, demonstrating that this residue is important for GTP discrimination and may participate in controlling transient substrate orientation during 3′3′3′-cAAG product formation (Figure 3.2F). Together, these results demonstrate key hydrogen bonding interactions in the *Ec*CdnD active site that control substrate recognition and nucleotide specificity.

Divergent CD-NTase enzymes exhibit distinct patterns of recognition in the acceptor and donor pockets

We next compared interactions in the *Vc*DncV and *Ec*CdnD acceptor and donor nucleotide pockets. We purified catalytically inactive *Vc*DncV D193N and *Ec*CdnD D71N variants with mutations that prevent reaction progression but do not disrupt nucleotide coordination (Kranzusch et al., 2014; Zhu et al., 2014), and used a thermo-denaturation assay to measure the ability of nucleotide substrates to increase enzyme stability. *Vc*DncV synthesizes 3′3′-cGAMP through a defined reaction order that first utilizes GTP in the donor nucleotide position to form a pppA[3′–5′]pG linear intermediate (Kranzusch et al., 2014; Zhu et al., 2014). In the presence of the correct initiating donor nucleotide GTP, we observed significant stabilization with the melting temperature (T_m) of *Vc*DncV shifting ~13°C (Figure 3.3A). In contrast, the T_m of *Vc*DncV shifts only 4°C in the presence of ATP (Figure 3.3A). We observed no synergistic effect when equimolar amounts of ATP and GTP are present, indicating enhanced thermostability is dependent on donor nucleotide interactions (Figure 3.2G). Nucleotides exhibit a similar ability to stabilize the T_m of *Ec*CdnD, with ATP resulting in a ~21°C shift and GTP resulting in a 9°C shift (Figure 3.3A, Figure 3.2G). Enhanced stability of the *Ec*CdnD–ATP complex suggests that ATP is selectively used as the first donor nucleotide during 3′3′3′-cAAG formation. Notably, disruption of the *Ec*CdnD donor pocket residue Y250 required for nucleobase stacking ablates all ATP-induced enhancement of enzyme stability (Figure 3.3B, Figure 3.2H). In contrast, mutations to the acceptor pocket and auxiliary metal binding site again demonstrate these interactions have a minor effect on thermo-stabilization (Figure 3.3B, Figure 3.2G,H). These results reveal that a key determinant of CD-NTase nucleotide selectivity is preferential interactions leading to stabilization of an enzyme–substrate pre-reaction state.

Building on previous work with nucleotide analog interference mapping (Dash et al., 2006; Launer-Felty and Strobel, 2018; Rausch et al., 2009; Wang et al., 2021), we next used a panel of 14 nucleotide analogs to determine how individual contacts impact CD-NTase

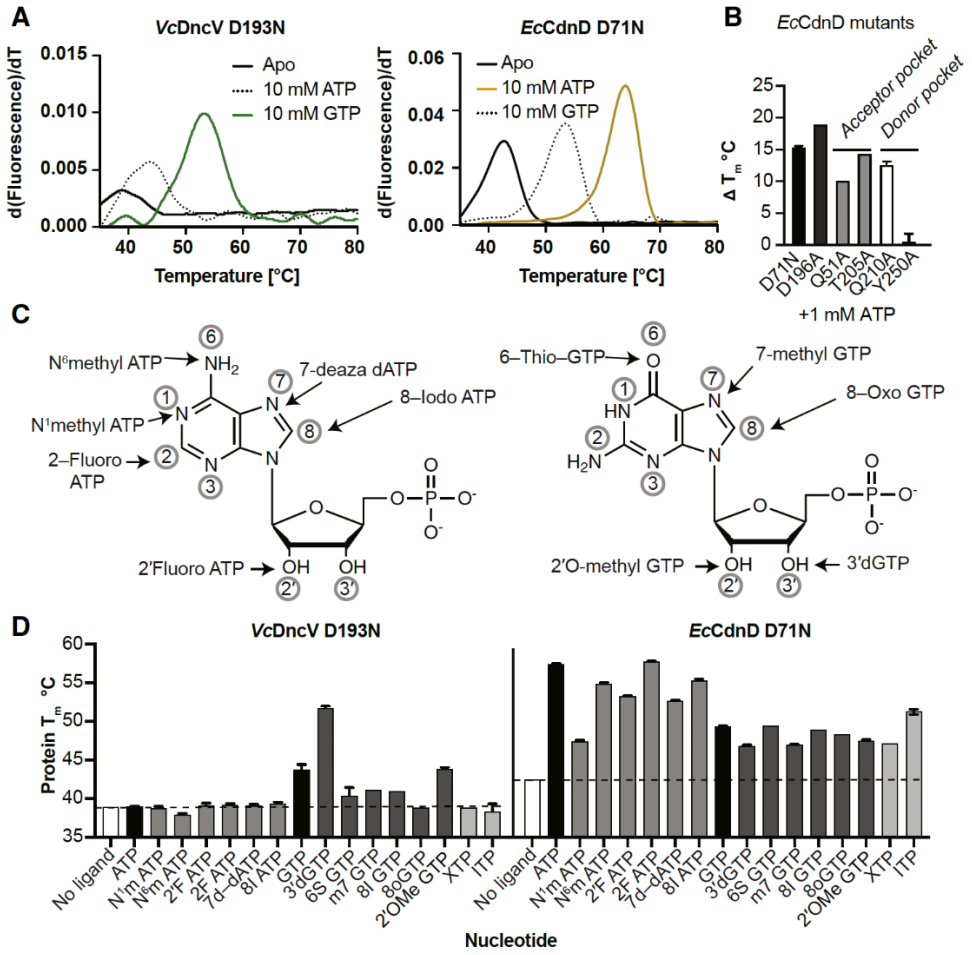


Figure 3.3. Stable CD-NTase–donor nucleotide complex formation is a key determinant of product specificity

(A) Left, Protein fluorescence over change in temperature for catalytically inactive *Vc*DnV D193N. Peak indicates protein melting temperature (T_m) with a rightward shift indicating greater stabilization. GTP, which initially binds in the donor pocket in the pre-reactive state, stabilizes *Vc*DncV to a greater extent than ATP. Right, T_m plot for catalytically inactive *Ec*CdnD D71N, showing that ATP stabilizes to a greater extent than GTP. Data are representative of three independent experiments. **(B)** ΔT_m of *Ec*CdnD mutants bound to 1 mM ATP. $\Delta T_m = T_m$(1 mM ATP) – T_m(Apo). D71N and D196A are capable of binding ATP. Y250A is not further stabilized upon addition of ATP and is likely incapable of binding ATP. Data are the mean of two independent experiments and error bars denote the standard deviation. **(C)** Schematic of purine nucleotides AMP and GMP with atoms numbered around the base. Chemically modified nucleotides and relevant nucleotide positions are indicated. **(D)** Comparison of protein T_m for *Vc*DncV D193N and *Ec*CdnD D71N apo or bound to 1 mM ATP, GTP, ITP, XTP, or modified nucleotide as indicated. Data plotted are the mean of two independent experiments and error bars denote the standard deviation.

substrate selectivity. The nucleotide panel includes ATP and GTP analogs with nucleobase and ribose chemical modifications (Figure 3.3C, Figure 3.2I, J) and reveals major features controlling CD-NTase specificity. First, the data demonstrate that both *Vc*DncV and *Ec*CdnD are tolerant to a wide variety of nucleotide modifications (Figure 3.3D). Chemical modifications to the Watson-Crick and Hoogsteen edges do not disrupt binding and enhancement of CD-NTase thermostability (Figure 3.3D, Figure 3.2K). Compared to previous studies with GGDEF- and DisA-family enzymes of cyclic dinucleotide synthases (Chan et al., 2004; Rosenberg et al., 2015; Witte et al., 2008), the wide tolerance of CD-NTases to nucleotide modifications is likely a result of the necessity to accommodate multiple substrate orientations within a single active site (Kranzusch, 2019; Whiteley et al., 2019). In agreement with this hypothesis, we observed that the *Ec*CdnD active site used for cyclic trinucleotide synthesis is more accommodating of modified nucleotides compared to the *Vc*DncV active site used for cyclic dinucleotide synthesis (Figure 3.3D).

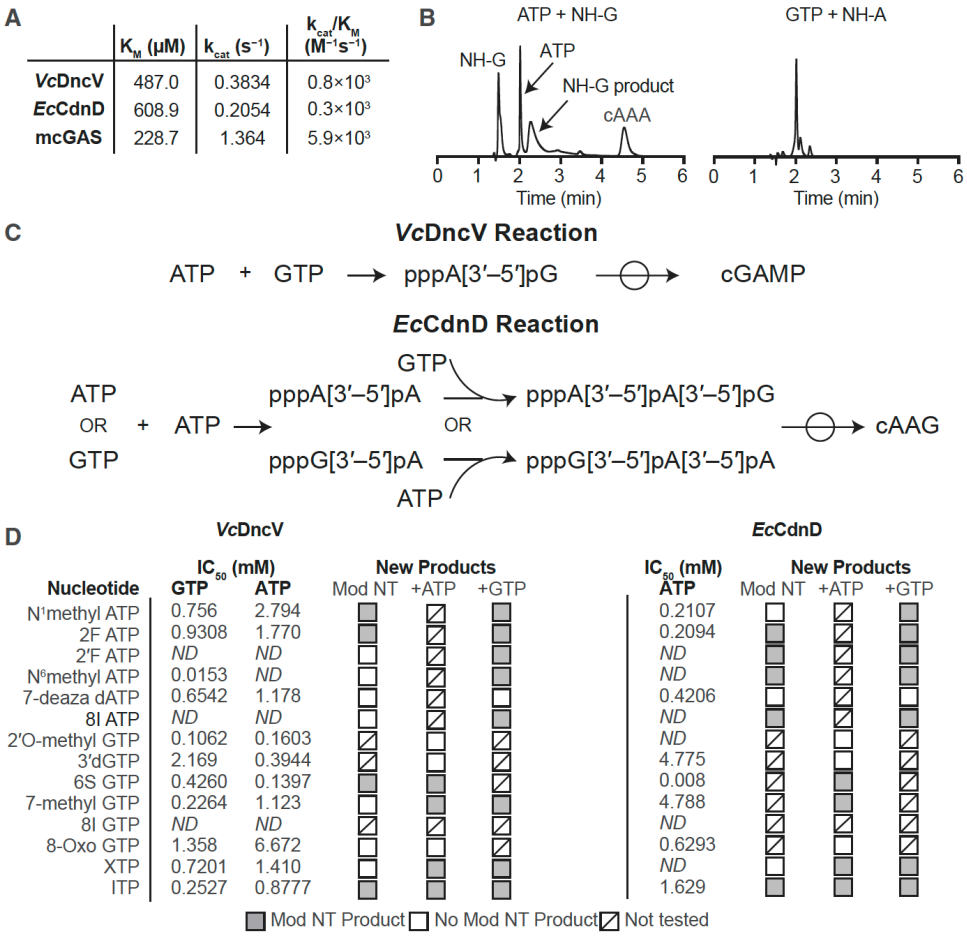
A second major feature of CD-NTase nucleotide specificity is preferential interaction with the nucleobase Watson-Crick edge. Both *Vc*DncV and *Ec*CdnD maintain specific interactions with the ATP N1 position and the GTP N1 and N2 positions. *Vc*DncV in complex with inosine triphosphate (ITP), which shares the N1 and O6 position with GTP but lacks the amine group at position 2, has a similar T_m to the *Vc*DncV–ATP complex (Figure 3.3D). The enhanced *Vc*DncV thermostability upon interaction with GTP is therefore dependent on hydrogen bond contacts with the N2 amine. *Ec*CdnD in complex with ITP results in a T_m shift that is intermediate between that observed with ATP or GTP, agreeing with the overall less pronounced ability of *Ec*CdnD to preferentially interact with GTP (Figure 3.3D). Xanthine triphosphate (XTP), which adds an additional oxygen atom to the N2 position and includes a hydrogen bond donor at both positions 1 and 3, destabilizes *Vc*DncV while slightly increasing the stability of *Ec*CdnD (Figure 3.3D, Figure 3.2K). For the ATP Watson-Crick edge, modifications to the N1 position destabilize the interaction with both *Vc*DncV and *Ec*CdnD while modifications to the N6 position destabilize interaction with *Vc*DncV but only have a minimal effect on *Ec*CdnD stability (Figure 3.3D, Figure 3.2K). Strikingly,

modification of the N2 position of ATP with 2-fluoro-ATP does not impair the ability of ATP to stabilize either enzyme (Figure 3.3D, Figure 3.2K). Overall, these interactions with modified nucleotides agree with contacts observed in the crystal structures of *Ec*CdnD and *Vc*DncV intermediate states (Figure 3.1) (Kranzusch et al., 2014; Zhu et al., 2014), and demonstrate specific CD-NTase donor and acceptor pocket interactions controlling nucleotide selection.

Lastly, nucleotide analogs provide insight into the importance of nucleobase positions that do not form clear contacts in existing CD-NTase structures. Substituting the GTP O6 hydrogen bond acceptor position with 6-thio-GTP results in a decrease in the stability of *Vc*DncV but not *Ec*CdnD (Figure 3.3D). Masking the N7 purine ring position destabilizes interaction with both CD-NTases suggesting transient recognition of this position could play a role in discriminating against pyrimidine nucleotides (Figure 3.3D, Figure 3.2K). Notably, chemical modifications at nucleobase position 8 reveal additional differences between *Vc*DncV and *Ec*CdnD with 8-iodo-GTP and 8-oxo-GTP stabilizing *Ec*CdnD while only 8-iodo-GTP stabilizes *Vc*DncV and 8-oxo-GTP has a specific destabilizing effect (Figure 3.3D). Likewise, *Vc*DncV tolerates ribose 2′ or 3′ OH modifications while *Ec*CdnD is destabilized (Figure 3.3D). Taken as a whole, detailed analysis of CD-NTase interactions with nucleotide analogs reveals key determinants of substrate specificity and serves as a guide for further structural analysis of CD-NTase enzymes.

Preferential selection of donor pocket nucleotides dictates product specificity

We next used high-performance liquid chromatography (HPLC) and mass spectrometry to analyze how nucleotide binding and stabilization influences CD-NTase product formation. Under optimized conditions for *in vitro* activation of bacterial CD-NTase enzymes (Whiteley et al., 2019), we observed that *Vc*DncV and *Ec*CdnD are ~10× slower than DNA-activated mouse cGAS or human cGAS (Figure 3.4A, Figure 3.5A, B) (Zhou et al., 2018). CD-NTase enzymes recognize an undefined cue required for activation *in vivo* (Cohen et al., 2019; Lowey et al., 2020; Ye et al., 2020), suggesting that in some cases, further increase in CD-NTase reaction rate may require



	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
***Vc*DncV**	487.0	0.3834	0.8×10^3
***Ec*CdnD**	608.9	0.2054	0.3×10^3
mcGAS	228.7	1.364	5.9×10^3

Figure 3.4. Nucleotide interference mapping analysis of CD-NTase reaction progression
(A) Comparison of enzyme kinetics between bacterial synthases *Vc*DncV and *Ec*CdnD, and mouse cGAS activated using equimolar 45 bp immunostimulatory DNA. The catalytic efficiency (k_{cat}/K_M) of the bacterial CD-NTases under these conditions is ~10× less than mcGAS. **(B)** High-performance liquid chromatography analysis of *Ec*CdnD reactions. Left, ATP and non-hydrolyzable Gpcpp (NH-G) form a product containing NH-G that cannot proceed to cyclization and cyclic tri-AMP (3′3′3′-cAAA). Right, GTP and non-hydrolyzable Apcpp (NH-A) do not react to form a major product containing NH-A. The carbon atom prevents Apcpp from donating a phosphate, and GTP cannot efficiently act as the initial phosphate donor. Reactions were eluted from a C18 column with mobile phase 94% NaH_2PO_4 and 6% Acetonitrile. Data are representative of three independent experiments. **(C)** Schematic comparison of reaction order for *Vc*DncV 3′3′-cGAMP synthesis and *Ec*CdnD 3′3′3′-cAAG synthesis. The circle symbol indicates cyclization of the linear intermediate into a final cyclic product. **(D)** Summary of *Vc*DncV and *Ec*CdnD reactions with modified nucleotides, or modified nucleotides with ATP or GTP as indicated. For each modified nucleotide reaction, IC_{50} values are calculated for 50% inhibition of c-di-AMP or c-di-GMP formation (*Vc*DncV) or 50% inhibition of 3′3′3′-cAAA formation (*Ec*CdnD), with increasing concentrations of modified nucleotides. See supplemental information for more detail. *ND*, not determined.

