



Mechanisms of nucleotide signaling in bacterial and metazoan antiviral defense

Citation

Govande, Apurva. 2021. Mechanisms of nucleotide signaling in bacterial and metazoan antiviral defense. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37370170>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

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

*Mechanisms of nucleotide signaling in bacterial and metazoan
antiviral defense*

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

Signature: 
Lee Gehrke (Aug 30, 2021 20:58 EDT)

Typed Name: Dr. Lee Gehrke

Signature: 
Ann Hochschild (Aug 26, 2021 01:11 EDT)

Typed Name: Dr. Ann Hochschild

Signature: 

Typed Name: Dr. Victoria D'Souza

Signature: 
Joshua Woodward (Aug 25, 2021 13:54 PDT)

Typed Name: Dr. Joshua Woodward

Date: August 19, 2021

Mechanisms of nucleotide signaling in bacterial and metazoan antiviral defense

A dissertation presented

by

Apurva Govande

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

August 2021

© 2021 Apurva Govande
All rights reserved

Mechanisms of nucleotide signaling in bacterial and metazoan antiviral defense

Abstract

Nucleotide-based antiviral signaling is broadly conserved between bacteria and metazoans. Mammalian cGAS and OAS1 are structurally homologous enzymes that share a pol- β -like nucleotidyltransferase fold and synthesize antiviral nucleotide signals. Mammalian cGAS synthesizes a cyclic dinucleotide using ATP and GTP with one 2'–5' linkage and one 3'–5' linkage, known as 2',3'-cGAMP, that signals downstream to the effector STING, while OAS1 synthesizes a linear polyadenylate molecule with 2'–5' linkages, known as 2'–5' oligoadenylate, that signals downstream to the effector RNase L. In bacteria, recent work has uncovered enzymes with structural homology to cGAS, named CD-NTase enzymes, that are diverse in primary sequence and are capable of synthesizing novel cyclic oligonucleotide signals out of all four RNA nucleotides. Using biochemical and structural biology approaches, we define conserved patterns of CD-NTase product formation enabling the preliminary prediction of novel nucleotide products based on CD-NTase primary sequence. Bacterial CD-NTase enzymes are contained in operons that encode effector proteins. Effector function is dependent on recognition of the correct nucleotide signal synthesized by the cognate CD-NTase within the same operon. Effectors are comprised of several different categories of enzymes, and the diverse mechanisms of activation and regulation are open areas of study. Using X-ray crystallography, negative-stain EM, and biochemical approaches, we describe the activation and regulation of a bacterial STING homolog and its ancient effector function of cleaving NAD⁺. In metazoans, increasing complexity results in pathways that encode signaling functions that activate several different effector pathways, such as STING regulated activation of IRF3 and NF- κ B, leading to interferon induction and inflammatory cytokine synthesis. Using enzymatically synthesized 2'–5' oligoadenylate and sepharose bead-mediated affinity purification and mass spectrometry, we isolate and characterize a novel downstream receptor of 2'–5' oligoadenylate, the metabolic enzyme Decr1. Altogether, functional and structural conservation of signaling pathways across bacteria and metazoans highlights broad patterns of antiviral defense.

Table of Contents

TITLE.....	II
COPYRIGHT.....	III
ABSTRACT	III
ACKNOWLEDGEMENTS.....	V
DEDICATION	VI
EPIGRAPH.....	VII
FIGURES AND TABLES	VIII
INTRODUCTION	1
CYCLIC NUCLEOTIDE SIGNALING	1
CYCLIC OLIGONUCLEOTIDES IN INNATE IMMUNITY.....	3
CD-NTASE ENZYME ACTIVATION.....	3
STRUCTURAL DETERMINANTS OF CD-NTASE FUNCTION.....	4
BACTERIAL CD-NTASE-CONTAINING OPERONS	7
DOWNSTREAM EFFECTORS AND FUNCTIONS	8
MOLECULAR PATTERNS	10
CHAPTER 1: MOLECULAR BASIS OF CD-NTASE NUCLEOTIDE SELECTION IN CBASS ANTI-PHAGE DEFENSE	12
ABSTRACT	13
INTRODUCTION	14
RESULTS AND DISCUSSION	16
CONCLUSIONS.....	43
METHODS	44
CHAPTER 2: STING CYCLIC DINUCLEOTIDE SENSING ORIGINATED IN BACTERIA	53
ABSTRACT	54
RESULTS AND DISCUSSION	55
METHODS	87
CHAPTER 3: RNASE L ACTIVATING 2'–5' OLIGOADENLYATES BIND DECR-1.....	97
ABSTRACT	98
INTRODUCTION	99
RESULTS AND DISCUSSION	101
CONCLUSIONS.....	107
METHODS	109
DISCUSSION.....	115
REFERENCES	123

Acknowledgements

While the writing of a dissertation and the credit of a PhD is often conferred on an individual, the experience truly is a feat that requires the willing participation and support of many, many people. Firstly, seeing through the formidable task of a PhD would not have been possible without the many moments of celebration, commiseration, and support of the friends I made throughout, so thank you Colin O'Leary, Kailey Slavik, James Eaglesham, and Ben Morehouse, for all the moments inside the lab and out that made this PhD possible.

Additionally, thank you to all members of the Kranzusch lab, past and present, for lovely discussions and ponderings about science and thoroughly engaging lab meetings, my DAC, Lee Gehrke, Sun Hur and Alan Engelman, for their guidance, and my exam committee for reading of this thesis.

I'd also like to thank my parents for their love, and all my friends back in California (and Zahraa, who left California for a bit), for encouragement and your never-ending enthusiasm for hearing about my work, no matter how obscure – it's nice to know that someone else can be briefly excited about the esoteric problems of a PhD. I would also like to thank my sister, Ashlesha, for being a constant source of fun whenever I needed a break.

Lastly, thank you to Philip, for taking me on as one of your first graduate students. I'm so fortunate to have learned crystallography, protein purification, and other biochemistry techniques, including the joys of radioactivity assays, directly from you. Since the time I joined in 2017, there have been many moments of cool science and genuinely exciting discoveries that I'm happy to have been a part of, two lab retreats and many celebrations, and of course a global pandemic that we're still crawling through. Thank you for a great start to my scientific career!

And thank you all!

Apurva

To my friends and family

In science it often happens that scientists say, 'You know that's a really good argument; my position is mistaken,' and then they would actually change their minds and you never hear that old view from them again. They really do it. It doesn't happen as often as it should, because scientists are human and change is sometimes painful. But it happens every day. I cannot recall the last time something like that happened in politics or religion. (1987) -- Carl Sagan

“Science never solves a problem without creating ten more.” —George Bernard Shaw

Figures and Tables

Figure I.1: Summary of a Cyclic oligonucleotide messenger-mediated response.....	2
Figure I.2: Type I and Type II CD-NTase activation.....	4
Figure I.3: The mechanism of reaction of CD-NTase enzymes follows a defined order.....	6
Figure I.4: CBASS operons are organized into three different types.....	7
Figure I.5: Conserved response architecture in metazoan and bacterial antiviral signaling.....	9
Figure 1.1: Crystal structure of the 3'3'3'-cAAG synthase EcCdnD in a pre-reactive state.....	17
Figure 1.2: <i>EcCdnD</i> and <i>VcDncV</i> structural comparison and complex formation with nucleotide substrates.....	23
Figure 1.3: Stable CD-NTase–donor nucleotide complex formation is a key determinant of product specificity.....	24
Figure 1.4: Measurement of <i>VcDncV</i> and <i>EcCdnD</i> reaction rates.....	26
Figure 1.5: Nucleotide interference mapping analysis of CD-NTase reaction progression.....	27
Figure 1.6: IC ₅₀ quantification of modified nucleotide incorporation correlates with modified nucleotide product formation.....	33
Figure 1.7: HPLC analysis of <i>VcDncV</i> reactions with modified nucleotides.....	37
Figure 1.8: Structural and sequence-based comparison of divergent bacterial CD-NTases reveals correlates of nucleotide selection.....	39
Figure 1.9: <i>BdCdnG</i> effector Cap5 responds to A-containing minor product.....	42
Table 1.1: Summary of crystallographic statistics.....	18
Table 1.2: Summary of mass spectrometry results.....	28
Figure 2.1: Structure of bacterial STING and definition of metazoan-specific insertions.....	54
Figure 2.2: Structural analysis of bacterial STING–cyclic dinucleotide complex formation.....	59
Figure 2.3: Bacterial STING systems signal through the 3'–5'-linked nucleotide second messenger c-di-GMP.....	61
Figure 2.4: Structural analysis of STING-associated CD-NTase enzymes.....	63
Figure 2.5: Biochemical analysis of c-di-GMP synthesis by bacterial STING-associated CD-NTases.....	65
Figure 2.6: Biochemical analysis of bacterial STING cyclic dinucleotide recognition specificity.....	68
Figure 2.7: Cyclic dinucleotide-recognition controls bacterial STING oligomerization and TIR NAD ⁺ cleavage activity.....	69
Figure 2.8: Bacterial STING activation of TIR NADase activity.....	71
Figure 2.9: Mutagenesis analysis of bacterial STING cyclic dinucleotide recognition and TIR activation.....	74
Figure 2.10: Conservation of oligomerization as a mechanism of STING activation.....	76
Figure 2.11: Structural basis of STING adaptation for signaling regulation.....	77
Figure 2.12: Filament formation is required for bacterial TIR-STING activation.....	79
Figure 2.13: Structural analysis of metazoan TIR-STING homologs.....	81
Figure 2.14: Structure-guided analysis of STING phylogenetic conservation and cyclic dinucleotide recognition.....	84
Table 2.1: Summary of data collection, phasing, and refinement statistics.....	55
Figure 3.1: Decr1 is a novel binding partner of 2'–5' OA.....	101
Figure 3.2: Mouse Decr1 binding to 2'–5' OA is dependent on the 5' triphosphate.....	102
Figure 3.3: 2'–5' OA binding obstructs co-factor NADPH binding pocket in Decr1 in a species-specific manner.....	105
Table 3.1: Summary of Crystallographic Statistics.....	104

Introduction

Cyclic nucleotide signaling

Cyclic nucleotides are widely used in bacteria and eukaryotes as both intra- and extra- cellular signals. Nucleotides act as diffusible second messengers that control multiple levels of a pathway to coordinate complex responses to stimuli. The original second messenger, cyclic AMP (cAMP), is widely recognized as a universal signaling molecule that is important in many cellular processes, initially characterized as the unique second messenger molecule involved in hormone-dependent signaling (SUTHERLAND and RALL 1958; RALL and SUTHERLAND 1958). Over the years, additional roles of cAMP discovered in a variety of biological processes solidified its role as a ubiquitous second messenger signaling molecule (Serezani et al. 2008; Ravnskjaer, Madiraju, and Montminy 2015). In bacteria, cAMP is best characterized as a regulatory molecule during glucose sensing and regulation of the lac operon, enabling sensitive control of bacterial growth dependent on the availability of sugar sources (Kimata et al. 1997). This signified it as an important signaling molecule in bacteria and eukaryotes. Cyclic GMP, on the other hand, was best known as a eukaryotic signaling molecule involved in nitric oxide receptor signaling (Friebe, Sandner, and Schmidtke 2020) but was later found to have signaling roles in bacteria as well, controlling nutrient sensing and cyst formation in the bacteria *Rhodospirillum centenum*, and others (Berleman, Hasselbring, and Bauer 2004).

Further identification of cyclic dinucleotides as major regulatory signals and second messenger molecules that control a wide array of cellular processes expanded the field of cellular signaling (Danilchanka and Mekalanos 2013). Cyclic di-GMP (c-di-GMP) is widespread in bacteria and is the best studied cyclic dinucleotide, with roles in regulating bacterial virulence, biofilm formation, motility, growth and cell cycle progression, synthesized by enzymes that contain diguanylate cyclase domains (Chan et al. 2004; Jenal, Reinders, and Lori 2017), and is degraded by phosphodiesterases (PDE) containing EAL or HD-GYP domains (Danilchanka and Mekalanos 2013). Several receptors responsive to c-di-GMP have been characterized. Another purine cyclic dinucleotide, cyclic di-AMP (c-di-AMP) maintains bacterial homeostasis, regulates DNA repair and potassium export, and binds to signal and regulate several proteins and riboswitches (Faozia et al. 2021; Kalia et al. 2013; Rosenberg et al. 2015). The best characterized c-di-AMP synthase is a diadenylate cyclase domain-containing protein from *Bacillus subtilis*, DisA (Witte et

al. 2008). Both the synthesis and degradation of c-di-AMP and c-di-GMP are tightly controlled (Jenal, Reinders, and Lori 2017; Gundlach et al. 2015). The synthesis and signaling capabilities of cyclic dinucleotides play an important role in regulating diverse cellular processes, while control of degradation ensures timely activation and cessation.

The discovery and identification of a mixed-base, hybrid cyclic dinucleotide comprised of one molecule of ATP and one molecule of GTP, named cyclic GMP-AMP (cGAMP), in both bacteria and eukaryotes hinted at the increasing complexity of cyclic oligonucleotide signals (Ablasser et al. 2013; Davies et al. 2012; Sun et al. 2013b; Gao, Ascano, Wu, et al. 2013). cGAMP in bacteria was thought to serve roles in virulence, while it served an innate immune function in mammalian cells (Davies et al. 2012; Ablasser et al. 2013; Sun et al. 2013a). In both bacteria and eukaryotes, cGAMP signals downstream to a receptor protein that carries out an effector function.

The architecture and regulation of an oligonucleotide second messenger-mediated response is broadly conserved between bacteria and eukaryotes. First, cyclic oligonucleotide synthases act as sensors of infection that are regulated by the presence of an activating cue or de-repression, leading to an active state capable of synthesizing a cyclic second messenger molecule. Second, the second messenger molecule binds to an effector protein, leading to higher-order oligomeric states and activation of the effector. Lastly, effector activation leads to a change in the cell that inhibits viral propagation, either through additional signaling or through cell death (Figure I.1). This framework emphasizes conserved features of several diverse pathways and enables a broader discussion of the evolution of cyclic oligonucleotide synthase enzymes in the context of innate immune signaling.

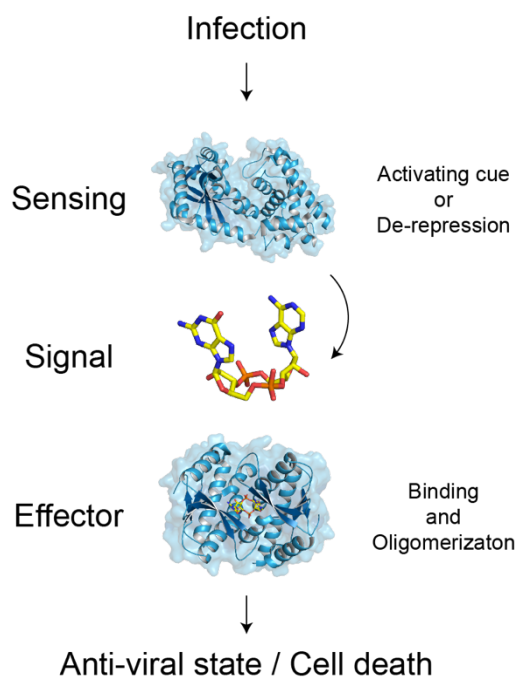


Figure I.1 Summary of a Cyclic oligonucleotide messenger-mediated response. Sensing of a cue, or de-repression, activates the synthesis of a cyclic oligonucleotide second messenger, which signals downstream to an effector, resulting in either an anti-viral state or cellular death.

Cyclic oligonucleotides in innate immunity

Eukaryotic cyclic GMP-AMP synthase (cGAS) is an innate immune sensor that upon sensing DNA undergoes a conformational change leading to activation, generating a mixed cyclic second messenger molecule out of one molecule of ATP and one molecule of GTP with one non-canonical linkage through the 2' phosphate, called 2'–5'/ 3'–5' cGMP-AMP (2',3'-cGAMP) (Ablasser et al. 2013; Civril et al. 2013; Gao, Ascano, Wu, et al. 2013). In bacteria, the first characterized example of a mixed-nucleotide synthase was discovered in the 7th pandemic strain of *Vibrio cholerae*, called Dinucleotide cyclase in *Vibrio* (VcDncV). VcDncV generates canonically linked 3'–5'/ 3'–5' cGMP-AMP (3',3'-cGAMP) and shares structural homology with mammalian cGAS (Davies et al. 2012; Kranzusch et al. 2014). The presence of the VcDncV-containing operon in the 7th pandemic strain of *Vibrio* and its absence from earlier strains suggested horizontal gene transfer events as the driver of acquisition, and the operon is contained within a mobile genetic element (Davies et al. 2012).

Recent work has expanded the field of bacterial cyclic oligonucleotide signaling to be one of the major mechanisms of anti-phage defense in bacteria (Whiteley et al. 2019; Lowey et al. 2020; Cohen et al. 2019). Bioinformatic and biochemical studies revealed a widespread network of cyclic oligonucleotide based signaling in bacteria in pathways that are analogous to eukaryotic antiviral immunity (Burroughs et al. 2015). This ancient evolutionary history reveals a much larger conserved picture of innate immune signaling that serves as the first arm of defense against invading pathogens. Bacterial homologs of cGAS contained within operons surrounded by transposon-like sequences began a new family of bacterial enzymes with cGAS and DncV as the founding members, called cGAS/DncV-like nucleotidyltransferase enzymes, or CD-NTase enzymes (Whiteley et al. 2019).

CD-NTase enzyme activation

CD-NTase enzymes responsive to an activating ligand are classified as Type I CD-NTases (Kranzusch 2019). Mammalian cGAS contains a positively charged groove situated behind the active site that binds the nucleic acid that acts as the activating cue (Civril et al. 2013). The positively charged groove is conserved within bacterial CD-NTase enzymes, suggesting that nucleic acid sensing and binding could be a conserved feature, however no currently characterized bacterial CD-NTase enzymes are responsive to nucleic acids. CD-NTases such as VcDncV that are catalytically active *in vitro* in the absence of activating

ligand are classified as Type II CD-NTases (Figure I.2). However, VcDncV is inhibited by binding to folate-like compounds, suggesting that Type II CD-NTase enzyme regulation could be dependent on binding events that suppress function (D. Zhu et al. 2014). Remaining questions in the field include the role of Type

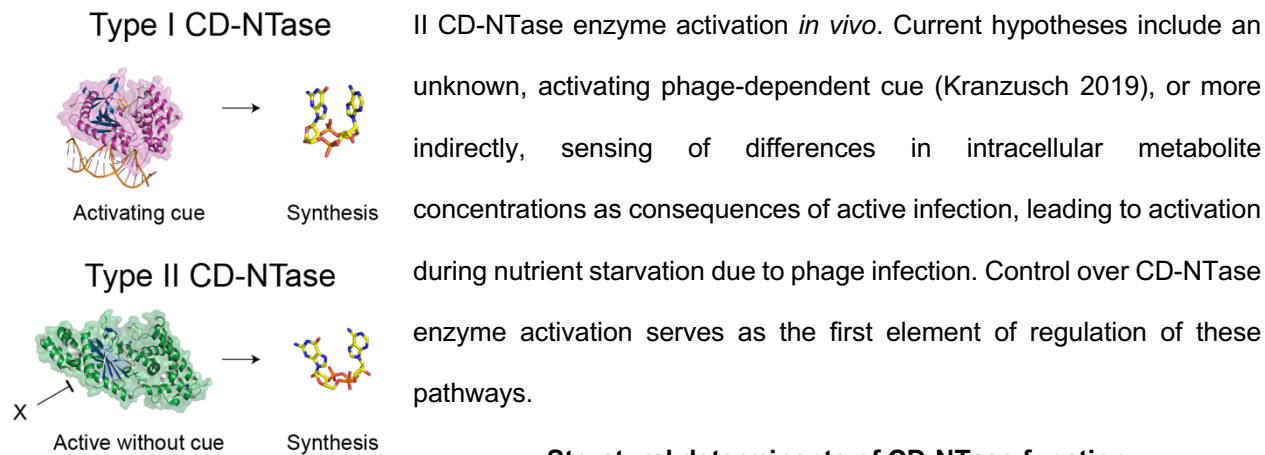


Figure I.2 Type I and Type II CD-NTase activation. Type I CD-NTases require an activating ligand for synthesis of a cyclic oligonucleotide signal. Type II CD-NTases are constitutively active *in vitro* but may be regulated by inhibitory signals *in vivo*.

II CD-NTase enzyme activation *in vivo*. Current hypotheses include an unknown, activating phage-dependent cue (Kranzusch 2019), or more indirectly, sensing of differences in intracellular metabolite concentrations as consequences of active infection, leading to activation during nutrient starvation due to phage infection. Control over CD-NTase enzyme activation serves as the first element of regulation of these pathways.

Structural determinants of CD-NTase function

Bacterial homologs of cGAS are incredibly diverse in primary sequence but are structurally homologous to mammalian cGAS, retaining a core nucleotidyltransferase fold that shares structural homology with the DNA polymerase- β -like superfamily. CD-NTase enzymes are structured with a cage-like architecture surrounding the active site, replacing the traditional binding site of a DNA or RNA template for templated polymerases (Kranzusch et al. 2014; Kottur and Nair 2018; Freudenthal et al. 2013; Yang, Lee, and Nowotny 2006). The active site is comprised of three conserved catalytic aspartates that coordinate two metal ions (Figure I.3 A). Synthesis of cyclic nucleotide second messenger molecules is accomplished through the formation of subsequent phosphodiester linkages through two-metal ion catalyzed synthesis (Yang, Lee, and Nowotny 2006) within a single active site and with a water molecule as the acting nucleophile. In contrast, canonically characterized c-di-AMP or c-di-GMP synthases rely on the formation of an active site at an oligomerization interface (Witte et al. 2008; Rosenberg et al. 2015; Chan et al. 2004).

The initial acceptor and donor nucleotides bind in distinct binding pockets in the CD-NTase active site designated as the “acceptor pocket” and “donor pocket” (Civril et al. 2013; Gao, Ascano, Wu, et al. 2013; D. Zhu et al. 2014; Kranzusch et al. 2014). The caged-like architecture of the active site supports a 180° rotation of the linear intermediate molecule for the subsequent formation of another phosphodiester

linkage to generate the final cyclic molecule, necessitating that both nucleotides be capable of binding within the acceptor pocket and donor pocket (Kranzusch et al. 2014). The binding of nucleotides within each of these pockets is dependent on specific protein–nucleotide contacts within the active site that break and reform between the formation of the linear intermediate and formation of the cyclic molecule (Figure I.3 B). However, CD-NTase second messenger synthesis proceeds through a defined reaction order, suggesting additional determinants of specificity within the acceptor and donor pockets driving product formation. For both di- and tri-oligonucleotide synthesizing CD-NTase enzymes, and likely for CD-NTases synthesizing larger oligomeric rings as well, the active site contains a single donor and acceptor pocket.

Mammalian cGAS synthesis of 2',3'-cGAMP proceeds through a pppG[2'–5']pA linear intermediate with the non-canonical bond forming first, with ATP acting as the initial phosphate donor and GTP acting as the initial phosphate acceptor and subsequent cyclization (Gao, Ascano, Wu, et al. 2013). VcDncV synthesis of 3',3'-cGAMP proceeds in the opposite order, with the formation of a pppA[3'–5']pG linear intermediate with ATP acting as the initial phosphate acceptor and GTP as the initial phosphate donor prior to subsequent cyclization (Figure I.3 C)(Kranzusch et al. 2014). VcDncV is also capable of synthesizing minor products c-di-AMP and c-di-GMP (Kranzusch et al. 2014; Davies et al. 2012). Key residues within the caged active site that affect the positioning of the bases and the 2' or 3' OH dictate the formation of an initial canonical versus non-canonical bond (Kranzusch et al. 2014).

A subset of CD-NTase enzymes contain an additional helical bundle, named the AGS-C. This c-terminal bundle is also found on some bacterial adenylate and guanylate cyclase enzymes and may act as an additional sensor domain, but the full significance and purpose of the AGS-C is not yet clear (Burroughs et al. 2015). Furthermore, a structurally related enzyme that shares the core nucleotidyltransferase fold, mammalian Oligoadenylate synthase (OAS), synthesizes a linear molecule comprised of 2'–5' linked adenine bases in response to sensing of double stranded RNA (dsRNA) (Kerr and Brown 1978; Kerr, Brown, and Hovanessian 1977; HOVANESEAN and KERR 1978). There are key sequence motifs that distinguish OAS proteins (Kjaer et al. 2009) and structural features that are unique to OAS (Lohöfener et al. 2015); however, the features that lead to the formation of a cyclic versus linear molecule are not yet well understood.

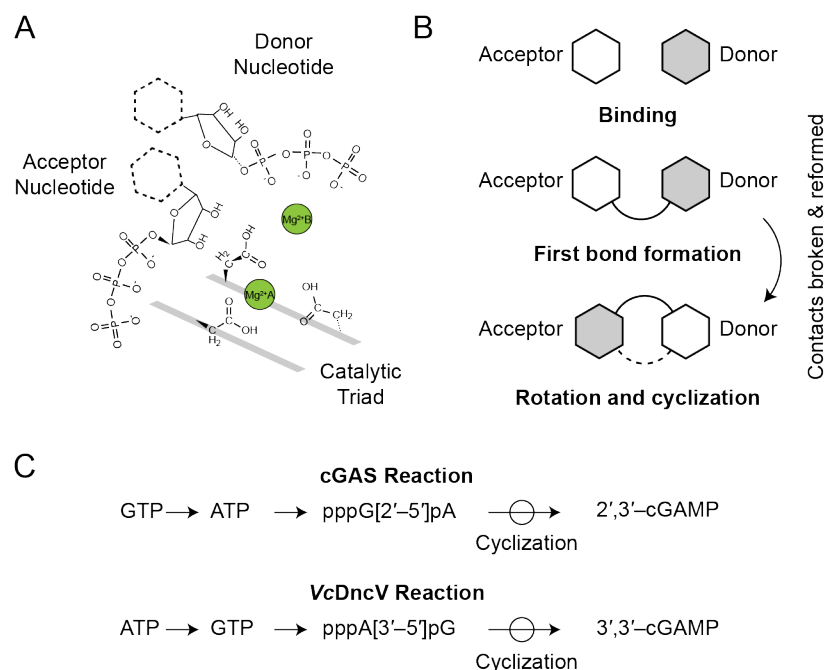


Figure I.3 The mechanism of reaction of CD-NTase enzymes follows a defined order. A) The catalytic triad coordinates two metal ions that orient the acceptor and donor nucleotide and water molecules for a nucleophilic attack for the generation of the first phosphodiester linkage. B) Acceptor and Donor nucleotides rotate 180° between the formation of the first bond and cyclization. C) cGAS and VcDncV synthesis reactions have opposite reaction orders driven by the initial recognition of nucleotides in the acceptor and donor pockets through protein–RNA contacts.

The structural determinants of CD-NTase function reveal a complex evolutionary history. The incredible sequence variation between structurally homologous bacterial CD-NTases supports a parallel diversification of possible cyclic oligonucleotide products beyond cGAMP, and indeed several different kinds of purine- and pyrimidine-containing cyclic oligonucleotides have been characterized (Whiteley et al. 2019; Lowey et al. 2020; Lau et al. 2020). Protein-RNA contacts along the lid region caging the active site are highly important for specific nucleotide selection, and single substitutions can affect the identity of the synthesized signaling molecule (Whiteley et al. 2019). *In vitro*, biochemical characterization of CD-NTase enzymes reveals a capacity for synthesizing major products and less-abundant minor products (Whiteley et al. 2019; Lowey et al. 2020). The determinants of how CD-NTase enzymes synthesize the biologically relevant signaling molecule is an ongoing topic of study, but recognition by the downstream receptor/effector provides one method of classifying the biologically relevant activating molecule synthesized by CD-NTase enzymes.

Bacterial CD-NTase-containing operons

Related bacterial genes are organized into operons. CD-NTase enzyme-containing defense operons contain second messenger responsive effectors that act as executors of cellular death (Cohen et al. 2019; Davies et al. 2012; Whiteley et al. 2019). Further bioinformatic work identified additional systems that were toxic to bacteria upon activation in a second-messenger dependent manner. These defense operons are named cyclic-oligonucleotide-based anti-phage signaling systems (CBASS) and function in anti-phage bacterial immunity (Cohen et al. 2019). CBASS operons likely function as abortive infection modules that execute altruistic bacterial cell death to prevent phage propagation and otherwise destruction of the bacterial colony (Millman et al. 2020; Lopatina, Tal, and Sorek 2020).

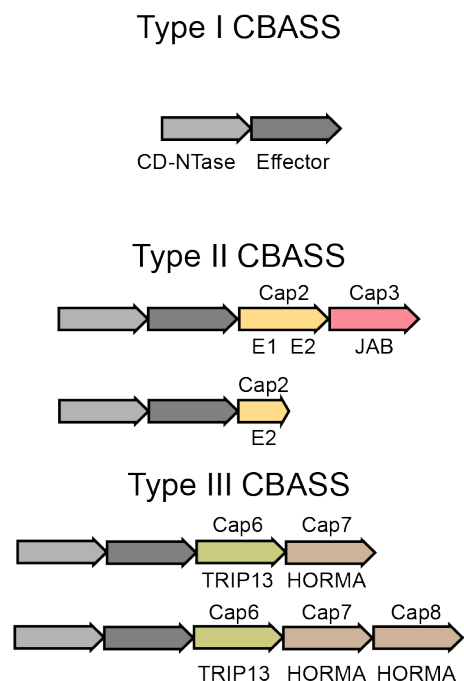


Figure I.4 CBASS operons are organized into three different types. The simplest CBASS operons, Type I, contain only the CD-NTase and effector proteins, while more complex operons contain a diverse array of ancillary genes, named CD-NTase associated proteins (Cap).

CBASS operons are organized into three types of increasing complexity (Figure I.4). The simplest CBASS operon is a Type I operon that contains a CD-NTase and an effector gene (Millman et al. 2020). More complex CBASS operons contain various ancillary genes, named CD-NTase associated proteins (Cap) in addition to the CD-NTase and effector (Lowey et al. 2020; Cohen et al. 2019), including JAB proteases/deubiquitinases, E1-E2 like ubiquitin ligase-like proteins, HORMA domain containing proteins, and TRIP-13 like proteins. In addition to a CD-NTase and effector, Type II CBASS operons contain an E1 E2 like protein and a JAB protease, but the JAB-protease like protein may be absent. Type III operons are comprised of Trip 13 and HORMA domain containing regulatory genes that are required for CD-NTase activation and regulation (Ye et al. 2020; Lau et al. 2020).

The abortive infection model of bacterial immunity operates in tandem with the CRISPR-based adaptive immune system of bacteria and provides the framework for a pan-bacteria immune system and survival of the bacterial colony. Effector activation can be modified by ancillary genes, but further work is required to

elucidate the roles of ancillary genes in the CBASS operons and their interplay with other aspects of bacterial immunity (Bernheim and Sorek, 2020).

Downstream effectors and functions

Both bacterial and mammalian CD-NTase dependent responses can be organized into similar response structures. The cGAS-STING innate immune axis depends on DNA-induced 2',3'-cGAMP synthesis that signals downstream to the effector, STING (Kranzusch et al. 2015; Shang et al. 2019; Ishikawa and Barber 2008; Zhong et al. 2008; Burdette et al. 2011). In mammals, STING binding to 2',3'-cGAMP initiates a conformational change and oligomerization of STING that leads to IRF-3 and TBK-1 recruitment through the C-terminal tail (CTT) (C. Zhang et al. 2019; Zhao et al. 2019; de Oliveira Mann et al. 2019). The CTT is a unique feature of vertebrate STING and likely evolved after the acquisition of the core STING domain (de Oliveira Mann et al. 2019). TBK1 phosphorylation of IRF-3 and its relocation to the nucleus initiates the production of interferon, which activates a potent anti-viral state in the cell through the induction of interferon stimulated genes (ISGs) (C. Zhang et al. 2019). STING signaling requires oligomerization of STING and relocation from the ER to the Golgi. After activation STING is rapidly degraded through initiation of autophagy, regulating the response (Gui et al. 2019). STING signaling and the onset of an anti-viral state in the cell constitutes the downstream effector function in this arm of the mammalian antiviral response, rather than a direct effector function (Figure I.5 A).

As part of the ongoing arms race between viruses and host cells, Vaccinia virus (VACV) encodes a specialized phosphodiesterase, named poxvirus immune nuclelease, or poxin (Eaglesham et al. 2019). Poxin specifically cleaves 2',3'-cGAMP during VACV infection and acts as a viral suppressor of the cGAS-STING innate immune pathway. Poxin homologs in moths, butterflies, and insect viruses again highlight the ancient evolutionary origins of innate immune signaling, however no mammalian homologs of poxin have been identified. Whether an alternative means of degrading 2',3'-cGAMP in the cytoplasm of cells exists in mammals, or whether 2',3'-cGAMP levels are regulated through an entirely different mechanism remains to be seen.

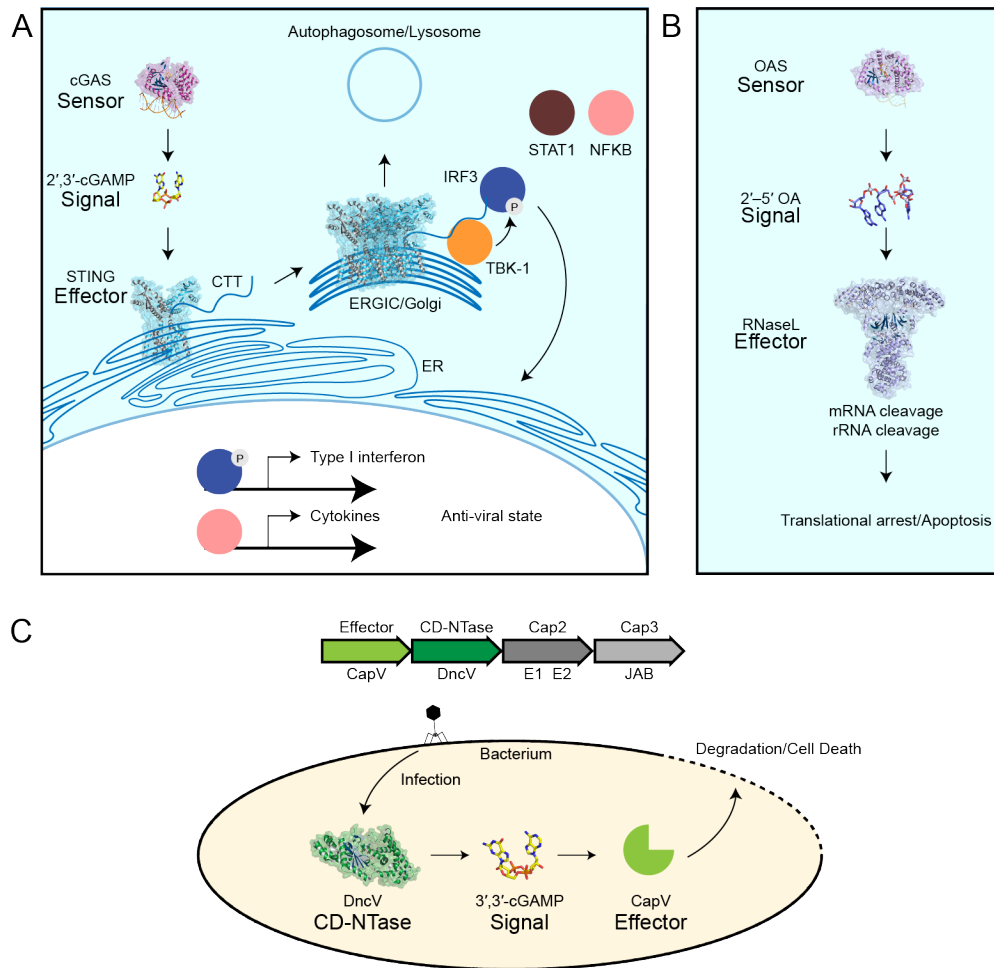


Figure I.5 Conserved response architecture in metazoan and bacterial antiviral signaling. A) The sensor, mammalian cGAS, signals downstream to effector STING through 2',3'-cGAMP, inducing an antiviral state in the cell. B) Sensor OAS signals downstream to effector RNase L through 2'-5' OA, inducing mRNA and rRNA cleavage, translational arrest, and apoptosis. C) Sensor DncV signals downstream to effector CapV through 3',3'-cGAMP, inducing phospholipid degradation of the inner membrane and cell death.