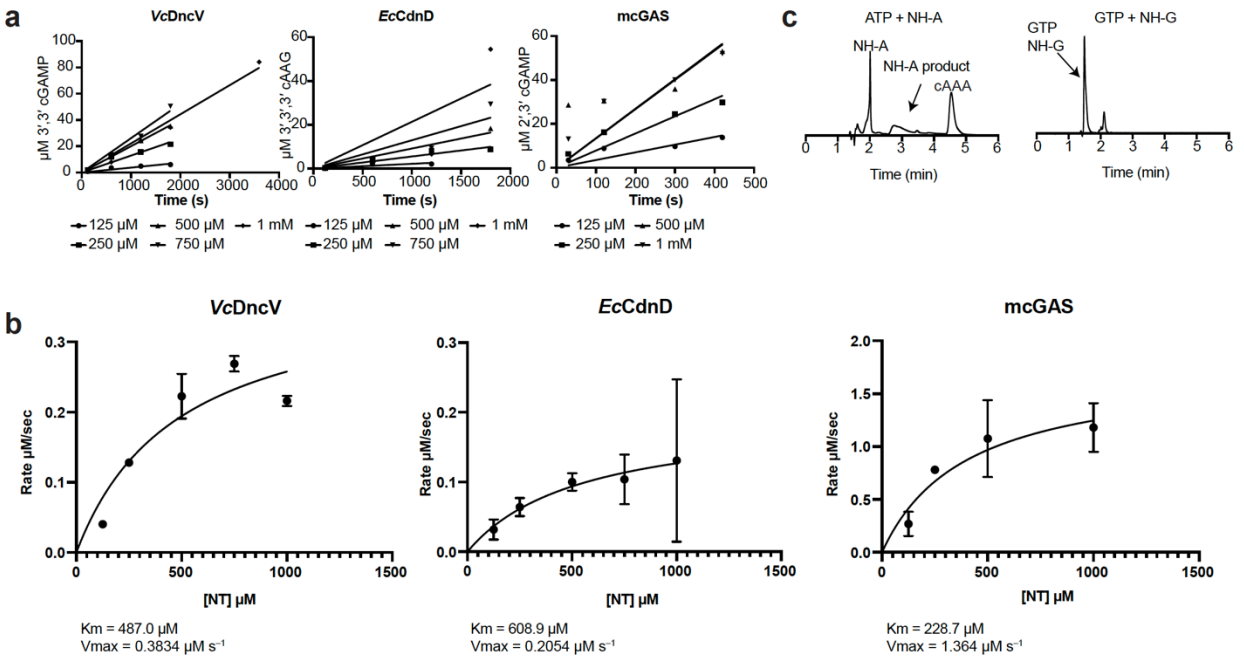


Figure 3.5. Measurement of *Vc*DncV and *Ec*CdnD reaction rates
(A) Best-fit lines and quantification for rates of product formation during VcDncV, EcCdnD and mcGAS reactions over time with increasing concentrations of ATP and GTP. Reactions were quantified by integrating product peaks for reactions analyzed by HPLC and comparing to a standard curve of chemically synthesized 2′3′-cGAMP, 3′3′-cGAMP, or 3′3′3′-cAAG. Data points are representative of two independent experiments for each concentration and time point indicated. **(B)** Michaelis-Menten kinetics plotted for VcDncV, EcCdnD, and mcGAS. Data points are the mean of two independently calculated rates of product formation at each concentration and error bars indicate standard deviation. **(C)** High-performance liquid chromatography analysis of EcCdnD reactions. Left, ATP and NH-A form a product containing NH-A and 3′3′3′-cAAA. Right, GTP and NH-G do not efficiently form a product. Reactions were eluted from a C18 column with mobile phase 94% NaH_2PO_4 and 6% acetonitrile. Data are representative of three independent experiments.

additional signals. Using nonhydrolyzable nucleotide analogs Apcpp or Gpcpp, we observed that the initial *Ec*CdnD reaction steps occur with ATP and Gpcpp but not with GTP and Apcpp (Figure 3.4B; Table 3.2). Furthermore, in the presence of ATP and Apcpp, *Ec*CdnD forms a linear reaction intermediate, while GTP and Gpcpp are not sufficient for *Ec*CdnD product formation (Figure 3.5C). These results confirm that *Ec*CdnD uses ATP as the initial donor nucleotide followed by selection of GTP and reveal that cyclic trinucleotide synthesizing CD-NTases also use a defined reaction order (Figure 3.4C) (Kranzusch et al., 2014; Whiteley et al., 2019).

As previously demonstrated for *Vc*DncV (Launer-Felty and Strobel, 2018), we observed

Table 3.2 Mass Spectrometry analysis of CD-NTase reactions

Nucleotides	Reaction step	Expected major product	Chemical formula	Product Molecular mass	Found Mass
*Ec*CdnD reaction order					
ATP + GTP	Complete	c[A(3',5')pA(3',5')pG(3',5')p]	$C_{30}H_{36}N_{15}O_{19}P_3$	1003.15	1003.153
ATP	Complete	c[A(3',5')pA(3',5')pG(3',5')p]	$C_{30}H_{36}N_{15}O_{18}P_3$	987.16	987.16
GTP	No product expected	c[G(3',5')pG(3',5')pG(3',5')p]	$C_{30}H_{36}N_{15}O_{21}P_3$	1035.14	522.991 1035.14
NH-A + ATP	Donor ATP	ppcpA(3',5')pA(3',5')pA, ppcpA(3',5')pA	$C_{31}H_{42}N_{15}O_{24}P_5$ $C_{21}H_{30}N_{10}O1_8P_4$	1163.12, 834.07	834.07
NH-A + GTP	No donor GTP	ppcpA(3',5')pG	$C_{21}H_{30}N_{10}O_{19}P_4$	850.06	Some 850.06, 363.058
NH-G + ATP	Donor ATP	ppcpG(3',5')pA(3',5')pA, ppcpG(3',5')pA	$C_{31}H_{42}N_{15}O_{25}P_5$ $C_{21}H_{30}N_{10}O_{19}P_4$	1179.12, 850.06	850.06, 521.011
*Vc*DncV					
N1 methyl ATP	VcDncV	c[N1 methyl A(3',5')pN1 methyl A(3',5')p]	$C_{22}H_{28}N_{10}O_{12}P_2$	686.14	N1 methyl ATP
N6 methyl ATP	VcDncV	c[N6 methyl A(3',5')pN6 methyl A(3',5')p]	$C_{22}H_{28}N_{10}O_{12}P_2$	686.14	N6 methyl ATP
2F ATP	VcDncV	c[2F A(3',5')p2F A(3',5')p]	$C_{20}H_{22}F_2N_{10}O_{12}P_2$	694.09	2F ATP
2'F ATP	VcDncV	c[2'F A(3',5')p2'F A(3',5')p]	$C_{20}H_{22}F_2N_{10}O_{12}P_2$	662.1	2'F ATP
7 deaza dATP	VcDncV	No product expected	$C_{11}H_{17}N_4O_{12}P_3$	490.01	7 deaza dATP
8I ATP	VcDncV	c[8I A(3',5')p8I A(3',5')p]	$C_{20}H_{22}I_2N_{10}O_{12}P_2$	909.9	8I ATP
6S GTP	VcDncV	c[6S G(3',5')p6S G(3',5')p]	$C_{20}H_{24}N_{10}O_{12}P_2S_2$	722.05	6S GTP
Methyl-7 GTP	VcDncV	No product expected	$C_{11}H_{18}N_5O_{14}P_3$	537.01	Methyl-7 GTP
8-oxo GTP	VcDncV	No product expected	$C_{10}H_{16}N_5O_{15}P_3$	538.99	8-oxo GTP
XTP	VcDncV	No product expected	$C_{10}H_{15}N_4O_{15}P_3$	523.97	XTP
ITP	VcDncV	c[I(3',5')pI(3',5')p]	$C_{20}H_{22}N_8O_{14}P_2$	660.07	ITP
N1 methyl ATP + GTP	VcDncV	c[N1 methyl A(3',5')pG(3',5')p]	$C_{21}H_{26}N_{10}O_{13}P_2$	688.12	N1 methyl ATP + GTP
N6 methyl ATP + GTP	VcDncV	c[N6 methyl A(3',5')pG(3',5')p]	$C_{21}H_{26}N_{10}O_{13}P_2$	688.12	N6 methyl ATP + GTP
2F ATP + GTP	VcDncV	c[2F A(3',5')pG(3',5')p]	$C_{20}H_{23}FN_{10}O_{13}P_2$	692.09	2F ATP + GTP
2'F ATP + GTP	VcDncV	c[2'F A(3',5')pG(3',5')p]	$C_{20}H_{23}FN_{10}O_{12}P_2$	676.1	2'F ATP + GTP
7 deaza dATP + GTP	VcDncV	No product expected	$C_{11}H_{17}N_4O_{12}P_3$	490.01	7 deaza dATP + GTP
8I ATP + GTP	VcDncV	c[8I A(3',5')pG(3',5')p]	$C_{20}H_{23}IN_{10}O_{13}P_2$	800	8I ATP + GTP
3'dGTP + ATP	VcDncV	No product expected	$C_{20}H_{26}N_{10}O_{14}P_2$	507	3'dGTP + ATP
6S GTP + ATP	VcDncV	c[A(3',5')p6S G(3',5')p]	$C_{20}H_{24}N_{10}O_{12}P_2S$	690.08	6S GTP + ATP
Methyl-7 GTP + ATP	VcDncV	c[A(3',5')pm7G(3',5')p]	$C_{21}H_{27}N_{10}O_{13}P_2+$	689.12	Methyl-7 GTP + ATP
8-oxo GTP + ATP	VcDncV	c[A(3',5')8oxoG(3',5')p]	$C_{20}H_{24}N_{10}O_{14}P_2$	690.09	8-oxo GTP + ATP
XTP + ATP	VcDncV	c[A(3',5')pX(3',5')p]	$C_{20}H_{23}N_9O_{14}P_2$	675.08	XTP + ATP
ITP + ATP	VcDncV	c[A(3',5')pI(3',5')p]	$C_{20}H_{23}N_9O_{13}P_2$	659.09	ITP + ATP
XTP + GTP	VcDncV	c[X(3',5')pG(3',5')p]	$C_{20}H_{23}N_9O_{15}P_2$	691.08	XTP + GTP
ITP + GTP	VcDncV	c[I(3',5')pG(3',5')p]	$C_{20}H_{23}N_9O_{14}P_2$	675.08	ITP + GTP
Methyl-7 GTP + GTP	VcDncV	c[G(3',5')pm7G(3',5')p]	$C_{21}H_{27}N_{10}O_{14}P_2$	705.1	Methyl-7 GTP + GTP

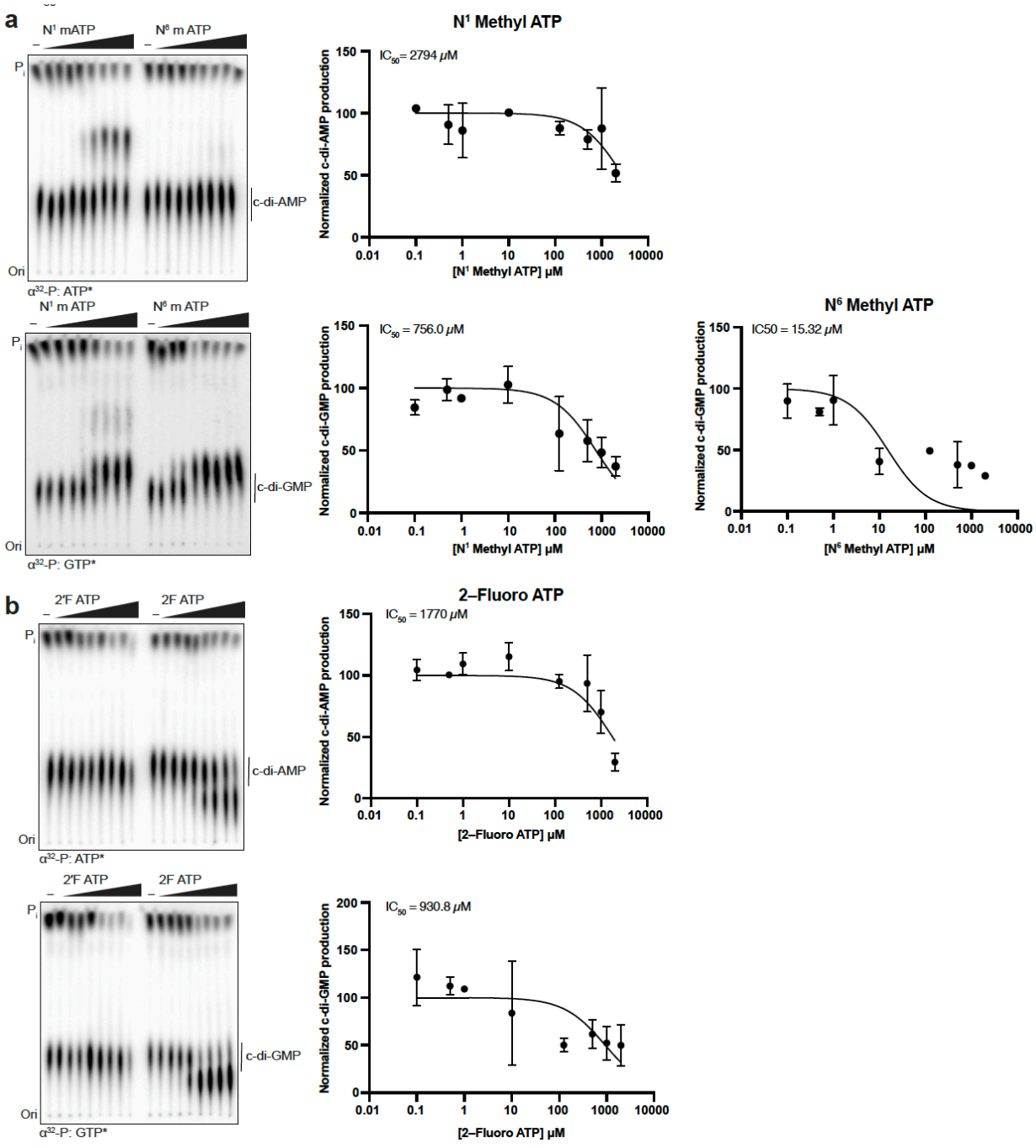
Table 3.2 Mass Spectrometry analysis of CD-NTase reactions, continued

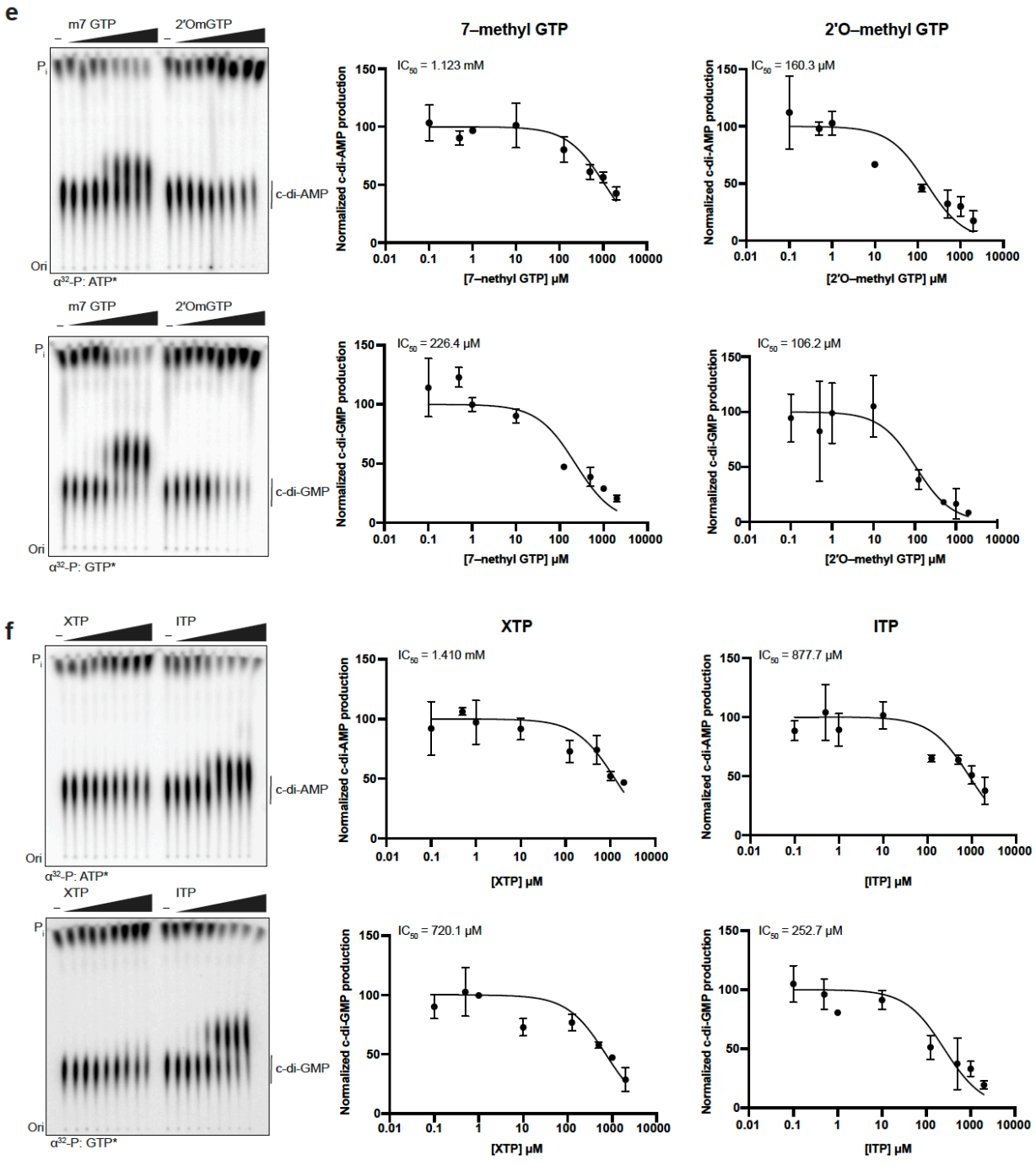
Nucleotides	Reaction step	Expected major product	Chemical formula	Product Molecular mass	Found Mass
*Ec*CdnD					
N1 methyl ATP	EcCdnD	c[N1 methyl A(3',5')pN1 methyl A(3',5')pN1 methyl A(3',5')p]	$C_{33}H_{42}N_{15}O_{18}P_3$	1029.205	N1 methyl ATP
N6 methyl ATP	EcCdnD	c[N6 methyl A(3',5')pN6 methyl A(3',5')pN6 methyl A(3',5')p]	$C_{33}H_{42}N_{15}O_{18}P_3$	1029.205	N6 methyl ATP
2F ATP	EcCdnD	c[2F A(3',5')p2F A(3',5')p2F A(3',5')p]	$C_{20}H_{22}F_2N_{10}O_{12}P_2$	1041.13	2F ATP
2'F ATP	EcCdnD	c[2'F A(3',5')p2'F A(3',5')p2'F A(3',5')p]	$C_{30}H_{33}F_3N_{15}O_{15}P_3$	993.145	2'F ATP
7 deaza dATP	EcCdnD	No product expected	$C_{11}H_{17}N_4O_{12}P_3$	490.01	7 deaza dATP
8I ATP	EcCdnD	c[8I A(3',5')p8I A(3',5')p8I A(3',5')p]	$C_{30}H_{33}I_3N_{15}O_{18}P_3$	1364.85	8I ATP
N1 methyl ATP + GTP	EcCdnD	c[N1 methyl A(3',5')pN1 methyl A(3',5')pG(3',5')p]	$C_{32}H_{40}N_{15}O_{19}P_3$	1031.184	N1 methyl ATP + GTP
N6 methyl ATP + GTP	EcCdnD	c[N6 methyl A(3',5')pN6 methyl A(3',5')pG(3',5')p]	$C_{32}H_{40}N_{15}O_{19}P_3$	1031.184	N6 methyl ATP + GTP
2F ATP + GTP	EcCdnD	c[2F A(3',5')p2F A(3',5')pG(3',5')p]	$C_{30}H_{34}F_2N_{15}O_{19}P_3$	1039.134	2F ATP + GTP
2'F ATP + GTP	EcCdnD	c[2'F A(3',5')p2'F A(3',5')pG(3',5')p]	$C_{30}H_{34}F_2N_{15}O_{17}P_3$	1007.14	2'F ATP + GTP
7 deaza dATP + GTP	EcCdnD	No product expected	$C_{11}H_{17}N_4O_{12}P_3$	490.01	7 deaza dATP + GTP
8I ATP + GTP	EcCdnD	c[8I A(3',5')p8I A(3',5')pG(3',5')p]	$C_{30}H_{34}I_2N_{15}O_{19}P_3$	1254.95	8I ATP + GTP
3'dGTP + ATP	EcCdnD	No product expected	$C_{10}H_{16}N_5O_{13}P_3$	507	3'dGTP + ATP
6S GTP + ATP	EcCdnD		$C_{30}H_{36}N_{15}O_{18}P_3S$	1019.13	6S GTP + ATP
Methyl-7 GTP + ATP	EcCdnD	c[A(3',5')pA(3',5')pMethyl-7 G(3',5')p]	$C_{31}H_{39}N_{15}O_{19}P_3+$	1018.175	Methyl-7 GTP + ATP
2'O-methyl GTP + ATP	EcCdnD	No product expected	$C_{11}H_{18}N_5O_{14}P_3$	537.01	2'O-methyl GTP + ATP
XTP	EcCdnD	c[X(3',5')pX(3',5')pX(3',5')p]	$C_{30}H_{33}N_{12}O_{24}P_3$	1038.09	XTP
ITP	EcCdnD	c[I(3',5')pI(3',5')pI(3',5')p]	$C_{30}H_{33}N_{12}O_{21}P_3$	990.11	ITP
XTP + GTP	EcCdnD	c[X(3',5')pX(3',5')pG(3',5')p]	$C_{30}H_{34}N_{13}O_{23}P_3$	1037.11	XTP + GTP
ITP + GTP	EcCdnD	c[I(3',5')pI(3',5')pG(3',5')p]	$C_{30}H_{34}N_{13}O_{21}P_3$	1005.121	ITP + GTP
XTP + ATP	EcCdnD	c[A(3',5')pA(3',5')pX(3',5')p]	$C_{30}H_{35}N_{14}O_{20}P_3$	1004.14	XTP + ATP
ITP + ATP	EcCdnD	c[A(3',5')pA(3',5')pI(3',5')p] ; c[A(3',5')pI(3',5')pI(3',5')p]	$C_{30}H_{34}N_{13}O_{20}P_3$ $C_{30}H_{35}N_{14}O_{19}P_3$	989.13, 988.14	ITP + ATP
8-oxo GTP + ATP	EcCdnD	c[A(3',5')pA(3',5')p8oxoG(3',5')p]	$C_{30}H_{36}N_{15}O_{20}P_3$	1019.15	8-oxo GTP + ATP

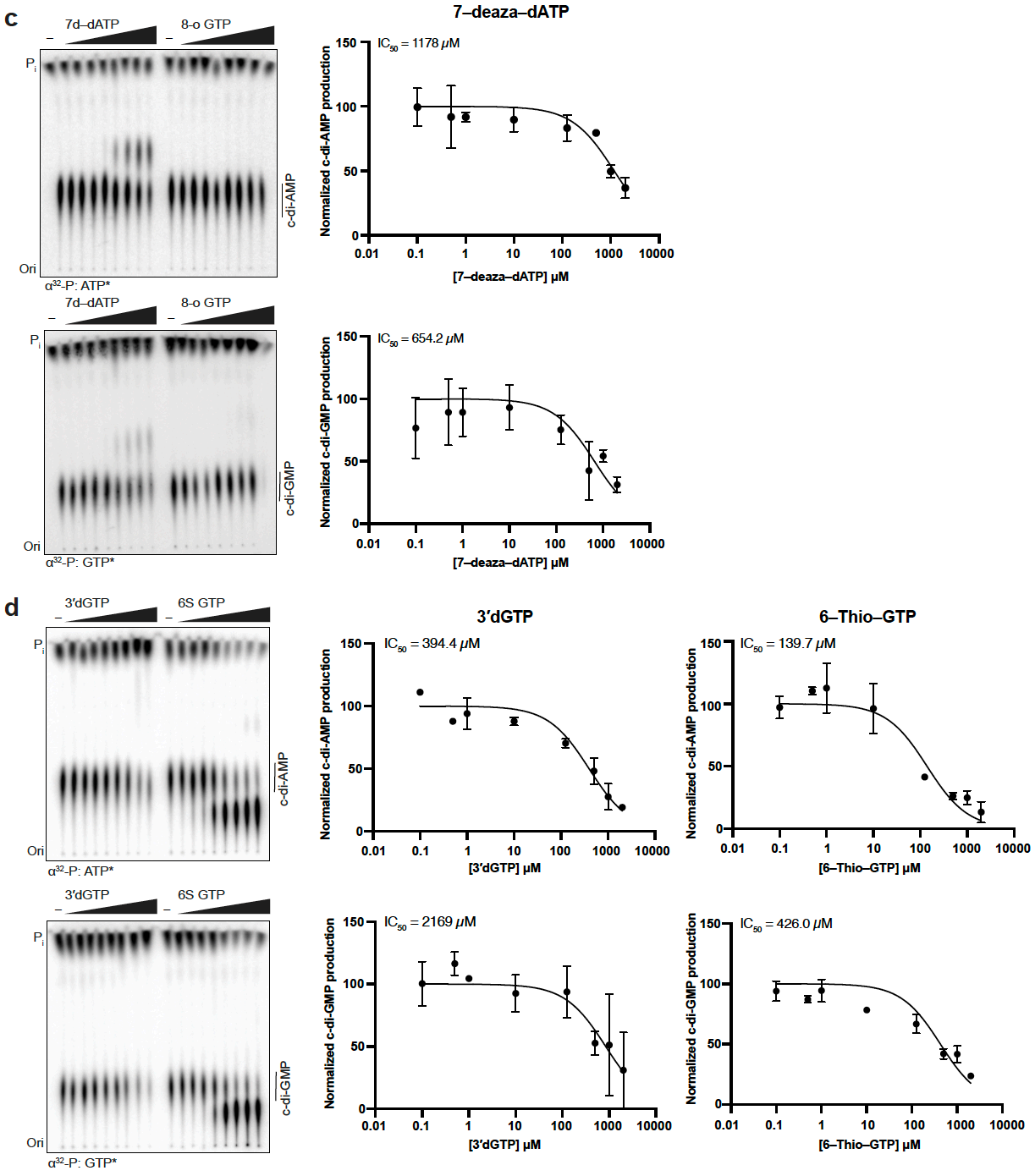
that both *Vc*DncV and *Ec*CdnD are capable of incorporating nucleotide analogs and synthesizing a diverse array of cyclic oligonucleotide species (Figure 3.4D, Table 3.2, Figure 3.8). For each nucleotide analog, we calculated the IC_{50} as the concentration of nucleotide required to decrease product formation by 50% (Figure 3.4D, Figure 3.6). In each case, nucleotide analog incorporation is more efficient in the presence of the correct ATP or GTP partner nucleotide indicating

Figure 3.6. IC50 quantification of modified nucleotide incorporation correlates with modified nucleotide product formation

(A–F) Thin-layer chromatography analysis and quantification of pixel intensity for each *Vc*DncV reaction with increasing concentrations of modified nucleotide, as indicated, and radiolabeled α^{32}P-ATP. Tested modified nucleotide concentrations: 0, 100 nM, 500 nM, 1 µM, 10 µM, 125 µM, 500 µM, 1 mM, 2 mM with 12.5 µM ATP in each reaction. For all modified nucleotides, experimental data is representative of two independent experiments, average is plotted and error bars denote the standard deviation. **(G–M)** Thin-layer chromatography analysis and quantification of pixel intensity for each *Ec*CdnD reaction with increasing concentrations of modified nucleotide, as indicated, and radiolabeled α^{32}P-ATP. Tested modified nucleotide concentrations: 0, 100 nM, 500 nM, 1 µM, 10 µM, 125 µM, 500 µM, 1 mM, 2 mM with 12.5 µM ATP in each reaction. For all modified nucleotides, experimental data is representative of two independent experiments, average is plotted and error bars denote the standard deviation.


a
N¹ mATP
N⁶ m ATP
Pi
Ori
c-di-AMP
α32-P: ATP*
N¹ m ATP
N⁶ m ATP
Pi
Ori
c-di-GMP
α32-P: GTP*
N¹ Methyl ATP
IC50= 2794 μM
Normalized c-di-AMP production
[N¹ Methyl ATP] μM
IC50 = 756.0 μM
Normalized c-di-GMP production
[N¹ Methyl ATP] μM
N⁶ Methyl ATP
IC50 = 15.32 μM
Normalized c-di-GMP production
[N⁶ Methyl ATP] μM
b
2'F ATP
2F ATP
Pi
Ori
c-di-AMP
α32-P: ATP*
2'F ATP
2F ATP
Pi
Ori
c-di-GMP
α32-P: GTP*
2–Fluoro ATP
IC50 = 1770 μM
Normalized c-di-AMP production
[2–Fluoro ATP] μM
IC50 = 930.8 μM
Normalized c-di-GMP production
[2–Fluoro ATP] μM


e
m7 GTP
2'OmGTP
P i
Ori
c-di-AMP
α32-P: ATP*
7–methyl GTP
2'O–methyl GTP
IC50 = 1.123 mM
IC50 = 160.3 μM
Normalized c-di-AMP production
[7–nethyl GTP] μM
[2'O–methyl GTP] μM
c-di-GMP
α32-P: GTP*
IC50 = 226.4 μM
IC50 = 106.2 μM
Normalized c-di-GMP production
f
XTP
ITP
IC50 = 1.410 mM
IC50 = 877.7 μM
[XTP] μM
[ITP] μM
IC50 = 720.1 μM
IC50 = 252.7 μM


c
7d–dATP
8-o GTP
P$_i$
Ori
c-di-AMP
α^{32}-P: ATP*
7d–dATP
8-o GTP
P$_i$
Ori
c-di-GMP
α^{32}-P: GTP*
7–deaza–dATP
IC$_{50}$ = 1178 μM
Normalized c-di-AMP production
[7–deaza–dATP] μM
IC$_{50}$ = 654.2 μM
Normalized c-di-GMP production
[7–deaza–dATP] μM
d
3′dGTP
6S GTP
P$_i$
Ori
c-di-AMP
α^{32}-P: ATP*
3′dGTP
6S GTP
P$_i$
Ori
c-di-GMP
α^{32}-P: GTP*
3′dGTP
IC$_{50}$ = 394.4 μM
Normalized c-di-AMP production
[3′dGTP] μM
6–Thio–GTP
IC$_{50}$ = 139.7 μM
Normalized c-di-AMP production
[6–Thio–GTP] μM
IC$_{50}$ = 2169 μM
Normalized c-di-GMP production
[3′dGTP] μM
IC$_{50}$ = 426.0 μM
Normalized c-di-GMP production
[6–Thio–GTP] μM