Comparable to the cGAS-STING innate immune axis, OAS proteins are ISGs that sense dsRNA and synthesize 2'-5' oligoadenylate (2'-5' OA) molecules and signal downstream to an effector ribonuclease that is activated upon binding to 2'-5' OA. Ribonuclease L (RNase L) oligomerizes and cleaves single stranded RNA in the cytoplasm, degrading messenger RNA transcripts and cleaving ribosomal RNA, leading to translational shut down and eventual apoptosis (Zilberstein et al. 1978; G. Li et al. 2004; Dong and Silverman 1995; Cooper et al. 2014; Catherine Bisbal and Silverman 2007). mRNA degradation and translational shutdown are the direct effector functions (Figure I.5 B). Functional OAS proteins are found in evolutionarily ancient organisms that predate RNase L, which evolved in jawed vertebrates, suggesting that alternative receptors of 2'-5' OA with possible novel effector functions existed

prior to RNase L evolution (Kjaer et al. 2009; Kuusksalu et al. 1995; Reintamm et al. 2008). 2'–5' OA is degraded by PDE-12 in cells and *in vitro* (Wood et al. 2015).

In bacteria, pioneering work with the *VcDncV* operon showed the CD-NTase-containing “defense island” encoded a phospholipase gene named *CapV* adjacent to *VcDncV* and responsive to 3',3'-cGAMP. *CapV* degrades phospholipids of the inner membrane in a 3',3'-cGAMP dependent manner, leading to bacterial cell death (Severin et al. 2018) (Figure 1.5 C). Several new studies have described effector functions in bacteria. New bacterial effectors within CBASS operons were identified through bioinformatical analysis or structural homology to genes of known function, followed by biochemical validation of effector function. Bacterial effectors regulated by CBASS operons encompass several classes of enzymes and are the focus of extensive characterization. Bacterial effector architecture consists of a nucleotide second messenger sensing domain and an executor domain. Several effectors consist of a conserved nucleotide sensing domain named SAVED that shares structural homology to CARF-domain containing CRISPR nucleases (Lowey et al. 2020). Currently identified executor domains include restriction endonucleases, HNH-nucleases, lipases, SLATT-containing domains, TM-containing domains, and TIR domains, among others (Millman et al. 2020).

Effectors within CBASS operons are highly specific to the second messenger molecule synthesized by their cognate CD-NTase and crosstalk between different second messengers is limited (Whiteley et al. 2019). Effector selectivity likely prevents aberrant activation of effector function that could lead to bacterial death and raises questions about the evolution of specificity of the system in response to phage-driven selection pressure. CBASS operons are acquired and lost rapidly through horizontal gene transfer, and as such the health and survival of the overall bacterial colony depends on the diversity of the systems encoded within subpopulations of bacteria (Bernheim and Sorek, 2020; Ofir *et al.*, 2018). Bacterially encoded endogenous phosphodiesterase proteins capable of restricting cyclic oligonucleotides, and phage encoded anti-CBASS phosphodiesterase or nuclease proteins are active areas of study.

Molecular patterns

Innate immune responses depend on the recognition of specific patterns that drive specific responses. Separate sensing and effector execution allows for both greater control over the initiation and execution of cellular defense and for the amplification of a signal. Both CD-NTase enzymes and effectors

are proteins that respond to specific molecular patterns mediated by both direct hydrogen bonding interactions as well as water mediated contacts. Overall, the specificity and functionality of each of these defense pathways is dependent on a series of molecular protein-RNA contacts that drive functional conformational changes or induce oligomerization to control activation, with the goal of preventing viral propagation.

A careful study of these molecular contacts and the evolutionary forces driving the acquisition and maintenance of these pathways allows us to probe these systems on a molecular level to explore both the mechanisms and limits of recognition and signaling. Unifying these different signaling pathways within a common scaffold of innate immune signaling helps highlight the evolutionary history while the differences hint at specialization that opens further avenues of study.

Chapter 1: Molecular basis of CD-NTase nucleotide selection in CBASS anti-phage defense

Apurva A. Govande^{1,2}, Brianna Lowey^{1,2}, James B. Eaglesham^{1,2}, Aaron T. Whiteley³, 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 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 may be accessed as follows:

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

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

Abstract

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; D. 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 *EcCdnD* 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 (*EcCdnD*) in complex with two molecules of ATP (1.45 Å) (Table S1). The *EcCdnD* 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 1.1a) (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, Ascano, Wu, et al. 2013; Kranzusch 2019; Kranzusch et al. 2014; Whiteley et al. 2019; D. Zhu et al. 2014). *EcCdnD* 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 1.1b, left). The *EcCdnD* donor ATP adopts a similar orientation to the donor nucleotide in structures of the 3'3'-cGAMP synthase *Vibrio cholerae* DncV (*VcDncV*) (Figure 1.1c, 1.2a inset) (Kato et al. 2015; Kranzusch et al. 2014; D. Zhu et al. 2014). In contrast, the ATP within the acceptor pocket of *EcCdnD* is held in an unexpected inverted orientation that is rotated 120° relative to the *VcDncV* acceptor nucleotide (Figure 1.1 b,c).

To further understand the positioning of the nucleotides in complex with *EcCdnD*, we screened related Clade D CD-NTases and determined a crystal structure of *Salmonella enterica* CdnD (2.60 Å, ~80% identical to *EcCdnD*, 0.48 Å Cα RMSD) in complex with two molecules of GTP (Table 1.1, Figure 1.2b). The *SeCdnD*–GTP structure reveals nearly identical nucleotide positioning with an inverted acceptor nucleotide (Figure 1.2b) suggesting that the *EcCdnD*–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 1.1 b, c). The inverted acceptor nucleotide is further coordinated by an auxiliary metal-binding site created between D196 and the nucleotide triphosphate (Figure 1.1d). *EcCdnD* and *SeCdnD* are most closely related to previous structures *Pseudomonas aeruginosa* CdnD (*PaCdnD*, 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 *PaCdnD* (Figure 1d, Figure S1c). 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 *EcCdnD* that allows the inverted ATP to extend from the active site (Figure 1.2 d,f). We mutated *EcCdnD* 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 1.1e). In contrast, a D71N mutation of the main active site results in complete abrogation of *EcCdnD* enzymatic activity (Figure 1.1e).

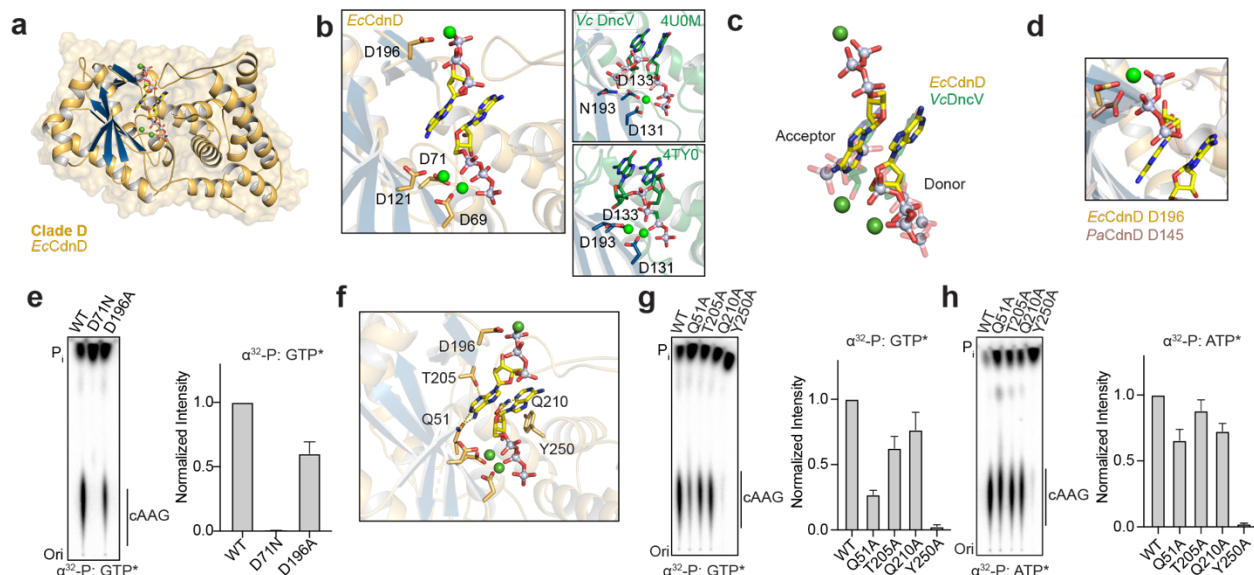


Figure 1.1. Crystal structure of the 3'3'3'-cAAG synthase *EcCdnD* in a pre-reactive state

a. Crystal structure of the Clade D CD-NTase from *Enterobacter cloacae* (*EcCdnD*) in complex with two molecules of ATP. *EcCdnD* shares all conserved structural features of the CD-NTase family of enzymes. **b.** Left: Zoom-in cutaway of the *EcCdnD* 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 *VcDncV* with catalytic aspartates D131 and D133 coordinating a single magnesium ion and ATP and GTP in an orientation conducive to bond formation. A *VcDncV* D193N mutation stalls the reaction in a pre-reactive state (PDB 4U0M). Bottom, active site of *VcDncV* 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 *EcCdnD* crystal structure, in yellow, and pre-reactive ATP and GTP from *VcDncV* crystal structure, in green (PDB 4U0M). Phosphate donor nucleotides are oriented similarly while the phosphate acceptor nucleotide in *EcCdnD* is inverted relative to the *VcDncV* acceptor nucleotide. **d.** Overlay highlighting *EcCdnD* D196 and coordination of a third metal ion compared to the crystal structure of a related Clade D CD-NTase from *Pseudomonas aeruginosa*, *PaCdnD*, 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 (TLC) analysis of *EcCdnD* 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 *EcCdnD* crystal structure. **g and h.** TLC analysis and quantification of nucleotide-coordinating residues. Q51A mutation affects the incorporation of GTP, but shows little defect in ATP incorporation and synthesizes minor product 3'3'3'-cAAA. Y250A mutation completely abrogates synthesis of 3'3'3'-cAAG. All TLC data are representative of three independent experiments, and error bars denote the standard deviation.

Table 1.1. Summary of crystallographic statistics

	<i>Enterobacter cloacae</i> CdnD	<i>Enterobacter cloacae</i> CdnD (SeMet)	<i>Salmonella enterica</i> CdnD	<i>Bacteroides fragilis</i> CdnB	<i>Bacteroides fragilis</i> CdnB (SeMet)
Data Collection					
Resolution (Å) ^a	1.45	1.498	2.59	1.76	3.095
Wavelength (Å)	0.97918	0.97949	0.97918	0.97918	0.97910
Space group	P 1 2 ₁ 1	P 1 2 ₁ 1	P 1 2 ₁ 1	P 2 ₁ 2 ₁ 2 ₁	P 3 ₁ 2 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	63.37 69.61 92.40	69.71 69.71 182.71
Unit cell: α, β, γ (°)	90.00 117.61 90.00	90.00 118.04 90.00	90.00 110.35 90.00	90.00 90.00 90.00	90.00 90.00 120.00
Molecules per ASU	2	2	2	2	2
Total reflections	987705	5739620	93109	273132	189210
Unique reflections	140248	128307	26345	41249	9979
Completeness (%) ^a	99.5 (97.8)	99.2 (92.9)	99.3 (96.4)	99.9 (99.1)	99.0 (100.0)
Multiplicity ^a	7.0 (6.8)	44.7 (31.1)	3.5 (3.5)	6.6 (4.5)	19.0 (18.1)
I/σI ^a	13.8 (1.4)	13.6 (0.9)	3.5 (1.1)	10.6 (1.2)	9.5 (1.9)
CC(1/2) ^b (%) ^a	99.9 (61.9)	99.9 (37.8)	98.5 (38.1)	99.7 (36.4)	99.5 (52.1)
Rpim ^c (%) ^a	2.5 (55.8)	4.2 (91.9)	13.5 (96.5)	4.1 (83.6)	9.2 (107.5)
Sites		12			6
Refinement					
Resolution (Å)	38.87 – 1.45		38.49 – 2.59	46.862 – 1.760	
Free reflections	2046		1983	2000	
R-factor / R-free	15.2/16.5		22.9/26.5	17.8/19.5	
Bond distance (RMS Å)	0.008		0.003	0.003	
Bond angles (RMS °)	1.04		0.66	0.79	
Structure/Stereochemistry					
No. atoms: protein	5615		5581	2356	
No. atoms: ligand	156		134	55	
No. atoms: water	806		91	394	
Average B-factor: protein	26.6		54.6	31.0	
Average B-factor: ligand	19.8		50.9	49.0	
Average B-factor: water	36.4		46.1	41.3	
Ramachandran plot: favored	97.96%		96.62%	98.26%	
Ramachandran plot: allowed	2.04%		3.67%	1.74%	
Ramachandran plot: outliers	0.00%		0.00%	0.00%	
Rotamer outliers	0.33%		1.15%	0.39%	
MolProbity ^d score	1.00		1.62	1.06	

Protein Data Bank ID	7LJL	7LJM	7LJO
----------------------	------	------	------

^a Highest resolution shell values in parentheses

^b (Karplus and Diederichs, 2012)

^c (Weiss, 2001)

^d (Chen et al., 2010)

	<i>Bradyrhizobium diazoefficiens</i> CdnG	<i>Bradyrhizobium diazoefficiens</i> CdnG (SeMet)
Data Collection		
Resolution (Å) ^a	1.598	3.685
Wavelength (Å)	0.97920	0.97920
Space group	P 1 2 ₁ 1	P 6 ₂
Unit cell: a, b, c (Å)	112.00 68.69 112.43	128.99 128.99 59.56
Unit cell: α, β, γ (°)	90.00 108.40 90.00	90.00 90.00 120.00
Molecules per ASU	2	2
Total reflections	725390	498887
Unique reflections	212025	6251
Completeness (%) ^a	99.0 (90.7)	99.6 (98.7)
Multiplicity ^a	3.4 (3.2)	79.8 (75.7)
I/σ ^a	9.1 (2.7)	9.0 (1.4)
CC(1/2) ^b (%) ^a	99.5 (76.1)	100.00 (75.0)
Rpim ^c (%) ^a	4.5 (32.6)	8.8 (270.3)
Sites		6
Refinement		
Resolution (Å)	37.94 – 1.59	
Free reflections	2018	
R-factor / R-free	19.7/21.7	
Bond distance (RMS Å)	0.006	
Bond angles (RMS °)	0.86	
Structure/Stereochemistry		
No. atoms: protein	10660	
No. atoms: ligand	132	
No. atoms: water	2038	
Average B-factor: protein	22.3	
Average B-factor: ligand	13.4	
Average B-factor: water	31.8	
Ramachandran plot: favored	99.10%	
Ramachandran plot: allowed	0.90%	
Ramachandran plot: outliers	0.00%	
Rotamer outliers	0.54%	
MolProbity ^d score	1.37	
Protein Data Bank ID	7LJN	

^a Highest resolution shell values in parentheses

^b (Karplus and Diederichs, 2012)

^c (Weiss, 2001)

^d (Chen et al., 2010)

We next analyzed nucleotide contacts in the active site of *EcCdnD* (Figure 1.1f). 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 1.1g, Figure 1.2 e), while a Q210A mutation has little effect on product formation. The inverted ATP in the *EcCdnD* 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 1.1f). Interestingly, *EcCdnD* Q51 is positioned to make Watson-Crick edge contacts that discriminate between ATP and GTP. We observed that an *EcCdnD* Q51A mutant enzyme exhibits a defect in incorporation of GTP resulting in increased synthesis of a minor reaction product 3'3'3'-cAAA (Figure 1.1h). Similarly, a Q51S (present in *E. coli* CdnC) or Q51T (present in *PaCdnD*) 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 1.2f). Together, these results demonstrate key hydrogen bonding interactions in the *EcCdnD* active site that control substrate recognition and nucleotide specificity.

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

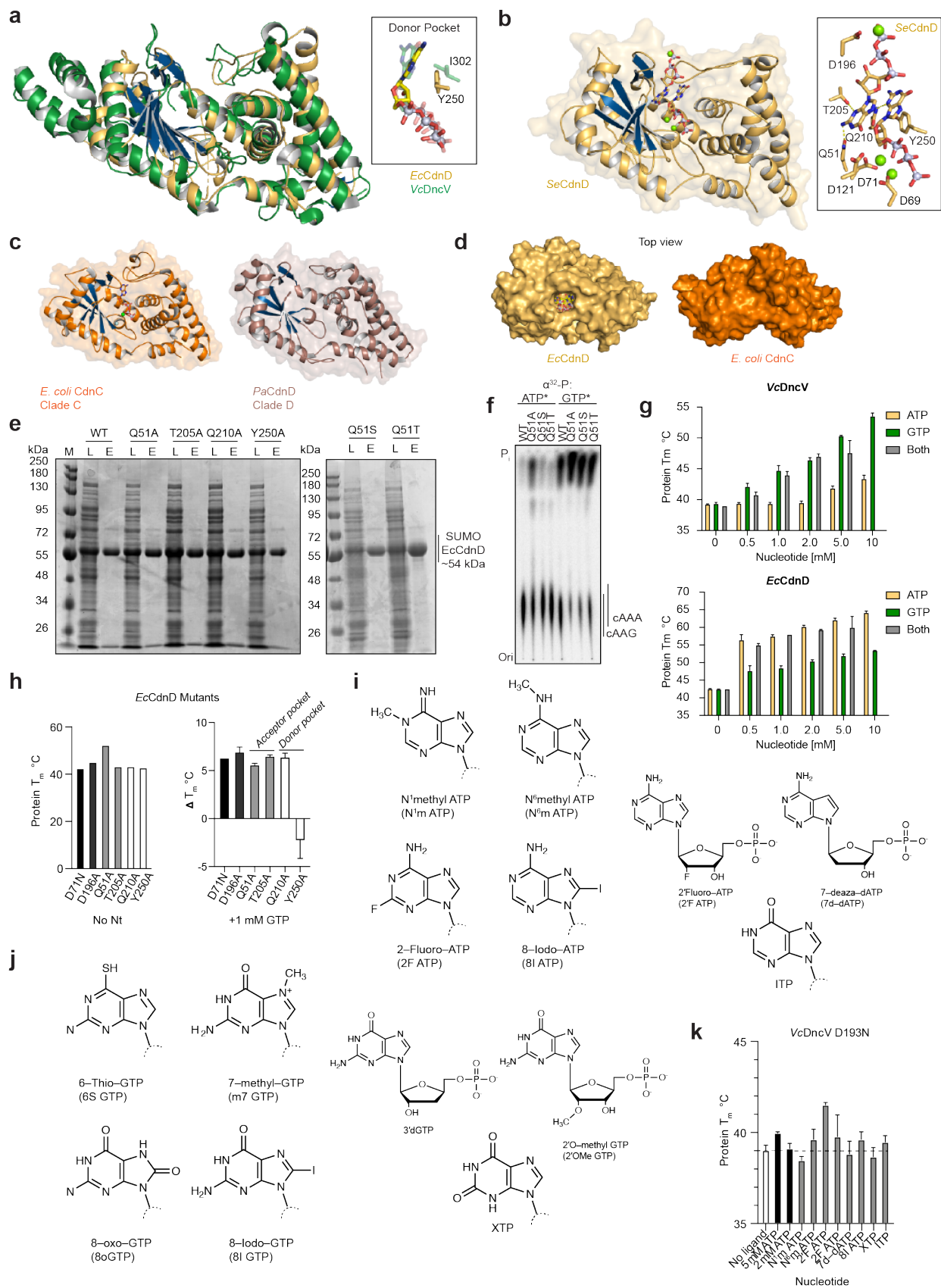
We next compared interactions in the *VcDncV* and *EcCdnD* acceptor and donor nucleotide pockets. We purified catalytically inactive *VcDncV* D193N and *EcCdnD* D71N variants with mutations that prevent reaction progression but do not disrupt nucleotide coordination (Kranzusch et al. 2014; D. Zhu et al. 2014), and used a thermo-denaturation assay to measure the ability of nucleotide substrates to increase enzyme stability. *VcDncV* 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; D. 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 *VcDncV* shifting $\sim 13^\circ\text{C}$ (Figure 1.3a). In contrast, the T_m of *VcDncV* shifts only 4°C in the presence of ATP (Figure 1.3a). We observed no synergistic effect when equimolar amounts of ATP and GTP are present, indicating enhanced thermostability is dependent on donor nucleotide interactions (Figure 1.2g). Nucleotides exhibit a similar ability to stabilize the T_m of *EcCdnD*, with ATP resulting in a $\sim 21^\circ\text{C}$ shift and GTP resulting in a 9°C shift (Figure 1.3a, Figure 1.2g). Enhanced stability

of the *EcCdnD*–ATP complex suggests that ATP is selectively used as the first donor nucleotide during 3'3'3'-cAAG formation. Notably, disruption of the *EcCdnD* donor pocket residue Y250 required for nucleobase stacking ablates all ATP-induced enhancement of enzyme stability (Figure 1.3b, Figure 1.2h). In contrast, mutations to the acceptor pocket and auxiliary metal binding site again demonstrate these interactions have a minor effect on thermo-stabilization (Figure 1.3b, Figure 1.2 g, 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; Z. Wang et al. 2021), we next used a panel of 14 nucleotide analogs to determine how individual contacts impact CD-NTase substrate selectivity. The nucleotide panel includes ATP and GTP analogs with nucleobase and ribose chemical modifications (Figure 1.3c, Figure 1.2i, j) and reveals major features controlling CD-NTase specificity. First, the data demonstrate that both *VcDncV* and *EcCdnD* are tolerant to a wide variety of nucleotide modifications (Figure 1.3d). Chemical modifications to the Watson-Crick and Hoogsteen edges do not disrupt binding and enhancement of CD-NTase thermostability (Figure 1.3d, Figure 1.2k). 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 *EcCdnD* active site used for cyclic trinucleotide synthesis is more accommodating of modified nucleotides compared to the *VcDncV* active site used for cyclic dinucleotide synthesis (Figure 1.3d).

Figure 1.2 *EcCdnD* and *VcDncV* structural comparison and complex formation with nucleotide substrates

a. Structural overlay and comparison of *EcCdnD* and *VcDncV*. *EcCdnD* shares the pol- β like palm domain and core CD-NTase fold. Inset, detail of phosphate donor nucleotide from *EcCdnD* and *VcDncV*. Y250 stacks the donor nucleotide in the *EcCdnD* structure, while I302 plays a similar role in *VcDncV*. **b.** Crystal structure of GTP bound *SeCdnD*, inset shows GTP nucleotides and conserved coordinating residues. **c.** Crystal structures of cyclic trinucleotide synthases *E. coli* CdnC (PDB 6P80) and *PaCdnD* (PDB 6P82). **d.** Left, top view surface representation of *EcCdnD* 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 *EcCdnD* mutants showing expression level and purity. L, clarified lysate. E, elution of final Ni-NTA purified protein with a 6 $\times$ 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 T_m for *VcDncV* D193N and *EcCdnD* 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.** *EcCdnD* mutant T_m without nucleotides or ΔT_m of *EcCdnD* mutants bound to 1 mM GTP. $\Delta T_m = T_m(1 \text{ mM GTP}) - T_m(\text{Apo})$. 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: N¹m ATP, N⁶m 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 second major feature of CD-NTase nucleotide specificity is preferential interaction with the nucleobase Watson-Crick edge. Both *VcDncV* and *EcCdnD* maintain specific interactions with the ATP N1 position and the GTP N1 and N2 positions. *VcDncV* 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

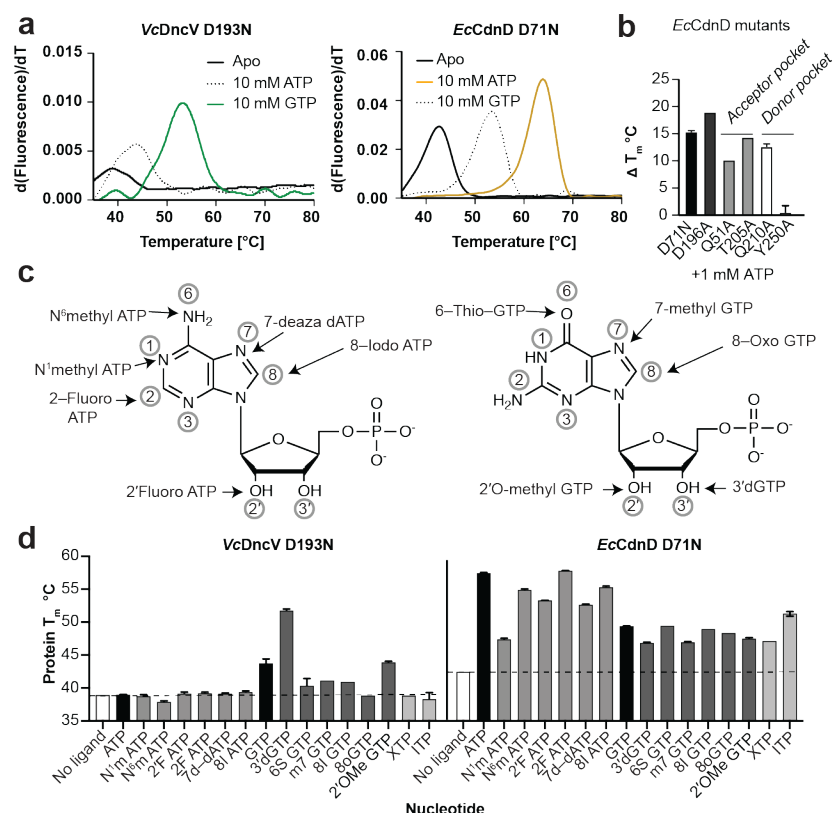


Figure 1.3. Stable CD-NTase-donor nucleotide complex formation is a key determinant of product specificity. **a.** Left, Protein fluorescence over change in temperature for catalytically inactive *VcDncV* D193N. Peak indicates protein melting temperature (T_m) with a rightward shift indicating greater stabilization. GTP, which initially binds in the donor pocket in the pre-reactive state, stabilizes *VcDncV* to a greater extent than ATP. Right, T_m plot for catalytically inactive *EcCdnD* D71N, showing that ATP stabilizes to a greater extent than GTP. Data are representative of three independent experiments. **b.** ΔT_m of *EcCdnD* mutants bound to 1 mM ATP. $\Delta T_m = T_m(1 \text{ mM ATP}) - T_m(\text{Apo})$. 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 *VcDncV* D193N and *EcCdnD* 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.

VcDncV-ATP complex (Figure 1.3d). The enhanced *VcDncV* thermostability upon interaction with GTP is therefore dependent on hydrogen bond contacts with the N2 amine. *EcCdnD* 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 *EcCdnD* to preferentially interact with GTP (Figure 1.3d). 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 *VcDncV* while slightly increasing the stability of *EcCdnD* (Figure 1.3d, Figure 1.2k). For the ATP Watson-Crick edge, modifications to the N1 position destabilize the interaction with both *VcDncV* and

EcCdnD while modifications to the N6 position destabilize interaction with *VcDncV* but only have a minimal effect on *EcCdnD* stability (Figure 1.3d, Figure 1.2k). Strikingly, modification of the N2 position of ATP with 2-fluoro-ATP does not impair the ability of ATP to stabilize either enzyme (Figure 1.3d, Figure 1.2k). Overall, these interactions with modified nucleotides agree with contacts observed in the crystal structures of *EcCdnD* and *VcDncV* intermediate states (Figure 1.1) (Kranzusch et al. 2014; D. 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 *VcDncV* but not *EcCdnD* (Figure 1.3d). 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 1.3d, Figure 1.2k). Notably, chemical modifications at nucleobase position 8 reveal additional differences between *VcDncV* and *EcCdnD* with 8-iodo-GTP and 8-oxo-GTP stabilizing *EcCdnD* while only 8-iodo-GTP stabilizes *VcDncV* and 8-oxo-GTP has a specific destabilizing effect (Figure 1.3d). Likewise, *VcDncV* tolerates ribose 2' or 3' OH modifications while *EcCdnD* is destabilized (Figure 1.3d). 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 *VcDncV* and *EcCdnD* are ~10× slower than DNA-activated mouse cGAS or human cGAS (Figure 1.4a, b, Figure 1.5a)(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 additional signals. Using nonhydrolyzable nucleotide analogs Apcpp or Gpcpp, we observed that the initial *EcCdnD* reaction steps occur with ATP and Gpcpp but not with GTP and Apcpp (Figure 1.5b; Table 1.2). Furthermore, in the presence of ATP and Apcpp, *EcCdnD* forms a linear reaction intermediate, while GTP and Gpcpp are not sufficient for *EcCdnD* product formation (Figure

1.4c). These results confirm that *EcCdnD* 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 1.5c) (Kranzusch et al. 2014; Whiteley et al. 2019).

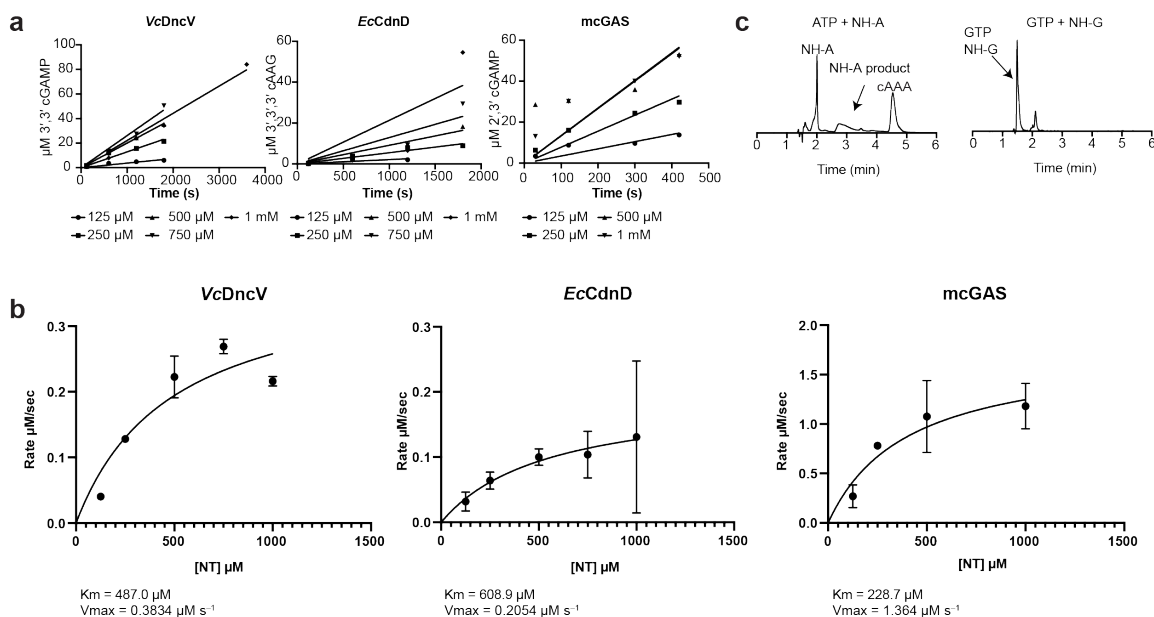


Figure 1.4 Measurement of *VcDncV* and *EcCdnD* reaction rates

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.

As previously demonstrated for *VcDncV* (Launer-Felty and Strobel 2018) we observed that both *VcDncV* and *EcCdnD* are capable of incorporating nucleotide analogs and synthesizing a diverse array of cyclic oligonucleotide species (Figure 1.5, Figure 1.6, Table 1.2). For each nucleotide analog, we calculated the IC_{50} as the concentration of nucleotide required to decrease product formation by 50% (Figure 1.5d, Figure 1.6). In each case, nucleotide analog incorporation is more efficient in the presence of the correct ATP or GTP partner nucleotide indicating maintenance of the preference for synthesizing mixed-base

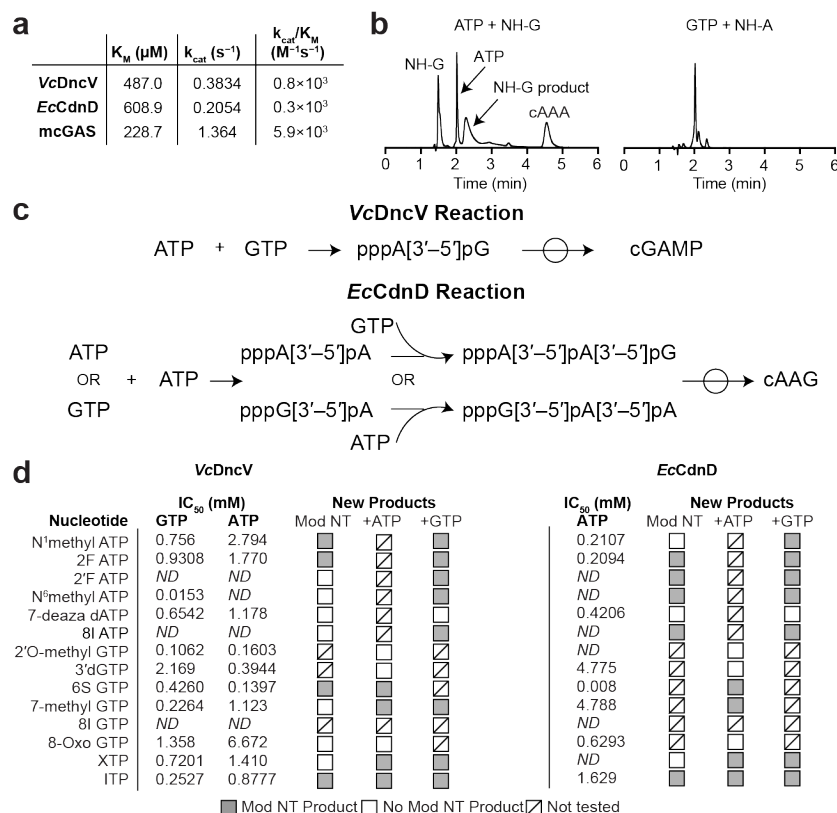


Figure 1.5 Nucleotide interference mapping analysis of CD-NTase reaction progression

a. Comparison of enzyme kinetics between bacterial synthases VcDncV and EcCdnD, 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 $\sim 10\times$ less than mcGAS. **b.** High-performance liquid chromatography analysis of EcCdnD 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 VcDncV 3'3'-cGAMP synthesis and EcCdnD 3'3'3'-cAAG synthesis. The circle symbol indicates cyclization of the linear intermediate into a final cyclic product. **d.** Summary of VcDncV and EcCdnD 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 (VcDncV) or 50% inhibition of 3'3'3'-cAAA formation (EcCdnD), with increasing concentrations of modified nucleotides. See supplemental information for more detail. ND, not determined.

confirm our thermostability findings and further explain the donor and acceptor pocket interactions controlling CD-NTase product specificity.