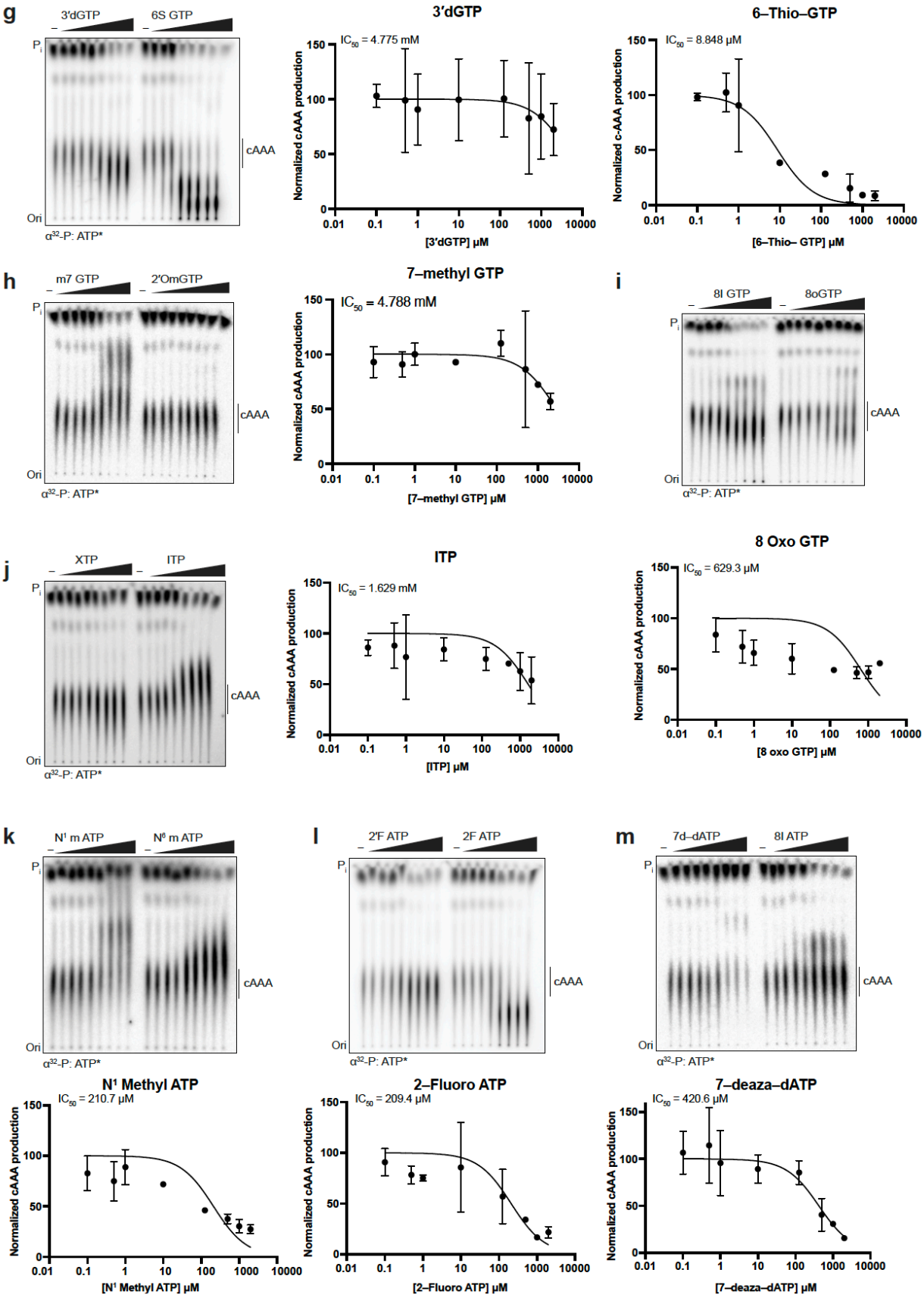
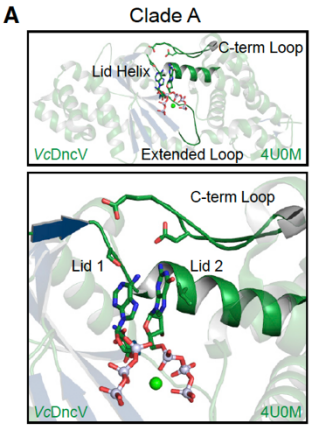

g
3′dGTP
6S GTP
P$_i$
cAAA
Ori
α^{32}-P: ATP*
3′dGTP
IC$_{50}$ = 4.775 mM
Normalized cAAA production
[3′dGTP] µM
6–Thio–GTP
IC$_{50}$ = 8.848 µM
Normalized c-AAA production
[6–Thio– GTP] µM
h
m7 GTP
2′OmGTP
7–methyl GTP
IC$_{50}$ = 4.788 mM
[7–methyl GTP] µM
i
8I GTP
8oGTP
j
XTP
ITP
ITP
IC$_{50}$ = 1.629 mM
[ITP] µM
8 Oxo GTP
IC$_{50}$ = 629.3 µM
[8 oxo GTP] µM
k
N^1 m ATP
N^6 m ATP
N^1 Methyl ATP
IC$_{50}$ = 210.7 µM
[N^1 Methyl ATP] µM
l
2′F ATP
2F ATP
2–Fluoro ATP
IC$_{50}$ = 209.4 µM
[2–Fluoro ATP] µM
m
7d–dATP
8I ATP
7–deaza–dATP
IC$_{50}$ = 420.6 µM
[7–deaza–dATP] µM

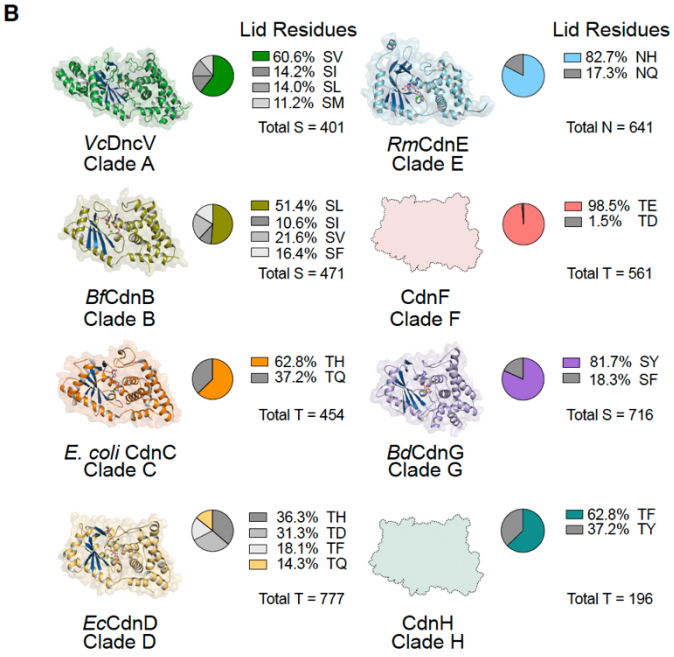
Figure 3.7. Structural and sequence-based comparison of divergent bacterial CD-NTases reveals correlates of nucleotide selection

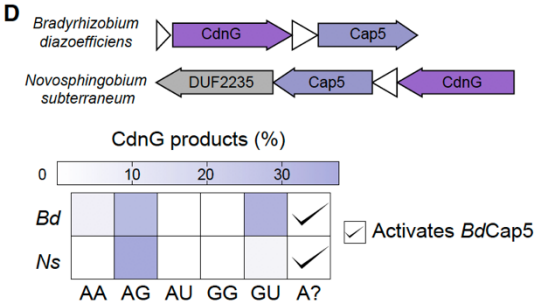
(A) *Vc*DncV structure 4U0M with highlighted structural features. Top, structure of *Vc*DncV showing the relative positioning of nucleotide coordinating structural features. Bottom, highlight of Lid Helix, lid residues 1 and 2 and the C-term loop with additional coordinating residues. **(B)** Comparative analysis of CD-NTase structure and distribution of lid residues 1 and 2, with pie charts showing pairwise occurrence of lid residues 1 and 2 tabulated using aligned sequences for each clade. For Clades A–E and G, the pie chart shows lid residue 1 as it appears in the representative CD-NTase for that clade (structure shown). For Clade F and H, the pie chart displays the most frequent lid residue 1. (*Vc, Vibrio cholerae; Bf, Bacteroides fragilis; E. coli, Escherichia coli; Ec, Enterobacter cloacae; Rm, Rhodothermus marinus; Bd, Bradyrhizobium diazoefficiens)* **(C)** Comparison of CD-NTases and products from different clades, highlighting residue XGS (ie, Q51 in *Ec*CdnD), active site, and lid residues 1 and 2. (*Fs, Flavobacteriaceae sp.; Lp, Legionella pneumophila; Ns, Novosphingobium subterraneum)*. **(D)** Top, *Bd* and *Ns* CBASS operon schematics containing a CD-NTase and a CD-NTase-associated protein (Cap) effector HNH-SAVED nuclease, described as Cap5. Bottom, summary heat map of major product formation and effector activation for *Bd*CdnG and *Ns*CdnG. **(E)** Graphical summary and model of the evolutionary relationship between CD-NTases, downstream effectors, and phage. CD-NTase product variability is balanced by effector selectivity. The influence of the effector is greater than the influence of the CD-NTase as effector activation is critical for an anti-phage response and survival of the bacterial colony.

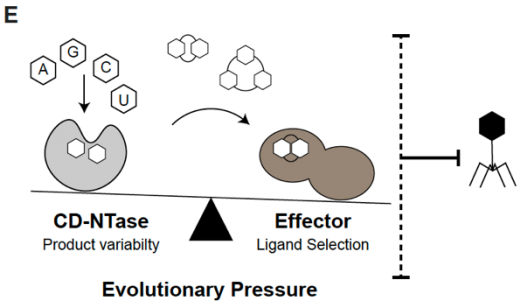


C

CD-NTase	Clade	XGSX	Active Site		Lid 1	Lid 2	Major Product
*Vc*DncV	A01	QGSF	MDIDD	IDV	S	V	3′3′-cGAMP
*Bf*CdnB	B04	QGSY	CDLDD	IDL	S	I	3′3′-cGAMP
E.coli CdnC	C01	SGSL	NDIDV	VDI	T	H	3′3′3′-cAAA
*Ec*CdnD	D02	QGSI	PDVDI	LDL	T	Q	3′3′3′-cAAG
*Fs*CdnE	E	QGSV	NDVDL	IDV	D	S	3′3′ c-di-GMP
*Em*CdnE	E01	QGSY	SDVDI	ADV	S	H	3′3′-cGAMP
*Rm*CdnE	E01	QGSY	SDVDV	ADV	N	H	3′3′ cUMP–AMP
*Lp*CdnE	E02	QGSY	SDVDI	ADV	N	N	3′3′ c-di-UMP
*Bd*CdnG	G10	QGSM	YDLDI	LDI	S	F	3′3′ cUMP–GMP
*Ns*CdnG	G10	QGSM	YDLDA	LDV	A	F	3′3′-cGAMP