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 VcDncV and 3'dGTP for EcCdnD; nucleotides capable of binding but not reacting (no modified product formation, $IC_{50} < 0.5$ mM) including 2'O-methyl GTP for VcDncV and N1 methyl ATP for EcCdnD; and nucleotides capable of binding and reacting (modified product formation, $IC_{50} < 0.5$ mM) including 2F ATP for both VcDncV and EcCdnD. Interestingly, we observed EcCdnD reactions with 6-thio-GTP exhibit enhanced product formation ($IC_{50} < 0.01$ mM), suggesting some modified nucleotides are more efficiently incorporated (Figure 1.5d, Figure 1.6 g). Together these results

Table 1.2

Nucleotides	Reaction step	Expected major product	Chemical formula	Product Molecular mass	Unreacted masses	Found Mass	Notes
ATP + GTP	Completion	c[A(3',5')pA(3',5')pG(3',5')p]	C30H36N15O19P3	1003.15	507.00, 522.99	1003.153	
ATP	Completion	c[A(3',5')pA(3',5')pG(3',5')p]	C30H36N15O18P3	987.16	507	987.16	
GTP	No product expected	c[G(3',5')pG(3',5')pG(3',5')p]	C30H36N15O21P3	1035.14	522.99	522.991, 1035.14	Abundant unreacted GTP and GDP masses found
NH-A + ATP	Donor ATP	ppcpA(3',5')pA(3',5')pA	C31H42N15O24P5	1163.12	505.02, 507.00	834.07	Not most abundant
		ppcpA(3',5')pA	C21H30N10O18P4	834.07			
NH-A + GTP	No donor GTP	ppcpA(3',5')pG	C21H30N10O19P4	850.06	505.02, 522.99	Some 850.06, 363.058	Product mass not the most abundant, GMP mass found
NH-G + ATP	Donor ATP	ppcpG(3',5')pA(3',5')pA	C31H42N15O25P5	1179.12	521.01, 507.00	850.06, 521.011	
		ppcpG(3',5')pA	C21H30N10O19P4	850.06			
Modified nucleotide	Enzyme	Expected major product	Chemical formula	Molecular mass	Found Mass	Notes	
N1 methyl ATP	VcDncV	c[N1 methyl A(3',5')pN1 methyl A(3',5')p]	C22H28N10O12P2	686.14	686.136	Unreacted mass also present	
N6 methyl ATP	VcDncV	c[N6 methyl A(3',5')pN6 methyl A(3',5')p]	C22H28N10O12P2	686.14	Not found	Differs from HPLC result	
2F ATP	VcDncV	c[2F A(3',5')p2F A(3',5')p]	C20H22F2N10O12P2	694.09	694.086		

2'F ATP	VcDncV	c[2'F A(3',5')p2'F A(3',5')p]	C20H22F2N10O10P2	662.1	Not found	Unreacted mass present
7 deaza dATP	VcDncV	No product expected	C11H17N4O12P3	490.01	490.006	Unreacted mass present
8I ATP	VcDncV	c[8I A(3',5')p8I A(3',5')p]	C20H22I2N10O12P2	909.9	Not found	Differs from HPLC result
6S GTP	VcDncV	c[6S G(3',5')p6S G(3',5')p]	C20H24N10O12P2S2	722.05	722.05	
Methyl-7 GTP	VcDncV	No product expected	C11H18N5O14P3	537.01	537.006	Unreacted mass present
8-oxo GTP	VcDncV	No product expected	C10H16N5O15P3	538.99	538.986	Unreacted mass present
XTP	VcDncV	No product expected	C10H15N4O15P3	523.97	523.975	Unreacted mass present
ITP	VcDncV	c[(3',5')pl(3',5')p]	C20H22N8O14P2	660.07	660.073	
N1 methyl ATP + GTP	VcDncV	c[N1 methyl A(3',5')pG(3',5')p]	C21H26N10O13P2	688.12	688.116	Unreacted mass also present
N6 methyl ATP + GTP	VcDncV	c[N6 methyl A(3',5')pG(3',5')p]	C21H26N10O13P2	688.12	688.116	
2F ATP + GTP	VcDncV	c[2F A(3',5')pG(3',5')p]	C20H23FN10O13P2	692.09	692.091	
2'F ATP + GTP	VcDncV	c[2'F A(3',5')pG(3',5')p]	C20H23FN10O12P2	676.1	676.096	
7 deaza dATP + GTP	VcDncV	No product expected	C11H17N4O12P3	490.01	490.006	Unreacted mass present
8I ATP + GTP	VcDncV	c[8I A(3',5')pG(3',5')p]	C20H23I2N10O13P2	800	674.1	Mass for cGAMP present
3'dGTP + ATP	VcDncV	No product expected	C20H26N10O14P2	507	506.996	Unreacted mass present
6S GTP + ATP	VcDncV	c[A(3',5')p6S G(3',5')p]	C20H24N10O12P2S	690.08	690.078	
Methyl-7 GTP + ATP	VcDncV	c[A(3',5')pm7G(3',5')p]	C21H27N10O13P2+	689.12	688.115	
8-oxo GTP + ATP	VcDncV	c[A(3',5')8oxoG(3',5')p]	C20H24N10O14P2	690.09	Not found	Unreacted mass not found
XTP + ATP	VcDncV	c[A(3',5')pX(3',5')p]	C20H23N9O14P2	675.08	675.084	Not most abundant
ITP + ATP	VcDncV	c[A(3',5')pl(3',5')p]	C20H23N9O13P2	659.09	659.089	

XTP + GTP	VcDncV	c[X(3',5')pG(3',5')p]	C20H23N9O15P2	691.08	691.079	
ITP + GTP	VcDncV	c[(3',5')pG(3',5')p]	C20H23N9O14P2	675.08	675.084	
Methyl-7 GTP + GTP	VcDncV	c[G(3',5')pm7G(3',5')p]	C21H27N10O14P2	705.1	705.118	

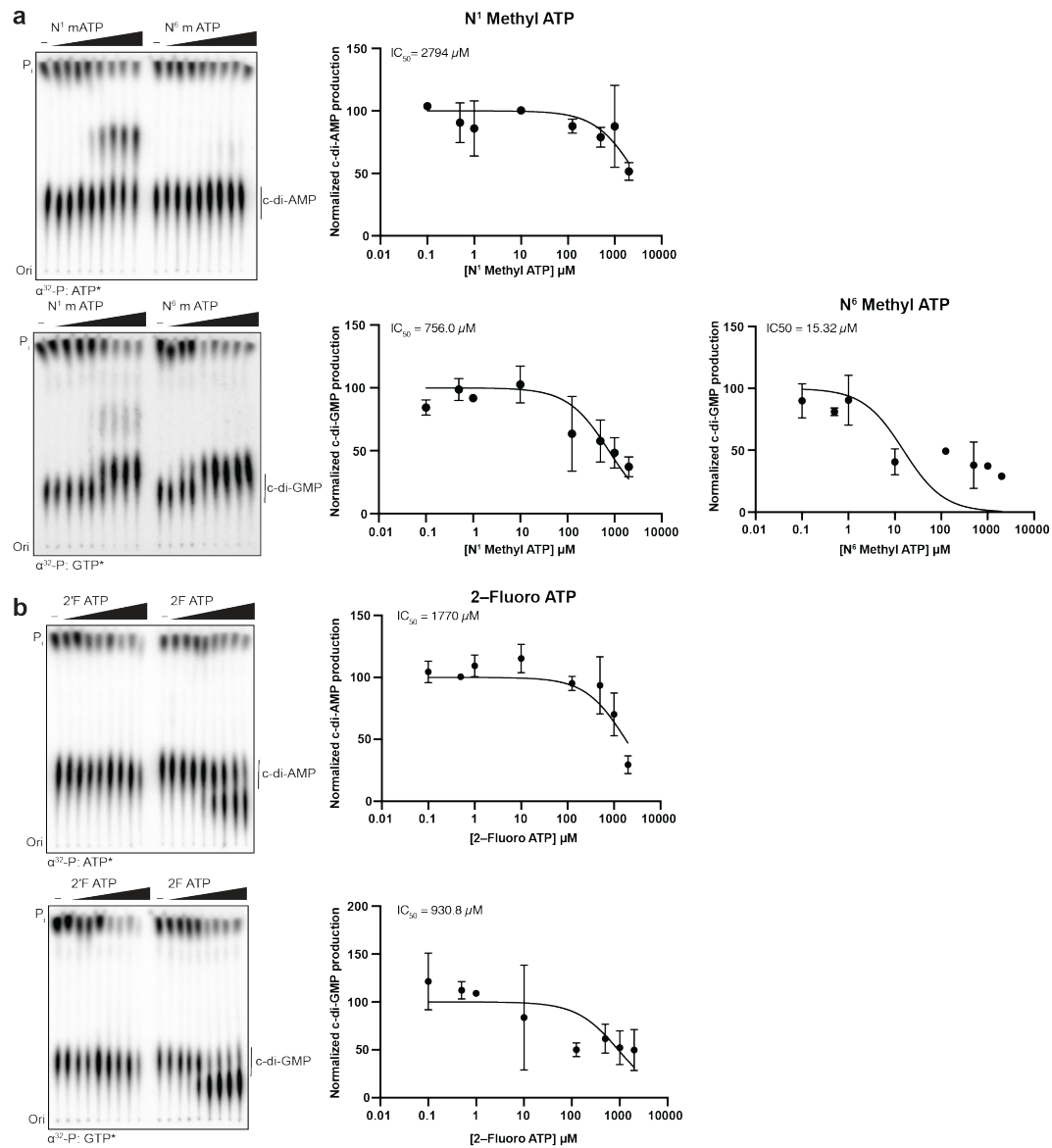
Modified nucleotide	Enzyme	Expected major product	Chemical formula	Molecular mass	Found Mass	Notes
N1 methyl ATP	EcCdnD	c[N1 methyl A(3',5')pN1 methyl A(3',5')pN1 methyl A(3',5')p]	C33H42N15O18P3	1029.205	521.011	Unreacted N1 methyl ATP present, no product mass
N6 methyl ATP	EcCdnD	c[N6 methyl A(3',5')pN6 methyl A(3',5')pN6 methyl A(3',5')p]	C33H42N15O18P3	1029.205	1029.205	
2F ATP	EcCdnD	c[2F A(3',5')p2F A(3',5')p2F A(3',5')p]	C20H22F2N10O12P2	1041.13	1041.129	
2'F ATP	EcCdnD	c[2'F A(3',5')p2'F A(3',5')p2'F A(3',5')p]	C30H33F3N15O15P3	993.145	993.1446	
7 deaza dATP	EcCdnD	No product expected	C11H17N4O12P3	490.01	490.006	Unreacted mass found
8I ATP	EcCdnD	c[8I A(3',5')p8I A(3',5')p8I A(3',5')p]	C30H33I3N15O18P3	1364.85	Not found	Some unreacted mass present, differs from TLC result
N1 methyl ATP + GTP	EcCdnD	c[N1 methyl A(3',5')pN1 methyl A(3',5')pG(3',5')p]	C32H40N15O19P3	1031.184	1031.184	
N6 methyl ATP + GTP	EcCdnD	c[N6 methyl A(3',5')pN6 methyl A(3',5')pG(3',5')p]	C32H40N15O19P3	1031.184	1031.184	
2F ATP + GTP	EcCdnD	c[2F A(3',5')p2F A(3',5')pG(3',5')p]	C30H34F2N15O19P3	1039.134	1039.134	
2'F ATP + GTP	EcCdnD	c[2'F A(3',5')p2'F A(3',5')pG(3',5')p]	C30H34F2N15O17P3	1007.14	1007.144	
7 deaza dATP + GTP	EcCdnD	No product expected	C11H17N4O12P3	490.01	490.006	Unreacted mass found
8I ATP + GTP	EcCdnD	c[8I A(3',5')p8I A(3',5')pG(3',5')p]	C30H34I2N15O19P3	1254.95	799.997	Mass of dinucleotide

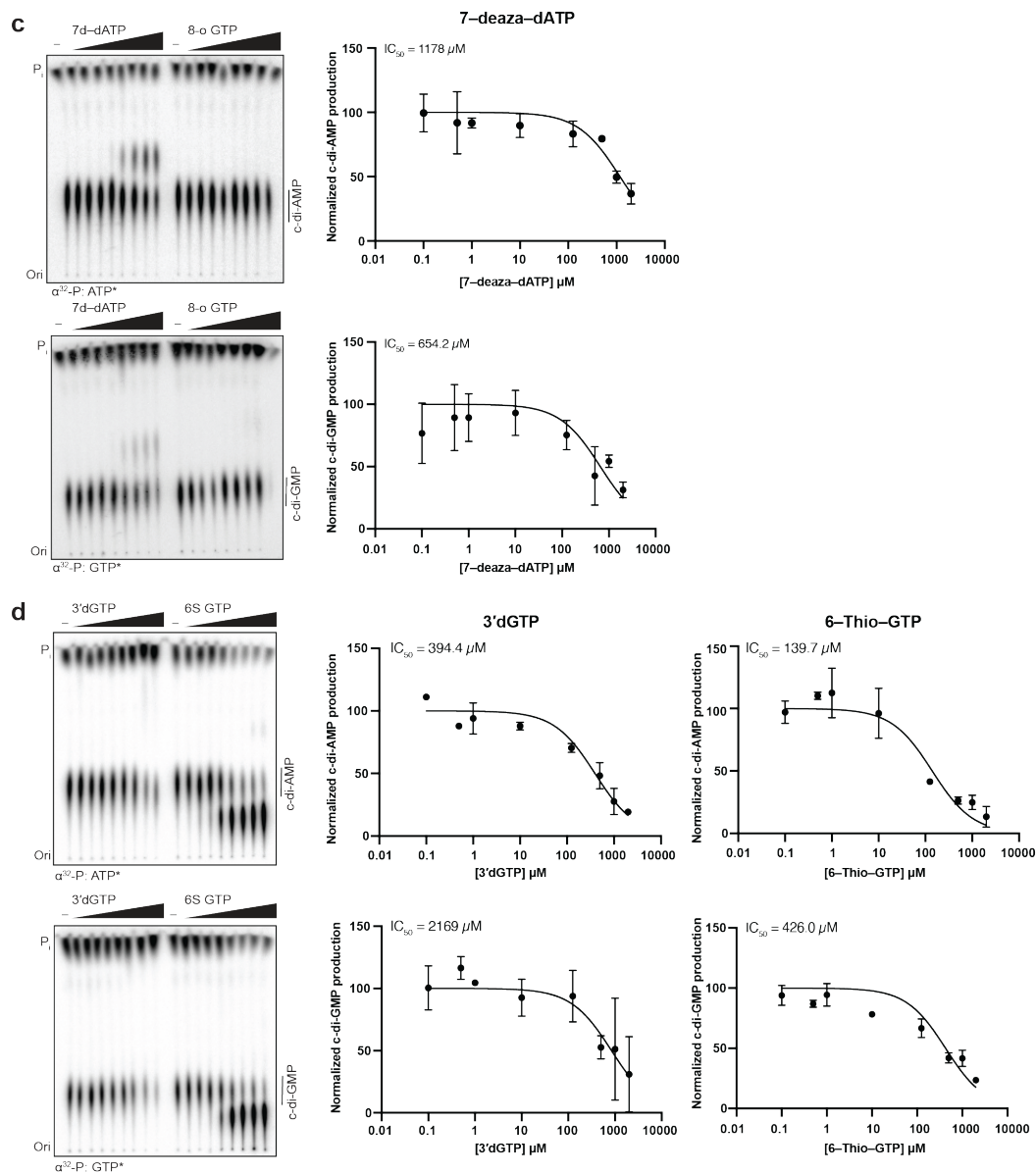
3'dGTP + ATP	EcCdnD	No product expected	C ₁₀ H ₁₆ N ₅ O ₁₃ P ₃	507	506.996	product found c[8l A(3',5')pG(3',5')p]
6S GTP + ATP	EcCdnD		C ₃₀ H ₃₆ N ₁₅ O ₁₈ P ₃ S	1019.13	1019.130	Unreacted mass found
Methyl-7 GTP + ATP	EcCdnD	c[A(3',5')pA(3',5')pMethyl-7 G(3',5')p]	C ₃₁ H ₃₉ N ₁₅ O ₁₉ P ₃ +	1018.175	1018.175	
2'O-methyl GTP + ATP	EcCdnD	No product expected	C ₁₁ H ₁₈ N ₅ O ₁₄ P ₃	537.01	537.006	Unreacted mass found
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	Not found	
ITP	EcCdnD	c[l(3',5')pl(3',5')pl(3',5')p]	C ₃₀ H ₃₃ N ₁₂ O ₂₁ P ₃	990.11	990.11	
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	691.079	Mass of dinucleotide product found c[X(3',5')pG(3',5')p]
ITP + GTP	EcCdnD	c[l(3',5')pl(3',5')pG(3',5')p]	C ₃₀ H ₃₄ N ₁₃ O ₂₁ P ₃	1005.121	1005.121	
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	1004.136	
ITP + ATP	EcCdnD	c[A(3',5')pA(3',5')pl(3',5')p] ; c[A(3',5')pl(3',5')pl(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	989.125594, 988.142	
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	Not found	Differs from TLC result

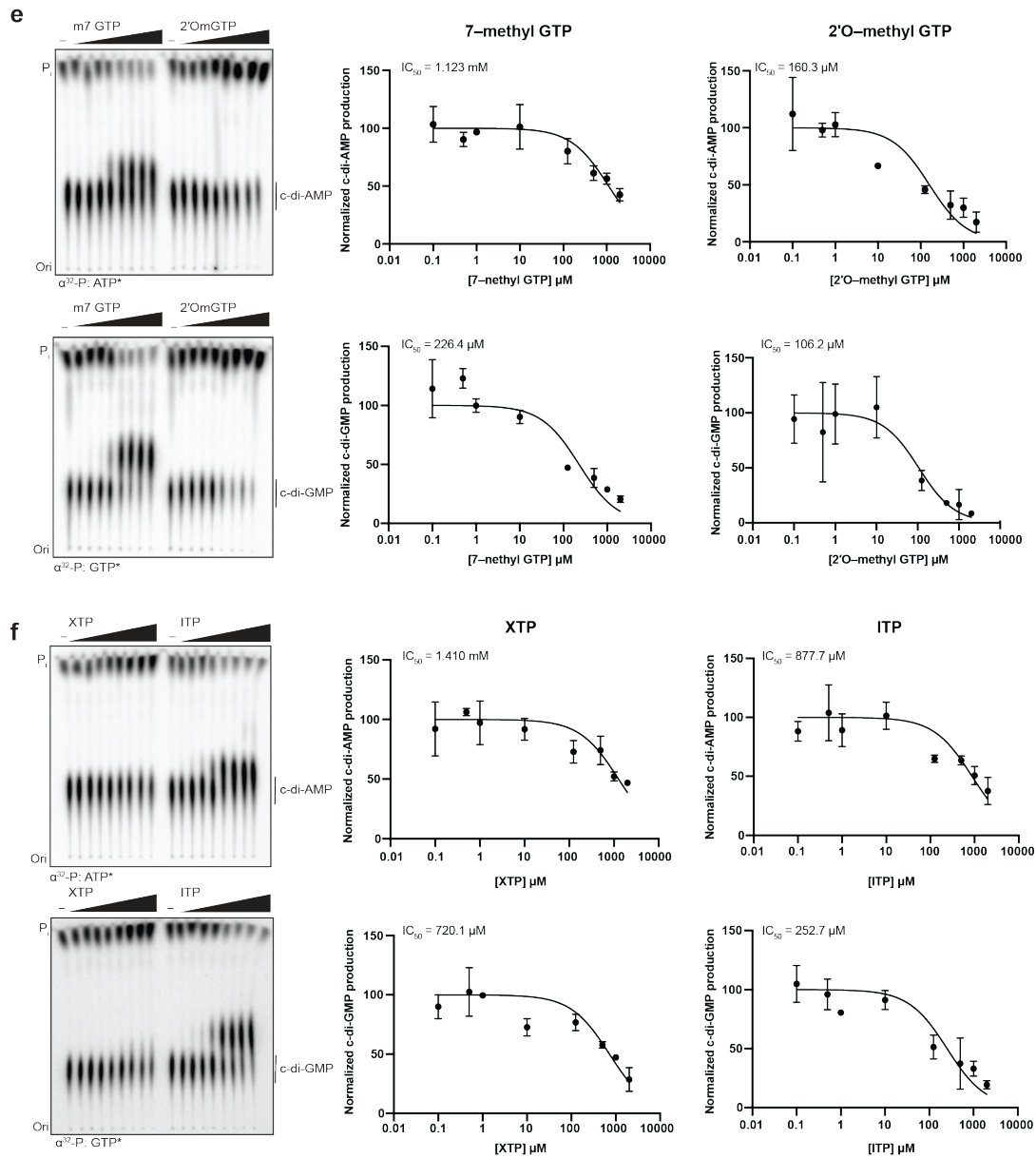
Figure 1.6 IC₅₀ quantification of modified nucleotide incorporation correlates with modified nucleotide product formation

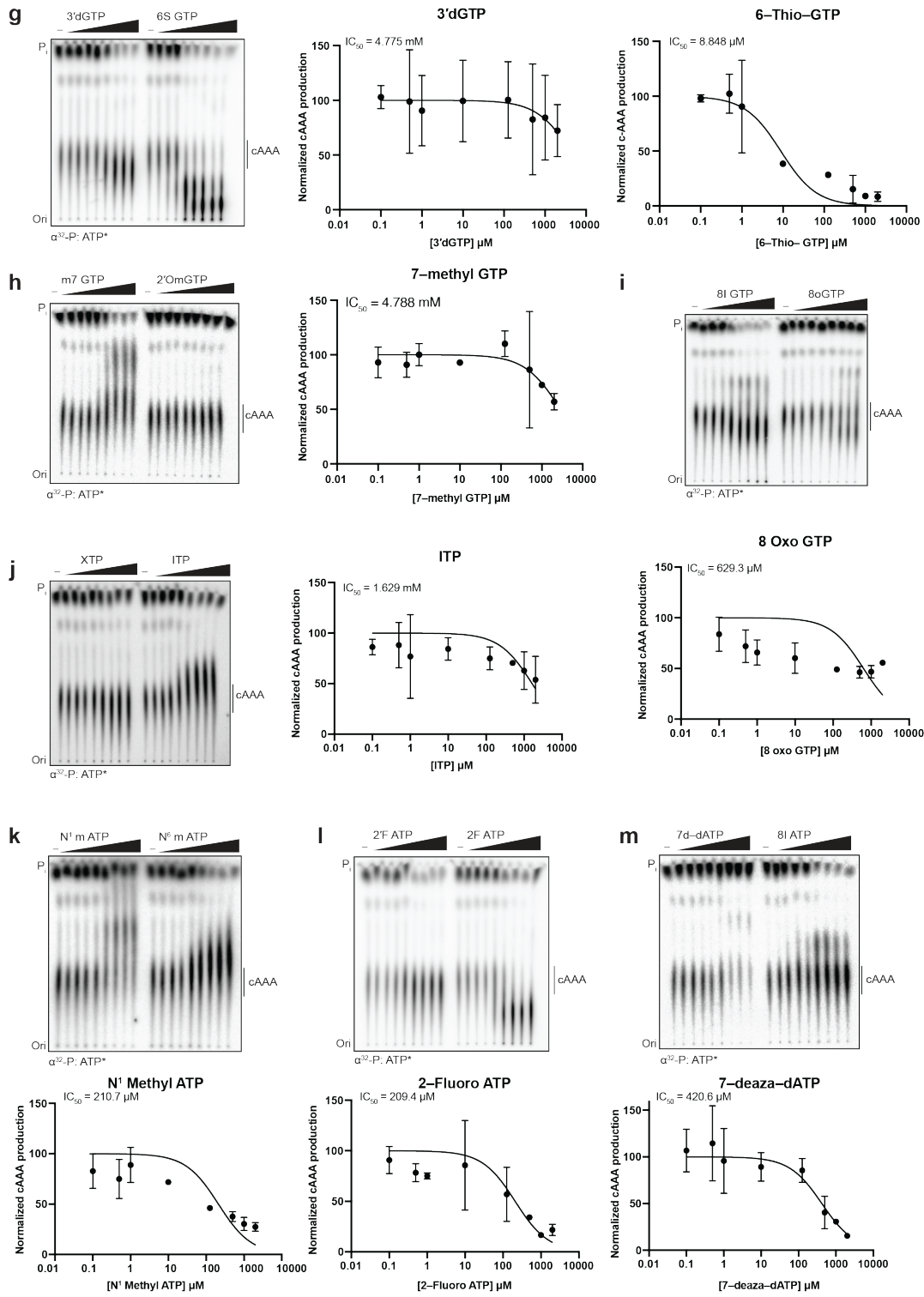
a–f. Thin-layer chromatography analysis and quantification of pixel intensity for each VcDncV reaction with increasing concentrations of modified nucleotide, as indicated, and radiolabeled $\alpha^{32}\text{P}$ -ATP. Tested modified nucleotide concentrations: 0, 100 nM, 500 nM, 1 μM , 10 μM , 125 μM , 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 EcCdnD reaction with increasing concentrations of modified nucleotide, as indicated, and radiolabeled $\alpha^{32}\text{P}$ -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.









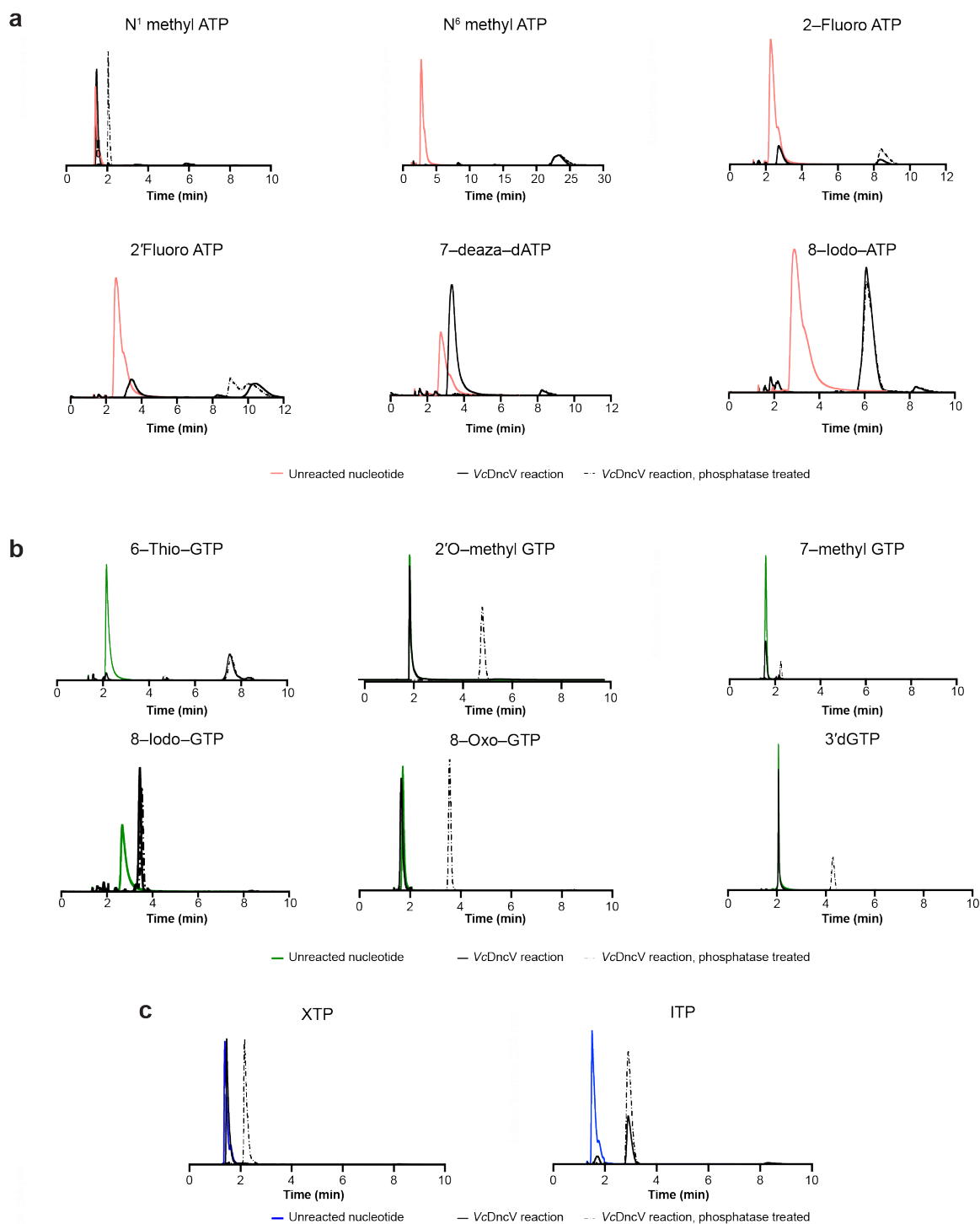


Figure 1.7 HPLC analysis of VcDncV reactions with modified nucleotides

a–c. High-performance liquid chromatography analysis of VcDncV reactions with each modified nucleotide, as indicated. Reactions were eluted from a C18 column with mobile phase 97% NaH₂PO₄ 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.

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 *EcCdnD* 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 (*BfCndB*, accession WP_032579276.1) and a 1.6 Å crystal structure of *Bradyrhizobium diazoefficiens* CdnG (*BdCdnG*, accession NP_766712.1) providing the first CD-NTase structures from Clades B and G (Figure 1.8b; Table 1.1). 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 1.8a). Lid residues correlate with major product synthesis (Figure 1.8 b, c), 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 1.8c) (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 1.8c).

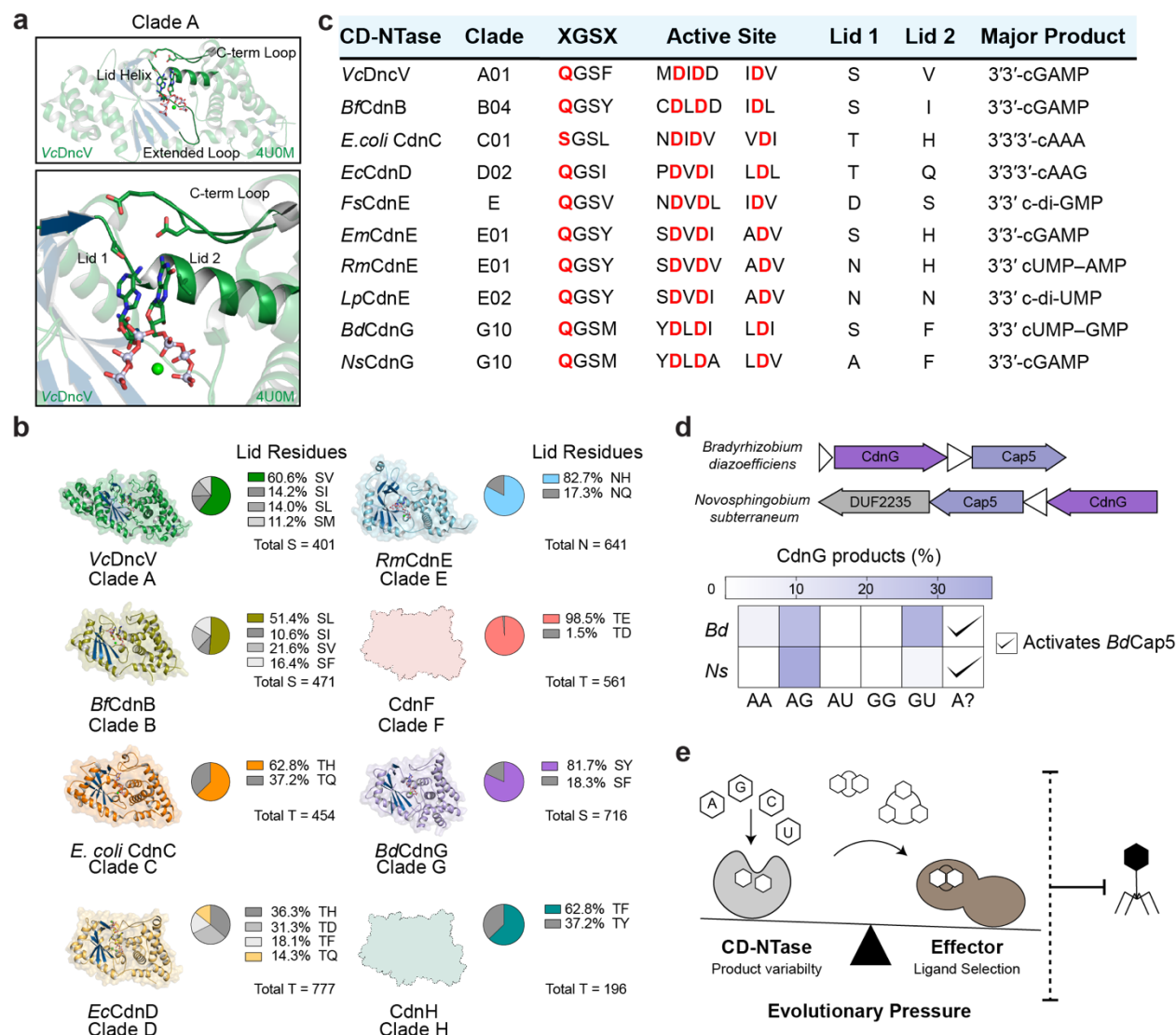


Figure 1.8 Structural and sequence-based comparison of divergent bacterial CD-NTases reveals correlates of nucleotide selection

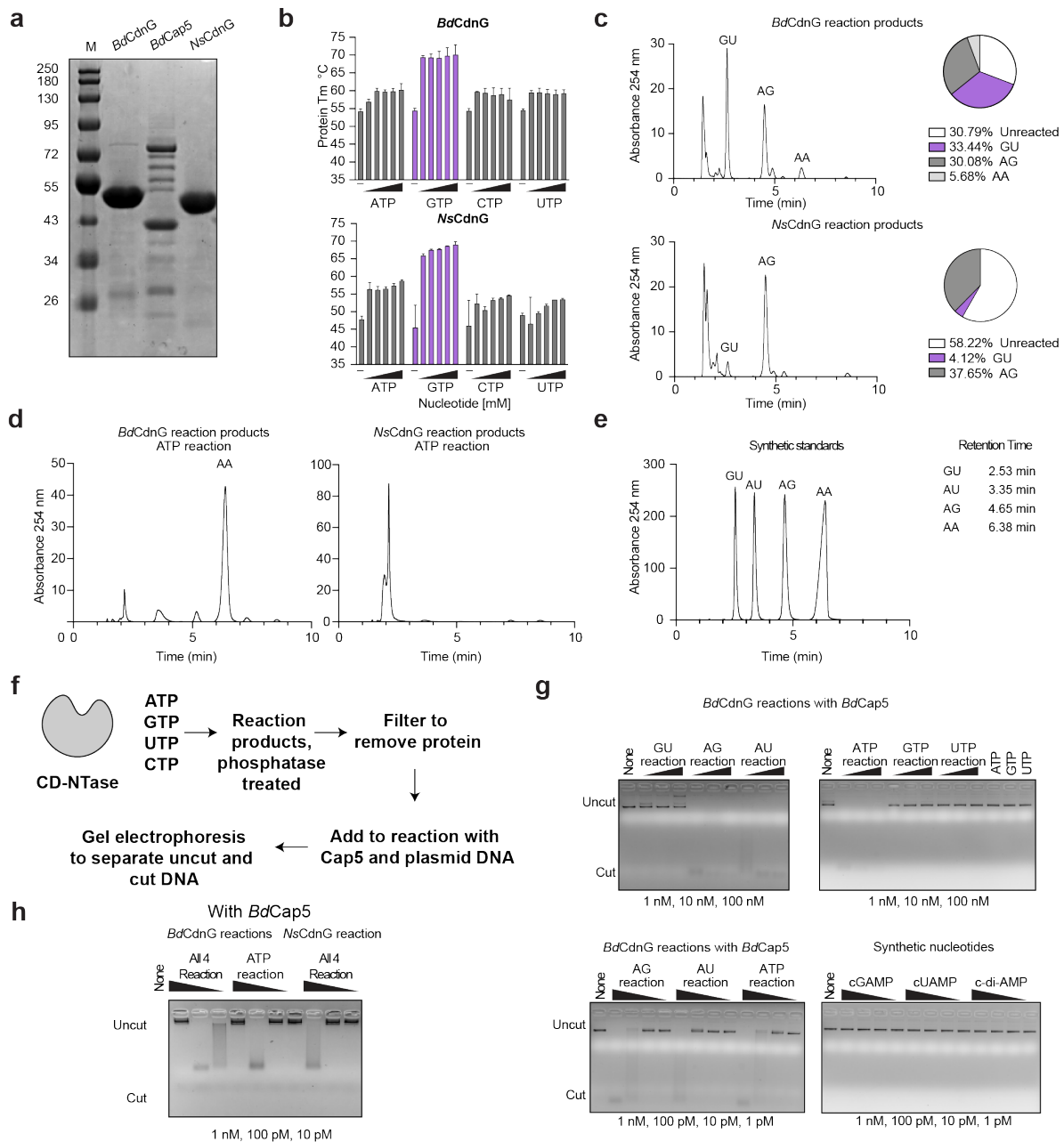
a. VcDncV structure 4U0M with highlighted structural features. Top, structure of VcDncV showing the relative positioning of nucleotide coordinating structural features. Bottom, highlight of Lid Helix, lid residues 1 and 2 and the C-term loop with additional coordinating residues. **b.** Comparative analysis 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. For more detail on other residues in Lid position 1 for each clade, consult Table S3. (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 EcCdnD), active site, and lid residues 1 and 2. For more detail see Table S3. (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 BdCdnG and NsCdnG. **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.

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 (Methods). To test these analyses, we experimentally verified the product specificity of uncharacterized Clade G CD-NTase enzymes *BdCdnG* and *Novosphingobium subterraneum* CdnG (*NsCdnG*) (Figure 1.9a). *BdCdnG* encodes a serine at lid helix position 1 similar to *VcDncV*, *BfCdnB*, and *EmCdnE* (Figure 1.8c), suggesting specific incorporation of GTP or ATP. We observed enhanced thermostability upon GTP addition over a range of concentrations for both *BdCdnG* and *NsCdnG* supporting a role for GTP as a donor nucleotide (Figure 1.9b). In further agreement with predicted CD-NTase product specificity, we observed that *in vitro* *BdCdnG* synthesizes 3'3'-cGMP–UMP, 3'3'-cGAMP and low levels of 3'3'-c-di-AMP, while *NsCdnG* synthesizes primarily 3'3'-cGAMP (Figure 1.8d, Figure 1.9 c–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 *BdCdnG* activity we next purified and tested activity of the *Bradyrhizobium diazoefficiens* effector CD-NTase-associated protein 5 (*BdCap5*) annotated as an HNH-*SAVED* nuclease (Burroughs et al. 2015; Lowey et al. 2020). *BdCdnG* reactions are capable of activating *BdCap5* and inducing promiscuous DNase activity and degradation of plasmid DNA (Figure 1.8d, Figure 1.9 f–h). However, *BdCdnG* reactions producing the major product 3'3'-cGMP–UMP fail to activate *BdCap5*. We observed that *BdCap5* responds to *BdCdnG* reactions containing only ATP (Figure 1.8d; Figure 1.9 g,h), as well as *NsCdnG* reactions, suggesting that *BdCap5* recognizes a minor product that is conserved between closely related CD-NTase enzymes (Figure 1.9h). *BdCap5* 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.

Figure 1.9 *BdCdnG* effector Cap5 responds to A-containing minor product

a. Coomassie stained gel showing expression level and purity of *BdCdnG*, *BdCap5*, and *NsCdnG*. **b.** Plot of T_m for *BdCdnG* and *NsCdnG* 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 *BdCdnG* and *NsCdnG* reactions. Top, *BdCdnG* 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, *NsCdnG* 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 *BdCdnG* and *NsCdnG* 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.** *BdCdnG* reactions with *BdCap5* effector assay as depicted in f. *BdCdnG* Reaction with indicated nucleotide combinations or single nucleotide reactions were incubated with *BdCap5* and empty pcDNA4 or pGEM plasmid DNA. Unreacted ATP, GTP, and UTP are included as controls. AG, AU, and ATP containing reactions activate *BdCap5* at concentrations as low as 100 pM, while synthetic 3'3'-cGAMP, 3'3'-cAMP-UMP, and c-di-AMP do not activate *BdCap5*. **h.** *BdCdnG* or *NsCdnG* reactions with *BdCap5*. Both *BdCdnG* and *NsCdnG* make an A-containing product that activates *BdCap5*.



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 *VcDncV* and *EcCdnD*, 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 1.8), 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 1.1) (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 1.8e). 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.

Methods

MATERIALS AVAILABILITY

This study did not generate new unique reagents.

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 Na₂PO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 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 Na₂PO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 85.6 mM NaCl, 2 mM MgSO₄, 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 (see Experimental Model and Subject Details) overnight at 37°C, and used to inoculate 1 L M9ZB media cultures (see Experimental Model and Subject Details) at OD₆₀₀ ~0.05. Cultures were grown at 37°C, 230 RPM until OD₆₀₀ 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 Na₂PO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 85.6 mM NaCl, 2 mM MgSO₄, 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 selenomethionine 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 *EcCdnD* 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 crystallized 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 crystallized as follows: 1) *EcCdnD* + 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 MgCl₂ 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 MgCl₂ and ATP prior to freezing in liquid nitrogen; 2) *SeCdnD* + GTP: Native protein was diluted to 8.25 mg mL⁻¹, supplemented with 10.5 mM MgCl₂ 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) *BfCdnB*: Native or selenomethionine-substituted protein was diluted to 8 mg m(Agilent)⁻¹, supplemented with 10 mM MgCl₂ 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) *BdCdnG* + GTP: Native or selenomethionine-substituted protein was diluted to 10 mg m(Agilent)⁻¹ with 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP, supplemented with 10.5 mM MgCl₂, 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). Selenomethionine-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 (V. B. 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 μ L reactions with 50 μ M enzyme, 12.5 μ M NTP and trace α^{32} P-ATP or α^{32} P-GTP, as indicated. VcDncV reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT and EcCdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM MgCl₂, 1 mM DTT. EcCdnD mutant characterization reactions were carried out for 30 min at 37°C before heat inactivation at 95°C for 2 min, and reactions for IC₅₀ 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 μ L of Quick CIP (NEB) at 37°C for 30 min. Reactions were analyzed by thin-layer chromatography by spotting 1 μ L of each phosphatase-treated reaction on a PEI-cellulose plate (Millipore) and developed in 1.5 M KH₂PO₄ 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 IC₅₀ calculation, pixel intensity corresponding to radiolabeled c-di-AMP, c-di-GMP (VcDncV), or 3'3'3'-cAAA (EcCdnD) was measured for each lane containing increasing amounts of unlabeled modified nucleotide (0, 0.1, 0.5, 1, 10, 125, 500, 1000, 2000 μ M) 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 μ L reactions containing 10 μ M protein and no nucleotide, or a range of nucleotide concentrations as indicated, 0.001 – 20 μ M. VcDncV reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT and EcCdnD reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM MgCl₂, 1 mM DTT. Protein and nucleotide solutions were incubated on ice for at least 45 min until measuring melting temperature of 10 μ L of reaction between 35–95°C over 3 min.

HPLC enzymatic reaction analysis

For HPLC analysis, CD-NTase reactions were performed in 20 μ L reactions with 50 μ M enzyme and 500 μ M NTPs. *VcDncV* reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM MgCl₂, 1 mM DTT. Reactions were incubated at 37°C overnight for ~16 h, heat inactivated at 95°C for 2 min, diluted to 200 μ L and filtered through a Millipore Amicon ultra-0.5 mL 30 kDa cut-off filter by centrifugation at 10,000 RPM (9350 $\times$ 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 $\times$ 150 mm, 3.5- μ m) and Agilent 1200 Infinity Series LC system. 20 μ L 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 NaH₂PO₄ pH 6.8 supplemented with 3% HPLC grade acetonitrile for *VcDncV* reactions and 6% acetonitrile for *EcCdnD* reactions. For quick-CIP treated reactions, 20 μ L of heat inactivated reaction was treated with 1 μ L 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 μ m 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 μ M protein (*VcDncV*, *EcCdnD*, or mcGAS) and nucleotide concentrations 125, 250, 500, 750 μ M, or 1 mM ATP and GTP. *VcDncV* and mcGAS reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM MgCl₂, 1 mM DTT. For each nucleotide concentration, 100 μ L reactions were incubated at 37°C and 20 μ L aliquots were collected at 2, 10, 20, 30, 60 min (*VcDncV*); 2, 10, 20, 30 min (*EcCdnD*); 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 μ L

aliquots were diluted to 200 μ L and filtered through a Millipore Amicon ultra-0.5 mL 30 kDa cut-off filter by centrifugation at 10,000 RPM ($9350 \times g$) in a table-top microcentrifuge for 10 min. 40 μ L of each sample was then injected and reactions were analyzed by HPLC as detailed above. Absorbance units were converted to μ M by comparing to a standard curve from 10–750 μ M 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. Michaelis-Menten kinetics were calculated using GraphPad Prism version 9.0.0.

***BdCap5* nuclease assay**

Nuclease assays were performed as previously described (Lowey et al. 2020). Briefly, 50 nM *BdCap5* protein was incubated with 10 ng μ L⁻¹ pGEM9z plasmid (Promega) or 10 ng μ L⁻¹ pcDNA4 plasmid in a 10 μ L reaction for 30 min at 37°C containing 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM MgCl₂, 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”. 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.

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. The work was funded by grants to P.J.K. from the Pew Biomedical Scholars Program, Burroughs Wellcome Fund PATH award, Mark Foundation for Cancer Research, Mathers Foundation, and Parker Institute for Cancer Immunotherapy. A.A.G. is supported by a National Science Foundation Graduate Research Fellowship, B.L. is supported as a Herchel Smith Graduate Research Fellow. 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).

Author Contributions

Apurva A. Govande: Designed and conceived experiments, performed nucleotide binding studies and data analysis, collected x-ray crystallographic data, completed molecular modeling, and wrote the manuscript.

Brianna Duncan-Lowey: Collected x-ray crystallographic data and completed molecular modelling

James B. Eaglesham: Collected x-ray crystallographic data and carried out molecular modelling

Aaron T. Whiteley: Assisted with initial enzyme activity experiments and data analysis

Philip J. Kranzusch: Designed and conceived experiments, collected x-ray crystallographic data, and wrote the manuscript

Chapter 2: STING cyclic dinucleotide sensing originated in bacteria

Benjamin R. Morehouse^{1,2}, Apurva A. Govande^{1,2}, Adi Millman³, Alexander F.A. Keszei⁴, Brianna Lowey^{1,2}, Gal Ofir³, Sichen Shao⁴, Rotem Sorek³, Philip J. Kranzusch^{1,2,5*}

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 76100, Israel

⁴ Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

⁵ Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA 02115, USA

The published manuscript may be accessed as follows:

Nature (2020) V 586(7829) p429-433

<https://www.nature.com/articles/s41586-020-2719-5>

Abstract

Stimulator of interferon genes (STING) is a receptor in human cells that senses foreign cyclic dinucleotides released during bacterial infection and endogenous cyclic GMP–AMP signaling during viral infection and antitumor immunity (Ishikawa and Barber 2008; Zhong et al. 2008; Burdette et al. 2011; Sun et al. 2013b; Ablasser and Chen 2019). STING shares no structural homology with other known signaling proteins (Ouyang et al. 2012; X. Zhang et al. 2013; Kranzusch et al. 2015; Shang et al. 2019), limiting functional analysis and preventing explanation for the origin of cyclic dinucleotide signaling in mammalian innate immunity. Here we discover functional STING homologues encoded within prokaryotic defense islands and reveal a conserved mechanism of signal activation. Crystal structures of bacterial STING define a minimal homodimeric scaffold that selectively responds to c-di-GMP synthesized by a neighboring cGAS/DncV-like nucleotidyltransferase (CD-NTase) enzyme. Bacterial STING domains couple cyclic dinucleotide recognition with protein filament formation to drive TIR effector domain oligomerization and rapid NAD⁺ cleavage. We reconstruct the evolutionary events following acquisition of STING into metazoan innate immunity and determine the structure of a full-length TIR-STING fusion from the Pacific oyster *C. gigas*. Comparative structural analysis demonstrates how metazoan-specific additions to the core STING scaffold enabled a switch from direct effector function to regulation of antiviral transcription. Together, our results explain the mechanism of STING-dependent signaling and reveal conservation of a functional cGAS-STING pathway in prokaryotic bacteriophage defense.

Results and Discussion

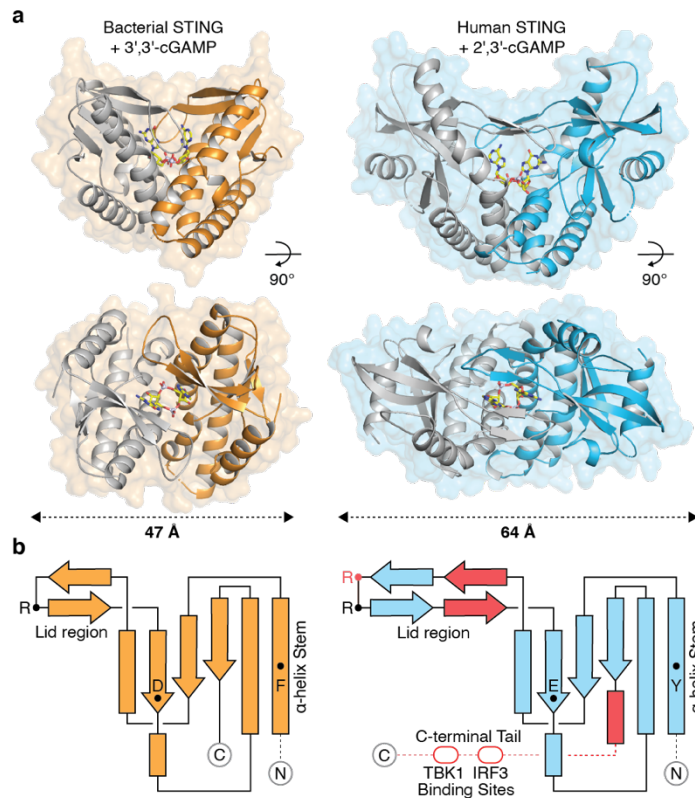


Figure 2.1 Structure of bacterial STING and definition of metazoan-specific insertions.

a) Crystal structure of a STING receptor from the bacterium *Flavobacteriaceae* sp. (orange) in complex with the cyclic dinucleotide 3',3'-cGAMP. The *Fs*STING–3',3'-cGAMP structure demonstrates a conserved mechanism of cyclic dinucleotide sensing shared between bacteria and human cells, and allows direct comparison with the human STING–2',3'-cGAMP complex (PDB 4KSY, blue). For clarity, one monomer of each homodimer is depicted in grey.

b) STING topology diagrams denoting α -helices (rectangles), β -strands (arrows), and residues important for ligand recognition (*Fs*STING F92, D169, R153; hSTING Y167, E260, R232 in red, R238). Bacterial STING receptors reveal a minimal protein architecture required for cyclic dinucleotide sensing and allow direct definition of the structural insertions (red) in metazoan STING sequences required for signaling in animal cells.

Bioinformatic analysis of prokaryotic bacteriophage defense islands revealed a group of divergent genes encoding the first candidate proteins outside of metazoan innate immunity with predicted homology to STING (Figure 2.2a) (Cohen et al. 2019). To understand a potential role for STING in bacterial signaling, we determined 1.8 Å and 2.8 Å crystal structures of candidate homologues from *Flavobacteriaceae* sp. (IMG Gene ID 2624319773) and *Capnocytophaga granulosa* (IMG Gene ID 2541326748) (Table 2.1). The bacterial structures exhibit clear homology to the

human STING cyclic dinucleotide binding domain and confirm this subset of defense island proteins as novel prokaryotic members of the STING family of receptors (Figure 2.1a). *Flavobacteriaceae* sp. STING (*Fs*STING) and *Capnocytophaga granulosa* STING (*Cg*STING) adopt a canonical V-shaped, homodimeric architecture with a hydrophobic α -helix stem similar to that observed in all structures of metazoan STING proteins (Figure 2.1b)(Ouyang et al.

2012; X. Zhang et al. 2013; Kranzusch et al. 2015; Shang et al. 2019). We determined the *Fs*STING

structure in complex with 3'-5' / 3'-5' cyclic GMP-AMP (3',3'-cGAMP), confirming bacterial STINGs are functional

Table 2.1 Summary of data collection, phasing and refinement statistics

	FsSTING-3',3'- cGAMP (6WT4)	FsSTING-3',3'- cGAMP SeMet	CgSTING Apo SeMet (6WT5)
Data collection			
Space group	I 4	I 4	P 4 ₁ 2 ₁ 2
Cell dimensions			
a, b, c (Å)	121.85, 121.85, 48.16	120.33, 120.33, 47.01	123.71, 123.71, 108.22
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.0, 90.0, 90.0
Resolution (Å)	44.79–1.78 (1.84–1.78)	43.78–1.95 (2.00–1.95)	49.26–2.80 (2.95–2.80)
R _{pim}	1.9 (53.1)	2.4 (38.7)	3.2 (70.2)
I/σ ²	17.3 (1.1)	13.8 (1.4)	15.4 (1.2)
Completeness (%)	99.9 (99.2)	99.6 (94.4)	100.0 (100.0)
Redundancy	13.8 (13.2)	7.7 (7.4)	27.0 (27.6)
Refinement			
Resolution (Å)	44.79–1.78	–	49.26–2.80
No. reflections			
Total	470286	190713	576011
Unique	34093	24780	21303
Free	2004	–	1999
R _{work} / R _{free}	19.8 / 23.0	–	23.5 / 28.1
No. atoms			
Protein	2365	–	4637
Ligand/ion	100 (3',3'-cGAMP, SO ₄)	–	–
Water	98	–	–
B-factors			
Protein	53.49	–	85.85
Ligand/ion	35.05	–	–
Water	48.10	–	–
R.m.s. deviations			
Bond lengths (Å)	0.003	–	0.002
Bond angles (°)	0.632	–	0.484

*All datasets were collected from individual crystals. *Values in parentheses are for the highest resolution shell.

	<i>C. gigas</i> STING Apo (6WT6)	<i>C. gigas</i> STING Apo SeMet	<i>C. gigas</i> STING–2',3'- cGAMP (6WT7)	<i>C. gigas</i> STING– 2',3'-cGAMP SeMet
Data collection				
Space group	P 4 ₂ 2 ₁ 2	P 4 ₂ 2 ₁ 2	P 3 ₂ 2 1	P 3 ₂ 2 1
Cell dimensions				
a, b, c (Å)	135.90, 135.90, 40.22	136.42, 136.42, 40.33	108.13, 108.13, 71.57	111.65, 111.65, 70.45
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	48.05–2.40 (2.49–2.40)	68.21–2.65 (2.78–2.65)	46.82–2.90 (3.08–2.90)	48.35–2.90 (3.08– 2.90)
R_{pim}	9.4 (69.1)	7.8 (40.9)	3.4 (40.8)	3.2 (73.4)
I/σI	6.8 (1.3)	5.7 (1.7)	12.7 (2.0)	13.1 (1.1)
Completeness (%)	99.9 (99.6)	100.0 (100.0)	100.0 (100.0)	100.0 (99.9)
Redundancy	16.0 (15.6)	14.4 (14.2)	12.0 (12.1)	20.5 (21.0)
Refinement				
Resolution (Å)	48.05–2.40	–	46.82–2.90	–
No. reflections				
Total	246284	167735	131783	236871
Unique	15377	11632	10981	11537
Free	1530	–	1106	–
R_{work} / R_{free}	21.6 / 24.2	–	23.1 / 25.3	–
No. atoms				
Protein	2470	–	2656	–
Ligand/ion	–	–	45	–
Water	61	–	–	–
B-factors				
Protein	51.46	–	90.2	–
Ligand/ion	–	–	65.4	–
Water	45.28	–	–	–
R.m.s. deviations				
Bond lengths (Å)	0.003	–	0.002	–
Bond angles (°)	0.568	–	0.44	–

*All datasets were collected from individual crystals. *Values in parentheses are for the highest resolution shell.

	FsCdnE Apo SeMet (6WT8)	CgCdnE Apo SeMet (6WT9)
Data collection		
Space group	P 2 ₁ 2 ₁ 2 ₁	C 1 2 1
Cell dimensions		
a, b, c (Å)	51.18, 78.10, 111.11	113.00, 48.88, 59.70
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 92.51, 90.00
Resolution (Å)	46..49–1.52 (1.55– 1.52)	44.85–2.30 (2.38– 2.30)
R_{pim} (%)	3.1 (70.1)	9.6 (45.7)
I/σI	16.3 (1.7)	17.0 (11.3)
Completeness (%)	99.7 (99.1)	99.6 (99.5)
Redundancy	17.6 (17.1)	10.4 (10.2)
Refinement		
Resolution (Å)	46..49–1.52	44.85–2.30
No. reflections		
Total	1215428	151480
Unique	69098	14608
Free	2000	1462
R_{work} / R_{free}	16.8 / 18.8	18.6 / 23.9
No. atoms		
Protein	2913	2475
Ligand/ion	–	–
Water	394	169
B-factors		
Protein	23.36	28.02
Ligand/ion	–	–
Water	34.60	29.51
R.m.s. deviations		
Bond lengths (Å)	0.014	0.002
Bond angles (°)	1.35	0.42

*All datasets were collected from individual crystals. *Values in parentheses are for the highest resolution shell.

Figure 2.2 Structural analysis of bacterial STING–cyclic dinucleotide complex formation.

a, Phylogenetic tree of all CBASS-associated bacterial STING homologues based on structure-guided sequence alignment and previous bioinformatic analysis (Cohen et al. 2019; J. Ghosh et al. 2010). STING homologues investigated in this study are highlighted in orange, and a star denotes determined STING crystal structures. All TM-STING fusions cluster together.

b, Crystal structure of a STING receptor from the bacterium *Capnocytophaga granulosa* (CgSTING) in the apo state reveals an open configuration with a solvent exposed cyclic dinucleotide-binding pocket at the dimeric interface (monomers in gold and grey for clarity). The CgSTING structure confirms that both divergent TM-STING and TIR-STING fusions are members of the same structurally conserved family of STING receptors.

c, Comparison of the CgSTING, FsSTING–3',3'-cGAMP, and human STING–2',3'-cGAMP structures demonstrates conservation of an open-to-closed β -strand lid movement upon ligand binding.

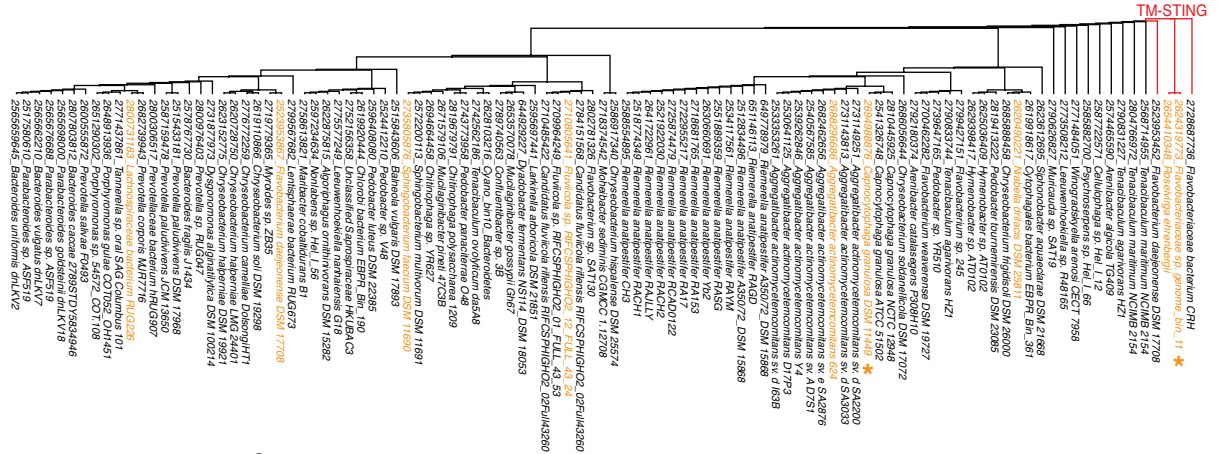
d, Overlay of the β -strand lid of CgSTING (grey) and FsSTING (orange) shows both inward translation and slight rotation resulting in a displacement of ~ 5 Å. R153 of FsSTING stacks between the bases of 3',3'-cGAMP while R151 is splayed away from ligand.

e, Comparison of the human STING and FsSTING lid region shows conserved contacts from β -strand arginine residues. Unlike in bacterial STING, human STING R232 makes an additional contact with the cyclic dinucleotide phosphodiester backbone that is critical for recognition of the 2'–5' linkage in 2',3'-cGAMP. See also Extended Data Figure 10b for detailed comparison.

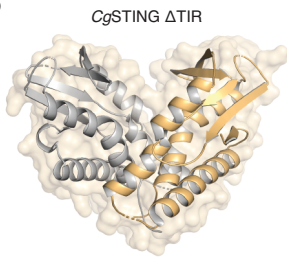
f, Modeling of 2',3'-cGAMP into the FsSTING–3',3'-cGAMP structure demonstrates an additional feature of bacterial STING preventing recognition of 2'–5'-linked cyclic dinucleotides. Although the overall cyclic dinucleotide conformation is shared between human and bacterial STING, the α -helix ending at P264 in human STING is a half turn longer in FsSTING (also ending in a proline) which places a conserved T173 residue in a position that occludes where the free 3'-OH of 2',3'-cGAMP would be positioned.

g, Structure-guided alignment of FsSTING and human STING cyclic dinucleotide binding domains. FsSTING and human STING exhibit no detectable sequence homology but share a conserved structural fold. Key residues involved in cyclic dinucleotide binding that are shared between bacterial and human STING are boxed in orange and human STING specific cyclic dinucleotide contacts are boxed in red. In FsSTING, D169 directly reads out the guanine base of c-di-GMP.

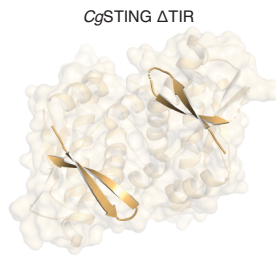
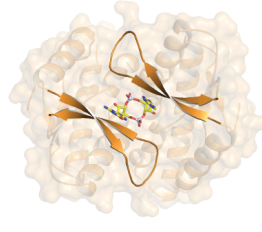
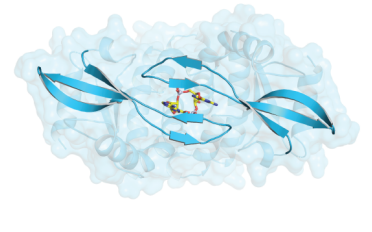
a



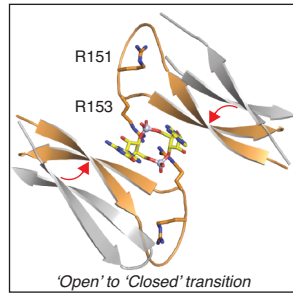
b



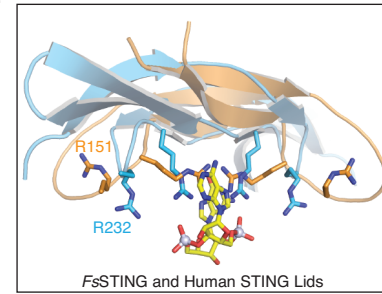
c

FsSTING Δ TM + 3',3'-cGAMPHuman STING Δ TM + 2',3'-cGAMP

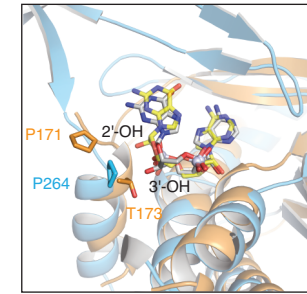
d



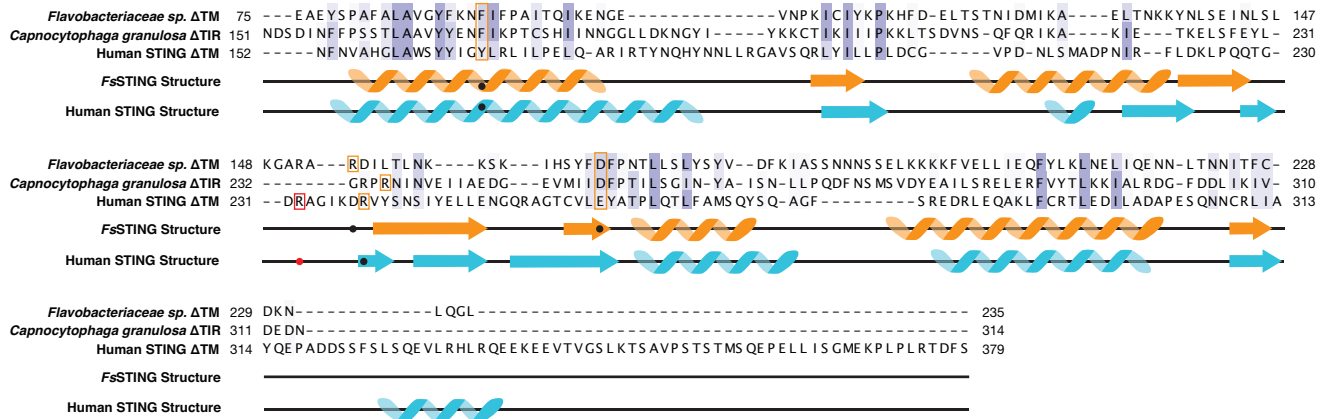
e



f



g



cyclic dinucleotide receptors (Figure 2.1a). Alignment of the *Fs*STING–3',3'-cGAMP complex with apo CgSTING reveals that cyclic dinucleotide binding induces rotation of the monomeric domains and results in closure of β -strands 2 and 3 to form a lid that seals a central cyclic dinucleotide binding pocket (Figure 2.1a; Figure 2.2 b–d). Although the overall architecture is conserved with metazoan STING, bacterial STING proteins are 20% smaller and strikingly compact. Metazoan insertions into the core bacterial STING fold include an extension within the β -strand lid domain, addition of a terminal α -helix required for induction of autophagy, and an unstructured C-terminal tail containing motifs required for kinase and transcription factor recruitment in vertebrates (Figure 2.1b) (Gui et al. 2019; C. Zhang et al. 2019; Zhao et al. 2019; de Oliveira Mann et al. 2019).

We analyzed bacterial STING sequences and found that 84% are encoded in Cyclic Oligonucleotide-Based Signaling System (CBASS) immunity operons (Figure 2.3a)(Cohen et al. 2019). Similar to cGAS-dependent sensing of viral replication in human cells (Sun et al. 2013b; Ablasser and Chen 2019), CBASS immunity relies on activation of a CD-NTase enzyme to initiate a second messenger-dependent antiviral effector response (Cohen et al. 2019; Whiteley et al. 2019; Ye et al. 2020). We cloned the adjacent *Flavobacteriaceae* sp. CD-NTase *CdnE* (CD-NTase in clade E) and observed that *FsCdnE* specifically synthesizes the cyclic dinucleotide c-di-GMP (Figure 2.3b) (Whiteley et al. 2019). *FsCdnE* c-di-GMP synthesis is constitutively active *in vitro*, in agreement with an emerging model where CBASS immune systems may function through inhibitory molecules that repress CD-NTase activation in the absence of phage infection (Cohen et al. 2019; Whiteley et al. 2019; Ye et al. 2020; Lowey et al. 2020). We confirmed the exclusive production of c-di-GMP by CdnE enzymes encoded within three additional divergent STING-containing CBASS operons (Figures 2.4a and 2.5 a–c).

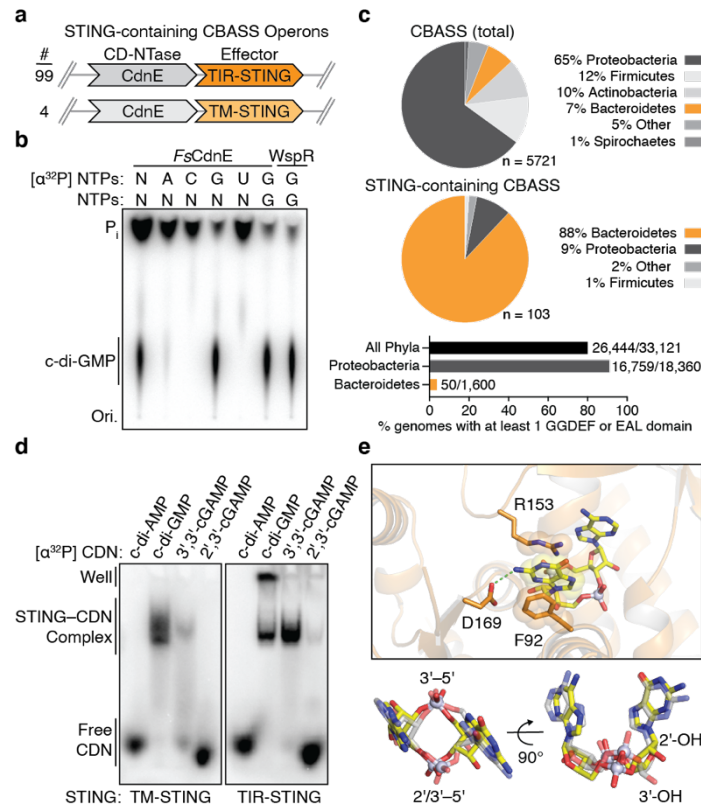


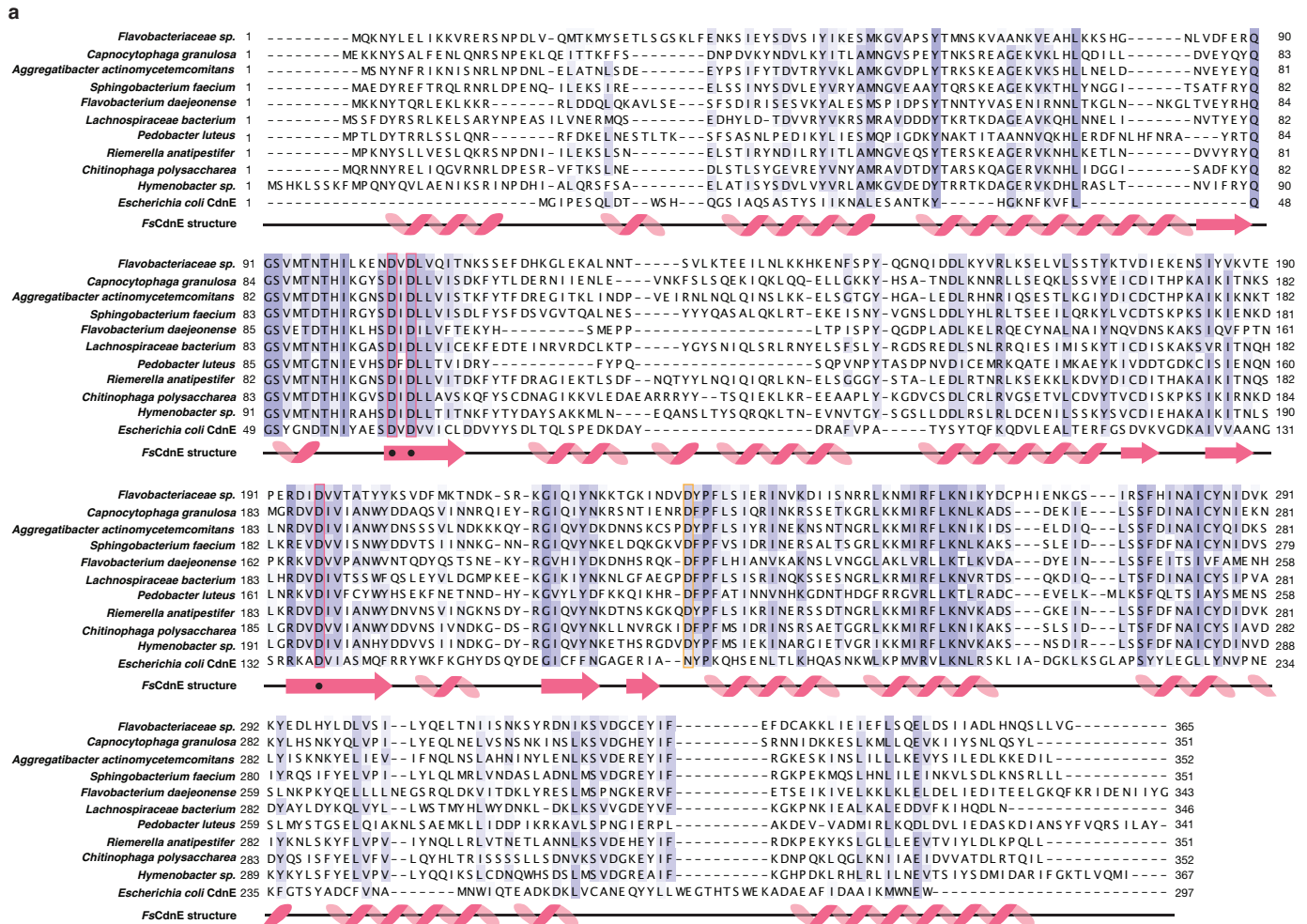
Figure 2.3 Bacterial STING systems signal through the 3'–5'-linked nucleotide second messenger c-di-GMP.

a) Schematic of STING-containing CBASS prokaryotic defense operons. STING is encoded downstream of a Clade E CD-NTase nucleotide second messenger synthase (CdnE) and exists as a fusion protein appended to Toll/Interleukin-1 receptor (TIR) or transmembrane (TM) effector modules with 99 and 4 sequences found, respectively. **b)** Thin-layer chromatography analysis of *Flavobacteriaceae* sp. CdnE (*FsCdnE*) nucleotide second messenger synthesis. CdnE enzymes in STING-containing CBASS operons require GTP for catalysis and specifically synthesize the 3'–5'-linked product c-di-GMP. Control c-di-GMP synthesis reactions performed with the GGDEF enzyme *P. aeruginosa* WspR. N, all four radiolabeled NTPs; P_i, inorganic phosphate; Ori., origin. Data are representative of 3 independent experiments. **c)** Analysis of the genomic context of all STING-containing CBASS operons. STING-containing CBASS operons in bacteria are occur mainly in *Bacteroidetes* (top). Sequenced *Bacteroidetes* genomes are generally devoid of canonical GGDEF and EAL c-di-GMP signaling domains (bottom) and 97 of 99 STING-containing bacteria lack any other predicted c-di-GMP signaling component. Loss of canonical c-di-GMP signaling provides an explanation for co-option of c-di-GMP as a CD-NTase immune signal. Numbers near bar graphs denote genomes containing at least one gene with a GGDEF or EAL domain, out of total genomes in the analyzed database. **d)** Electrophoretic mobility shift assay monitoring bacterial STING–cyclic dinucleotide complex formation. Bacterial TM-STING and TIR-STING proteins preferentially recognize c-di-GMP and exhibit weaker affinity for 3',3'-cGAMP. Bacterial STING receptors are unable to recognize the mammalian second messenger 2',3'-cGAMP. TM-STING (*R. ehrenbergii*, Δ TM) and TIR-STING (*S. faecium*, full-length). Data are representative of 3 independent experiments. **e)** Zoom-in cutaway of the cyclic dinucleotide binding pocket in the *FsSTING*–3',3'-cGAMP structure. Top, *FsSTING* makes sequence-specific contacts to the guanine base consistent with preferential c-di-GMP recognition. Bottom, bacterial STING recognizes 3',3'-cGAMP (yellow) in a compact conformation similar to 2',3'-cGAMP (grey) in complex with human STING (PDB 4KSY).