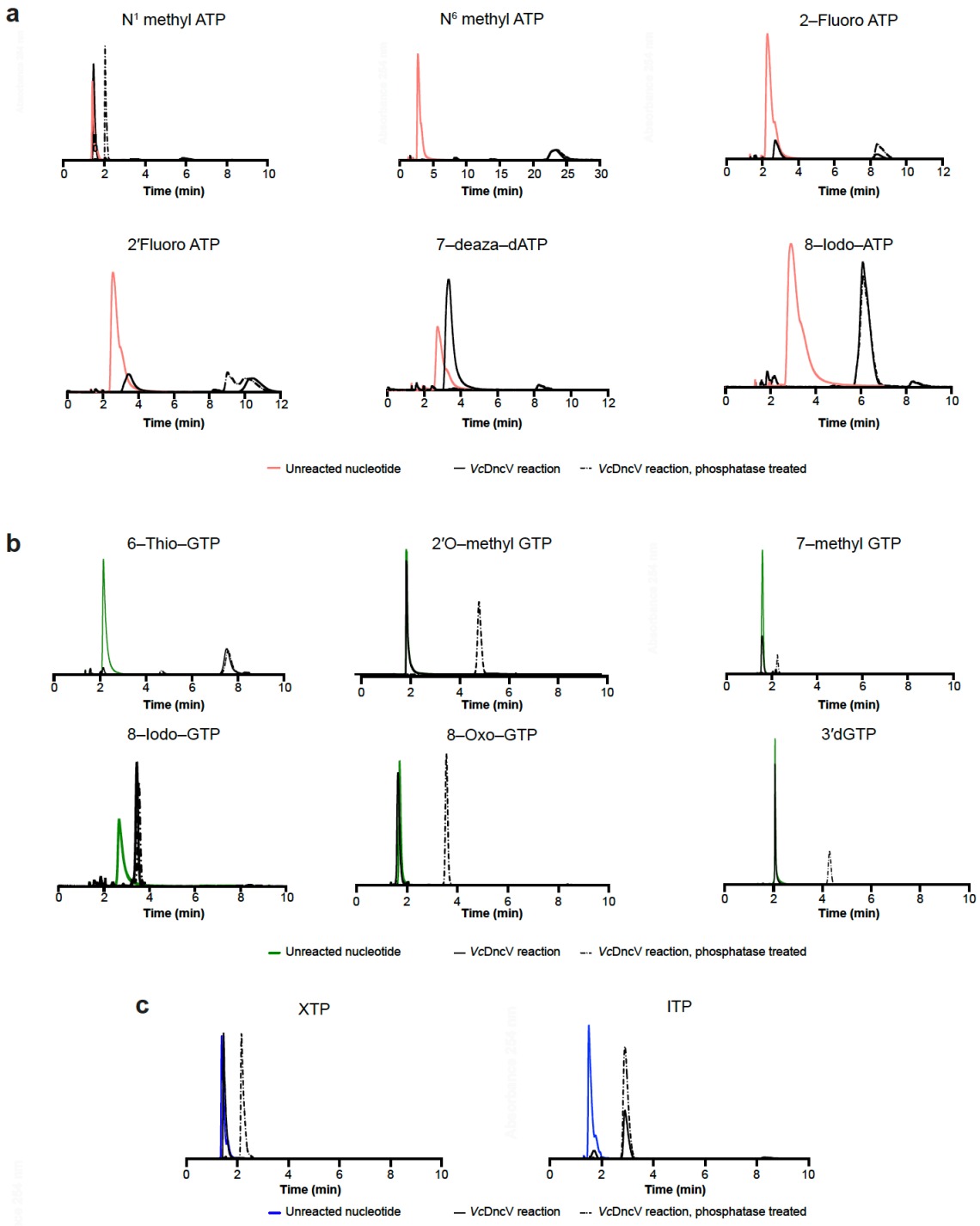


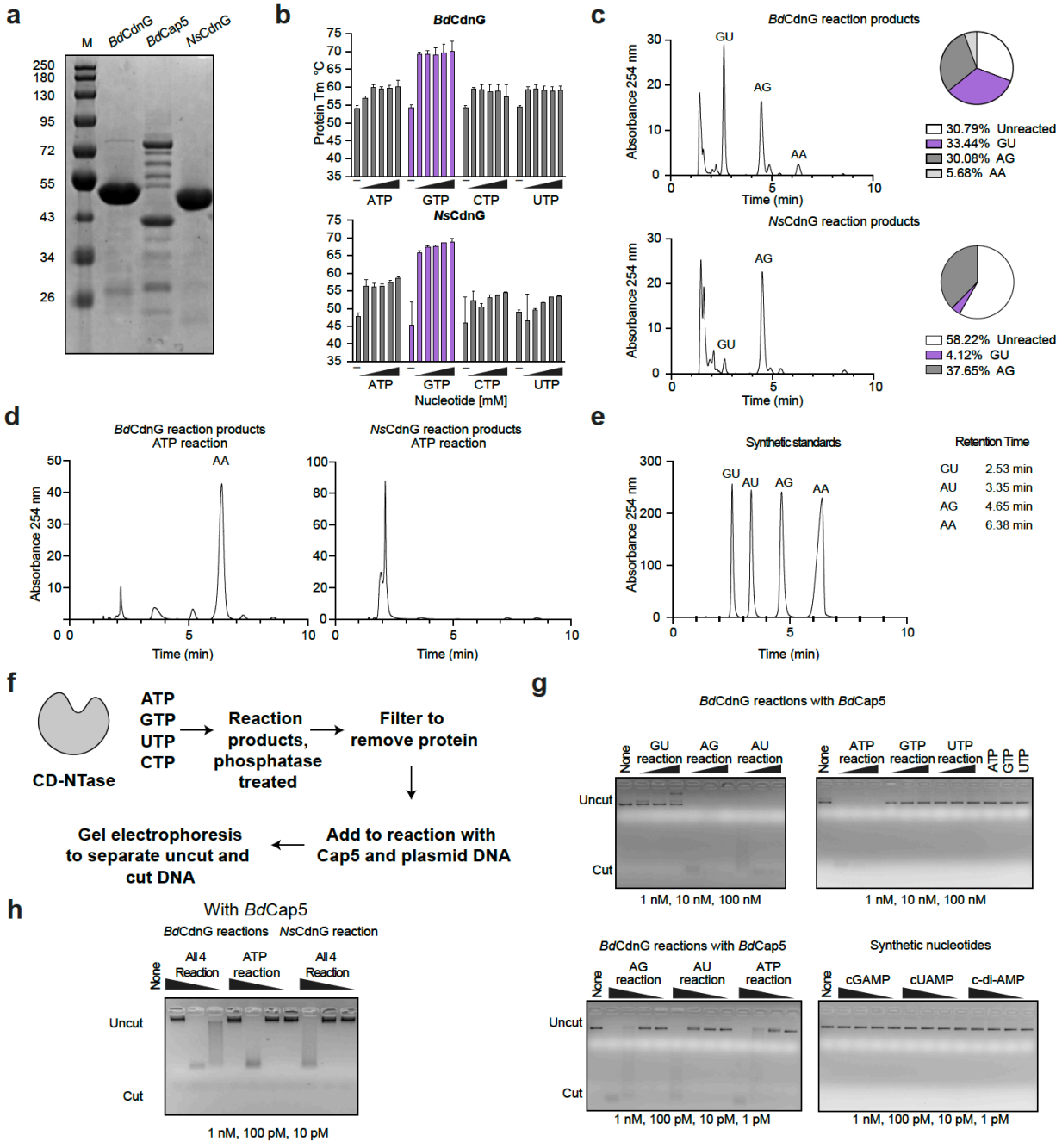
Figure 3.8. HPLC analysis of *Vc*DncV reactions with modified nucleotides
(A-C) High-performance liquid chromatography analysis of *Vc*DncV reactions with each modified nucleotide, as indicated. Reactions were eluted from a C18 column with mobile phase 97% NaH_2PO_4 and 3% acetonitrile. Reactions were treated with phosphatase as indicated. Cyclic products are phosphatase resistant and elute similarly to untreated reactions. Unreacted nucleotide is phosphatase-susceptible and treatment results in increased retention time.

maintenance of the preference for synthesizing mixed-base products. The reactions with nucleotide analogs largely fall into three groups. Nucleotide analogs not capable of binding CD-NTases (no modified product formation, IC_{50} >0.5 mM), including 7-deaza dATP for *Vc*DncV and 3′dGTP for *Ec*CdnD; nucleotides capable of binding but not reacting (no modified product formation, IC_{50} <0.5 mM) including 2'O-methyl GTP for *Vc*DncV and N1 methyl ATP for *Ec*CdnD; and nucleotides capable of binding and reacting (modified product formation, IC_{50} <0.5 mM) including 2F ATP for both *Vc*DncV and *Ec*CdnD. Interestingly, we observed *Ec*CdnD reactions with 6-thio-GTP exhibit enhanced product formation (IC_{50} <0.01 mM), suggesting some modified nucleotides are more efficiently incorporated (Figure 3.4D, Figure 3.6G). Together these results confirm our thermostability findings and further explain the donor and acceptor pocket interactions controlling CD-NTase product specificity.

CD-NTase structural analysis defines correlates of nucleotide second messenger specificity

Bacterial CD-NTase enzymes are remarkably diverse and cluster into eight major clades designated A–H (Millman et al., 2020; Whiteley et al., 2019). We used the *Ec*CdnD structure and existing CD-NTase crystal structures from Clades A, C, and E to compare conservation of nucleotide coordinating residues. As additional support for this analysis, we determined a 1.8 Å crystal structure of *Bacteroides fragilis* CdnB (*Bf*CndB, accession WP_032579276.1) and a 1.6 Å crystal structure of *Bradyrhizobium diazoefficiens* CdnG (*Bd*CdnG, accession NP_766712.1) providing the first CD-NTase structures from Clades B and G (Figure 3.7B; Table 3.1). Structural analysis demonstrates that each CD-NTase shares a conserved helix that lines the active-site lid and positions residues above the acceptor and donor nucleotide pockets, an unstructured region extending between the active-site GS-motif and first strand of the catalytic triad, and a second unstructured region of variable length that spans between two alpha-helices as part of the C-terminal lobe (Figure 3.7A). Lid residues correlate with major product synthesis (Figure 3.7B, C)

Figure 3.9. *Bd*CdnG effector Cap5 responds to A-containing minor product
(A) Coomassie stained gel showing expression level and purity of *Bd*CdnG, *Bd*Cap5, and *Ns*CdnG. (B) Plot of T_m for *Bd*CdnG and *Ns*CdnG with titrated ATP and GTP at 0, 0.5, 1, 2, 5,10 mM total nucleotide concentration. Bars are the mean of two independent experiments with error bars indicating standard deviation. (C) High-performance liquid chromatography analysis of *Bd*CdnG and *Ns*CdnG reactions. Top, *Bd*CdnG reactions containing ATP, GTP, UTP and CTP makes predominantly 3′3′-cGMP–UMP, followed by 3′3′-cGAMP and c-di-AMP. Top right, pie chart showing the breakdown of product formation calculated by integrating the area under the curve of each product peak. Unreacted peaks were integrated and summed to calculate bulk total unreacted amount. Bottom, *Ns*CdnG predominantly makes 3′3′-cGAMP and some 3′3′-cGMP–UMP. Bottom right, pie chart showing the breakdown of product formation calculated by integrating the area under the curve of each product peak. Data is representative of two independent experiments. (D) High-performance liquid chromatography analysis of *Bd*CdnG and *Ns*CdnG reactions containing only ATP. (E) High-performance liquid chromatography analysis of synthetic chemical standards of 3′3′-cGMP–UMP, 3′3′-cUMP–AMP, 3′3′-cGAMP, and c-di-AMP. (F) Schematic of Cap5 effector assay. (G) *Bd*CdnG reactions with *Bd*Cap5 effector assay as depicted in F. *Bd*CdnG Reaction with indicated nucleotide combinations or single nucleotide reactions were incubated with *Bd*Cap5 and empty pcDNA4 or pGEM plasmid DNA. Unreacted ATP, GTP, and UTP are included as controls. AG, AU, and ATP containing reactions activate *Bd*Cap5 at concentrations as low as 100 pM, while synthetic 3′3′-cGAMP, 3′3′-cAMP–UMP, and c-di-AMP do not activate *Bd*Cap5. (H) *Bd*CdnG or *Ns*CdnG reactions with *Bd*Cap5. Both *Bd*CdnG and *Ns*CdnG make an A-containing product that activates *Bd*Cap5.


a
M
BdCdnG
BdCap5
NsCdnG
250
180
130
95
72
55
43
34
26
b
BdCdnG
Protein Tm °C
ATP
GTP
CTP
UTP
NsCdnG
Nucleotide [mM]
c
BdCdnG reaction products
Absorbance 254 nm
GU
AG
AA
Time (min)
30.79% Unreacted
33.44% GU
30.08% AG
5.68% AA
NsCdnG reaction products
58.22% Unreacted
4.12% GU
37.65% AG
d
BdCdnG reaction products
ATP reaction
NsCdnG reaction products
ATP reaction
e
Synthetic standards
Retention Time
GU 2.53 min
AU 3.35 min
AG 4.65 min
AA 6.38 min
f
CD-NTase
ATP
GTP
UTP
CTP
Reaction products, phosphatase treated
Filter to remove protein
Add to reaction with Cap5 and plasmid DNA
Gel electrophoresis to separate uncut and cut DNA
g
BdCdnG reactions with BdCap5
None
GU reaction
AG reaction
AU reaction
ATP reaction
GTP reaction
UTP reaction
Uncut
Cut
1 nM, 10 nM, 100 nM
1 nM, 100 pM, 10 pM, 1 pM
Synthetic nucleotides
cGAMP
cUAMP
c-di-AMP
h
With BdCap5
BdCdnG reactions
NsCdnG reaction
All4 Reaction
ATP reaction
1 nM, 100 pM, 10 pM

supporting previous results that more closely related CD-NTase enzymes share similar nucleotide second messenger specificity (Whiteley et al., 2019). Clade E CD-NTases exhibit inter-clade variability and synthesize 3′3′-cUMP–AMP, 3′3′-cGAMP, and 3′3′-c-di-GMP as antiviral signals (Figure 3.7C), (Morehouse et al., 2020; Whiteley et al., 2019). In further support of subclade designations indicating product specificity, each of these Clade E enzymes contains substitutions within the conserved active-site lid helix demonstrating that contacts along this helix are a critical determinant for nucleobase incorporation (Figure 3.7C).

Previous large-scale analysis of nucleic acid binding proteins demonstrates a close relationship between side-chain identity and nucleobase interactions (Nobeli et al., 2001). Using experimentally determined CD-NTase product specificity and known protein–nucleobase interaction patterns as a guide, we next analyzed all CD-NTase enzyme sequences and developed rules to enable prediction of CD-NTase product specificity. To test these analyses, we experimentally verified the product specificity of uncharacterized Clade G CD-NTase enzymes *Bd*CdnG and *Novosphingobium subterraneum* CdnG (*Ns*CdnG) (Figure 3.9A). *Bd*CdnG encodes a serine at lid helix position 1 similar to *Vc*DncV, *Bf*CdnB, and *Em*CdnE (Figure 3.7C), suggesting specific incorporation of GTP or ATP. We observed enhanced thermostability upon GTP addition over a range of concentrations for both *Bd*CdnG and *Ns*CdnG supporting a role for GTP as a donor nucleotide (Figure 3.9B). In further agreement with predicted CD-NTase product specificity, we observed that *in vitro Bd*CdnG synthesizes 3′3′-cGMP–UMP, 3′3′-cGAMP and low levels of 3′3′-c-di-AMP, while *Ns*CdnG synthesizes primarily 3′3′-cGAMP (Figure 3.7D, Figure 3.9C–E).

The anti-phage activity of CBASS operons depends upon nucleotide second messenger-mediated activation of associated effector proteins (Cohen et al., 2019; Lowey et al., 2020; Ye et al., 2020). To further characterize *Bd*CdnG activity we next purified and tested activity of the *Bradyrhizobium diazoefficiens* effector CD-NTase-associated protein 5 (*Bd*Cap5) annotated as an HNH-SAVED nuclease (Burroughs et al., 2015; Lowey et al., 2020). *Bd*CdnG reactions are capable of activating *Bd*Cap5 and inducing promiscuous DNase activity and degradation of

plasmid DNA (Figure 3.7D, Figure 3.9F–H). However, *Bd*CdnG reactions producing the major product 3′3′-cGMP–UMP fail to activate *Bd*Cap5. We observed that *Bd*Cap5 responds to *Bd*CdnG reactions containing only ATP (Figure 3.7D; 3.9G,H), as well as *Ns*CdnG reactions, suggesting that *Bd*Cap5 recognizes a minor product that is conserved between closely related CD-NTase enzymes (Figure 3.9H). *Bd*Cap5 specificity for a signal distinct from 3′3′-cGMP–UMP agrees with previous observations that CBASS effector proteins can adapt to sense minor CD-NTase reaction products (Lowey et al., 2020). Together, these results reveal conserved patterns controlling CD-NTase product specificity and highlight that CBASS signaling networks rely on the specificity of both CD-NTase nucleobase incorporation and binding of the associated effector proteins to control antiviral defense.

Conclusions

Our structural and biochemical analysis of bacterial CD-NTase function reveals new insight into nucleotide second messenger product formation during CBASS anti-phage defense. Through detailed comparison of *Vc*DncV and *Ec*CdnD, we demonstrate that both cyclic dinucleotide and cyclic trinucleotide-synthesizing CD-NTase enzymes share a conserved mechanism of product formation. Universal aspects of CD-NTase nucleotide second messenger formation include use of a defined reaction order and specific nucleotide recognition through contacts conserved in the active-site lid. Using complementary mutational analysis and detailed assessment of a panel of modified nucleotide analogs, we demonstrate that in addition to major contacts with the active-site lid, CD-NTase enzymes achieve specificity through the net effect of minor contacts that result in formation of a stabilized enzyme–substrate complex. Our data further establish that CD-NTase-nucleotide contacts occur as a generalizable pattern conserved within discrete enzyme subclades (Figure 3.7), supporting that closely related CD-NTase enzymes produce a common nucleotide second messenger product (Whiteley et al., 2019). Although product identification requires detailed analysis for each CBASS operon, our results provide the foundation to begin predicting nucleotide specificity from primary CD-NTase protein sequence

alone.