Figure 2.4 Structural analysis of STING-associated CD-NTase enzymes.

a) Sequence and secondary structure alignment of STING-associated CD-NTases reveals the extent of homology between CdnE homologues from unrelated bacterial strains. Highlighted positions include active-site residues (pink box), and an aspartic acid substitution at a position known to be involved in nucleotide substrate selection (orange box) that is unique to CD-NTases in STING-containing CBASS operons (Whiteley et al. 2019). A divergent *E. coli* CdnE that synthesizes cyclic UMP–AMP is included for comparison.

b) Crystal structures of *FsCdnE* and *CgCdnE* from STING-containing CBASS operons allow direct comparison with previously determined bacterial and human CD-NTase structures. The *FsCdnE* and *CgCdnE* structures are most closely related to the Clade E CD-NTase structure from *Rhodothermus marinus* CdnE (*RmCdnE*). *RmCdnE* PDB 6E0L (Whiteley et al. 2019), *V. cholerae* DncV PDB 4TY0 (Kranzusch et al. 2014), human cGAS PDB 6CTA (DNA omitted for clarity) (Zhou et al. 2018).



b

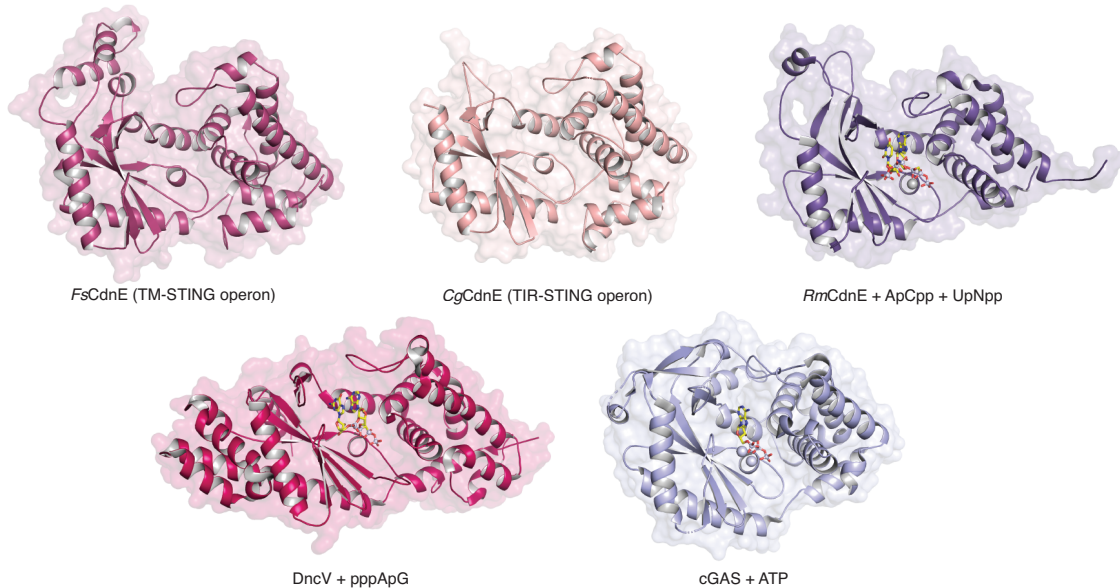
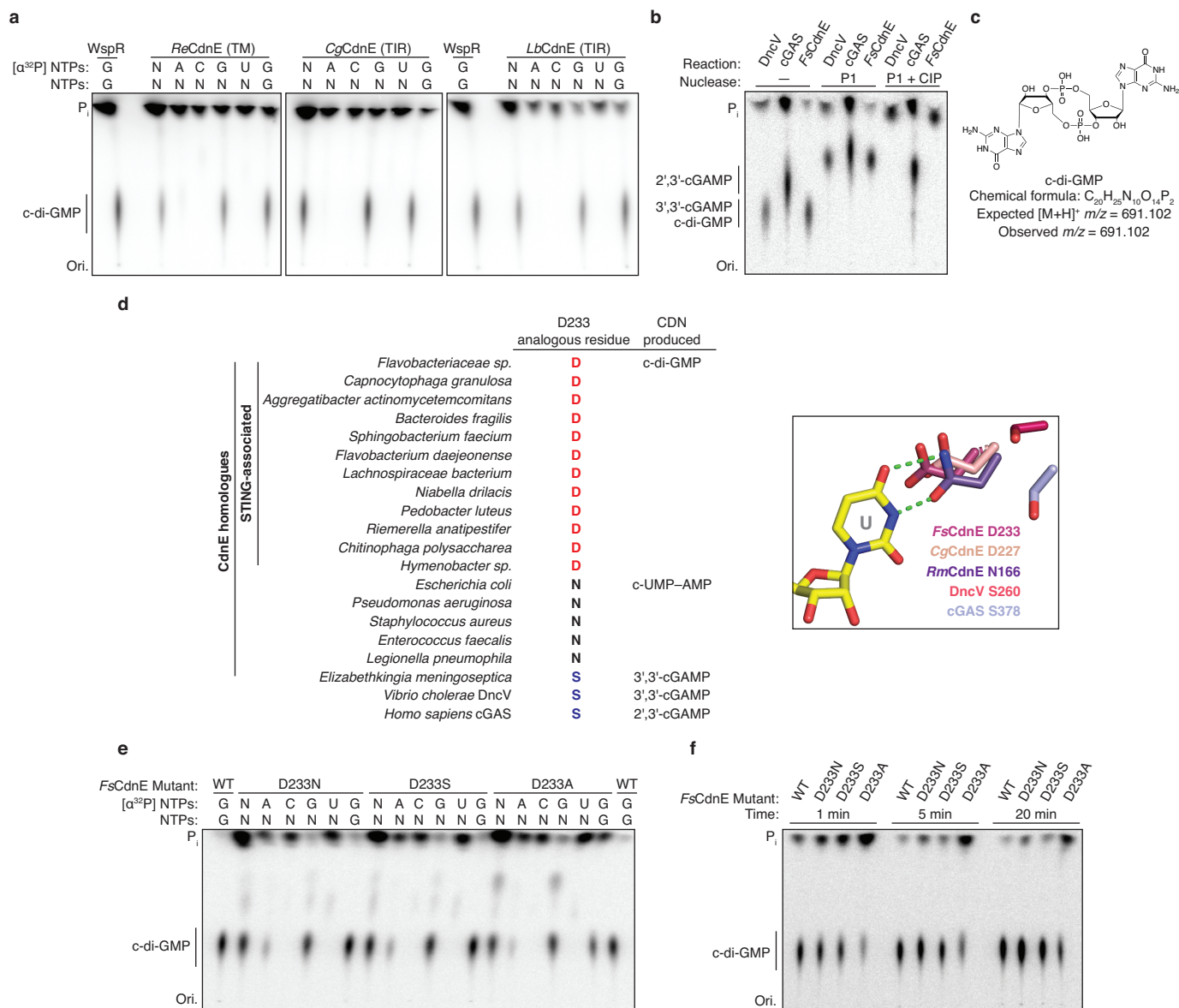


Figure 2.5 Biochemical analysis of c-di-GMP synthesis by bacterial STING-associated CD-NTases.

a) In addition to *FsCdnE* (see Figure 2.3b), CdnE homologues from three divergent STING-containing CBASS operons were purified and tested for cyclic dinucleotide synthesis specificity using $\alpha^{32}\text{P}$ -radiolabeled NTPs and thin-layer chromatography. Deconvolution experiments show a single major product requiring only GTP which migrates identically to c-di-GMP synthesized by the GGDEF enzyme WspR. All reactions were treated with alkaline phosphatase to remove exposed phosphates. Of note, only two bacterial genomes encoding a STING-containing CBASS operon retain proteins with a predicted canonical GGDEF c-di-GMP signaling domain. The exceptions are *Chlorobi bacterium EBPR_Bin_190* which contains a single GGDEF domain that is fused to a SLATT domain and may be part of a CBASS-like system (Cohen et al. 2019) and a *Lachnospiraceae bacterium RUG226* genome that encodes many GGDEF genes. The *Lachnospiraceae bacterium RUG226* CdnE retains exclusive production of c-di-GMP suggesting the CdnE-STING system is sequestered in this bacterium or that an unknown mechanism may exist to prevent toxic STING activation. *ReCdnE*, *Roseivirga ehrenbergii*; *CgCdnE*, *Capnocytophaga granulosa*; *LbCdnE*, *Lachnospiraceae bacterium*; N, all four rNTPs; P_i , inorganic phosphate; Ori., origin. Data are representative of 2 independent experiments. **b)** Nuclease treatment confirms that the *FsCdnE* enzymatic product contains only canonical 3'–5' phosphodiester bonds. The $\alpha^{32}\text{P}$ -GTP cyclic dinucleotide product is susceptible to cleavage by nuclease P1 resulting in release of GMP as a new species which migrates further up the TLC plate. Further digestion with calf-intestinal phosphatase (CIP) removes all exposed phosphates resulting in complete loss of a labeled product spot. *DncV* (Dinucleotide cyclase in *Vibrio*) derived 3',3'-cGAMP is similarly susceptible to complete digestion by P1 and CIP treatment whereas 2',3'-cGAMP synthesized by mouse cGAS is only partially digested due to the presence of the non-canonical 2'–5' bond. Data are representative of 2 independent experiments. **c)** High-resolution mass spectrometry analysis confirms the identity of the major *FsCdnE* enzymatic product as canonical c-di-GMP. Chemically synthesized c-di-GMP was used for direct spectral comparison. **d)** Sequence alignment and zoom-in inset of the active-site of the *RmCdnE* structure in complex with nonhydrolyzable UTP and ATP analogs (PDB 6E0L) demonstrating a contact in the CD-NTase lid domain known to control nucleobase sequence specificity (Whiteley et al. 2019). *RmCdnE* synthesizes cyclic UMP–AMP and uses N166 to specifically contact the uridine Watson-Crick edge. In contrast, *FsCdnE* and *CgCdnE* contain an aspartic acid substitution at this position and synthesize c-di-GMP, and *Vibrio cholerae* *DncV* and human cGAS contain a serine substitution at this position and synthesize 3',3'-cGAMP and 2',3'-cGAMP. 93% of STING-associated CD-NTase enzymes (96 of 103) contain a conserved aspartic acid at the *FsCdnE* D233 position consistent with strict specificity of c-di-GMP as the nucleotide second messenger controlling bacterial STING activation. *RmCdnE* PDB 6E0L (Whiteley et al. 2019), *V. cholerae* *DncV* PDB 4TY0 (Kranzusch et al. 2014), human cGAS PDB 6CTA (DNA omitted for clarity) (Zhou et al. 2018) **e,f)** Mutational analysis of the importance of D233 in *FsCdnE* c-di-GMP synthesis activity. D233 substitutions do not disrupt the overall ability of *FsCdnE* to selectively synthesize c-di-GMP, but a D233A substitution causes a mild reduction in nucleobase selectivity and efficiency of c-di-GMP synthesis. These results are consistent with a role for D233 in nucleobase selection but demonstrate full selectivity is achieved by additional contacts in the active site pocket. Data are representative of 3 independent experiments.



Additionally, we determined 1.5 Å and 2.3 Å crystal structures for two CdnE homologues revealing an active enzyme conformation and unique substitutions in the nucleobase acceptor pocket consistent with adaptation for c-di-GMP synthesis (Figures 2.4b and 2.5 d–f) (Whiteley et al. 2019). c-di-GMP is a common nucleotide second messenger used to regulate bacterial growth and intracellular signaling (Jenal, Reinders, and Lori 2017), and it is therefore difficult to conceive how bacteria could distinguish these functions of c-di-GMP from induction of CBASS immunity that results in rapid bacterial death (Cohen et al. 2019; Lowey et al. 2020; Ye et al. 2020). To explain the unexpected role of c-di-GMP in CBASS immunity, we analyzed the genomic context of all STING-containing CBASS operons and discovered that these systems are encoded almost exclusively in bacteria devoid of canonical GGDEF and EAL c-di-GMP signaling domains suggesting complete co-option of normal c-di-GMP function for a role in STING activation (Figure 2.3c).

Bacterial STINGs bind c-di-GMP with nanomolar affinity and exhibit a clear preference for canonical 3'–5'-linked cyclic dinucleotides (Figure 2.3d; Figure 2.6 a,b). In the *Fs*STING–3',3'-cGAMP structure, the cyclic dinucleotide backbone is coordinated with N91 and N172, and each nucleobase is sandwiched between a stacking interaction formed with F92 and R153 (Figure 2.3e). In agreement with the strong preference of bacterial STING for c-di-GMP, a universally conserved D169 residue reads out the guanosine nucleobase by making a sequence-specific contact to the N2 position (Figure 2.3e). Using mutational analysis, we verified the importance of residues in the STING cyclic dinucleotide binding pocket and confirmed that conserved nucleobase contacts are required for c-di-GMP recognition (Figures 2.8 and 2.9 a–e). Human STING does not form prominent sequence-specific contacts with cyclic dinucleotides. Instead, human STING residue R232 makes additional contact to the phosphodiester backbone and is critical for high-affinity recognition of the mammalian cGAS product 2'–5' / 3'–5' cGAMP (2',3'-cGAMP)(X. Zhang et al. 2013; Gao, Ascano, Zillinger, et al. 2013). A striking feature of bacterial STING receptors is an inability to recognize mammalian 2',3'-cGAMP (Figure 2.3d; Figure 2.6 a, b). Cyclic dinucleotides occupy a similar compact conformation in both bacterial STING and human STING, but in bacterial STING the R232-equivalent position R151 is flipped outwards and does not contact the cyclic dinucleotide backbone (Figure 2.2e). Additionally, a universally conserved T173 residue in bacterial STING located beneath the cyclic dinucleotide binding pocket reduces the space that would be necessary to accommodate a free 3'-OH

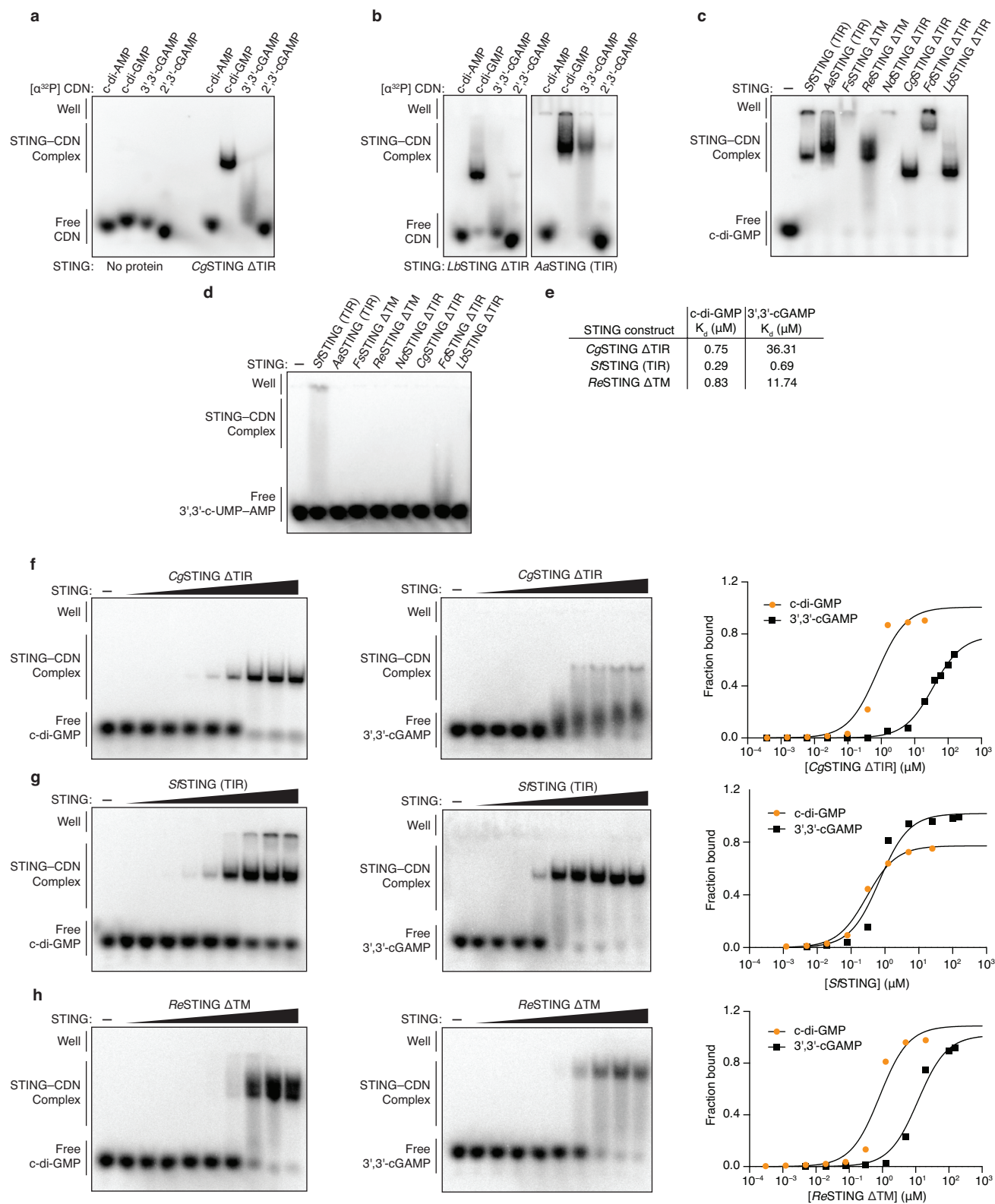
Figure 2.6 Biochemical analysis of bacterial STING cyclic dinucleotide recognition specificity.

a,b) Electrophoretic mobility shift assay (EMSA) of purified bacterial STING proteins with radiolabeled cyclic dinucleotide (CDN) ligands. Bacterial STING receptors specifically recognize c-di-GMP and have a weak ability to bind 3',3'-cGAMP. No interaction was observed with c-di-AMP or 2',3'-cGAMP. *Capnocytophaga granulosa* (CgSTING), *Lahnospiraceae bacterium* STING (*Lb*STING), *Aggregatibacter actinomycetemcomitans* STING (*Aa*STING). Data are representative of 2 independent experiments.

c) EMSA analysis of a diverse panel of bacterial STING homologues demonstrates conservation of c-di-GMP binding in both TM-STING and TIR-STING CBASS immunity. *Nd*STING (*Niabella drilacis*) and *Fd*STING (*Flavobacterium daejeonense*). Higher-order complex formation visible as well-shifted complexes is consistent with STING oligomerization results (see Figure 3f and Extended Data Figure 7). Data are representative of 3 independent experiments.

d) EMSA analysis of diverse bacterial STING homologues broadly demonstrates no interaction with the 3',3'-c-UMP-AMP second messenger synthesized by the divergent CD-NTase *E. coli* CdnE¹⁵ and further confirms the specificity of c-di-GMP signaling in bacterial STING-containing CBASS operons. Data are representative of 2 independent experiments.

e,f,g,h) EMSA analysis and quantification of the affinity of bacterial STING homologues for c-di-GMP and 3',3'-cGAMP. Signal intensity analysis is plotted as fraction bound (shifted / total signal) as a function of increasing protein concentration and fit to a single binding isotherm. CgSTING and ReSTING have a >10-fold preference for c-di-GMP while SfSTING has a similar apparent affinity for c-di-GMP and 3',3'-cGAMP. Data are representative of 2 independent experiments.



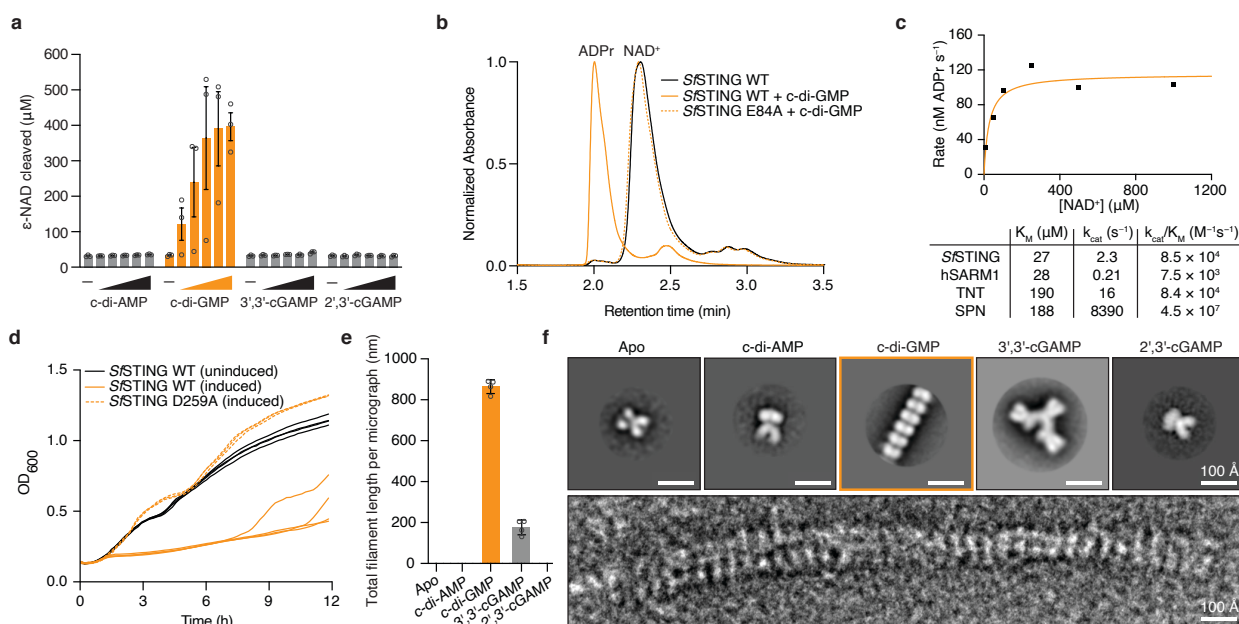


Figure 2.7 Cyclic dinucleotide-recognition controls bacterial STING oligomerization and TIR NAD⁺ cleavage activity.

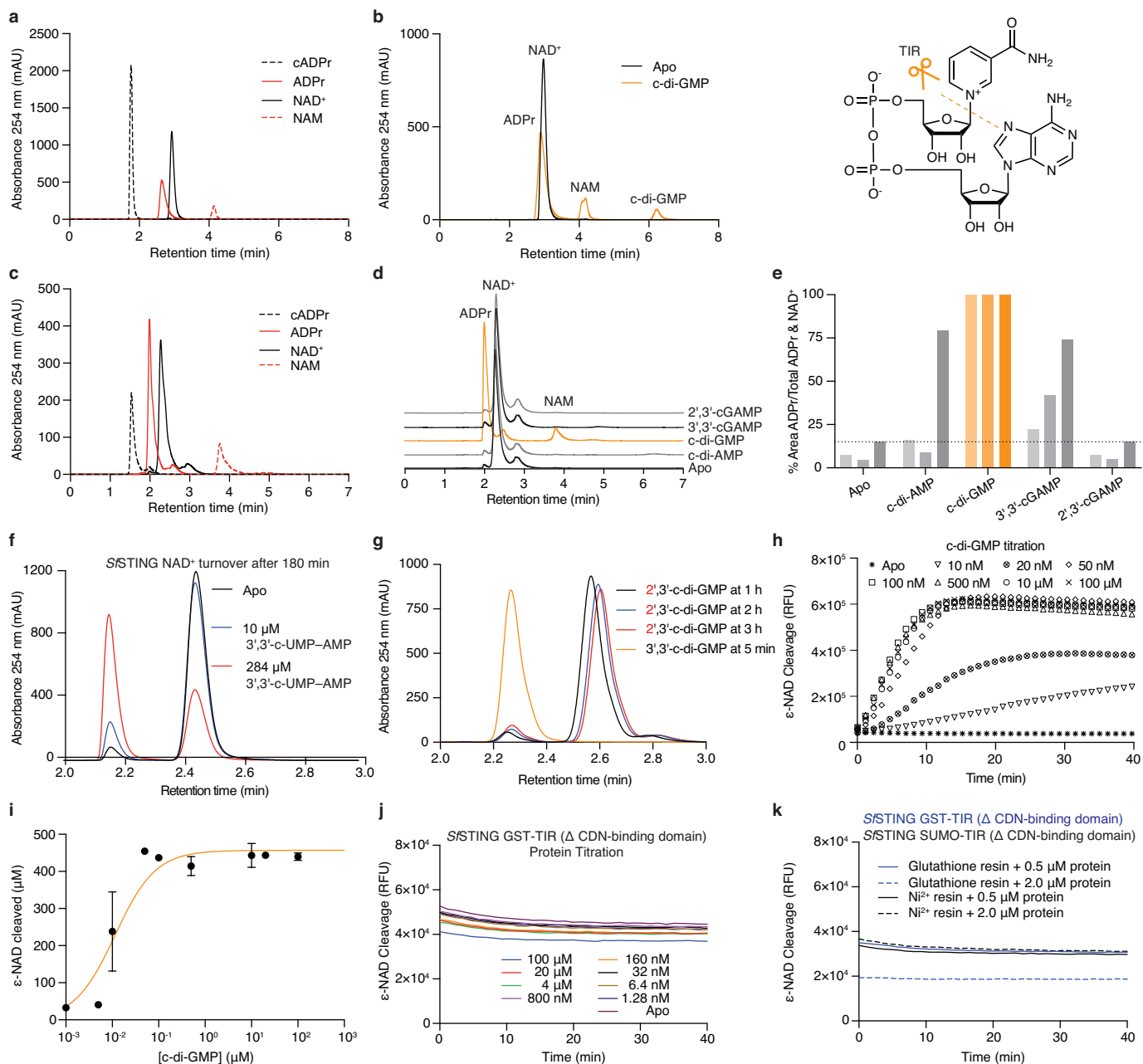
a) Analysis of *S. faecium* TIR-STING (SfSTING) NAD⁺ cleavage activity using the fluorescent substrate ε-NAD. TIR-STING activity is potently stimulated in the presence of c-di-GMP (0, 5, 10, 15, 20 nM, 20 μM). Data are mean ± s.e.m. for n = 3 biological replicates. **b)** HPLC analysis of TIR-STING NAD⁺ cleavage. SfSTING cleaves NAD⁺ into the products ADPr and NAM, and activity is strictly dependent on the TIR active-site residue E84. Data are representative of 3 independent experiments. **c)** Quantification of SfSTING activity and comparison with previously characterized NADase and glycosyl hydrolase enzymes human SARM1 (hSARM1) (Loring et al. 2020), *Mycobacterium tuberculosis* TNT (Tak et al. 2019), and *Streptococcus pyogenes* SPN (J. Ghosh et al. 2010). Data are mean for n = 2 independent replicates and are representative of 3 independent experiments. **d)** Analysis of SfSTING toxicity in *E. coli* cells expressing normal c-di-GMP signaling enzymes. Expression of SfSTING from an arabinose-inducible promoter (induced) results in a potent growth arrest phenotype. Toxicity is lost with a D259A mutation to the SfSTING cyclic dinucleotide-binding domain that inhibits c-di-GMP recognition. Each line represents the average of 2 technical replicates for each of four separately outgrown colonies. Data are representative of 2 independent experiments. **e)** Negative stain electron microscopy (EM) analysis and quantification demonstrates that SfSTING forms stable filaments only in the presence of c-di-GMP. Data are mean ± s.d. for quantification of n = 4 groups of 10 micrograph images. **f)** Electron microscopy analysis of SfSTING filament formation. Top, 2D class averages demonstrate that SfSTING forms well-ordered filaments in the presence of c-di-GMP and smaller fragmented assemblies in the presence of 3',3'-cGAMP. Bottom, SfSTING–c-di-GMP filaments can be >30 nm in length. 2D class averages were derived from particles selected from 75 micrographs for each condition.

within 2',3'-cGAMP (Figure 2.2 f). We observed that bacterial STING is >1000× more sensitive to c-di-GMP than a synthetic analog with a 2'–5' linkage (2',3'-c-di-GMP) (Figure 2.8g), further confirming the strict specificity of bacterial STING for 3'–5'-linked cyclic dinucleotides.

Activation of CBASS immunity induces bacterial growth arrest or cell death to destroy virally infected cells and limit phage propagation (Cohen et al. 2019; Ye et al. 2020; Lowey et al. 2020). Bacterial STING domains occur primarily as fusions to a Toll/Interleukin-1 receptor (TIR) adaptor domain,

Figure 2.8 Bacterial STING activation of TIR NADase activity.

a) HPLC analysis of chemical standards separated with an ammonium acetate:methanol gradient elution used to analyze bacterial TIR-STING activity (see Methods). The NAD⁺ and ADPr peaks have overlapping bases under these conditions. cADPr, cyclic adenosine diphosphate-ribose; ADPr, adenosine diphosphate-ribose; NAD⁺, β-nicotinamide adenine dinucleotide; NAM, nicotinamide. **b)** HPLC analysis of SfSTING NAD⁺ cleavage activity with gradient elution. SfSTING at 500 nM protein with 2 μM c-di-GMP converts 500 μM NAD⁺ into ADPr and NAM in 30 min at ambient temperature. SfSTING does not generate any cyclized product and is therefore a standard glycosyl hydrolase. Inset: Schematic of NAD⁺ cleavage reaction. **c)** HPLC analysis of chemical standards separated with an alternative isocratic elution strategy (see Methods) that results in clearer separation of NAD⁺ and ADPr peaks. **d)** HPLC analysis of SfSTING NAD⁺ cleavage activity with isocratic elution. SfSTING NAD⁺ cleavage activity requires specific activation with c-di-GMP (30 min reactions). **e)** HPLC analysis of SfSTING NAD⁺ cleavage activity and cyclic dinucleotide agonist specificity. Each reaction was tested with 500 nM SfSTING, 500 μM cyclic dinucleotide, and 500 μM NAD⁺ and sampled at 45, 90 or 180 min (gradient coloring in bars). SfSTING preferentially responds to c-di-GMP, but 3',3'-cGAMP and c-di-AMP can function as weak agonists. Data are representative of 3 independent experiments. **f)** HPLC analysis of SfSTING NAD⁺ cleavage activity in the presence of 3',3'-c-UMP-AMP. 3',3'-c-UMP-AMP is a >1000× weaker agonist than c-di-GMP. Data are representative of 3 independent experiments. **g)** HPLC analysis of SfSTING NAD⁺ cleavage activity in the presence of a synthetic c-di-GMP analog with a noncanonical 2'-5' linkage (2',3'-c-di-GMP). 2',3'-c-di-GMP is not capable of stimulating robust SfSTING activation even at very high concentrations (250 μM vs. 250 nM canonical c-di-GMP), confirming the specificity of bacterial STING for 3'-5'-linked cyclic dinucleotides. Data are representative of 3 independent experiments. **h)** Plate reader analysis of SfSTING NAD⁺ cleavage activity using the fluorescent substrate ε-NAD. ε-NAD increases in fluorescence intensity after cleavage. SfSTING exhibits rapid catalysis with complete turnover at 500 nM protein with 500 nM c-di-GMP after 10 min at 25°C. No background activity is observed in the absence of ligand. Data are representative of 3 independent experiments. **i)** Plate reader analysis of SfSTING NAD⁺ cleavage activity in the presence of 500 nM protein with increasing c-di-GMP concentration reveals that low nM c-di-GMP levels are sufficient to induce ε-NAD cleavage. Greater c-di-GMP concentrations are required for maximal activity consistent with binding data and the higher amount of c-di-GMP required for complete stabilization of the SfSTING-c-di-GMP complex (See Extended Data Figure 4e-h). Saturation occurs above 100 nM c-di-GMP with 40 min reactions. Data are ± s.d. of n = 3 technical replicates and are representative of 3 independent experiments. **j,k)** TIR-domain NAD⁺ cleavage activity requires protein oligomerization. For other systems (Horsefield et al.; Wan et al.), TIR activation has been observed at very high *in vitro* protein concentrations or in the presence of affinity resins as an artificial oligomerization-inducing matrix (Horsefield et al.; Wan et al.). We used a GST-TIR construct to express the SfSTING TIR domain in absence of the STING cyclic dinucleotide binding domain and observed that even at >200× the concentrations for which the full-length protein shows c-di-GMP induced activity, or in the presence of multivalent affinity resin, no NAD⁺ cleavage activity occurs. These results demonstrate NAD⁺ cleavage activity specifically requires STING cyclic dinucleotide recognition for activation. Data are representative of 2 independent experiments.



or more rarely appended to predicted transmembrane segments (Figure 2.3a; Figure 2.2a). TIR domains can function as NAD⁺ (β-nicotinamide adenine dinucleotide) hydrolases in plant and animal immunity (Essuman et al. 2018; Horsefield et al. 2019; Wan et al. 2019). We therefore tested a full-length TIR-STING fusion from *Sphingobacterium faecium* (*Sf*STING; IMG Gene ID 2735805876) for catalytic function and observed rapid hydrolysis of NAD⁺ to nicotinamide and adenine diphosphate-ribose (Figure 2.7a, b; Figure 2.8 a–d). *Sf*STING potently responds to c-di-GMP and weakly to 3',3'-cGAMP, and catalysis is abolished by mutating a conserved TIR active-site glutamic acid (Figure 2.7b; Figure 2.9 f, g). Bacterial TIR-STING activation requires <100 nM c-di-GMP and results in cleavage of NAD⁺ at 10× the rate observed for plant and animal TIR proteins involved in immune defense (Figure 2.7c; Figure 2.8h, i). We expressed *Sf*STING in *E. coli* cells that synthesize c-di-GMP and observed potent growth inhibition (Figure 2.7d). *Sf*STING-induced toxicity is lost with a D259A mutation that prevents c-di-GMP recognition and can be partially overcome with nicotinamide supplementation (Figure 2.7d; Figure 2.9 h, i), further supporting a role for bacterial STING in CBASS-mediated immunity as a c-di-GMP-responsive NADase effector.

Cyclic dinucleotide recognition drives bacterial STING oligomerization and formation of ordered filaments readily observable with negative stain electron microscopy (Figure 2.7 e,f; Figure 2.10). In the presence of activating c-di-GMP, *Sf*STING assembles into filaments typically 25–30 nm in length that exhibit four-fold symmetry and contain an ordered array of parallel-stacked homodimers (Figure 2.7f; Figure 2.10 j–l). Electron microscopy and biochemical analysis of *Sf*STING in the presence of the weak agonist 3',3'-cGAMP reveals partial filament formation and STING complexes in a tetramer-like conformation, indicating that cyclic dinucleotide-stabilization of STING tetramers is likely required to seed filament growth (Figure 2.7 e, f; Figure 2.10j–l). Higher-order oligomerization is a known requirement for TIR activation (Horsefield et al. 2019; Wan et al. 2019), and mutations to the *Sf*STING cyclic dinucleotide binding domain that disrupt c-di-GMP-induced oligomerization prevent activation of TIR-STING NAD⁺ cleavage (Figure 2.9 b, f, g). Remarkably, bacterial STING filament formation is consistent with a recently proposed model of human STING activation where parallel-stacking of STING homodimers initiates oligomerization and

Figure 2.9 Mutagenesis analysis of bacterial STING cyclic dinucleotide recognition and TIR activation.

a) Table of bacterial STING cyclic dinucleotide contacts tested with mutagenesis analysis. Residues were selected according to contacts observed in the structure of the *Fs*STING–3',3'-cGAMP complex (*Fs*STING residues listed in parentheses) and tested in *Sf*STING to allow analysis of the impact on both c-di-GMP binding and NADase activity. **b)** Plate reader analysis of mutant *Sf*STING NAD⁺ cleavage activity in the presence of 500 nM protein and increasing c-di-GMP concentration. Mutant *Sf*STING variants with cyclic dinucleotide binding pocket mutations require 10–1000× greater c-di-GMP concentration for NADase activation. Data are ± s.d. of n = 3 technical replicates and are representative of 3 independent experiments. **c,d,e)** EMSA analysis demonstrating that *Sf*STING R234A and D259A mutations reduce stable c-di-GMP complex formation compared to *Sf*STING WT (see Figure 2d). *Sf*STING D259A protein titration and quantification confirms significantly reduced affinity for c-di-GMP. Data are representative of 3 independent experiments. **f,g)** HPLC and plate reader analysis of mutant *Sf*STING NAD⁺ cleavage activity in the presence of 500 nM protein and 10 μM c-di-GMP (HPLC) or 500 nM protein and ± 20 μM c-di-GMP (plate reader). Mutation of residues responsible for ligand recognition in the cyclic dinucleotide binding domain (R234 and D259) and catalysis in the TIR enzymatic domain (E84) disrupts *Sf*STING NADase activity and explains loss of *E. coli* toxicity observed in Figure 3d. Note, 100 μM of c-di-GMP was used for *Sf*STING D259A plate reader NAD⁺ cleavage analysis to confirm complete loss of c-di-GMP-induced activation. HPLC data are representative of 3 independent experiments. Plate reader data are ± s.d. of n = 3 technical replicates and are representative of 3 independent experiments. **h,i)** Analysis of *Sf*STING toxicity in *E. coli* cells expressing normal c-di-GMP signaling enzymes with and without nicotinamide (NAM) supplementation. NAM supplementation is sufficient to partially alleviate *Sf*STING WT-induced NADase toxicity. Each line represents the average of 2 technical replicates for each of four separately outgrown colonies. Data are representative of 2 independent experiments.

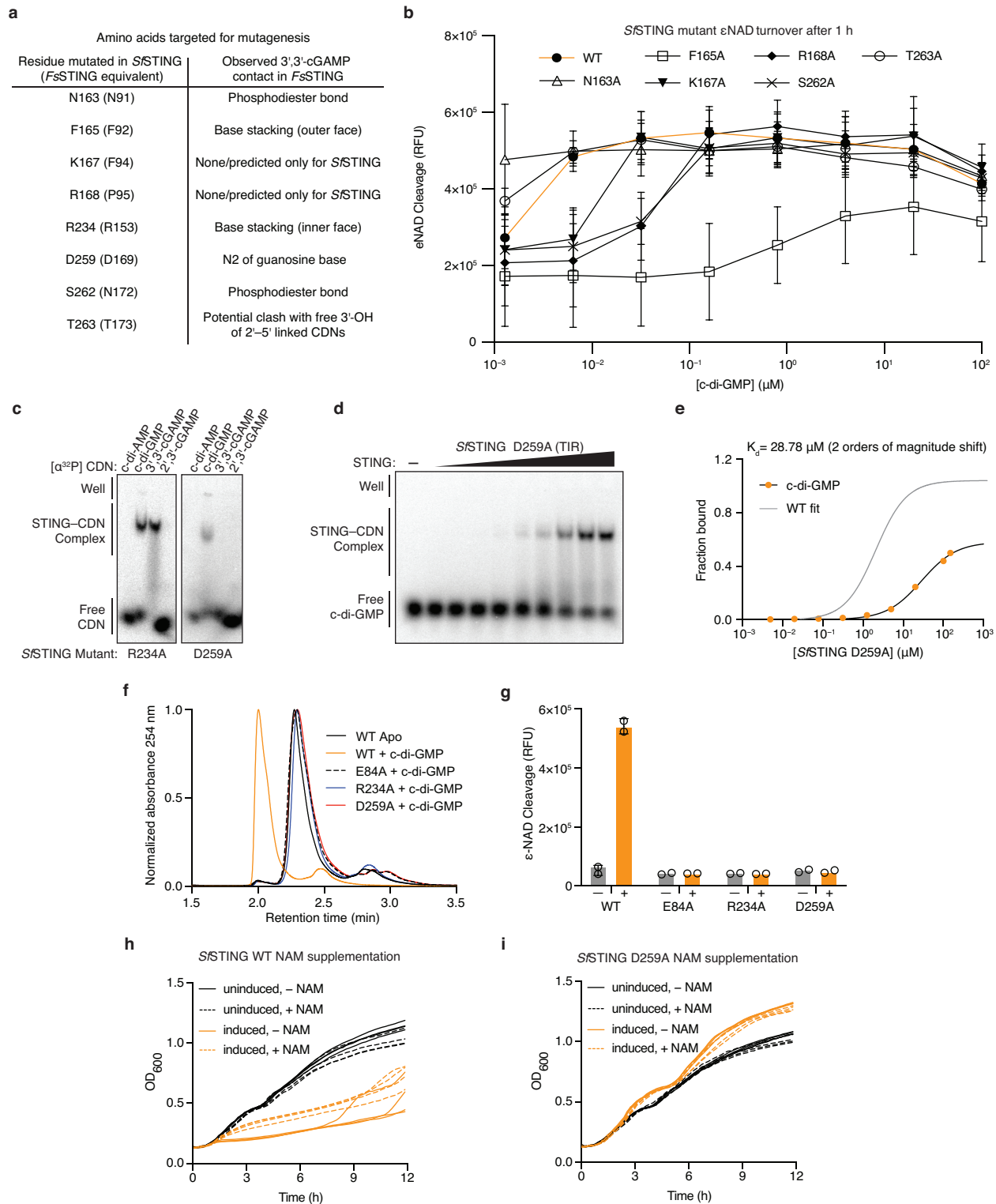
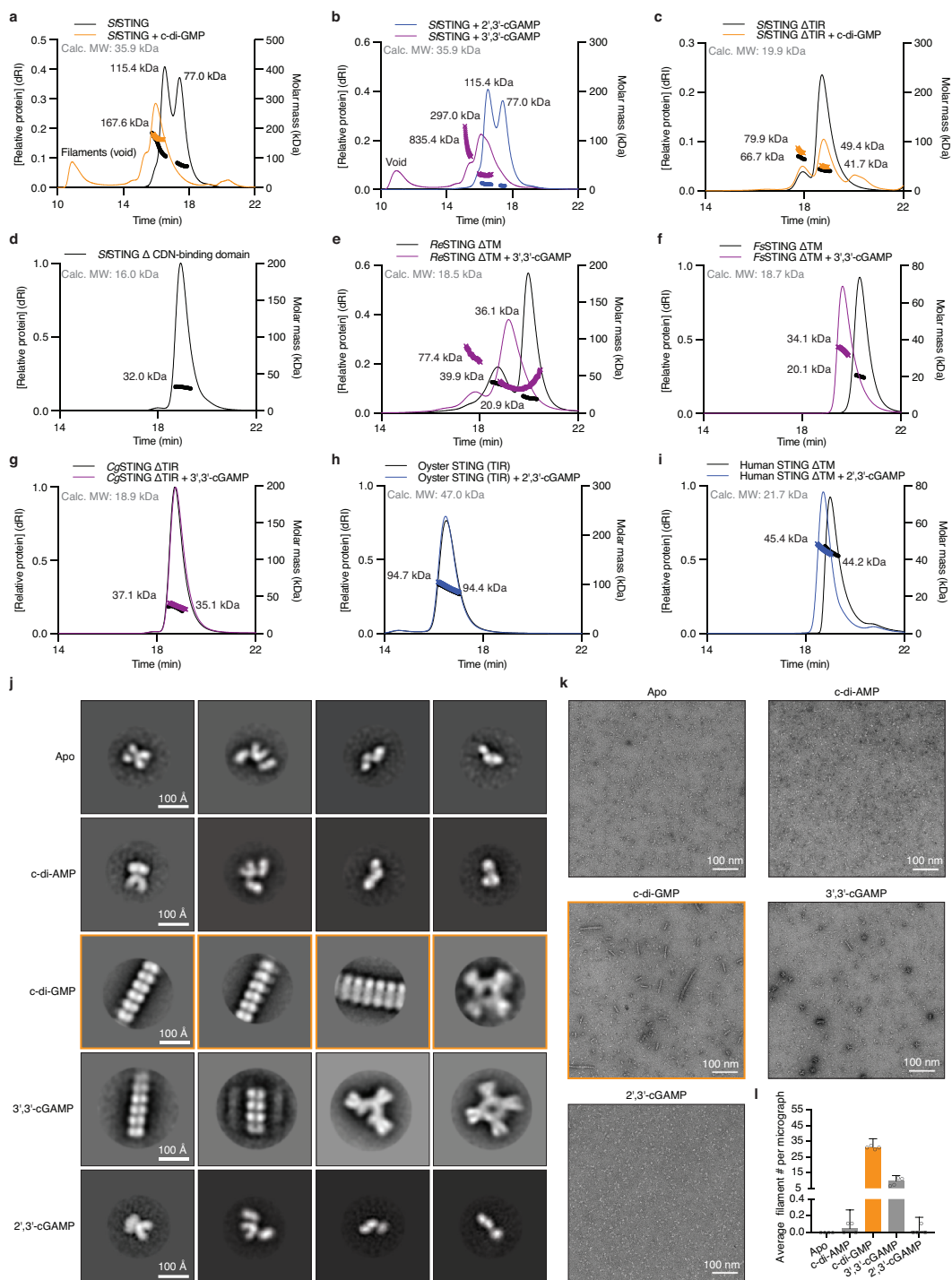


Figure 2.10 Conservation of oligomerization as a mechanism of STING activation.

a–i) STING SEC-MALS analysis. **a,b**, Full-length *Sf*STING changes oligomeric state in the presence of c-di-GMP or the weak agonist 3',3'-cGAMP, and does not change oligomeric state in the presence of 2',3'-cGAMP. **c**, With the TIR domain removed (Δ TIR), *Sf*STING no longer forms higher order complexes but notably remains dimeric in the apoprotein form. Bacterial TIR-STING proteins therefore appear to require the TIR domain to maintain stable higher-order oligomerization suggesting that intermolecular contacts are made with both TIR and STING domains. **d**, A TIR-only construct of *Sf*STING with the cyclic dinucleotide binding domain removed (Δ CDN) elutes as a single species which is consistent with the molecular weight for a homodimer. **h**, human STING (Δ TM) is a dimer in solution with or without 2',3'-cGAMP, confirming that TM contacts are required for oligomerization and filament formation (C. Zhang et al. 2019; Ergun et al. 2019). Nearly all tested bacterial and metazoan STING constructs migrate as dimers in solution consistent with the cyclic dinucleotide binding domain forming a constitutive homodimeric complex for ligand recognition. Two exceptions include *Re*STING (Δ TM) and *Fs*STING (Δ TM) which form a mixture of monomeric and dimeric states in the absence of ligand and dimers or tetramers in the presence of 3',3'-cGAMP. These results indicate that alternative oligomerization events may be required for activation of bacterial TM-STING effector function. **j**) Negative stain electron microscopy 2D class averages for *Sf*STING (E84A mutant) alone or in the presence of cyclic dinucleotide ligands. Stable STING filament formation requires c-di-GMP. 2D class averages were derived from particles selected from 75 micrographs for each condition. **k**) Representative micrograph images reveal extensive filament formation of varying length and orientation in the presence of c-di-GMP. Apo, c-di-AMP, and 2',3'-cGAMP micrographs lack filaments. Images are each representative of $n = 75$ micrographs for each condition. **l**) Particle counting analysis of micrograph images shows that c-di-GMP induces more filament formation than 3',3'-cGAMP and stable filament formation does not occur in the presence of c-di-AMP or 2',3'-cGAMP. Data are mean $\pm$ s.d. for quantification of $n = 4$ groups of 10 micrograph images each.



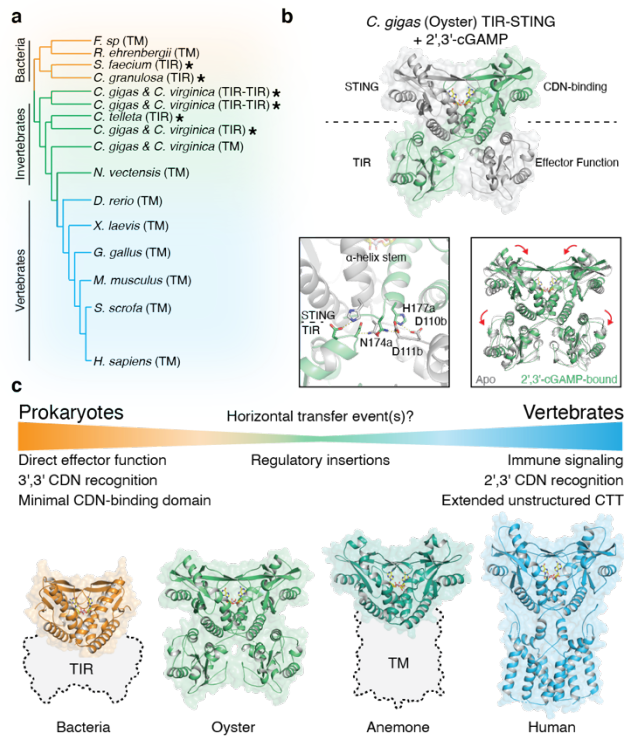


Figure 2.11 Structural basis of STING adaptation for signaling regulation.

a) Schematic derived from a structure-guided alignment of all known bacterial and metazoan STING sequences. Black stars denote TIR-STING fusions encoded in bacterial and metazoan genomes. Oysters encode several TIR-STING variants including genes predicted to contain multiple TIR domains (TIR-TIR). **b)** Crystal structure of the full-length *C. gigas* (oyster) TIR-STING receptor (Genbank sequence XP_011430837.1) in complex with 2',3'-cGAMP. Top, oyster TIR-STING adopts a domain-swapped conformation with TIR domains appended across a central dimeric axis. Bottom left, 2',3'-cGAMP recognition results in new interactions between the STING domain α -helix stem of one monomer and the TIR domain of the other monomer. Bottom right, Ligand binding induces STING domain lid closure and a downward rotation that repositions the appended TIR domains. **c)** Comparative structural analysis of STING adaptation suggests a model for the origin of cyclic dinucleotide sensing in innate immunity. STING evolved as a c-di-GMP sensor in prokaryotic bacteriophage defense, and acquisition in early metazoan cells may have allowed recognition of bacterial cyclic dinucleotides. Metazoan-specific structural insertions adapted STING for recognition of endogenous 2',3'-cGAMP signaling and enabled antiviral and anti-tumor immune signaling in vertebrate cells. Structures shown: bacterial STING-3',3'-cGAMP (*Fs*STING), oyster STING-2',3'-cGAMP (*C. gigas*), sea anemone STING-2',3'-cGAMP (*N. vectensis* PDB 5CFQ), human STING-2',3'-cGAMP (*H. sapiens* PDB 6NT5 modeled with 6NT7).

recruitment of the kinase TBK1 (C. Zhang et al. 2019; Ergun et al. 2019). Using the chicken STING cryo-electron microscopy structure as a guide (Shang et al. 2019), we constructed a model of bacterial STING oligomerization and identified surfaces in the bacterial STING domain at the beginning of helix $\alpha 2$ and end of helix $\alpha 4$ predicted to mediate oligomerization (Figure 2.12a). Mutations to these surfaces do not disrupt SfSTING c-di-GMP-binding but prevent oligomerization into filaments (Figure 2.12 b, c). In the absence of filament formation all SfSTING NADase activity is lost (Figure 2.12d) confirming that c-di-GMP-induced oligomerization is essential for TIR domain activation and effector function. These results define a core role of filament formation in STING activation and suggest that the primordial mechanism of STING signaling is cyclic dinucleotide-dependent effector oligomerization.

High-resolution crystal structures of *Fs*STING and *Cg*STING enabled construction of a structure-guided phylogenetic alignment of 103 bacterial and 492 metazoan STING proteins for comparative analysis of STING

Figure 2.12 Filament formation is required for bacterial TIR-STING activation.

a) Model of bacterial STING oligomerization and identification of surfaces involved in c-di-GMP-mediated filament formation. Electron microscopy analysis of *Sf*STING in the presence of c-di-GMP reveals filament formation likely occurs through parallel stacking of the homodimeric cyclic dinucleotide binding domain (see Figure 2.10). To construct a potential model of this interaction, we used the X-ray crystal structure of the human STING–2',3'-cGAMP complex (PDB 4KSY) (Kranzusch et al. 2015) and the cryo-electron microscopy structure of the chicken STING tetramer (PDB 6NT8) (Shang et al. 2019) (top) as guides to position the *Fs*STING–3',3'-cGAMP complex structure into a tetrameric conformation. The resulting model predicts that oligomerization-mediating surfaces in *Sf*STING include T200–N204, L275–F284, and G306–A310. **b)** EMSA analysis of *Sf*STING variants indicates that mutations to the predicted oligomerization surfaces do not prevent c-di-GMP recognition. *Sf*STING mutants tested include R307E/A309R, L201R/D203R, and the loop L275–Q282 replaced with a short GlySer linker (GSGGS). Data are representative of 2 independent experiments. **c)** Electron microscopy analysis of *Sf*STING variants in the presence of c-di-GMP. Mutations to the *Sf*STING surfaces identified in the structural model prevent all observable cyclic dinucleotide-induced filament formation supporting their predicted role in mediating oligomerization and bacterial STING filament formation. Images are each representative of $n = 5$ micrographs for each condition. **d)** HPLC analysis of mutant *Sf*STING NAD^+ cleavage activity in the presence of 500 nM protein and 250 μM c-di-GMP for 3 h. In the absence of cyclic dinucleotide-mediated oligomerization all *Sf*STING NADase activity is lost, confirming the requirement of filament formation in TIR domain activation. Data are representative of 3 independent experiments.

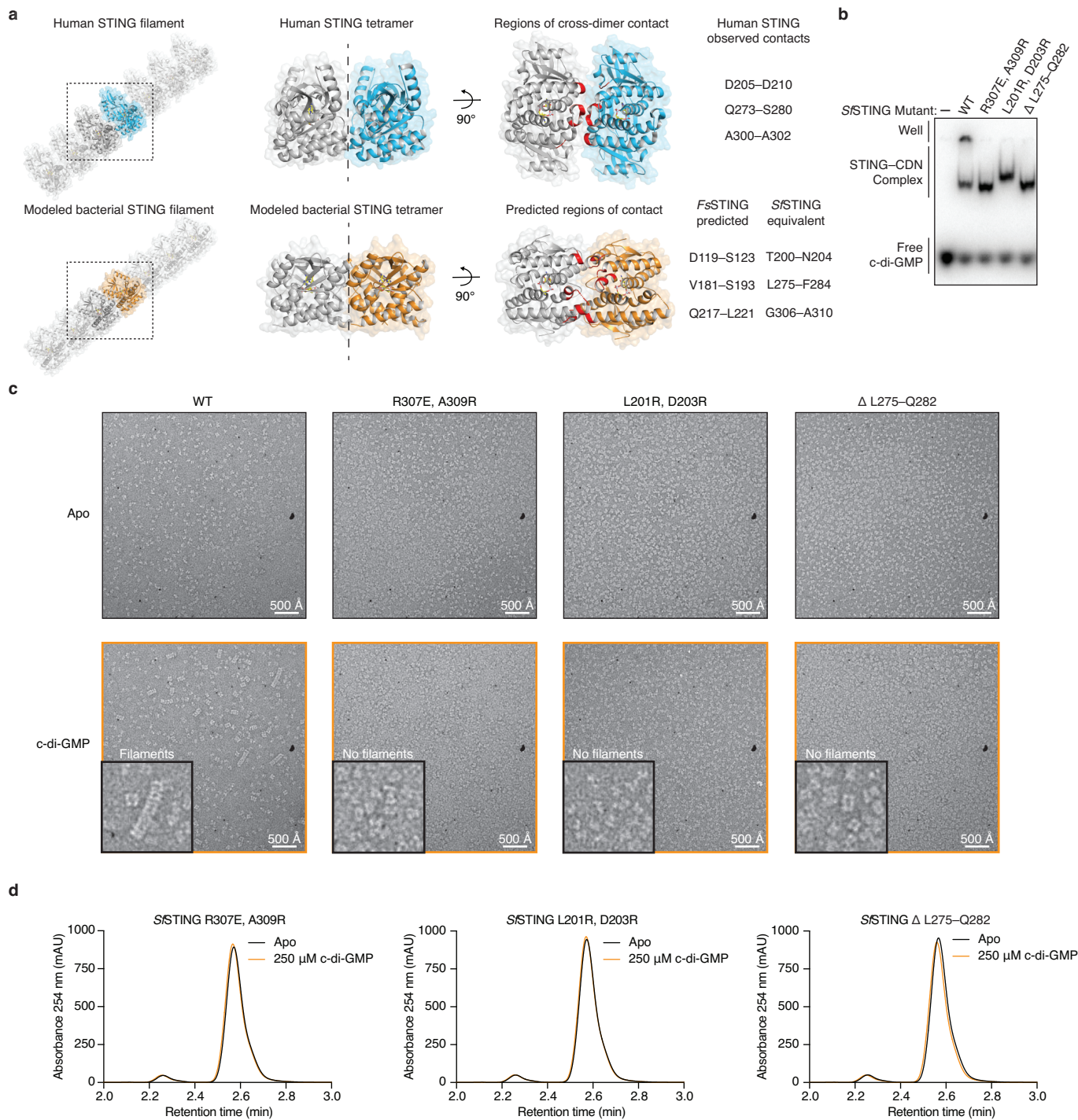
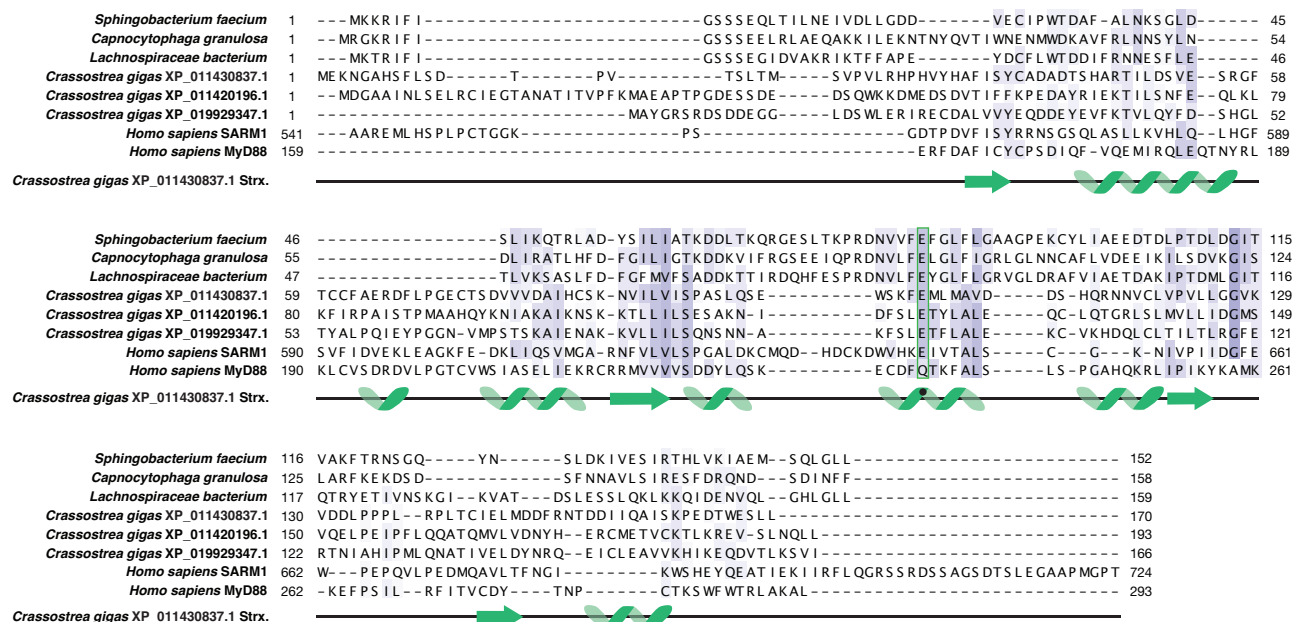


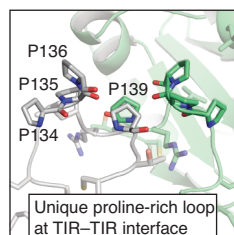
Figure 2.13 Structural analysis of metazoan TIR-STING homologs.

a) Structure-guided alignment of the TIR domain in oyster TIR-STING with reference bacterial and metazoan TIR-domain-containing proteins. SARM1 is an example of a human TIR domain that catalyzes NAD⁺ cleavage and MyD88 is an example of a human TIR domain that signals through protein–protein interaction. The catalytic glutamate responsible for supporting NAD⁺ cleavage is conserved at the same spatial position among bacterial and oyster TIRs but is mutated in MyD88 (green box). However, it is not currently possible to predict from structure or sequence alone if TIR domains have enzymatic activity. **b)** Distinct from other TIR domain structures, the TIR domain in oyster TIR-STING contains a proline-rich loop region at the interface suggesting a specific role in dimer stabilization. **c)** Superposition of a homology model of the *Sf*STING TIR-domain with the TIR-domain of oyster TIR-STING shows the predicted catalytic glutamates for both proteins occupy distinct locations in the TIR fold. **d)** Superposition of a homology model of the *Sf*STING TIR-domain compared to the crystal structure of human SARM1 bound to ribose implies that different NAD⁺ binding pockets may exist between bacterial and eukaryotic TIRs as previously suggested (Wan et al.). **e)** Superposition of a homology model of the *Sf*STING TIR-domain with the bacterial TIR domain from *P. dentrificans* shows a high degree of similarity. No crystal structures are available for bacterial TIR domains in an active state preventing identification of a specific mechanism of catalytic activation. **f)** EMSA analysis of oyster TIR-STING and mouse STING demonstrates a wide preference for cyclic dinucleotide interactions and clear ability to recognize the mammalian cGAS product 2',3'-cGAMP. Data are representative of 3 independent experiments. **g,h)** Oyster TIR-STING which binds all tested CDNs does not exhibit NAD⁺ cleavage activity even at 10× the protein and ligand concentrations used to achieve robust activity with bacterial TIR-STING. We tested four other oyster TIR-STING homologues and observed no cyclic dinucleotide-stimulated NAD⁺ cleavage activity. These results support a potential switch in TIR-dependent protein–protein interactions to control downstream signaling similar to the TIR domain in human MyD88. Data are representative of 2 independent experiments.

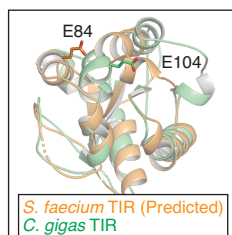
a



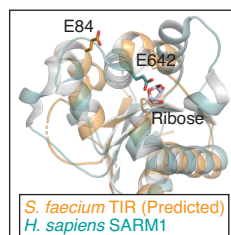
b



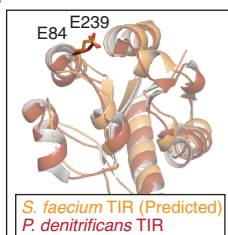
c



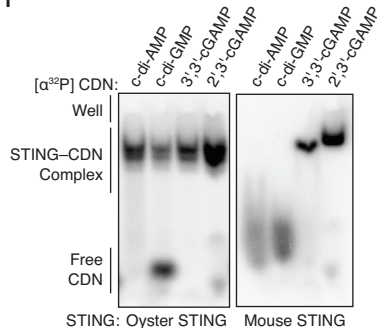
d



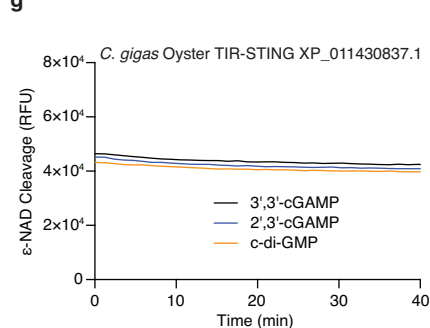
e



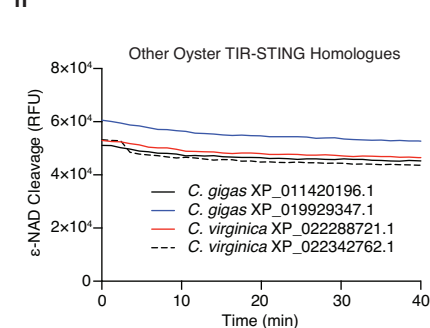
f



g



h



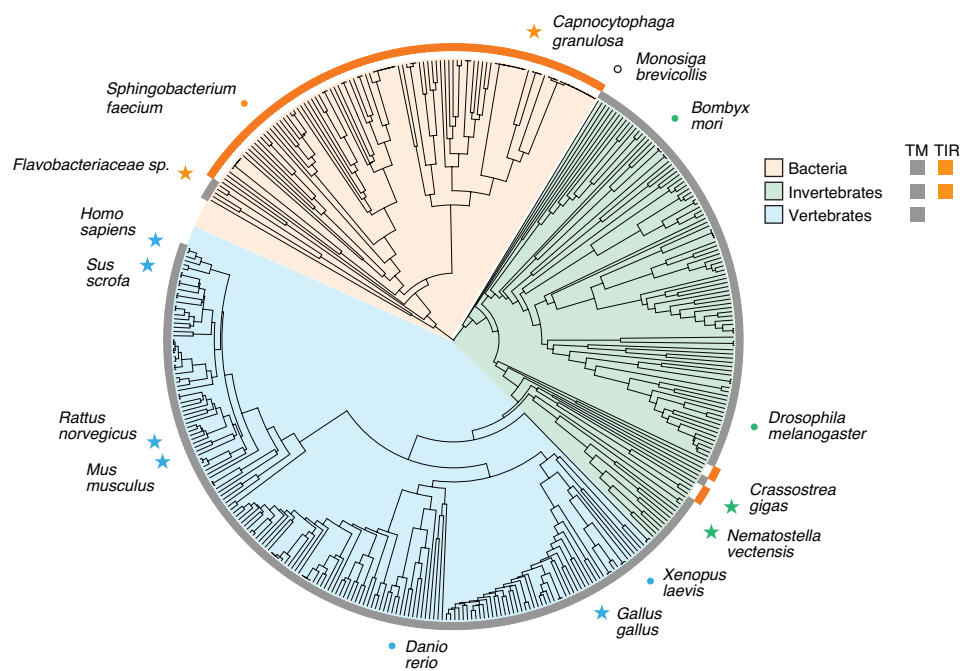
function and adaptation (Figure 2.11a; Figures 2.13 a, 2.14 a). STING distribution is most consistent with a bacterial origin and acquisition into an early metazoan ancestor (Figure 2.14a). The distinct cyclic dinucleotide binding pocket residues required for preferential recognition of c-di-GMP or 2',3'-cGAMP are fixed and nearly exclusive to bacterial or metazoan STING sequences. Strict kingdom-specific conservation further supports 2',3'-cGAMP as the dominant functional ligand throughout metazoan STING signaling, and suggests a clear transition occurred from c-di-GMP signaling in bacteria to noncanonical cyclic dinucleotide signaling in animals (Figure 2.14b) (Kranzusch et al. 2015; Gui et al. 2019; Margolis, Wilson, and Vance 2017). We identified several instances of invertebrate metazoan TIR-STING fusions with a predicted architecture similar to bacterial STING, including within oyster genomes previously noted to encode dramatic expansions of innate immune receptors and predicted CD-NTase enzymes (Margolis, Wilson, and Vance 2017; J. Wang et al. 2012; Kranzusch et al. 2015). To further understand adaptation of STING to signaling in animal cells, we determined a 2.4 Å crystal structure of metazoan TIR-STING from the Pacific oyster *C. gigas* (oyster TIR-STING). The full-length oyster TIR-STING structure reveals a domain-swapped configuration with a linker region that intertwines to connect the appended TIR effector module with the STING dimerization α -helix stem (Figure 2.11b). We next determined a 2.9 Å crystal structure of oyster TIR-STING in complex with 2',3'-cGAMP confirming the phylogenetic analysis of strict retention of 2',3'-cGAMP-interacting residues in invertebrate metazoan STING receptors (Figure 2.11b; Figures 2.13f and 2.14b). Distinct from recent cryo-electron microscopy analysis of full-length human STING with the N-terminal transmembrane domain, no 180° rotation is observed in the linker connecting the oyster STING and TIR domains in the presence of 2',3'-cGAMP (Shang et al. 2019). Instead, closure of oyster STING around 2',3'-cGAMP re-aligns inter-monomer contacts between the adjacent STING and TIR domains, inducing a downward 4° rotation in TIR domain orientation (Figure 2.11b). We did not observe oyster TIR-STING catalytic NAD⁺ cleavage activity in the presence of cyclic dinucleotides, suggesting additional requirements for enzyme activation or that reorientation of the oyster STING TIR domain may instead facilitate protein–protein interactions similar to the signaling adaptor function of TIR domains in human MyD88 and TRIF (Figure 2.13b–h) (Toshchakov and Neuwald 2020). These results further explain adaptation of STING in metazoans and provide a molecular mechanism for how STING cyclic dinucleotide sensing is structurally communicated to appended effector modules.

Our data define a conserved mechanism of STING-dependent signaling shared in bacteria and human cells, and support a unified model to explain the emergence of cyclic dinucleotide sensing in animal innate immunity (Figure 2.11c) (Kranzusch et al. 2015; Gui et al. 2019; Margolis, Wilson, and Vance 2017). Bacterial STING proteins function as c-di-GMP receptors that control oligomerization-dependent activation of appended effector domains, and are frequently encoded in the genomes of *Bacteroidetes* species that grow as commensals enriched in human and animal microbiota (Wexler and Goodman 2017). Interestingly, c-di-GMP was the first discovered ligand of human STING (Burdette et al. 2011), and human STING recognition of bacterial cyclic dinucleotides is critical for immune detection of intracellular pathogens such as *Listeria monocytogenes* (Woodward, Lavarone, and Portnoy 2010; Dey et al. 2015; Sixt et al. 2017). Acquisition of STING as a c-di-GMP receptor could therefore provide an immediate selective advantage enabling animal cells to sense intracellular bacteria through detection of essential cyclic dinucleotide second messengers conserved in prokaryotes. Our structural analysis defines cyclic dinucleotide binding pocket mutations that adapted metazoan STING for recognition of endogenous 2',3'-cGAMP signaling, and explains how emergence of metazoan-specific insertions enabled autophagy and type I interferon-dependent effector responses. Together with previous identification of diverse cGAS homologues and 2'–5'-linked signals in prokaryotic antiviral immunity (Cohen et al. 2019; Whiteley et al. 2019; Lowey et al. 2020), functional conservation of STING in bacteria reveals that each of the core protein components controlling human cGAS-STING signaling arose from an ancient mechanism of bacteriophage defense.

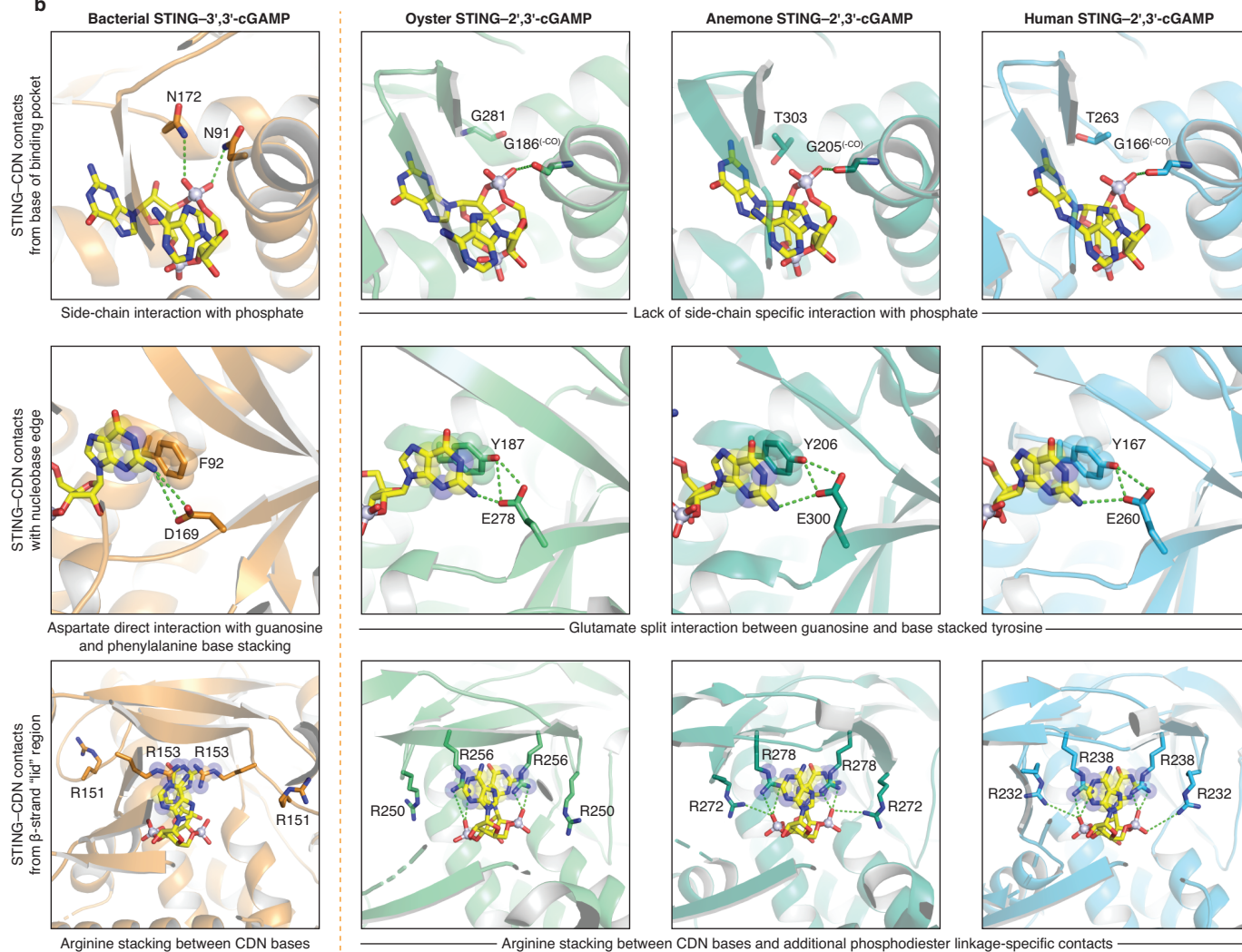
Figure 2.14 Structure-guided analysis of STING phylogenetic conservation and cyclic dinucleotide recognition.

a) Structure-guided alignment and phylogenetic tree of STING proteins across bacterial and metazoan kingdoms. Bacterial STING homologues form a distinct cluster separate from all metazoan STING sequences and are mostly represented by TIR-STING fusions. A single STING-domain containing protein was identified in the choanoflagellate *Monosiga brevicollis* (denoted as an open black circle as this species is outside of the kingdom metazoan), no STING-domain containing proteins were found in archaea. TIR-STING fusions are rare in eukaryotes and cluster among invertebrate metazoans. No TIR-STING examples occur in vertebrates. Specific species of interest are highlighted to show the breadth of sequence diversity and stars mark proteins with available structures. **b)** Direct comparison of bacterial, oyster, anemone, and human STING crystal structures reveals conservation of specific cyclic dinucleotide contacts and critical differences in phosphodiester linkage recognition. Stacking interactions formed with the cyclic dinucleotide nucleobase face, aromatic side chains at the top of the α -helix stem, and arginine residues extending downward from the lid region are major conserved features shared between bacterial and metazoan STING proteins. Nucleotide-specific contacts are divergent between distinct STING structures, but notably the critical D169 guanosine N2-specific contact present in bacterial STING is conserved with a glutamic acid side chain in nearly all metazoan STING proteins. A critical feature absent in bacterial STING receptors is additional arginine-specific contacts to the phosphodiester backbone. The human STING R232 side chain contact known to be critical for high-affinity interactions with 2',3'-cGAMP is conserved throughout metazoan STINGs representing a unique adaptation not found within bacterial STING receptors. Bacterial STING–3',3'-cGAMP (*Fs*STING), oyster STING–2',3'-cGAMP (*C. gigas*), sea anemone STING–2',3'-cGAMP (*N. vectensis* PDB 5CFQ) (Kranzusch et al. 2015), human STING–2',3'-cGAMP (*H. sapiens* PDB 4KSY) (X. Zhang et al. 2013).

a



b



Methods

Phylogenetic distribution of STING-domain proteins in bacterial genomes

CBASS operons and STING proteins were taken from the list identified in (Cohen et al. 2019; Millman et al. 2020) and supplemented by additional bacterial STING homologues identified using the “top IMG homologue hits” function in the Integrated Microbial Genomes (IMG) database (I. M. A. Chen et al. 2019). To assess the prevalence of c-di-GMP utilization in different phyla, Pfam annotation data of all genes in 38,167 bacterial and archaeal genomes were downloaded from the IMG database (I. M. A. Chen et al. 2019) in October 2017. Genomes were supplemented with 44 genomes manually identified with STING-containing operons that were absent from the October 2017 dataset. Genes annotated with the diguanylate cyclase GGDEF domain (pfam00990) or the diguanylate phosphodiesterase EAL domain (pfam00563) were counted for each phylum represented by at least 200 genomes in the database.