The capacity for CD-NTases to synthesize diverse nucleotide second messengers is a key aspect of CBASS immunity. CD-NTases are highly specific yet exhibit a remarkable ability to adapt within divergent defense operons. Single substitutions in residues surrounding the active site are sufficient to direct synthesis of alternative nucleotide second messenger signals (Figure 3.1) (Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019). CD-NTase plasticity creates a capacity for rapid adaptation under evolutionary pressure. In contrast, the effector proteins executing bacterial cell death in CBASS operons exhibit strict specificity and typically only respond to a single nucleotide second messenger signal (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019). The strict specificity in CBASS effector activation likely functions as a brake to restrain CD-NTase evolution and maintain a discrete signaling pathway necessary for an efficient response to phage infection (Figure 3.7E). Overall, our work further establishes the determinants of CD-NTase specificity controlling nucleotide second messenger formation and demonstrates how antiviral signaling systems are balanced for both functional defense and adaptability.

Author Contributions

Apurva A. Govande: Designed experiments, conducted enzymes activity and nucleotide binding experiments, collected X-ray crystallography data and completed molecular modeling, and wrote the manuscript.

Brianna Duncan-Lowey: Assisted with enzyme activity and nucleotide binding experiments, collected X-ray crystallography data and completed molecular modeling.

James B. Eaglesham: Collected X-ray crystallography data.

Aaron T. Whiteley: Assisted with initial enzyme activity experiments.

Philip J. Kranzusch: Designed experiments, collected X-ray crystallography data, and wrote the manuscript.

Specific Acknowledgements

The authors are grateful to A. Lee, K. Chat, and members of the Kranzusch lab for helpful discussion, A. Engelman, D. Pearson, M. Danielsen, and D. Malhiero for assistance with thermal denaturation assays and mass spectrometry. X-ray data were collected at the Lawrence Berkeley National Laboratory Advanced Light Source beamline 8.2.2 (DE-AC02-05CH11231) and at the Northeastern Collaborative Access Team beamlines 24-ID-C and 24-ID-E (P30 GM124165).

METHODS

DATA AND CODE AVAILABILITY

CD-NTase structures have been deposited in the RCSB Protein Data Bank under accession numbers PDB 7LJL, 7LJM, 7LJN, and 7LJO at https://www.rcsb.org/

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strain *E. coli* MG1655 was used for cloning and *E. coli* BL21-DE3 RIL (Agilent) was used for protein expression throughout the study. *E. coli* MG1655 was grown in minimal LB medium overnight at 37°C for DNA preparation. *E. coli* BL21-DE3 RIL starter cultures were grown in MDG media starter culture (0.5% glucose, 25 mM Na_2PO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.25% aspartic acid, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals) and used to inoculate 1 L M9ZB media cultures (0.5% glycerol, 1% CAS amino acids, 47.8 mM Na_2PO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 2 mM $MgSO_4$, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, and trace metals) for protein expression induced with 0.5 mM IPTG.

METHOD DETAILS

Protein expression and purification

Recombinant CD-NTase proteins were cloned from synthetic DNA (IDT) into a custom pET vector as a 6×His-SUMO2-tagged fusion protein and expressed in the *E. coli* strain BL21-DE3 RIL (Agilent) harboring the rare tRNA plasmid pRARE2 as previously described (Lowey et al., 2020; Zhou et al., 2018). Briefly, transformed colonies were grown in a 30 mL MDG media starter culture overnight at 37°C, and used to inoculate 1 L M9ZB media cultures at OD_{600} ~0.05. Cultures were grown at 37°C, 230 RPM until OD_{600} reached ~2.2, chilled on ice for 10 min, induced with 0.5 mM IPTG, and incubated at 16°C, 230 RPM for ~16 h before harvest. Cultures for production of selenomethionine-substituted (SeMet) protein were grown as previously described (Eaglesham

et al., 2019). Briefly, SeMet protein was grown in overnight MDG starter cultures and used to inoculate 1 L induction cultures of M9 medium (47.8 mM Na_2PO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 2 mM $MgSO_4$, 100 mg mL^{-1} ampicillin, 34 mg mL^{-1} chloramphenicol, 1 µg mL^{-1} thiamine-HCl, 0.4% glucose, and trace metals). SeMet M9 cultures were grown to an OD of ~0.6 before addition of 50 mg L^{-1} leucine, isoleucine, and valine (VWR), 100 mg L^{-1} phenylalanine, lysine, and threonine (VWR), and 75 mg L^{-1} selenomethionine (Acros Organics) for selenomethinonine labeling and suppression of methionine biosynthesis. Cultures were grown for an additional 20 min at 37°C with shaking at 230 RPM and then chilled in an ice bath for 20 min. Cultures were induced with addition of 0.5 mM IPTG and incubated overnight at 16°C with shaking at 230 RPM.

Cultures were pelleted and lysed by sonication in 1× Lysis Buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 30 mM imidazole, 10% glycerol). Recombinant protein was purified with Ni-NTA resin (QIAGEN) using gravity flow chromatography. Ni-NTA resin was washed with 1× Lysis Buffer supplemented to 1 M NaCl and subsequently eluted with 1× Lysis Buffer supplemented to 300 mM imidazole. For all crystallography experiments, purified protein was treated with recombinant human SENP2 protease (D364–L589, M497A) to remove the SUMO2 tag and dialyzed for ~16 h against Dialysis Buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM DTT). Protein was then concentrated using a 30K-cutoff concentrator (Millipore) and purified by size-exclusion chromatography on either a 16/600 Superdex 200 column or 16/600 Superdex 75 column with Gel Filtration Buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM TCEP). Purified fractions were concentrated to >30 mg mL^{-1}, aliquoted in 40 µL and flash frozen in liquid nitrogen and stored at −80°C. For *Ec*CdnD mutant proteins, following Ni-NTA purification the SUMO2 tag was left on, protein was dialyzed for ~16 h against Dialysis Buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM DTT) and then samples were concentrated to >30 mg mL^{-1}, aliquoted in 40 µL and flash frozen in liquid nitrogen and stored at −80°C.

Crystallization and structure determination

CD-NTase proteins were crystalized at 18°C in a hanging-drop format using 0.4 µL drops in 96-well trays or 2 µL drops in 15-well Easy-Xtal trays (Qiagen). Briefly, each CD-NTase was crystalized as follows: 1) *Ec*CdnD + ATP: Native or selenomethionine-substituted protein was diluted to 10 mg mL^{-1} with 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP, supplemented with 10.5 mM $MgCl_2$ and 5 mM ATP, mixed 1:1 with reservoir solution (100 mM MES pH 6.5, 19–21% PEG-1000) and allowed to grow for 3 days. Crystals were further optimized with micro-seeding, grown for 2 days and then harvested by cryoprotecting with reservoir solution supplemented with 17.5% ethylene glycol $MgCl_2$ and ATP prior to freezing in liquid nitrogen; 2) *Se*CdnD + GTP: Native protein was diluted to 8.25 mg mL^{-1}, supplemented with 10.5 mM $MgCl_2$ and 5 mM GTP, mixed 1:1 with reservoir solution (0.1 M MES pH 6.5, 25% PEG-3000) and allowed to grow for 5 days in a 96-well tray and then harvested by cryo-protecting with NVH oil prior to freezing in liquid nitrogen; 3) *Bf*CdnB: Native or selenomethionine-substituted protein was diluted to 8 mg m(Agilent) $^{-1}$, supplemented with 10 mM $MgCl_2$ and 5 mM ATP, 0.5 mM Gpcpp, mixed 1:1 with reservoir solution (10 mM TCEP-HCl, not adjusted for pH, 22% PEG-3350) and allowed to grow for 24 days. Crystals were harvested by cryoprotecting with reservoir solution supplemented with 25% ethylene glycol prior to freezing in liquid nitrogen; 4) *Bd*CdnG + GTP: Native or selenomethionine-substituted protein was diluted to 10 mg m^{-1} with 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP, supplemented with 10.5 mM $MgCl_2$, 5 mM GTP, 0.5 mM Upnpp, mixed with reservoir solution (1:1, 0.032 M Sodium Citrate pH 7.0, 0.1 M HEPES pH 7.0, 10% w/v PEG-5000 MME) and allowed to grow for 6 days. Crystals were harvested by cryoprotecting with NVH oil prior to freezing in liquid nitrogen.

X-ray data were collected at the Lawrence Berkeley National Laboratory Advanced Light Source beamline 8.2.2 (DE-AC02-05CH11231) supported in part by the ALS-ENABLE program (P30 GM124169-01), and at the Northeastern Collaborative Access Team beamlines 24-ID-C and 24-ID-E (P30 GM124165), and used a Pilatus detector (S10RR029205), an Eiger detector

(S10OD021527) and the Argonne National Laboratory Advanced Photon Source (DE-AC02-06CH11357). X-ray data were processed using XDS and AIMLESS with the SSRL autoxds script (A. Gonzalez, Stanford SSRL). Selenomethione-substituted crystals were used to experimentally determine phases for each structure using AutoSol in PHENIX (Liebschner et al., 2019). Model building was performed in Coot (Emsley and Cowtan, 2004), and refinement was performed in PHENIX. Statistics were analyzed as described in Table S1 (Chen et al., 2010; Karplus and Diederichs, 2012; Weiss, 2001).

Thin-layer chromatography

CD-NTase reactions for nucleotide second messenger product characterization were performed in 20 µL reactions with 50 µM enzyme, 12.5 µM NTP and trace α^{32}P-ATP or α^{32}P-GTP, as indicated. *Vc*DncV reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and *Ec*CdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT. *Ec*CdnD mutant characterization reactions were carried out for 30 min at 37°C before heat inactivation at 95°C for 2 min, and reactions for IC_{50} calculations were carried out for 15 min at 37°C before heat inactivation at 95°C for 2 min. Prior to thin-layer chromatography analysis, all reactions were treated with 0.5 µL of Quick CIP (NEB) at 37°C for 30 min. Reactions were analyzed by thin-layer chromatography by spotting 1 µL of each phosphatase-treated reaction on a PEI-cellulose plate (Millipore) and developed in 1.5 M KH_2PO_4 pH 3.8 until buffer had reached ~2 cm from the top of the plate. Plates were dried and exposed on a phosphor-screen and subsequently imaged with a Typhoon Trio Variable Mode Imager (GE Healthcare).

Quantification of thin-layer chromatography experiments was carried out using ImageQuant TL v8.2.0.0 software. Following background subtraction, spots were converted to pixel intensity and normalized to wildtype enzyme where indicated. For IC_{50} calculation, pixel intensity corresponding to radiolabeled c-di-AMP, c-di-GMP (*Vc*DncV), or 3′3′3′-cAAA (*Ec*CdnD) was measured for each lane containing increasing amounts of unlabeled modified nucleotide (0,

0.1, 0.5, 1, 10, 125, 500, 1000, 2000 μM) and normalized to reactions containing no modified nucleotide. Inhibition data were fit using a non-linear regression curve ([Inhibitor] vs normalized response) in GraphPad Prism version 9.0.0.

Thermal denaturation assay

Thermal denaturation assays were carried out using a Tycho NT.6 (NanoTemper) with 30 μL reactions containing 10 μM protein and no nucleotide, or a range of nucleotide concentrations as indicated, 0.001 – 20 μM. *Vc*DncV reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and *Ec*CdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT. Protein and nucleotide solutions were incubated on ice for at least 45 min until measuring melting temperature of 10 μL of reaction between 35–95°C over 3 min.

HPLC enzymatic reaction analysis

For HPLC analysis, CD-NTase reactions were performed in 20 μL reactions with 50 μM enzyme and 500 μM NTPs. *Vc*DncV reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and *Ec*CdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT. Reactions were incubated at 37°C overnight for ~16 h, heat inactivated at 95°C for 2 min, diluted to 200 μL and filtered through a Millipore Amicon ultra-0.5 mL 30 kDa cut-off filter by centrifugation at 10,000 RPM (9350 × g) in a table-top microcentrifuge for 10 min prior to analysis by HPLC. Reactions were analyzed by HPLC with a C18 column (Zorbax Bonus-RP 4.6×150 mm, 3.5-μm) and Agilent 1200 Infinity Series LC system. 20 μL of filtered and diluted reaction was injected and eluted at 50°C with a flowrate of 1 mL min^{-1} using a mobile phase of 50 mM NaH_2PO_4 pH 6.8 supplemented with 3% HPLC grade acetonitrile for *Vc*DncV reactions and 6% acetonitrile for *Ec*CdnD reactions. For quick-CIP treated reactions, 20 μL of heat inactivated reaction was treated with 1 μL of Quick-CIP (NEB) at 37°C for 1 h prior to dilution and filtering.

Mass spectrometry analysis

CD-NTase reactions samples for mass spectrometry analysis were prepared identically to HPLC reactions. Sample analysis was carried out by MS-Omics as follows: Samples were diluted 1:3 in 10% ultra-pure water and 90% acetonitrile containing 10 mM ammonium acetate at pH 9 and then filter through a Costar Spin-X centrifuge tube filter 0.22 µm Nylon membrane. The analysis was carried out using a Thermo Scientific Vanquish LC coupled to Thermo Q Exactive HF MS. An electrospray ionization interface was used as ionization source. Analysis was performed in positive ionization mode.

CD-NTase enzyme kinetics

Reactions to measure CD-NTase product formation Michaelis-Menten kinetics were assembled on ice with 1 µM protein (*Vc*DncV, *Ec*CdnD, or mcGAS) and nucleotide concentrations 125, 250, 500, 750 µM, or 1 mM ATP and GTP. *Vc*DncV and mcGAS reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and *Ec*CdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT. For each nucleotide concentration, 100 µL reactions were incubated at 37°C and 20 µL aliquots were collected at 2, 10, 20, 30, 60 min (*Vc*DncV); 2, 10, 20, 30 min (*Ec*CdnD); or 0.5, 2, 5, 7, 10 min (mcGAS) and immediately heat inactivated at 95°C for 2 min followed by quenching on ice for 1 min. 20 µL aliquots were diluted to 200 µL and filtered through a Millipore Amicon ultra-0.5 mL 30 kDa cut-off filter by centrifugation at 10,000 RPM (9350 × g) in a table-top microcentrifuge for 10 min. 40 µL of each sample was then injected and reactions were analyzed by HPLC as detailed above. Absorbance units were converted to µM by comparing to a standard curve from 10–750 µM of chemically synthesized 3′3′-cGAMP, 2′3′-cGAMP, or 3′3′3′-cAAG (Biolog Life Sciences). Data were fit by linear regression and non-linear curve fitting Michalis-Menten kinetics were calculated using GraphPad Prism version 9.0.0.

*Bd*Cap5 nuclease assay

Nuclease assays were performed as previously described (Lowey et al., 2020). Briefly, 50 nM *Bd*Cap5 protein was incubated with 10 ng μL^{-1} pGEM9z plasmid (Promega) or 10 ng μL^{-1} pcDNA4 plasmid in a 10 µL reaction for 30 min at 37°C containing 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, and 1 mM TCEP. Nuclease reactions were supplemented with filtered CD-NTase product reactions or chemically synthesized nucleotide second messengers at 1 pM, 10 pM, 100 pM, 1 nM, 10 nM or 100 nM as indicated. Reactions were separated on a 2% TAE agarose gel containing 10 mg mL^{-1} ethidium bromide. Gels were run at 120 V for 30 min and then imaged with a ChemiDoc MP Imaging System.

Sequence-based determination of CD-NTase product specificity

Download “Supplementary Table 3, All clades, Nucleotide coordinating regions” from https://doi.org/10.1016/j.celrep.2021.109206. The table is organized by Clade and NCBI protein accession number for each CD-NTase as originally presented in Whiteley et al, Supplementary Table 2. Using the sequence or NCBI accession number, identify a protein of interest and check if a related CD-NTase enzyme has already been experimentally analyzed, indicated by highlighted rows. Experimentally determined CD-NTase products are the best guide for determination of CBASS operon nucleotide second messenger specificity. If no previous biochemical analysis exists for a closely related CD-NTase, ensure that the GS loop (XGSX) and active site residues (XDXDX, XDX) are conserved. Next, observe Lid position 1 (above the acceptor pocket) and lid position 2 (above the donor pocket). These positions are indicated in Columns G and H of Supplementary Table 3. These residues are not sufficient for selection but correlate highly with specific product formation.