Protein expression and purification

Recombinant STING homologues were generated and purified as previously described (Whiteley et al. 2019; Zhou et al. 2018). Briefly, synthetic DNA constructs (Integrated DNA Technologies) were cloned via Gibson assembly into a modified pET16 vector for expression of recombinant N-terminal 6×His-, 6×His-SUMO2, or 6×His-MBP-SUMO2 fusion proteins in BL21-CodonPlus(DE3)-RIL *E. coli* (Agilent). Transformed bacteria were grown overnight in MDG media before inoculation in M9ZB media for large-scale protein expression (2–3× 1 L flasks, grown at 37°C with 230 rpm shaking). Once M9ZB cultures reached OD₆₀₀ ~2.5, flasks were placed on ice for 20 min to slow bacterial growth. Cultures were induced with 500 µM final IPTG concentration and incubated at 16°C for ~20 h at 230 rpm. Cultures were harvested by centrifugation and the pellets were washed once with chilled PBS before flash freezing with liquid nitrogen and storage at –80°C. Functional bacterial TIR-STING constructs were expressed in cultures additionally supplemented with 10–30 mM nicotinamide to limit TIR toxicity.

Bacterial pellets were 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). Clarified lysates were purified by gravity-flow over Ni-NTA resin (Qiagen). Resin was washed with 1× lysis buffer supplemented to 1 M NaCl, and recombinant protein was eluted with 1× lysis buffer supplemented to 300 mM imidazole. Recombinant human SENP2 protease

(D364–L589, M497A) was incubated with purified samples overnight during dialysis at 4°C against dialysis buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 10% glycerol, 1 mM DTT) to cleave the SUMO2 tag. Proteins were concentrated with 30 kDa-cutoff Amicon centrifuge filters (Millipore) before loading onto a 16/600 Superdex 200 size exclusion column equilibrated in gel filtration buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM TCEP). Protein purity was assessed by SDS-PAGE with Coomassie staining before concentrating samples to $>10 \text{ mg ml}^{-1}$. Final proteins samples were flash frozen in liquid nitrogen and stored at -80°C .

Synthetic cyclic dinucleotide standards

Synthetic cyclic dinucleotide ligands used for structural biology and biochemistry experiments were purchased from Biolog Life Science Institute: 3',3'-c-di-AMP (cat no. C 088), 3',3'-c-di-GMP (cat no. C 057), 3',3'-cGAMP (cat no. C 117), 2',3'-cGAMP (cat no. C 161), 3',3'-c-UMP–AMP (cat no. C 357), 2',3'-c-di-GMP (cat no. C 182).

Protein crystallization and structure determination

Crystals for all proteins were initially grown at 18°C using the hanging drop vapor diffusion method. Concentrated protein stocks were thawed from -80°C on ice and diluted in buffer (25 mM HEPES-KOH pH 7.5, 70 mM KCl, 1 mM TCEP) to final concentration. STING–cyclic dinucleotide complexes were formed by incubating protein with ligand on ice for 20 min before setting trays. In all cases, optimized crystals were obtained using EasyXtal 15-Well hanging drop trays (Qiagen) in 2 μl drops mixed 1:1 over a 350 μl reservoir of mother liquor after 1–3 days of growth at 18°C . Final optimized crystal growth conditions were as follows: Crystals of native and selenomethionine-substituted *Fs*STING bound to 3',3'-cGAMP grew at 10 mg ml^{-1} with 0.5 mM 3',3'-cGAMP in 2 M ammonium sulfate, 0.2 M sodium acetate pH 4.5 and were cryoprotected with NVH oil (Hampton). Crystals of selenomethionine-substituted *Cg*STING grew at 18 mg ml^{-1} with 0.5 mM c-di-GMP in 0.1 M HEPES-NaOH pH 7.5, 20% PEG-10,000 and were cryoprotected with mother liquor supplemented with 20% ethylene glycol (*Cg*STING continually precipitated upon incubation with c-di-GMP likely resulting in specific crystallization of the soluble apo form. Crystals of selenomethionine-substituted *Fs*CdnE grew at 7 mg ml^{-1} with 10.5 mM MgCl_2 and 0.5 mM GpCpp in 0.1 M sodium acetate pH 4.6, 2 M

sodium formate and were cryoprotected by supplementing reservoir solution with 25% ethylene glycol and 0.5 mM GpCpp. Crystals of selenomethionine-incorporated CgCdnE grew at 7 mg ml⁻¹ with 10.5 mM MgCl₂ and 0.5 mM GpCpp and grew in 0.2 M sodium thiocyanate, 20% PEG-3350 and were cryoprotected with NVH oil (Hampton). No GpCpp density is visible in either of the FsCdnE or CgCdnE crystal structures. Crystals of apo oyster TIR-STING (*Crassostrea gigas* XP_011430837.1) grew at 7 mg ml⁻¹ in 0.2 M MgCl₂, 0.1 M Tris pH 8.5, and 16% PEG-4000 and were cryoprotected by supplementing mother liquor with 20% glycerol. Crystals of oyster TIR-STING bound to 2',3'-cGAMP grew at 7 mg ml⁻¹ protein with 0.5 mM 2',3'-cGAMP in 0.2 M ammonium citrate pH 7.0, 20% PEG-3350 and were cryoprotected with mother liquor supplemented with 20% ethylene glycol and 0.5 mM 2',3'-cGAMP. X-ray diffraction data were collected with single crystals at the Advanced Photon Source (beamlines 24-ID-C and 24-ID-E) with a wavelength of 0.97918 Å and temperature of 80°K.

Data were processed with XDS and AIMLESS³⁷ using the SSRL autoxds script (A. Gonzalez, Stanford SSRL). Experimental phase information for all proteins was determined using data collected from selenomethionine-substituted crystals. Anomalous sites were identified and an initial map generated with AutoSol within PHENIX 1.17 (ref 38). Iterative model building and refinement was performed using Coot 0.8.9 (ref 39) and PHENIX. Final structures were refined to stereochemistry statistics for Ramachandran plot (favored / allowed), rotamer outliers, and MolProbity score as follows: FsSTING–3',3'-cGAMP, 97.49% / 2.51%, 1.12%, and 1.26; CgSTING apo, 97.23% / 2.77%, 1.90%, and 1.55; *C. gigas* STING apo, 95.68% / 4.32%, 0.72%, and 1.58; *C. gigas* STING–2',3'-cGAMP, 97.25% / 2.75%, 0.67%, and 1.38; FsCdnE apo, 98.86% / 1.14%, 0.30%, and 0.97; CgCdnE, 98.63% / 1.37%, 0.70%, and 1.09. See Supplementary Table 1 and 'Data availability' for deposited PDB codes. All structure figures were generated with PyMOL 2.3.4.

Structure-guided alignment of bacterial and metazoan STING domains

To guide alignment and phylogenetic analysis of all STING family receptors, the bacterial FsSTING, CgSTING, and oyster STING structures were superposed with human STING and representatives from all previously published metazoan STING structures (Ouyang et al. 2012; X. Zhang et al. 2013; Kranzusch et al. 2015; Shang et al. 2019; C. Zhang et al. 2019; Gao, Ascano, Zillinger, et al. 2013; Ergun et al. 2019; Y. H. Huang et al. 2012; Shang et al. 2012; Shu et al. 2012; Yin et al. 2012; Cavlar et al. 2013; H. Zhang et

al. 2015; Cong et al. 2019) using the secondary-structure matching algorithm in Coot (Emsley and Cowtan 2004; Krissinel and Henrick 2004). A sequence alignment was extracted from the superposed structures according to C-alpha position, and extended to include a list of all known STING protein sequences using PROMALS3D (Pei and Grishin 2014). Briefly, the list of bacterial STING protein sequences was prepared from analysis of bacterial defense islands as described above and included 103 sequences. Sequences were aligned with MAFFT (Kato, Rozewicki, and Yamada 2018) and manually trimmed in Jalview (Waterhouse et al. 2009) based on boundaries of the *Fs*STING crystal structure to remove effector domain sequences and obtain an alignment of the STING cyclic dinucleotide binding domain. Metazoan STING sequences were obtained from the Interpro (IPR033952) and PFam (PF15009) databases and trimmed to the cyclic dinucleotide binding domain based on boundaries of the human STING and oyster STING crystal structures resulting in 492 unique sequences. The output PROMALS3D alignment was inspected to ensure accurate secondary structure matching between bacterial and metazoan STING and includes 34 gap-free sites. A STING family phylogenetic tree was then calculated using the MAFFT server and visualized with iTOL (Letunic and Bork 2019) for analysis.

Cyclic dinucleotide synthesis and thin-layer chromatography analysis

CdnE homologues were tested for cyclic dinucleotide synthesis capability using $\alpha^{32}\text{P}$ -labeled NTPs and product resolution with thin-layer chromatography as previously described (Whiteley et al. 2019). In a 20 μl final reaction volume, 5 μM enzyme was incubated with 25 μM of each cold NTP (ATP, GTP, CTP, UTP; 100 μM total) and with trace radiolabeled NTP (~ 1 μCi) as indicated for 3 h at 37°C in buffer containing 50 mM Tris-HCl pH 7.5, 50 mM KCl, 10 mM $\text{Mg}(\text{OAc})_2$, and 1 mM TCEP. Reactions were treated with Calf Intestinal Phosphatase (New England Biolabs) at 37°C for 30 min to remove excess starting nucleotides. 0.5 μl of sample was separated on a PEI-Cellulose F TLC plate (EMD Biosciences) using a running solvent of 1.5 M KH_2PO_4 pH 3.8 for 0.5–1 h. Plates were dried for 1 h at ambient temperature before exposure to a phosphor-screen and imaging with a Typhoon Trio Variable Mode Imager (GE Healthcare). Control cyclic dinucleotides were generated with recombinant purified *Mus musculus* cGAS, *Vibrio cholerae* DncV, *Bacillus thuringiensis* DisA, *Escherichia coli* CdnE, and *Pseudomonas aeruginosa* WspR (with D70E constitutively activating mutation) as previously described (Whiteley et al. 2019; Kulesekara et al. 2006).

cGAS reactions were additionally supplemented with 5 μ M ISD45 dsDNA for enzyme activation (Zhou et al. 2018). *E. coli* CdnE reactions to generate 3',3'-c-UMP-AMP were conducted at pH 9.4 with 50 mM CAPSO.

STING-cyclic dinucleotide complex formation and electrophoretic mobility shift assay

STING interactions with cyclic dinucleotide ligands were monitored by electrophoretic mobility shift assay as previously described⁸. 10 μ l reactions were generated in reaction buffer (5 mM Mg(OAc)₂, 50 mM Tris-HCl pH 7.5, 50 mM KCl, and 1 mM TCEP) with 20 μ M final protein concentration and $\sim$ 1 μ M α ³²P-labeled cyclic dinucleotides ($\sim$ 0.1 μ Ci). Protein titration reactions utilized serial dilutions of stock protein to final concentrations ranging from 0.3 nM to 150 μ M with $\sim$ 250 nM cyclic dinucleotide. Reactions were incubated at 25°C for 5 min prior to resolution on a 7.2 cm 6% nondenaturing polyacrylamide gel run at 100 V for 45 min in 0.5 $\times$ TBE buffer. The gel was fixed for 15 min in a solution of 40% ethanol and 10% acetic acid before drying at 80°C for 1 h. The dried gel was exposed to a phosphor-screen and imaged on a Typhoon Trio Variable Mode Imager (GE Healthcare). Signal intensity was quantified using ImageQuant 5.2 software.

TIR NAD⁺ cleavage activity analysis with fluorescent ϵ -NAD

Plate reader reactions were prepared in a 96-well plate format in 50 μ l final volume with reaction buffer (20 mM HEPES-KOH pH 7.5, 100 mM KCl), 500 μ M ϵ -NAD, 500 nM protein, and between 1 nM – 100 μ M cyclic dinucleotide as indicated. Briefly, a master-mix was prepared on ice containing each ligand and protein was added immediately before beginning analysis. 96-well plates were read continuously over 1 h using a Synergy H1 Hybrid Multi-Mode Reader (BioTek) in fluorescence mode at 410 nm after excitation at 300 nm. Reactions were performed in technical duplicate and data are representative of independent biological replicates. Error bars indicate standard error of the mean for biological replicates. TIR oligomerization dependence experiments were performed in the absence of cyclic dinucleotide using glutathione or nickel-NTA resin at a 1:1 ratio as previously described (Horsefield et al. 2019; Wan et al. 2019) Resin was pre-equilibrated and re-suspended in 100 μ l of buffer containing 500 nM protein, 20 mM HEPES-KOH pH 7.5, 100 mM KCl, and 500 μ M ϵ -NAD, and 50 μ l of mixture was used for analysis.

HPLC TIR NAD⁺ cleavage activity analysis

High-pressure liquid chromatography was used to measure TIR-STING NAD⁺ cleavage activity and to directly observe product formation. Reactions were prepared in 20 mM HEPES-KOH pH 7.5, 100 mM KCl at a final concentration of 500 nM enzyme with or without addition of 10 μ M of cyclic dinucleotides or as indicated. TIR-STING and cyclic dinucleotide ligands were incubated on ice for 10 min prior to addition of NAD⁺ at 500 μ M. Reactions were incubated for 1 h at 25°C, and then heat-inactivated at 95°C for 1 min and incubation on ice for at least 5 min. Reactions were filtered through Millipore Amicon Ultra-0.5 ml filters with a 30 kDa cutoff by centrifugation for 10 min at 9,300 $\times$ g to remove protein before HPLC analysis. Products were separated and analyzed by HPLC with absorbance monitoring at 254 nm. Samples were injected onto a C18 column (Zorbax Bonus-RP 4.6x150 mm, 3.5- μ m) attached to an Agilent 1200 Infinity Series LC system. Two separate elution strategies were utilized: 1) Isocratic elution at 40°C with a flow rate of 1 ml min⁻¹ with 50 mM NaH₂PO₄ pH 6.8 supplemented with 3% acetonitrile. 2) Gradient elution at 50°C with a flow rate of 1 ml min⁻¹ using solvent A (10 mM NH₄OAc) and solvent B (100% methanol), and a gradient from 5–100% solvent B over 12 min.

Reactions to measure discontinuous kinetics were assembled on ice in reaction buffer (20 mM HEPES-KOH pH 7.5, 100 mM KCl) with final concentration of 50 nM protein, 2 μ M c-di-GMP, and indicated concentrations of NAD⁺. Reactions were started simultaneously and quenched after 10, 30, 120, 300, and 600 s. Products were analyzed by HPLC as described above, product ADPr peaks were integrated for each time point, and concentrations were calculated according to a standard curve. Data were analyzed using GraphPad Prism 8.4.2 software and the kinetics data fitted using a Michaelis-Menten model to calculate K_M and V_{max} . Results are representative of two independent biological replicates and plotted with errors bars denoting the standard deviation.

STING toxicity analysis in *E. coli*

SfSTING and mutant constructs were cloned into pBAD33 for arabinose-inducible expression in *E. coli* strain MG1655. Cells were transformed by electroporation, four colonies for each construct were sequence verified and used to inoculate individual LB liquid cultures supplemented with 30 μ g ml⁻¹ chloramphenicol

for 6 h at 37°C with 200 rpm shaking. Cultures were diluted 1:10 into fresh LB medium with 30 $\mu\text{g ml}^{-1}$ chloramphenicol, and 180 μl of diluted cultures were divided into wells in a 96-well plate. Plate cultures were supplemented with 0.2% arabinose and 30 mM nicotinamide as indicated to a final volume of 200 μl per well. OD₆₀₀ was followed using a TECAN Infinite 200 plate reader with measurement every 10 min.

Electron microscopy sample preparation, data collection, and processing

Samples for negative-stain electron microscopy analysis were prepared by diluting purified *Sf*STING (E84A mutation) protein alone, or with an equimolar ratio cyclic dinucleotide as indicated, to a concentration of 50 μM in gel filtration buffer (20 mM HEPES-KOH pH 7.5, 150 mM KCl, 1 mM TCEP). Protein samples were further serially diluted in gel filtration buffer to a final concentration of 0.026 mg ml^{-1} . For protein–ligand mixtures, gel filtration buffer was supplemented with 1 μM of ligand for each serial dilution step. 3 μl of the diluted sample was applied onto a glow-discharged (30 s, 30 mA) 400-mesh copper grid (Electron Microscopy Sciences) coated with an ~10 nm layer of continuous carbon (Safematic CCU-010), followed by a 30 s absorption step and side blotting to remove bulk solution. The grid was immediately stained with 1.5% uranyl formate and then blotted from the side. The staining procedure was repeated twice with a 30 s incubation with uranyl formate before the final blotting step. The grid was air dried prior to imaging.

Electron microscopy Images for 2D classification were collected using an FEI Tecnai T12 transmission electron microscope operating at 120 keV and equipped with a Gatan 4K $\times$ 4K CCD camera at a nominal magnification of 67,000 $\times$, pixel size of 1.68 Å, at a nominal defocus range of 1.0–2.0 μm . Electron microscopy images for *Sf*STING oligomerization mutants (R307E, A309R; L201R, D203R; Δ L275–Q282) were collected using a Philips CM10 transmission electron microscope operating at 100 keV and equipped with a Gatan UltraScan 894 (2k $\times$ 2k) CCD camera at a nominal magnification of 52,000 $\times$, pixel size of 2.06 Å, at a nominal defocus range of 1.5 μm .

Micrographs were converted to JPEG format using e2proc2d.py script (Tang et al. 2007). All image processing was performed in RELION-3⁵⁴. After CTF estimation with GCTF (K. Zhang 2016), particle picking was carried out with gautomatch (Kai Zhang, <https://www.mrc-lmb.cam.ac.uk/kzhang/>) followed by manual review. Where present, fibers were picked separately and treated as independent datasets.

Particles were extracted with a 224 pixel box size (~375 Å) and subjected to one round of 2D reference-free classification to generate 2D class averages.

Size-exclusion chromatography with multi-angle light scattering

Purified protein samples for SEC-MALS analysis were diluted in ice-cold running buffer (150 mM KCl, 20 mM HEPES-KOH pH 7.5, and 1 mM TCEP) to a final concentration of 2 mg ml⁻¹. Samples including ligand were prepared with 100 µM cyclic dinucleotide and incubated on ice for at least 5 min. All samples were subjected to brief centrifugation (21,000 × g, 5 min, 4°C) to remove precipitated protein before injection onto an SRT SEC-300 column (SEPAX). Refractive index (dRI) was measured with a Wyatt Optilab T-rex Refractive Index Detector and protein concentration estimated assuming a dn/dc of 0.185. The system was also equipped with a Wyatt Dawn Heleos II Multi-Angle Light Scattering detector used to determine molar mass. All analysis was carried out using ASTRA 7 software and figures produced with GraphPad Prism 8.4.2.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Acknowledgements

The authors are grateful to J. Morehouse, A. Lee, K. Chat, R. Vance, and members of the Kranzusch lab for helpful comments and discussion, Kelly Arnett and the Harvard University Center for Macromolecular Interactions, the Molecular Electron Microscopy Suite at Harvard Medical School, and the Harvard Center for Mass Spectrometry. The work was funded by the Richard and Susan Smith Family Foundation (P.J.K., S.S.), DFCI-Novartis Drug Discovery Program (P.J.K.), the Parker Institute for Cancer Immunotherapy (P.J.K.), a Cancer Research Institute CLIP Grant (P.J.K.), a V Foundation V Scholar Award (P.J.K.), the Pew Biomedical Scholars program (P.J.K.), Vallee Foundation (S.S.), the Ariane de Rothschild Women Doctoral Program (A.M.), the Israeli Council for Higher Education via the Weizmann Data Science Research Center (A.M.), 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.), the Knell Family Center for Microbiology (R.S.). B.R.M. is supported as a Ruth L. Kirschstein NRSA Postdoctoral Fellow NIH F32GM133063, A.A.G. is supported by a United States National Science Foundation Graduate Research Fellowship, B.L. is supported as a Herchel Smith Graduate Research Fellow, G.O. is supported by a Weizmann Sustainability and Energy Research Initiative (SAERI) doctoral fellowship. X-ray data were collected 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).

Author Contributions

Benjamin R. Morehouse: Designed and conceived experiments, performed structural and biochemical experiments, performed NAD⁺ cleavage assays, performed STING oligomerization assays, and wrote the manuscript.

Apurva A. Govande: Assisted with structural and biochemical experiments and performed NAD⁺ cleavage assays.

Adi Millman: Performed gene identification and phylogenetic analysis.

Alexander F. A. Keszei: Performed electron microscopy experiments and analysis.

Brianna Lowey: Performed STING oligomerization assays

Gal Ofir: Performed STING toxicity analysis

Sichen Shao: Performed electron microscopy experiments and analysis

Rotem Sorek: Performed STING toxicity analysis and designed and conceived experiments

Philip Kranzusch: Designed and conceived experiments, assisted with structural and biochemical experiments, and wrote the manuscript.

Chapter 3: RNase L activating 2'–5' oligoadenlyates bind Decr-1

Apurva A. Govande^{1, 2}, Philip J. Kranzusch^{1,2,3}

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

Abstract

Oligoadenylate synthase (OAS) is an interferon stimulated cytoplasmic sensor of double-stranded RNA (dsRNA) that upon activation, synthesizes a linear second messenger molecule 2'–5' oligoadenylate (2'–5' OA), composed of adenine bases linked through a non-canonical 2' OH to 5' phosphate linkage. 2'–5' OA binds to and activates RNase L, a latent endoribonuclease present in the cytoplasm that oligomerizes and cleaves single stranded RNA (ssRNA), including mRNA and ribosomal RNA (rRNA), leading to translational arrest and eventual apoptosis. RNase L is the only characterized downstream receptor of 2'–5' OA and is activated by 2'–5' OA greater than three linked adenine bases in length. Active OAS genes are present in organisms as old as sponges, while RNase L is present beginning in jawed vertebrates. There is no structural or functional conservation of RNase L activity in more ancient organisms, suggesting that alternate receptors of 2'–5' OA evolved in organisms lacking RNase L that may be conserved. Using 2'–5' OA conjugated beads and mass spectrometry, we identify and determine the structure of a novel binding partner of 2'–5' OA, Decr1, a metabolic enzyme previously uncharacterized as a receptor of 2'–5' OA that is well-conserved in ancient organisms.

Introduction

The OAS-RNase L system is a robust, interferon induced antiviral pathway in cells that is a sensor for dsRNA and functionally active against several RNA viruses (Y. Li et al. 2016; Rice, Roberts, and Kerr 1984). Oligoadenylate synthase (OAS) proteins are cytosolic double-stranded (dsRNA) sensors that detect viral infection and enzymatically synthesize the nucleotide second messenger 2'–5' oligoadenylate (2'–5' OA)(Hovanessian, Brown, and Kerr 1977; Kerr, Brown, and Hovanessian 1977; Kerr and Brown 1978; Donovan et al. 2015; Donovan, Dufner, and Korennykh 2013). Humans encode three catalytically active OAS enzymes, OAS1, OAS2, and OAS3, that vary in size and subcellular localization (Chebaths et al. 1987) and OASL, a catalytically inactive homolog that has been implicated in regulating RIG-I activity (Hartmann et al. 1998; J. Zhu et al. 2014; Ibsen et al. 2015). Immune crosstalk between RIG-I and the OAS1-RNase L axis is required for apoptosis after RIG-I activation (Boehmer et al. 2021).

In human cells, OAS1, OAS2, and OAS3 are required for antiviral innate immunity against a wide range of viruses with OAS3 particularly important for RNase L activation (A. Ghosh, Sarkar, and Sen n.d.; Y. Li et al. 2016; Kwon et al. 2013; RJ et al. 2009). The OAS family of proteins have undergone duplication and diversification events that result in greater functional specificity beyond RNase L activation, such as the importance of ER membrane association for catalytically inactive OAS1b in mice and its co-localization with ABCF3 (Peterson et al. 2019), and similarly, a recent publication exploring the localization of human OAS1 p46 to selective membrane compartments associated with the ER necessary for its antiviral activity (Soveg et al. 2021), both highlighting the diversity of localization and function within the OAS family.

Following dsRNA sensing, OAS proteins undergo a conformation change resulting in catalytic activation and synthesis of linear 2'–5' OA signaling molecules composed of adenine nucleotides linked through noncanonical 2'–5' phosphodiester bonds (Lohöfener et al. 2015). The product 2'–5' OA molecules vary in size between 3- and 30-nucleotides (Morin et al. 2010) with chains between three to five bases primarily responsible for activating the receptor RNase L in the cytosol of cells (Kerr and Brown 1978; Silverman et al. 1988). OAS1 and OAS2, and OAS3 have differential subcellular localization and produce 2'–5' OA chains of differing lengths. Once bound to 2'–5' OA, activated RNase L then promiscuously cleaves single stranded RNA resulting in degradation of viral and host mRNA transcripts and cellular ribosomal RNA. RNase L activation induces translational arrest and apoptosis leading to broad antiviral defense (G.

Li et al. 2004, 2003; Manivannan, Siddiqui, and Malathi 2020). RNase L is not activated by 2'–5' OA shorter than three bases in length (Morin et al. 2010), and there is no defined role of 2'–5' OA for chains longer than 5 bases.

OAS proteins are evolutionarily ancient and widely conserved in metazoans. In contrast, RNase L is limited to jawed vertebrates and is surprisingly absent in invertebrates and primitive metazoan lineages (Kuusksalu et al. 1995) and no ancient functional homologs of RNase L have been characterized. In mice, OAS proteins have undergone recent duplication events, leading to an expanded panel of OAS functions and OAS-like proteins (Perelygin et al. 2002; Scherbik et al. 2007; Peterson et al. 2019; Madden et al. 2019), including OAS2, OAS3, OASL, and eight paralogs of OAS1 (Elkhateeb et al. 2016). Evidence of evolutionary pressure in the OAS family, such as duplication and diversification events (Elkhateeb et al. 2016), positive selection at key nucleic acid binding regions (Hancks et al. 2015) as well as the absence of RNase L in primitive organisms suggests a wider role for 2'–5' OA signaling independent of RNase L, and highlights its cellular importance and its ongoing relevance (Reintamm et al. 2008; Kuusksalu et al. 1995). Here, we developed an unbiased approach to identify RNase L-independent roles of 2'–5' OA and discovered alternative 2'–5' OA binding partners by affinity purification and high-resolution mass spectrometry. Our results identify new 2'–5' OA binding partners and suggest auxiliary roles for OAS-signaling pathways.

Results and Discussion

Identification of Decr1, a novel binding partner of 2'–5' OA

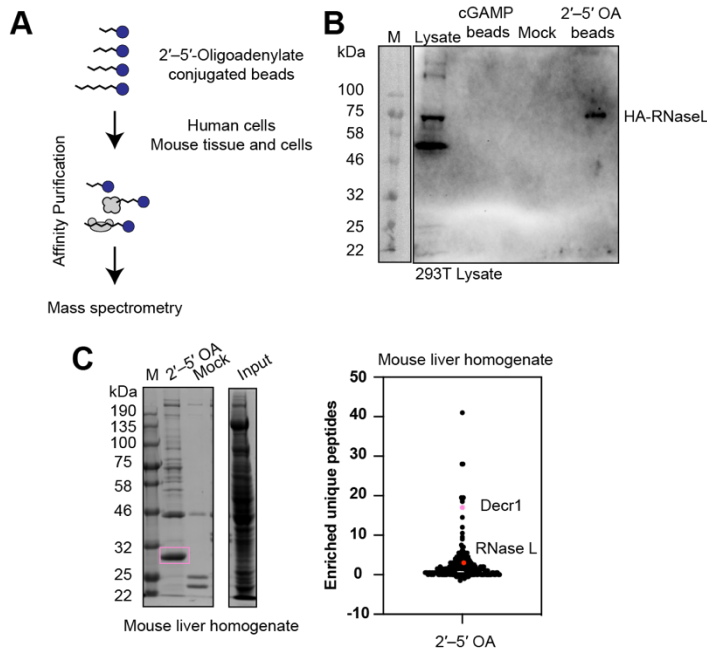


Figure 3.1. Decr1 is a novel binding partner of 2'–5' OA.

A) Schematic of 2'–5' OA affinity purification from cell lysates and tissue homogenates. Sepharose beads with an extended linker were coupled with enzymatically synthesized 2'–5' OA and used to purify novel binding partners from human cell lysates and homogenized mouse tissue. Each affinity purification was performed three times with each type of bead. B) Affinity purification of overexpressed, HA-tagged human RNase L from 293T cell lysate with 2'–5' OA beads, cGAMP beads, or mock control beads. RNase L specifically binds to 2'–5' OA beads and not cGAMP beads or control beads. Data is representative of two independent replicates. C) Coomassie stained gel of affinity purification using 2'–5' OA beads or mock beads from homogenized mouse liver, left. Decr1 band is highlighted in pink. Number of enriched unique peptides identified using mass spectrometry, right. Enriched unique peptides were determined by subtracting total unique peptides from mock control purification from total unique peptides from 2'–5' OA affinity purification for each identified protein. Decr1 and RNase L are highlighted by colored dots. Mass spectrometry results are representative of two independent experiments, including affinity purification and mass spectrometry data.

To identify novel binding partners of 2'–5' OA, we performed unbiased affinity purification from homogenized mouse liver, lung, and heart lysates with enzymatically synthesized 2'–5' OA covalently conjugated to Sepharose beads with a 12-carbon linker (Figure 3.1a) (McFarland et al. 2017; Poulsen et al. 2015; Morin et al. 2010) (see methods for additional detail). The extended linker allows for flexibility in 2'–5' OA conformation and orientation. To validate

our 2'–5' OA-conjugated Sepharose beads, we performed an affinity purification with 293T lysate and confirmed the specific 2'–5' OA pull-down of an overexpressed HA-tagged RNase L construct, enriched over cyclic GMP-AMP (cGAMP) beads or control beads (Figure 3.1b). We identified 2,4-Dienoyl-CoA reductase 1 (Decr1) as the highly enriched protein in 2'–5' OA pull-downs from mouse liver homogenate (Figure 3.1c).

Decr1 is a mitochondrially associated metabolic enzyme necessary for β -oxidation

and metabolism of polyunsaturated fatty acid CoA esters containing trans double bonds in both even- and odd-numbered positions, and requires NADPH cofactor binding for substrate processing (Helander et al.

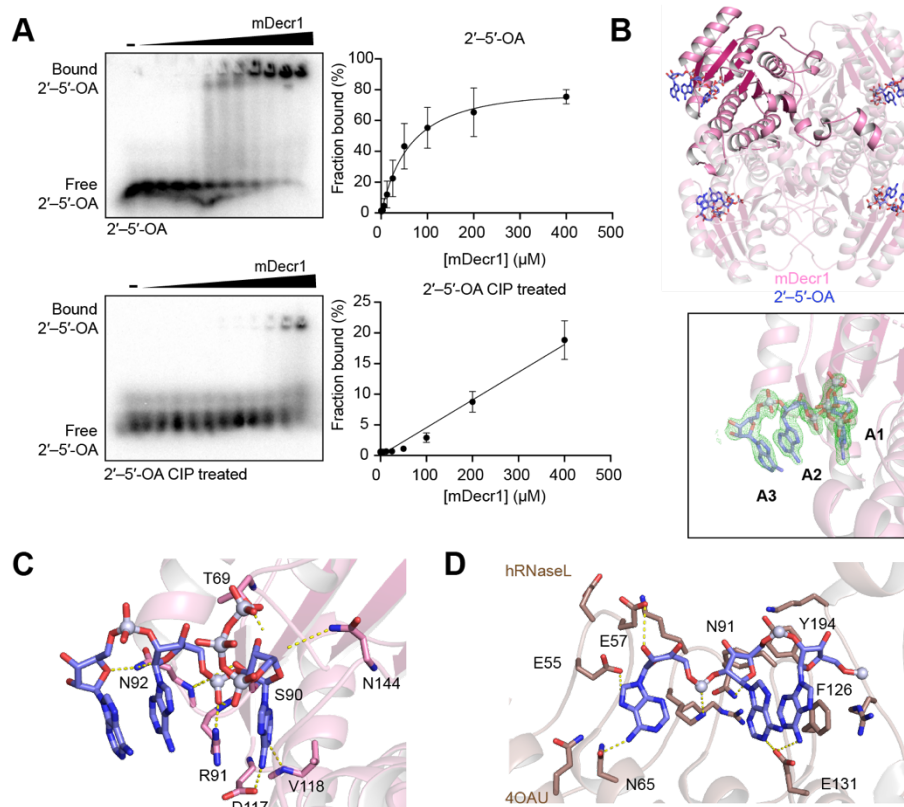


Figure 3.2 Mouse Decr1 binding to 2'-5' OA is dependent on the 5' triphosphate. A) Electrophoretic mobility shift assays (EMSA) with mouse Decr1 and radiolabeled, enzymatically synthesized 2'-5' OA with and without the 5' triphosphate. The mDecr1-2'-5' OA complex travels slowly through the gel and separates from unbound, free 2'-5' OA. Quantification of the bound fraction over total 2'-5' OA signal indicates Decr1 binding to 2'-5' OA depends on the 5' triphosphate, and binding is reduced when the 2'-5' OA is treated with phosphatase (CIP). Data is representative of three independent experiments; error bars indicate standard deviation. B) 1.35 Å structure of mouse Decr1 bound to four molecules of 2'-5' OA with clear density for three bases. Decr1 is a homotetrameric protein with four available binding sites, occupied by 2'-5' OA in our structure. One monomeric unit is highlighted in dark pink. C) 2'-5' OA is coordinated by sidechain residues and specific protein-RNA contacts that facilitate binding. The A1 base of 2'-5' OA is coordinated by D117 and the backbone of V118. Specific contacts to the 2' phosphate by R91 explain the specificity of the 2'-5' linkage recognition, and 2'-5' OA adopts a loosely helical conformation overall enabled by stacking interactions between the three bases and R91. D) Three bases of 2'-5' OA bound to human RNase L (PDB: 4OAU). The third base is held away from the first two and the conformation of 2'-5' OA is distorted from helical base stacking, in contrast to Decr1.