Briefly:

- S, Lid position 1 – A/G selection
- N, Lid position 1 – A/U selection
- T, Lid position 1 – A selection
- A, Lid position 1 – A/G selection
- Q, Lid position 2 – G selection
- H, Lid position 2 – A selection
- Y, Lid position 2 – A/G/U selection
- F, Lid position 2 – A/G/U selection
- S, Lid position 2 – G selection
- L/V/I, Lid position 2 – A/G selection
- D, Either position – G selection
- M, Lid position 1 – unknown
- C, Lid position 1 - unknown
- E, Lid position 2, unknown
- R, Either position – unknown

Next, SI Table 3 highlights the conserved extended loops in the active site and the C-terminal lobe region and number of residues in each loop. Highlighted in red are C-term loop residues located above the active site and extending into the active site pocket for CD-NTase enzymes with determined structures. These residues can be used as an additional guide for CD-NTase enzymes within a subclade that share high sequence similarity within the C-term loop. CD-NTase sequences without this loop are shorter sequences that lack the C-terminal region. These extended loops may suggest specificity for a cyclic trinucleotide or larger oligonucleotide product. For any CD-NTase, a complete analysis of product formation requires analysis of effector selectivity. CD-NTase enzymes are evolutionarily constrained by effector recognition of the second messenger. Effector activation provides an additional confirmation of the major product for a CD-NTase as part of a functional CBASS system. For further reference, please consult (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details for each experiment can be found in figure legends, and outlined in the methods section. Data are plotted with error bars that indicate the standard deviation, as indicated.

Discussion

Cyclic oligonucleotide-based antiphage signaling systems (CBASS) are widespread immune systems that encode diverse effector proteins that are active upon binding to nucleotide second messengers to induce cell death that halts phage replication. In these studies, we examined the synthesis of nucleotide second messengers by cGAS/DncV-like nucleotidyltransferases and defined the molecular mechanisms controlling activation of downstream effector proteins.

Nucleotide second messenger synthesis by CD-NTases

Here we explore the diversity of products synthesized by cGAS/DncV-like nucleotidyltransferases. While this family of enzymes have divergent primary amino acid sequences, with less than 20% identity between most members, the overall protein secondary structure is highly conserved (Whiteley et al., 2019). CD-NTases are non-templated polymerases that use amino acid side-chain contacts to select the nucleotides being incorporated into the final second messenger product (Govande et al., 2021). CD-NTases synthesize a wide variety of products, including directing incorporation of different nucleobases, distinct ring sizes, and alternative nucleotide phosphodiester linkages.

We used structural and biochemical methods to characterize CD-NTases and the nucleotide second messengers they synthesize. We identified *Ya*CdnE as a clade E CD-NTase that synthesizes cyclic dinucleotides containing uracil, a specificity that is matched by its cognate effector *Ya*Cap15. We also characterized *Ab*CdnD, which produces cyclic oligonucleotides with non-canonical 2′–5′ linkages. Previously, bacterial CD-NTases were thought to only produce second messengers with canonical 3′–5′ bonds, and non-canonical 2′–5′ linkages were thought to have evolved specifically in mammalian isomers of nucleotide second messengers (Ablasser et al., 2013). We found that the bacterial *Ab*CdnD synthesizes a nucleotide second messenger with non-canonical linkages, and that the cognate effector, *Ab*Cap4 cannot be activated by only

canonically-linked second messengers. These findings expand the known diversity of nucleotide second messengers used within CBASS immunity that activate downstream effectors.

We also report the structures of clade D CD-NTases (CdnD) from *Enterobacter cloacae* and *Salmonella enterica*, which synthesize the nucleotide second messenger cyclic AMP–AMP–GMP (cAAG). CD-NTases form the first bond between nucleotides by binding NTPs in well-defined donor and acceptor pockets (Figure 4.1 A).

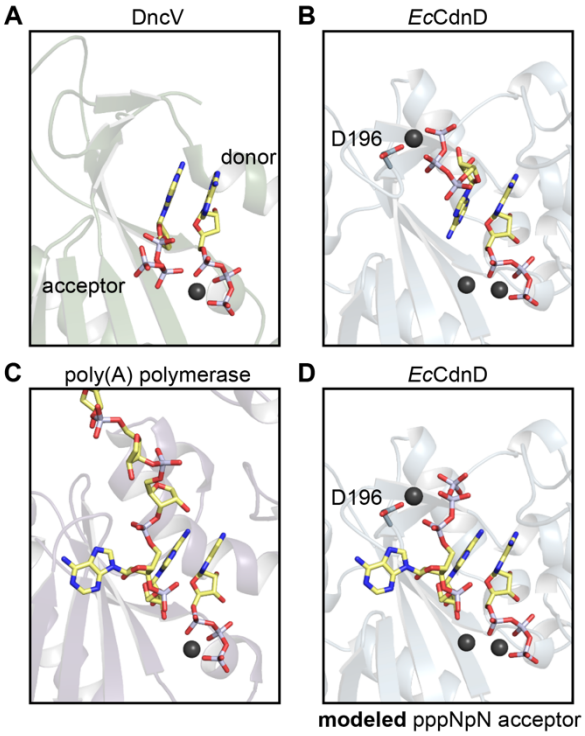


Figure 4.1. Synthesis of cyclic trinucleotides by *Ec*CdnD
(A–C) Experimentally determined structures of DncV (4XJ4), *Ec*CdnD bound to two NTPs (7LJL), and poly(A) polymerase (2Q66) showing nucleotides within the active site. (D) Modeled nucleotides showing the proposed accommodation of a pppNpN intermediate, poised to add a third nucleotide from the donor pocket.

In both the *Ec*CdnD and *Se*CdnD structures, we observed the same unexpected arrangement of nucleotides within the active site, with the second NTP is inverted and positioned above the canonical acceptor pocket coordinated by a magnesium ion positioned by D196 (Figure 4.1B). CD-NTases synthesize dinucleotides by linking two NTPs to form a linear intermediate (pppNpN), then flipping over the linear intermediate to form another phosphodiester bond using the same active site. A cyclic trinucleotide synthase must form three bonds, first by forming the usual pppNpN intermediate, then adding a third nucleotide in a pppNpNpN intermediate before finally cyclizing the molecule (see Figure 3.4C). While CdnD likely forms the first bond using the canonical donor and acceptor pockets, the new arrangement of nucleotides may illuminate how this active site is able to synthesize cyclic trinucleotides. Other polymerases, like poly(A) polymerase, must similarly form longer chains of nucleotides (Figure 4.1C). When compared to

structures of poly(A) forming a longer chain of AMP molecules, the unexpected configuration of the nucleotides within the *Ec*CdnD active site provides insight into how CD-NTases may accommodate the pppNpN intermediate during incorporation of the third nucleotide (Figure 4.1D). The triphosphate of the acceptor nucleotide is positioned in a way that could accommodate a $pppN_1pN_2$, where the NTP captured within the structure sits in the N_1 position (modeled in Figure 4.1D). This would allow for the triphosphate of the $pppN_1pN_2$ intermediate to be coordinated by D196 and magnesium ion as observed in the structure, while also positioning the N_2 nucleotide in the acceptor position. Future biochemical and structural studies will be needed to understand how CD-NTases with similar active sites are able to synthesize a wide diversity of nucleotide second messengers.

Phages have evolved many ways to overcome host immune systems, though the mechanisms used by phage to overcome CBASS immunity are currently unexplored. While there are no known phage nucleases that degrade CBASS second messengers, there are phage-encoded nucleases that degrade nucleotide second messengers to subvert type III CRISPR immunity (Athukoralage et al., 2020). It will therefore be interesting to explore how phages overcome CBASS immunity. Phage nucleases that are specific for one or a subset of nucleotide second messengers may provide pressure for CBASS to diversify the second messengers being used, with non-canonical linkages and larger ring sizes representing a host adaptation to restore signaling and immune defense.

Effectors of cell death in CBASS immunity

CBASS operons encode many types of effectors that are activated by nucleotide second messengers. Here we biochemically and structurally characterize two of the most prevalent classes of effectors, nucleases and transmembrane domain-containing proteins. We present crystal structures of the nucleotide second messenger-binding domains and define the mechanism of effector activation and induced cell death.

First, we characterized the “SAVED” ligand-binding domain of Cap4 effectors. We determined the first structures of SAVED domain-contain proteins, uncovering an unexpected connection to effector proteins in CRISPR immunity. We found that SAVED domains are highly selective cyclic oligonucleotide-binding domains, which specifically recognize the second messenger synthesized by their cognate CD-NTase. SAVED domains are encoded in many CBASS effectors; we characterized a structurally distinct effector that contains a SAVED domain, Cap5, and found it was similarly sensitive to activation by nucleotide second messengers. We also characterized the ligand-binding domain of Cap15, a transmembrane effector, and found a minimal β-barrel domain that similarly binds nucleotide second messenger signals. We also show that both Cap4 and Cap15 effectors oligomerize upon binding to nucleotide second messengers. Both SAVED domains and β-barrel ligand-binding domains form higher-order oligomers after binding to an activating nucleotide second messenger, allowing the effector to become activated.

To understand how activation of CBASS leads to cell death, we characterized the effector functions of Cap4 and Cap15. Cap4 contains a restriction endonuclease-like domain, which becomes activated to indiscriminately degrade double-stranded DNA, which eventually causes cell death. Cap15 contains transmembrane domains that disrupt the inner membrane upon activation, which leads to cell death through the rupture of the membrane. As we show here, these effectors are both sufficient to induce cell death in the absence of phage infection, using overexpression of CD-NTases to trigger activation of second messenger synthesis. It would be of interest to study the activation of these effectors in the context of phage infection, particularly the Cap15 transmembrane effectors. Phage tightly regulate the lysis of the host cell, expressing proteins to regulate three steps of cell lysis (Figure 4.2). Holins or pinholins accumulate in the membrane until they are triggered, forming either lesions large enough to leak proteins or pinholes that destroy the proton gradient. This releases lysozyme into the periplasm which can then

degrade the peptidoglycan, either by allowing the lysozyme to pass from the cytoplasm through large holes or by disrupting the membrane potential to release and activate signal anchor release (SAR) endolysins. Once the peptidoglycan is degraded, spanins can fuse the inner and outer membrane to lyse the cell (Cahill and Young, 2019). The timing of each of these steps is tightly regulated and could be disrupted by Cap15 activation to further exert anti-phage defense. It will be interesting to understand how activation of transmembrane and phospholipase effectors, which also target the membrane, may interrupt or prematurely trigger lysis of the cell when phage proteins are present.

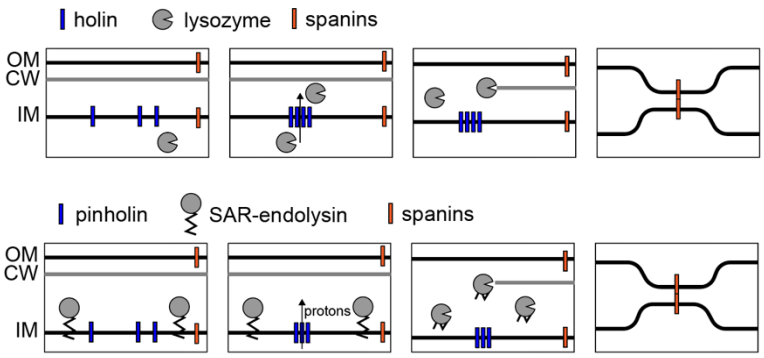


Figure 4.2. Phage lysis of bacterial hosts
Schematic showing lysis of bacterial host cells through phage holin/lysozyme (top) or pinholin/SAR-endolysin systems. OM: outer membrane, CW: cell wall/peptidoglycan, IM: inner membrane

These studies of CBASS effectors have uncovered some general principles by which most classes of effectors function. First, effectors target conserved bacterial features to induce cells death, like nucleic acid or cell membranes. Unlike some antiphage systems, for example restriction enzymes that target a specific DNA sequence within the phage genome, these effector targets cannot easily be altered to overcome restriction. While the aspect of phage infection that is sensed by CD-NTases remains unknown, we show that a wide variety of phages can be detected. This suggests that both the sensing of phage infection and the effector functions that are induced rely on broadly conserved aspects, likely making evasion of CBASS immunity challenging. Second, each effector fulfills two main functions: sensing nucleotide second messengers and inducing cell death. Here we characterize SAVED and β-barrel domains as nucleotide second-messenger binding domains and restriction endonuclease-like and transmembrane domains as cell death-inducing effector domains. We describe variations in

domain architecture that mix and match the sensor and effector domains, for example Cap14, which contains a SAVED domain and transmembrane domain. This modular architecture contributes to the wide diversity of CBASS effectors found encoded in CD-NTase operons. There remain many effectors with uncharacterized domains, which makes the characterization of these effectors to find new ligand-binding and cell death-inducing domains an interesting aspect of future studies.

References

Ablasser, A., and Chen, Z.J. (2019). cGAS in action: Expanding roles in immunity and inflammation. Science *363*.

Ablasser, A., Goldeck, M., Cavlar, T., Deimling, T., Witte, G., Rohl, I., Hopfner, K.P., Ludwig, J., and Hornung, V. (2013). cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. Nature *498*, 380-384.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Athukoralage, J.S., McMahon, S.A., Zhang, C., Gruschow, S., Graham, S., Krupovic, M., Whitaker, R.J., Gloster, T.M., and White, M.F. (2020). An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. Nature *577*, 572-575.

Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G.*, et al.* (2021). Prokaryotic viperins produce diverse antiviral molecules. Nature *589*, 120-124.

Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. Nat Rev Microbiol *18*, 113-119.

Burroughs, A.M., Zhang, D., Schaffer, D.E., Iyer, L.M., and Aravind, L. (2015). Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. Nucleic Acids Res *43*, 10633-10654.

Cahill, J., and Young, R. (2019). Phage Lysis: Multiple Genes for Multiple Barriers. Adv Virus Res *103*, 33-70.

Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U., and Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. Proc Natl Acad Sci U S A *101*, 17084-17089.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.

Chen, Z.H., and Schaap, P. (2012). The prokaryote messenger c-di-GMP triggers stalk cell differentiation in Dictyostelium. Nature *488*, 680-683.

Civril, F., Deimling, T., de Oliveira Mann, C.C., Ablasser, A., Moldt, M., Witte, G., Hornung, V., and Hopfner, K.P. (2013). Structural mechanism of cytosolic DNA sensing by cGAS. Nature *498*, 332-337.

Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. Nature *574*, 691-695.

Corrigan, R.M., and Grundling, A. (2013). Cyclic di-AMP: another second messenger enters the fray. Nat Rev Microbiol *11*, 513-524.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res *14*, 1188-1190.

Dal Peraro, M., and van der Goot, F.G. (2016). Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol *14*, 77-92.

Danilchanka, O., and Mekalanos, J.J. (2013). Cyclic dinucleotides and the innate immune response. Cell *154*, 962-970.

Dash, C., Fisher, T.S., Prasad, V.R., and Le Grice, S.F. (2006). Examining interactions of HIV-1 reverse transcriptase with single-stranded template nucleotides by nucleoside analog interference. J Biol Chem *281*, 27873-27881.

Daugherty, M.D., and Malik, H.S. (2012). Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet *46*, 677-700.

Davies, B.W., Bogard, R.W., Young, T.S., and Mekalanos, J.J. (2012). Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for V. cholerae virulence. Cell *149*, 358-370.

Dewey, J.S., Savva, C.G., White, R.L., Vitha, S., Holzenburg, A., and Young, R. (2010). Micron-scale holes terminate the phage infection cycle. Proc Natl Acad Sci U S A *107*, 2219-2223.

Diner, E.J., Burdette, D.L., Wilson, S.C., Monroe, K.M., Kellenberger, C.A., Hyodo, M., Hayakawa, Y., Hammond, M.C., and Vance, R.E. (2013). The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. Cell Rep *3*, 1355-1361.

Dodds, P.N., and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet *11*, 539-548.

Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. Science *359*.

Dou, J., Vorobieva, A.A., Sheffler, W., Doyle, L.A., Park, H., Bick, M.J., Mao, B., Foight, G.W., Lee, M.Y., Gagnon, L.A.*, et al.* (2018). De novo design of a fluorescence-activating beta-barrel. Nature *561*, 485-491.

Duncan-Lowey, B.L., McNamara-Bordewick, N.K., Tal, N., Sorek, R., and Kranzusch, P.J. Effector-mediated membrane disruptuion controls cell death in CBASS antiphage defense. Submitted.

Eaglesham, J.B., Pan, Y., Kupper, T.S., and Kranzusch, P.J. (2019). Viral and metazoan poxins are cGAMP-specific nucleases that restrict cGAS-STING signalling. Nature *566*, 259-263.

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

Gabler, F., Nam, S.Z., Till, S., Mirdita, M., Steinegger, M., Soding, J., Lupas, A.N., and Alva, V. (2020). Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Curr Protoc Bioinformatics *72*, e108.

Gaffney, B.L., Veliath, E., Zhao, J., and Jones, R.A. (2010). One-flask syntheses of c-di-GMP and the [Rp,Rp] and [Rp,Sp] thiophosphate analogues. Org Lett *12*, 3269-3271.

Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B.L., Zillinger, T., Serganov, A.A., Liu, Y., Jones, R.A., Hartmann, G.*, et al.* (2013a). Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. Cell *153*, 1094-1107.

Gao, P., Ascano, M., Zillinger, T., Wang, W., Dai, P., Serganov, A.A., Gaffney, B.L., Shuman, S., Jones, R.A., Deng, L.*, et al.* (2013b). Structure-function analysis of STING activation by c[G(2',5')pA(3',5')p] and targeting by antiviral DMXAA. Cell *154*, 748-762.

Govande, A.A., Duncan-Lowey, B.L., Eaglesham, J.B., Whiteley, A.T., and Kranzusch, P.J. (2021). Molecular basis of CD-NTase nucleotide selection in CBASS anti-phage defense. Cell Reports *35 (9)*, 109206.

Holleufer, A., Winther, K.G., Gad, H.H., Ai, X., Chen, Y., Li, L., Wei, Z., Deng, H., Liu, J., Frederiksen, N.A.*, et al.* (2021). Two cGAS-like receptors induce antiviral immunity in Drosophila. Nature.

Hornung, V., Hartmann, R., Ablasser, A., and Hopfner, K.P. (2014). OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids. Nat Rev Immunol *14*, 521-528.

Hsiao, J.J., Potter, O.G., Chu, T.W., and Yin, H. (2018). Improved LC/MS Methods for the Analysis of Metal-Sensitive Analytes Using Medronic Acid as a Mobile Phase Additive. Anal Chem *90*, 9457-9464.

Jenal, U., Reinders, A., and Lori, C. (2017). Cyclic di-GMP: second messenger extraordinaire. Nat Rev Microbiol *15*, 271-284.

Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D.J. (2019). CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. Mol Cell *75*, 944-956 e946.

Jia, N., and Patel, D.J. (2021). Structure-based functional mechanisms and biotechnology applications of anti-CRISPR proteins. Nat Rev Mol Cell Biol *22*, 563-579.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A.*, et al.* (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030-1033.

Kato, K., Ishii, R., Hirano, S., Ishitani, R., and Nureki, O. (2015). Structural Basis for the Catalytic Mechanism of DncV, Bacterial Homolog of Cyclic GMP-AMP Synthase. Structure *23*, 843-850.

Kawato, T., Mizohata, E., Shimizu, Y., Meshizuka, T., Yamamoto, T., Takasu, N., Matsuoka, M., Matsumura, H., Kodama, T., Kanai, M.*, et al.* (2015). Structure-based design of a streptavidin mutant specific for an artificial biotin analogue. J Biochem *157*, 467-475.

Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science *357*, 605-609.

Ko, T.P., Wang, Y.C., Tsai, C.L., Yang, C.S., Hou, M.H., and Chen, Y. (2021). Crystal structure and functional implication of a bacterial cyclic AMP-AMP-GMP synthetase. Nucleic Acids Res *49*, 4725-4737.

Kranzusch, P.J. (2019). cGAS and CD-NTase enzymes: structure, mechanism, and evolution. Curr Opin Struct Biol *59*, 178-187.

Kranzusch, P.J., Lee, A.S.Y., Wilson, S.C., Solovykh, M.S., Vance, R.E., Berger, J.M., and Doudna, J.A. (2014). Structure-guided reprogramming of human cGAS dinucleotide linkage specificity. Cell *158*, 1011-1021.

Kranzusch, P.J., Wilson, S.C., Lee, A.S., Berger, J.M., Doudna, J.A., and Vance, R.E. (2015). Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling. Mol Cell *59*, 891-903.

Kristiansen, H., Gad, H.H., Eskildsen-Larsen, S., Despres, P., and Hartmann, R. (2011). The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. J Interferon Cytokine Res *31*, 41-47.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol *305*, 567-580.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lau, R.K., Ye, Q., Birkholz, E.A., Berg, K.R., Patel, L., Mathews, I.T., Watrous, J.D., Ego, K., Whiteley, A.T., Lowey, B.*, et al.* (2020). Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity. Mol Cell *77*, 723-733 e726.

Launer-Felty, K.D., and Strobel, S.A. (2018). Enzymatic synthesis of cyclic dinucleotide analogs by a promiscuous cyclic-AMP-GMP synthetase and analysis of cyclic dinucleotide responsive riboswitches. Nucleic Acids Res *46*, 2765-2776.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J.*, et al.* (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861-877.

Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. Annu Rev Virol *7*, 371-384.

Lowey, B., and Kranzusch, P.J. (2020). CD-NTases and nucleotide second messenger signaling. Curr Biol *30*, R1106-R1108.

Lowey, B., Whiteley, A.T., Keszei, A.F.A., Morehouse, B.R., Mathews, I.T., Antine, S.P., Cabrera, V.J., Kashin, D., Niemann, P., Jain, M.*, et al.* (2020). CBASS Immunity Uses CARF-Related Effectors to Sense 3'-5'- and 2'-5'-Linked Cyclic Oligonucleotide Signals and Protect Bacteria from Phage Infection. Cell *182*, 38-49 e17.

Margolis, S.R., Wilson, S.C., and Vance, R.E. (2017). Evolutionary Origins of cGAS-STING Signaling. Trends Immunol *38*, 733-743.

Martin, M. (2011). Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnetjournal *17*, 10-12.

Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P. (2009). Enumeration of bacteriophages using the small drop plaque assay system. Methods Mol Biol *501*, 81-85.

McFarland, A.P., Luo, S., Ahmed-Qadri, F., Zuck, M., Thayer, E.F., Goo, Y.A., Hybiske, K., Tong, L., and Woodward, J.J. (2017). Sensing of Bacterial Cyclic Dinucleotides by the Oxidoreductase RECON Promotes NF-kappaB Activation and Shapes a Proinflammatory Antibacterial State. Immunity *46*, 433-445.

McMahon, S.A., Zhu, W., Graham, S., Rambo, R., White, M.F., and Gloster, T.M. (2020). Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. Nat Commun *11*, 500.

Millman, A., Melamed, S., Amitai, G., and Sorek, R. (2020). Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. Nat Microbiol *5*, 1608-1615.

Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., Lopez-Mendez, B., She, Q., and Montoya, G. (2019). Structure of Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. Nat Commun *10*, 4302.

Morehouse, B.R., Govande, A.A., Millman, A., Keszei, A.F.A., Lowey, B., Ofir, G., Shao, S., Sorek, R., and Kranzusch, P.J. (2020). STING cyclic dinucleotide sensing originated in bacteria. Nature *586*, 429-433.

Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R., and Ban, N. (2009). The structure of a cytolytic alpha-helical toxin pore reveals its assembly mechanism. Nature *459*, 726-730.

Niewoehner, O., Garcia-Doval, C., Rostol, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., and Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. Nature *548*, 543-548.

Niewoehner, O., and Jinek, M. (2016). Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA *22*, 318-329.

Nobeli, I., Laskowski, R.A., Valdar, W.S., and Thornton, J.M. (2001). On the molecular discrimination between adenine and guanine by proteins. Nucleic Acids Res *29*, 4294-4309.

Ouyang, S., Song, X., Wang, Y., Ru, H., Shaw, N., Jiang, Y., Niu, F., Zhu, Y., Qiu, W., Parvatiyar, K.*, et al.* (2012). Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. Immunity *36*, 1073-1086.

Pang, T., Fleming, T.C., Pogliano, K., and Young, R. (2013). Visualization of pinholin lesions in vivo. Proc Natl Acad Sci U S A *110*, E2054-2063.

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

Rausch, J.W., Chelico, L., Goodman, M.F., and Le Grice, S.F. (2009). Dissecting APOBEC3G substrate specificity by nucleoside analog interference. J Biol Chem *284*, 7047-7058.

Rosenberg, J., Dickmanns, A., Neumann, P., Gunka, K., Arens, J., Kaever, V., Stulke, J., Ficner, R., and Commichau, F.M. (2015). Structural and biochemical analysis of the essential diadenylate cyclase CdaA from Listeria monocytogenes. J Biol Chem *290*, 6596-6606.

Rostol, J.T., Xie, W., Kuryavyi, V., Maguin, P., Kao, K., Froom, R., Patel, D.J., and Marraffini, L.A. (2021). The Card1 nuclease provides defence during type III CRISPR immunity. Nature *590*, 624-629.

Samson, J.E., Magadan, A.H., Sabri, M., and Moineau, S. (2013). Revenge of the phages: defeating bacterial defences. Nat Rev Microbiol *11*, 675-687.

Severin, G.B., Ramliden, M.S., Hawver, L.A., Wang, K., Pell, M.E., Kieninger, A.K., Khataokar, A., O'Hara, B.J., Behrmann, L.V., Neiditch, M.B.*, et al.* (2018). Direct activation of a phospholipase by cyclic GMP-AMP in El Tor Vibrio cholerae. Proc Natl Acad Sci U S A *115*, E6048-E6055.

Slavik, K.M., Morehouse, B.R., Ragucci, A.E., Zhou, W., Ai, X., Chen, Y., Li, L., Wei, Z., Bahre, H., Konig, M.*, et al.* (2021). cGAS-like receptors sense RNA and control 3'2'-cGAMP signaling in Drosophila. Nature.

Sonnhammer, E.L., Eddy, S.R., and Durbin, R. (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins *28*, 405-420.

Steinegger, M., and Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol *35*, 1026-1028.

Stern, A., and Sorek, R. (2011). The phage-host arms race: shaping the evolution of microbes. Bioessays *33*, 43-51.

Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z.J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science *339*, 786-791.

Sutterlin, H.A., Shi, H., May, K.L., Miguel, A., Khare, S., Huang, K.C., and Silhavy, T.J. (2016). Disruption of lipid homeostasis in the Gram-negative cell envelope activates a novel cell death pathway. Proc Natl Acad Sci U S A *113*, E1565-1574.

Tamulaitiene, G., Jovaisaite, V., Tamulaitis, G., Songailiene, I., Manakova, E., Zaremba, M., Grazulis, S., Xu, S.Y., and Siksnys, V. (2017). Restriction endonuclease AgeI is a monomer which dimerizes to cleave DNA. Nucleic Acids Res *45*, 3547-3558.

Tanaka, K., Caaveiro, J.M., Morante, K., Gonzalez-Manas, J.M., and Tsumoto, K. (2015). Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. Nat Commun *6*, 6337.

Terwilliger, T.C. (1999). Reciprocal-space solvent flattening. Acta Crystallogr D Biol Crystallogr *55*, 1863-1871.

Uehara, T., Parzych, K.R., Dinh, T., and Bernhardt, T.G. (2010). Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. EMBO J *29*, 1412-1422.

Wang, X., and Montero Llopis, P. (2016). Visualizing Bacillus subtilis During Vegetative Growth and Spore Formation. Methods Mol Biol *1431*, 275-287.

Wang, Z., Zhao, C., Wang, C., Zhang, H., Ma, D., Zhang, Q., Wen, X., Li, L., and Xi, Z. (2021). Synthesis and biological evaluation of all possible inosine-mixed cyclic dinucleotides that activate different hSTING variants. Bioorg Med Chem *29*, 115899.

Watanabe, N., Takasaki, Y., Sato, C., Ando, S., and Tanaka, I. (2009). Structures of restriction endonuclease HindIII in complex with its cognate DNA and divalent cations. Acta Crystallogr D Biol Crystallogr *65*, 1326-1333.

Weiss, M.S. (2001). Global indicators of X-ray data quality. Journal of applied crystallography *34*, 130-135.

Whiteley, A.T., Eaglesham, J.B., de Oliveira Mann, C.C., Morehouse, B.R., Lowey, B., Nieminen, E.A., Danilchanka, O., King, D.S., Lee, A.S.Y., Mekalanos, J.J*., et al.* (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature *567*, 194-199.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer).

Witte, G., Hartung, S., Buttner, K., and Hopfner, K.P. (2008). Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. Mol Cell *30*, 167-178.

Ye, Q., Lau, R.K., Mathews, I.T., Birkholz, E.A., Watrous, J.D., Azimi, C.S., Pogliano, J., Jain, M., and Corbett, K.D. (2020). HORMA Domain Proteins and a Trip13-like ATPase Regulate Bacterial cGAS-like Enzymes to Mediate Bacteriophage Immunity. Mol Cell *77*, 709-722 e707.

Yin, Q., Tian, Y., Kabaleeswaran, V., Jiang, X., Tu, D., Eck, M.J., Chen, Z.J., and Wu, H. (2012). Cyclic di-GMP sensing via the innate immune signaling protein STING. Mol Cell *46*, 735-745.

Young, I., Wang, I., and Roof, W.D. (2000). Phages will out: strategies of host cell lysis. Trends Microbiol *8*, 120-128.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J Struct Biol *193*, 1-12.

Zhang, X., Shi, H., Wu, J., Zhang, X., Sun, L., Chen, C., and Chen, Z.J. (2013). Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. Mol Cell *51*, 226-235.

Zhou, W., Whiteley, A.T., de Oliveira Mann, C.C., Morehouse, B.R., Nowak, R.P., Fischer, E.S., Gray, N.S., Mekalanos, J.J., and Kranzusch, P.J. (2018). Structure of the Human cGAS-DNA Complex Reveals Enhanced Control of Immune Surveillance. Cell *174*, 300-311 e311.

Zhu, D., Wang, L., Shang, G., Liu, X., Zhu, J., Lu, D., Wang, L., Kan, B., Zhang, J.R., and Xiang, Y. (2014). Structural biochemistry of a Vibrio cholerae dinucleotide cyclase reveals cyclase activity regulation by folates. Mol Cell *55*, 931-937.

Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kubler, J., Lozajic, M., Gabler, F., Soding, J., Lupas, A.N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol *430*, 2237-2243.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.