1997). We purified recombinant mouse Decr1 (mDecr1) and measured interaction *in vitro* using an electrophoretic mobility shift assay. RNase L is known to bind 2'-5' OA with ~10 nM affinity, and high affinity RNase L-2'-5' OA interaction requires specific 5' triphosphate recognition (Dong and Silverman 1995; Han et al. 2012). We observed that mDecr1 directly binds radiolabeled 2'-5' OA *in vitro* and forms a stable protein-RNA complex with ~50–60 μM affinity (Figure 3.2a), and that mDecr1 interactions with 2'-5' OA are also diminished in the absence of a 5' triphosphate (Figure 3.2a). Together these results confirm mDecr1 as a specific 2'-5' OA-binding protein.

Molecular basis of Decr1 recognition of 2'–5' OA

To explain the molecular basis of mDecr1 2'–5' OA recognition, we determined a 1.4 Å crystal structure of the mDecr1–2'–5' OA complex (Figure 3.2b; Table 3.1). mDecr1–2'–5' OA crystals were grown in the presence of enzymatically synthesized 2'–5' OA using human OAS1 reactions that produce chains between 3–10 nucleotides (Ibsen et al. 2014; Rebouillat et al. n.d.; Morin et al. 2010). Similar to previous structures of human Decr1 (hDecr1) (Alphey et al. 2005), mDecr1 forms a homotetrameric enzyme with four outward facing active sites. Surprisingly, 2'–5' OA binds directly in the mDecr1 catalytic active site with clear electron density for 2'–5' OA observed in each protomer (Figure 3.2b). mDecr1 residues T69 and N144 coordinate the 5' end of 2'–5' OA making direct contact with the 5' triphosphate and the ribose of the A1 nucleotide (Figure 3.2c). mDecr1–2'–5' OA contacts additionally occur between D117 and the A1 adenine N1 and N6 positions. mDecr1 residues S90 and R91 coordinate the first 2'–5' phosphodiester bond and explain specific interaction with 2'–5'-linked nucleic acid (Figure 3.2c), while residue N92 contacts both the A2 and A3 adenine nucleobases. Three bases are ordered and appear with clear electron density (Figure 3.2c). The remaining bases of the 2'–5' OA chain extend away from the active-site, demonstrating that like RNase L, mDecr1 recognizes the first three bases of 2'–5' OA chains (Morin et al. 2010).

Comparing 2'–5' OA binding and contacts in our structure with the structure of 2'–5' OA bound to human RNase L (Han et al. 2014) shows 2'–5' OA bound to mDecr1 exhibits a loose helical conformation while 2'–5' OA bound to RNase L is more distorted (Figure 3.2d). 2'–5' OA binds RNase L at the dimer interface made up by two RNase L monomers and upon binding, RNase L forms higher order oligomers (Han et al. 2012; H. Huang et al. 2014). RNase L charged side chain E131 contacts the A1 N6 position and the N1 position of A2, while the N6 and C8 positions of A3 are coordinated by N65 and E57 (Figure 3.2d), and in contrast to mDecr1, both the first and second phosphate linkage make hydrogen bond contacts with RNase L sidechain residues. The extensive coordination of 2'–5' OA by RNase L correlates with its high affinity and canonical activation as the downstream receptor of the OAS innate immune axis (Kerr and Brown 1978).

Table 3.1. Summary of Crystallographic Statistics

<i>Mus musculus</i> Decr1	
Data Collection	
Resolution (Å) ^a	1.35
Wavelength (Å)	0.97918
Space group	C 2 2 2 ₁
Unit cell: a, b, c (Å)	82.38 117.19 232.25
Unit cell: α, β, γ (°)	90.00 90.00 90.00
Molecules per ASU	4
Total reflections	1669847
Unique reflections	244100
Completeness (%) ^a	99.8 (97.6)
Multiplicity ^a	6.8 (5.7)
<i>I</i> / σ ^a	10.6 (1.2)
CC(1/2) ^b (%) ^a	99.8 (97.6)
R _{pim} ^c (%) ^a	3.1 (55.9)
Sites	
Refinement	
Resolution (Å)	67.39 – 1.35
Free reflections	2000
R-factor / R-free	15.1/16.1
Bond distance (RMS Å)	0.008
Bond angles (RMS °)	1.06
Structure/Stereochemistry	
No. atoms: protein	18412
No. atoms: ligand	
No. atoms: water	891
Average B-factor: protein	26.6
Average B-factor: ligand	19.8
Average B-factor: water	36.4
Ramachandran plot: favored	97.53%
Ramachandran plot: allowed	2.47%
Ramachandran plot: outliers	0.00%
Rotamer outliers	1.2%
MolProbity ^d score	1.42
Protein Data Bank ID	

^a Highest resolution shell values in parentheses^b (Karplus and Diederichs 2012)^c (Weiss 2001)^d (V. B. Chen et al. 2010)

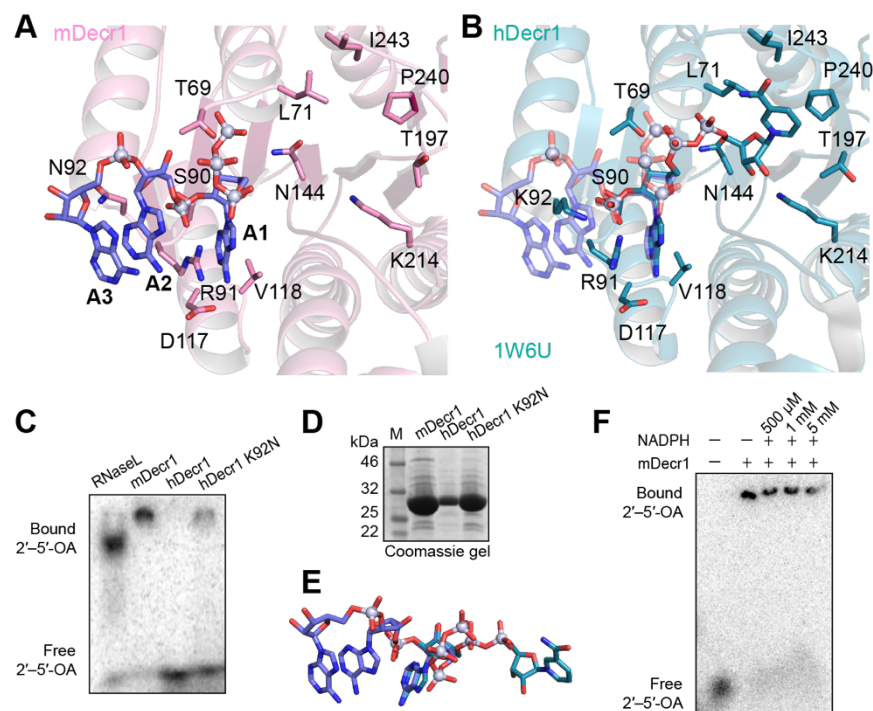


Figure 3.3. 2'-5' OA binding obstructs co-factor NADPH binding pocket in Decr1 in a species-specific manner. A, B) Comparison of mouse Decr1 bound to 2'-5' OA, A, and human Decr1 bound to NADPH (PDB: 1W6U), B. Mouse Decr1 and human Decr1 are 85% identical in primary sequence and side chain residues in the substrate and NADPH binding site are well conserved. 2'-5' OA binding extends out of the NADPH binding pocket, while NADPH bound to the human Decr1 structure is coordinated deeper in the binding pocket. Overlaying 2'-5' OA with the human structure reveals a steric clash between lysine 92 and 2'-5' OA. Mouse Decr1 encodes an asparagine at that position. C) EMSA with radiolabeled 2'-5' OA and mouse Decr1, human Decr1, or human mutant K92N. The K92N mutation is capable of partially restoring human Decr1 binding to 2'-5' OA. D) Coomassie stained protein expression gel of mouse Decr1, human Decr1 and human Decr1 mutant K92N. E) Overlay of 2'-5' OA and NADPH as they appear in either mouse or human Decr1-bound structures. 2'-5' OA and NADPH share a 2' phosphate structural motif. The first adenine base and first 2'-5' linkage of 2'-5' OA overlaps exactly with the adenine base and 2' phosphate of NADPH. F) Pre-incubation with high concentrations of NADPH does not appreciably prevent binding to trace radiolabeled 1 μ M 2'-5' OA, indicating that 2'-5' OA is capable of inhibiting NADPH binding in the absence of substrate.

2'–5' OA prevents Decr1 from binding co-factor NADPH

We compared the mDecr1–2'–5' OA structure with the structure of hDecr1 in complex with NADPH and hexanoyl-CoA (Alphey et al. 2005). mDecr1 and hDecr1 share 85% identity at the amino acid level and residues in the catalytic pocket are highly conserved (Figure 3.3a, b). hDecr1 residues S90, R91 and D117 coordinate the NADPH adenine base and 2' phosphate making near identical contacts as those observed between mDecr1 and 2'–5' OA (Figure 3.3b). Despite the high level of conservation between hDecr1 and mDecr1, overlay of the two bound structures reveals a steric clash between hDecr1 residue K92 and the 2'–5' OA A2 nucleobase (Figure 3.3b). In mDecr1, an Asn is encoded at position 92 that accommodates the extended nucleotide chain of 2'–5' OA and enables binding (Figure 3.3a, b). We tested interaction between hDecr1 and 2'–5' OA *in vitro* and confirmed that the wildtype protein is unable to form a complex with 2'–5' OA (Figure 3.3c), and this specific lack of binding is not due to defects in stability of the human Decr1 constructs (Figure 3.3d). An hDecr1 K92N substitution permits hDecr1 to weakly bind 2'–5' OA, confirming a key role for this Decr1 residue in controlling species-specific interactions with 2'–5' OA (Figure 3.3c). Notably, the chemical structures of NADPH and 2'–5' OA contain identical 2' phosphate and adenine base moieties, explaining how contacts in the Decr1 catalytic active site coordinate similar interactions with both NADPH and 2'–5' OA (Figure 3.3e). Decr1 affinity for co-factor NADPH is ~3.5 μ M (Fillgrove and Anderson n.d.), but strikingly, pre-incubation of mDecr1 with high concentrations of NADPH does not prevent radiolabeled 2'–5' OA binding (Figure 3.3f). Together, our results identify Decr1 as a species-specific binding partner of 2'–5' OA and suggest a role as an auxiliary second-level arm of the innate immune axis dependent on 2'–5' OA binding.

Conclusions

Our results define mouse Decr1 as novel 2'–5' OA-binding protein and suggest auxiliary functions for OAS proteins and 2'–5' OA signaling in animal cells (Figure 3.1). The micromolar binding affinity between 2'–5' OA and mouse Decr1, compared to the nanomolar binding affinity between RNase L and 2'–5' OA (Han et al. 2012) may reflect its role as a secondary arm of 2'–5' OA binding (Figure 3.2) and a hierarchy of signaling outcomes beginning top down with RNase L. Some viruses encode inhibitors of RNase L that impair either RNase L function or binding to 2'–5' OA (Drappier and Michiels 2015) while other viruses upregulate the cellular inhibitor of RNase L, RLI, during infection (C Bisbal et al. 1995; Martinand et al. 1998). In the event of RNase L inhibition, this secondary arm of 2'–5' OA effects may become more pronounced and required for an alternative means of cellular metabolic shutdown.

Our results suggest that 2'–5' OA blocks NADPH binding to Decr1 in the absence of substrate and could conceivably prevent its enzymatic activity. Knock out or inhibition of Decr1 leads to an accumulation of poly-unsaturated fatty acids (PUFAs), ER and mitochondrial stress and eventual ferroptosis (Blomme et al. 2020; Nassar et al. 2020), and could provide an alternative means of cell death and metabolic shutdown during viral infection. The accumulation of lipids and PUFAs is detrimental to enveloped viruses and may hinder RNA viral replication in membrane-bound compartments (Leu, Lin, and Hsu 2004; Heaton and Randall 2011), and these cumulative antiviral effects could make up for a defective RNase L response. Decr1 is well-conserved across more ancient organisms and lower metazoans. It has not escaped our notice that the effects of this secondary arm of OAS-dependent 2'–5' OA signaling may be indicative of more ancient pathways in organisms that do not encode RNase L, and more modern organisms may have lost these auxiliary features. As such, exploring alternative roles of 2'–5' OA in invertebrates or lower metazoans may provide further insight into RNase L independent roles of 2'–5' OA.

Our mass spectrometry results identified several other proteins that bind to or are dependent on NADPH. NADPH is a wide-spread enzymatic co-factor and regulates several different processes in the cell (VanLinden, Skoge, and Ziegler 2015; Pollak, Niere, and Ziegler 2007; Ju et al. 2020; Breitenbach et al. 2018) The shared 2' phosphate between NADPH and 2'–5' OA is an important feature of Decr1-2'–5' OA binding. NADPH-dependent enzymes often rely on contacts with the 2' phosphate to distinguish between NAD⁺ and NADP⁺, suggesting that this is an important chemical motif that is specifically coordinated by

NADPH-dependent enzymes. An alternative, low-level inhibition of NADPH-binding enzymes could be a tertiary function of 2'–5' OA signaling. However, widespread binding to a pool of NADPH-binding enzymes and other enzymes that bind a similar chemical motif could also limit the extent of RNase L recognition of 2'–5' OA and regulate its activity.

Furthermore, our results with mouse Decr1 indicate species-specific binding to 2'–5' OA and the ongoing evolution of the system (Figure 3.3). Mice contain eight paralogs of OAS1, a-h, but only OAS1a and OAS1g are catalytically active (Perelygin et al. 2002; Kakuta, Shibata, and Iwakura 2004). The recent duplication and diversification of function within the OAS family of proteins in mice illustrate that the OAS innate immune axis is a mutable and rapidly evolving antiviral pathway. The role of alternative binding partners in a species-specific manner reflects a continuing pressure to adapt to diverse viral threats and suggests the parallel evolution of novel OAS functions and novel binding partners in a species-specific manner. Altogether, our identification and validation of Decr1 as an alternate binding partner of 2'–5' OA emphasizes the possibility of diverse antiviral pathways controlled by a diffusible second messenger and highlights areas for future focus.

Methods

Enzymatic synthesis of 2'–5' OA

Large-scale enzymatic synthesis of 2'–5' OA was carried out using final concentration 100 nM recombinant human OAS1 and 1.25 mM ATP in reaction buffer containing 10 mM Tris-HCl at pH 7.5, 25 mM NaCl, 11 mM MgCl₂, 1 mM DTT, and 250 nM HMW-Poly I:C in a final volume of 500 ml and incubated at 37°C for ~17 h. The next day, the reaction was filtered through a .22-micron filter and purified using anion ion exchange chromatography (IEX) and eluted with 2 M NH₄CH₃CO₂ by FPLC. Purified and eluted reactions were pooled and lyophilized to powder over 48 h.

Generation of 2'–5' OA conjugated beads

Enzymatically synthesized and lyophilized 2'–5' OA was conjugated to GE Epoxy-Activated Sepharose 6B (GE Healthcare, Fisher Scientific 45-000-085, 15 g) by incubating ~100 mg lyophilized 2'–5' OA resuspended in 1 ml of 1×PBS pH 7.5 with washed and prepared Epoxy resin. 0.5 g resin was prepared by loading into a gravity flow column (2 cm diameter, 14 cm length (~40 ml volume)) and washing with 5× column volumes (CV) of 0.22-micron filtered sterile water to rehydrate resin, followed by 1×CV of 1×PBS pH 7.5. Liquid was allowed to drain out by gravity flow without drying out beads, and 1 ml of resuspended 2'–5' OA was allowed to flow through once. The flow-through was collected and reapplied to the resin, columns were capped and sealed with parafilm and incubated in a 55°C water bath for 48 h. After 48 h of coupling, liquid was allowed to drain out of the column and a 10 ml syringe was used to remove remaining liquid without over-drying resin. To verify coupling of resin and 2'–5' OA, a 1:1 slurry of beads and 1×PBS pH 7.5 was prepared and absorbance at 260 nm was measured prior to blocking resin using 5 ml of 50 mM Ethanolamine pH 8.0, ensuring resin was completely covered, and incubation at 45°C for ~4 h.

After Ethanolamine blocking, ethanolamine was allowed to drain out and remaining liquid was removed using a 10 ml syringe without over-drying resin. Column was washed with 2×CV of filtered sterile water followed by 1×CV of 1×PBS pH 7.5. Beads were removed from column and stored as a 50% slurry in 1× PBS pH 7.5 supplemented with 0.02% (v/v) sodium azide. Mock control beads were prepared identically and concurrently, substituting 1×PBS pH 7.5 for 2'–5' OA.

Affinity purification and mass spectrometry

5 CBL6 mice (Jackson labs), 2 injected with PBS and 3 injected with 200 µg vaccine-grade HMW Poly I:C (Invivogen) were euthanized ~24 h after injection. Livers, hearts and lungs from control mice and treated mice were removed and rinsed with 1×PBS pH 7.5 before snap freezing and weighing tissue. Frozen tissues were homogenized separately using a mortar and pestle. Tissue homogenate was weighed and resuspended in 3× w/v PBS pH 7.5, rocked on a nutator at 4°C for 30 m, pulsed with a 10 ml syringe and needle, rocked again on a nutator at 4°C for 10 m, and spun down at 20,000 RPM for 30 m at 4°C in oak ridge tubes. Clarified lysate was filtered through a 0.22-micron filter using a syringe, and subsequently incubated with 20 µl of washed (3× with 1×PBS pH 7.5) 2'–5' OA 50% resin slurry or control slurry (10 µl of packed beads). Lysates and resin were incubated at 4°C for 2 h while rocking.

After incubation, resin was washed 2× with 1×PBS pH 7.5 followed by 1× with 1 M NaCl and 1× with 1 mM ATP followed by 3× with 1×PBS pH 7.5. Washed resin was incubated with SDS-PAGE loading buffer and heated at 95°C for 5 min prior to loading onto a 10% polyacrylamide gel. The entire lane was analyzed by mass spectrometry. In-gel mass spectrometry sample quantification was performed by the Harvard Medical School Taplin Mass Spectrometry facility. Unique peptides from mouse tissue affinity purification were compared against the mouse genome (*Mus musculus*) and common contaminants were removed from the final data set.

RNase L 293T cell expression and Western Blot

HEK293T cells were transfected with a pcDNA4-HA-human RNase L construct using PEI. The next day, cells were rinsed with 1× HBBS, lysed in cell lysis buffer (150 mM KCl, 0.5% NP-40, 50 mM HEPES KOH pH 7.5, 1 mM DTT, 1 mM leupeptin, 1 mM pepstatin) and collected in a 1.5 ml tube. Cell lysate was spun at 10,000 RPM for 5 min at 4°C. Clarified lysate was incubated with 20 µl of washed 50% 2'–5' OA resin slurry (10 µl of packed resin) and rotated at 4°C for ~2 h. Next, resin was washed 5× with 1×PBS pH 7.5, 25 µl of SDS-PAGE loading dye was added directly to resin and samples were heated at 95°C for 5 min and resolved on a 10 % polyacrylamide denaturing gel. Protein was transferred onto a PVDF membrane and blocked in a 5% Milk solution in 1×TBS-T. Membrane was incubated with primary antibody (Mouse anti-HA. 1:20,000 dilution, 5mL of 1% milk in TBST) overnight, washed 3× with 1×TBS-T, and incubated

with 1:50,000 anti-mouse HRP secondary in 1% Milk TBST for 1 hour prior to imaging using Biorad Chemidoc.

Protein expression and purification

Recombinant proteins (human OAS1, human Decr1, human RNase L, mouse Decr1) were cloned from synthetic DNA (IDT) into a custom pET vector containing a 6×His-SUMO2 solubility tag. The tagged protein was expressed in *E. coli* strain BL21-RIL (Agilent) containing the rare tRNA plasmid pRARE2 as previously described (Lowey et al. 2020; Zhou et al. 2018). Transformed colonies were grown in a 30 mL MDG media starter culture (0.5% glucose, 25 mM Na₂PO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 0.25% aspartic acid, 100 mg ml⁻¹ ampicillin, 34 mg ml⁻¹ chloramphenicol, and trace metals) overnight at 37°C, and used to inoculate 1 L M9ZB media cultures (0.5% glycerol, 1% CAS amino acids, 47.8 mM Na₂PO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 85.6 mM NaCl, 2 mM MgSO₄, 100 mg ml⁻¹ ampicillin, 34 mg ml⁻¹ chloramphenicol, and trace metals) at OD₆₀₀ ~0.05. Cultures were grown at 37°C, 230 RPM until OD₆₀₀ ~2.5, chilled on ice for 10 min, induced with 0.5 mM IPTG, and incubated at 16°C for shaking at 230 RPM for ~16 h before harvest.

Cultures were pelleted, lysed and sonicated in 1× Lysis Buffer containing 20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 30 mM imidazole, and 10% glycerol. Protein was purified from sonicated lysates using Ni-NTA resin (Qiagen) using gravity flow. Ni-NTA resin was washed with 1× Lysis Buffer supplemented to 1 M NaCl and eluted with 1× Lysis Buffer supplemented to a final concentration of 300 mM imidazole. For crystallography, 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 containing 20 mM HEPES-KOH pH 7.5, 250 mM KCl, and 1 mM DTT. The next morning, protein was concentrated using a 30K-cutoff concentrator (Millipore) and purified by size exclusion chromatography (SEC) using a 16/600 Superdex 200 column with Gel Filtration Buffer containing 20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM TCEP. Purified SEC fractions were pooled and concentrated to >30 mg/ml, aliquoted in 50 µl, flash frozen in liquid nitrogen, and stored at -80°C. Protein for biochemical assays was purified similarly except the human SENP2 protease step was omitted, resulting in tagged protein.

Crystallography and structure determination

Recombinant purified mouse Decr1 was crystallized in hanging-drop format at 18°C using 0.4 μ l drops in 96 well trays or 2 μ l drops in 15-well Easy-Xtal trays (Qiagen). Mouse Decr1 was diluted to 10 mg/ml in 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP and supplemented with final concentration $\sim$ 2 mM enzymatically synthesized 2'–5' OA resuspended in PBS in each drop, mixed with reservoir solution (34% PEG-300, 100 mM Tris-HCl pH 8.6) and allowed to grow overnight. Crystals were harvested using reservoir solution supplemented with 10% Ethylene Glycol and $\sim$ 2 mM 2'–5' OA.

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 beamline 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). Data were processed using PHENIX (Liebschner et al. 2019) and the mouse Decr1 structure was determined using molecular replacement with the human Decr1 structure as a model (PDB: 1W6U). Model building was performed in Coot (Emsley and Cowtan 2004), and refinement was performed in PHENIX.

Electrophoretic mobility shift assays

To validate mouse Decr1 complex formation and binding to 2'–5' OA, recombinant protein and enzymatically synthesized, radiolabeled 2'–5' OA were incubated *in vitro*. 10 μ l reactions in binding buffer containing 10 mM Tris-HCl pH 7.5, 50 mM KCl, and 1 mM DTT and 1 μ M $\alpha^{32}\text{P}$ -radiolabeled 2'–5' OA and either 400 μ M mouse Decr1 or human Decr1, or serial dilutions (400 nM, 700 nM, 1 μ M, 3 μ M, 6.25 μ M, 12.5 μ M, 25 μ M, 50 μ M, 100 μ M, 200 μ M, 400 μ M) of mouse Decr1. Reactions were incubated at 25°C for 30 min prior to resolution on a 7.2 cm, 6% nondenaturing polyacrylamide gel run at 100 V for 30 min in 0.5 $\times$ TBE buffer. The gel was dried at 80°C for 1 h prior to exposure to a phosphor-screen for $\sim$ 5 hours to overnight and imaged using a Typhoon Trio Variable Mode Imager (GE Healthcare).

Quantification of mouse Decr1-2'–5' OA complex formation was carried out using ImageQuant TL v8.2.0.0 software. Pixel intensity for each mouse Decr1-2'–5' OA complex and total pixel intensity for each

lane was calculated, and percent binding was determined by dividing pixel intensity for each mouse Decr1-2'-5' OA complex by total pixel intensity for each lane. Data were fit using a non-linear regression curve for normalized one-site binding in GraphPad Prism version 9.0.0.

For NADPH preincubation experiments, 400 μ M mouse Decr1 was pre-incubated with indicated concentrations of NADPH for ~10 min prior to addition of radiolabeled 2'-5' OA and incubation of the total mixture for ~30 min at room temperature prior to resolving on a 6% nondenaturing polyacrylamide gel and all additional steps as detailed above.

Acknowledgements

We thank members of the Kranzusch lab for helpful comments and discussion, R. Tomaino and the Harvard Medical School Taplin Mass Spectrometry Facility for data collection and analysis, K. Wucherpennig for assistance with mouse protocols, advice, and discussion, and L. Ferrari De Andrade for assistance with mouse work. The work was funded by the Richard and Susan Smith Family Foundation (P.J.K., S.S.) and the Parker Institute for Cancer Immunotherapy (P.J.K.). A.A.G. is supported by a United States National Science Foundation Graduate Research Fellowship.

Author Contributions

Apurva Govande: Conceived and designed experiments, wrote Decr1 chapter

Philip J Kranzusch: Conceived and designed experiments, wrote Decr1 chapter

Discussion

Bacterial CD-NTase evolution and product formation

The 3',3'-cGAMP synthase DncV, one of the founding members of the CD-NTase family of enzymes, is a relatively recent acquisition in the *Vibrio cholerae* 7th pandemic strain El Tor as part of two major genomic changes from previous strains — the acquisition of *Vibrio* 7th pandemic island-1 and -2 (VSP-1, VSP-2) surrounded by transposable genetic elements. Davies et al identified that VSP-1 was responsive to the ToxT master regulator, and isolated DncV as the driver of chemotactic virulence dependent on its synthesis of the hybrid cyclic dinucleotide, 3',3'-cGAMP (Davies et al. 2012). Davies et al also pointed out that several other bacteria encoded genes predicted to be DncV homologs surrounded by integrative and conjugative elements, proposing that they could enable recombining into new genomes.

The discovery and characterization of the CD-NTase family of genes (Whiteley et al. 2019), the discovery that DncV and 3',3'-cGAMP regulated the phospholipase gene directly upstream in VSP-1 (Severin et al. 2018), and the further characterization of the CD-NTase containing operons as part of the repertoire of anti-phage genes in bacteria (Cohen et al. 2019) nucleated a field of study, and cyclic oligonucleotide based anti-phage signaling systems (CBASS) entered the catalogue of anti-phage pathways in bacteria. The use of both purine and pyrimidine nucleotides as substrates, canonical and non-canonical linkages, and the formation of different sized oligomeric rings, allows for the possible synthesis of many different signaling molecules, increasing the functional space of signaling and recognition.

Our work highlighted the evolutionary flexibility of the overall CD-NTase system. Within CBASS operons, effector proteins specifically respond to the nucleotide signal synthesized by their cognate CD-NTase, and in the case where a single bacterium encodes two CBASS systems, there is no crosstalk between CD-NTases and the alternate effector, accentuating the specificity of effector proteins as receptors (Whiteley et al. 2019). For the abortive infection module to work properly, effectors must recognize the correct signaling molecule and execute the effector function preventing phage replication. Evolutionary pressure from phage infection thus acts on effectors to maintain the overall function of the signaling pathway. The relative plasticity of bacterial CD-NTase enzymes and their capacity to alter product specificity with a single substitution (Whiteley et al. 2019) constitutes a rapid degree of evolution that is stabilized by the selectivity of the effector. Comparably in mammalian synthases such as cGAS and OAS, this pattern

of rapid evolution and positive selection is driven by viral evolutionary pressure (Hancks et al. 2015), suggesting that in bacteria, similar evolutionary arms races between CD-NTases and phage are counteracted by the balancing pressure of the effectors.

Our work exploring the universal determinants of CD-NTase cyclic oligonucleotide synthesis demonstrated that CD-NTase enzymes rely on the formation of a thermostable complex with the initial donor nucleotide, providing a means for determining both the order of cyclic oligonucleotide synthesis reactions and the determinants of nucleotide recognition and selection. We showed that encoded within the remarkable diversity of CD-NTases are sequence features in the lid helix that are conserved within clades and play a role in nucleotide selection. These lid helix residues are located where the DNA or RNA template sits in templated polymerases, and their position in CD-NTase enzymes emphasizes their role in nucleotide selection. Lid residues loosely assort by subclade and serve as the first signposts to look for patterns of interactions that define CD-NTase products and allow for product prediction based on primary sequence features (Chapter 1).

Furthermore, we showed that effector recognition and activation is an important means of identifying the correct activating product of a CD-NTase (Chapter 1). CD-NTase enzymes that make multiple products in the presence of all four nucleotides often have a “major product” that is more abundant, and a “minor product” synthesized in lower amounts, sometimes undetectable by biochemical assays and in insufficient quantities for identification by mass spectrometry. However, multiple examples of effectors recognizing CD-NTase “minor products” (Chapter 1) (Lowey et al. 2020) highlight the high affinity of binding between the effector and the correct activating product, and underscore remaining questions in the field. Is the “minor product” a consequence of genetic drift and evolutionary pressure as the CD-NTase mutates but the effector retains selectivity for a different product? Is it a method of modulating and regulating the CBASS operon to tone down robust activation of the effector? Or is it due to the lack of a “phage infection” environment that is not replicated in our *in vitro* conditions, suggesting that some phage-dependent cue alters the specificity and perhaps the conformation of CD-NTase enzymes, only when required? In support of a phage-dependent cue, in our *in vitro* conditions, the kinetics of 3',3'-cGAMP synthesis by DncV are slower than expected compared to mammalian cGAS (Chapter 1). A non-specific, phage-dependent

activating cue could rapidly increase the rate of synthesis of the effector-activating product as a first line of defense.

CBASS operons are frequently found in pathogenic bacteria. Early work with DncV established the first example of cyclic oligonucleotide signaling and effector activation, and points toward areas of future study. The defense island containing DncV in the 7th pandemic strain of *Vibrio cholerae* El Tor rapidly became the dominant strain, suggesting that acquisition of the CBASS operon containing DncV conferred a fitness advantage (Davies et al. 2012). Interestingly, the DncV-associated phenotype characterized by Davies et al, whereupon DncV activity inhibits chemotaxis, increases gut colonization of infant mice and pathogenicity of *Vibrio* may be due to the interactions between bacteria and phage within intestines, allowing *Vibrio* to escape gut phage and successfully colonize the gut. Further study of the trade-offs between selective fitness advantage against phage and the cost of maintaining the CBASS operon, and how that affects the pathogenicity of disease-causing bacteria is an open question in the field.

The CBASS operon containing DncV and the 3',3'-cGAMP responsive phospholipase also encodes additional CD-NTase associated proteins, including a protein with homology to E1-E2 ligases, and a protease that are required for protection against certain kinds of phage (Cohen et al. 2019). Whether these additional proteins play a role during phage recognition or defense, whether they modulate the activity of DncV, and whether they have a role in the pathogenicity of *Vibrio cholerae* are remaining questions in the field. Relatedly, a role for the ancillary genes present in CBASS operons was typified by characterization of CBASS operons containing HORMA-like domains, showing that the HORMA-domain containing proteins are important modulators of the overall system and regulate CD-NTase activity (Lau et al. 2020; Ye et al. 2020).

In contrast to CRISPR-based immunity which is analogous to mammalian adaptive immunity and relies on pre-organized recognition sequences, CBASS operons act as innate immune sensors, serving as a first line of defense against phage that the colony may not have previously encountered and deploying abortive infection as an effective strategy for preventing phage spread. Cyclic oligonucleotide signals as second messenger molecules amplify danger signals, facilitate downstream activation of effectors that drive an anti-phage response, and enable regulation and evolution at every level of the signaling response — from initial sensing and synthesis to effector activation and functional execution. Future work remains to

reconcile the anti-phage role of CBASS systems with conferred fitness advantages and fitness costs, the intersection of phage defense and other threats, as well as the role of CBASS within the repertoire of other bacterial anti-phage systems.

CBASS Effectors and shared evolutionary lineages

The conserved architecture of innate immune signaling pathways, and the structural similarities between bacteria and metazoan synthases and effectors suggests a shared evolutionary thread. However, the structural similarities between bacterial and metazoan effectors conceal both ongoing functional diversification and ancestral functions that were lost over evolutionary time.

Our work is the first characterization of a bacterial homolog of STING (Chapter 2), surprisingly portraying the ancient evolutionary conservation of cyclic oligonucleotide based innate immune pathways and effectively reconstituting the cGAS-STING innate immune axis within bacteria. Beyond the structural homology, there are key differences that emphasize the different functionality of the pathway in bacterial CBASS operons. Our work identified the first reported STING homolog that does not recognize 2',3'-cGAMP, and instead responds to c-di-GMP with nanomolar affinity within a subset of bacteria that do not encode traditional GGDEF and EAL c-di-GMP signaling domains. While the core structural features of STING are conserved between bacterial and metazoan STING, including the architecture of the ligand binding region and lid, the contacts that specify recognition of either 2',3'-cGAMP by human STING and c-di-GMP by bacterial STING are distinct. Furthermore, acquired features of metazoan STING include an additional alpha helix required for induction of autophagy, and the C-terminal tail necessary for induction of an interferon response in vertebrates (Gui et al. 2019; de Oliveira Mann et al. 2019). This further specialization of STING features from metazoans to vertebrates highlights the overall modular plasticity of innate immune pathways within a conserved signaling architecture.

Mammalian STING contains four N-terminal transmembrane helices that anchor it in the ER membrane and serve as oligomerization scaffolds during STING signaling (Shang et al. 2019; Ergun et al. 2019). A small number of identified bacterial STING proteins are fused to two transmembrane helices, but a far greater number are fused to a Toll/Interleukin-1 receptor (TIR) domain. TIR domains in mammals act as a signaling scaffold during innate immune signaling, but cleave NAD⁺ (β-nicotinamide adenine

dinucleotide) in plant immunity and during the elimination of damaged or unhealthy axons in humans (Essuman et al. 2018; Horsefield et al. 2019; Wan et al. 2019). NAD⁺ cleavage is an ancestral function of TIR domains reliant on a catalytic glutamate, found in TIR-domain containing proteins in both prokaryotes and archaea (Essuman et al. 2018). All characterized TIR domains with NAD⁺ cleaving enzymatic function rely on oligomerization to facilitate cleavage, but the TIR-STING fusion uniquely relies on ligand-induced oligomerization. Rapidly induced function could be one advantage of a ligand-dependent system that is regulated to turn on only when it is required, as opposed to a slower, oligomerization or crowding dependent activation, as seen in the plant TIR domains (Wan et al. 2019). While the enzymatic activity of TIR proteins is lost in metazoan innate immune signaling, the capacity to oligomerize is repurposed as a scaffold for further signaling, such as with the TIR domains of Toll-like receptors.

The CD-NTase upstream of the STING homolog synthesizes 3',3' c-di-GMP. Our work showed that bacterial STING is incredibly specific for canonically linked 3',3' c-di-GMP, once again highlighting the selectivity of effector proteins for their cognate ligand. Bacterial TIR-STING homologs cleave NAD⁺ in response to c-di-GMP signaling, leading to overall NAD⁺ depletion and toxicity. Upon sensing and binding to c-di-GMP, the TIR-STING oligomerizes into active filaments necessary for NAD⁺ cleavage. Bacterial STING activation is regulated by both binding to c-di-GMP and oligomerization into filaments, analogous to the oligomerization of human STING at the ER and ERGIC/Golgi (Ergun et al. 2019). Weak activation of STING by 3',3'-cGAMP results in partial filament formation and partial activation, but at much higher relative concentrations of 3',3'-cGAMP. The specificity for 3',3' c-di-GMP over 3',3'-cGAMP and 2',3'-cGAMP is a distinct feature of bacterial STING signaling, suggesting an evolutionary shift toward cGAMP signaling in metazoans mediated by key conserved residues that differ between bacterial STING homologs and mammalian STING.

In mammals, cGAS-STING signaling is a major antiviral pathway that serves to induce interferon in infected cells, a more indirect STING effector function compared to NAD⁺ cleavage. The discovery of bacterial cGAS homologs, and now bacterial STING homologs contained within bacterial defense operons suggests an ancient origin for these innate immune pathways. Conserved TIR-STING fusions within invertebrates such as oysters that do not have functional NAD⁺ cleavage suggests acquisition followed by adaptation and provides a model for the acquisition and specialization seen within mammalian STING

signaling. Cyclic oligonucleotide recognition, oligomerization, and additional modular elements of vertebrate STING allows for differentiation and evolution of specificity.

Novel receptors of 2'–5' OA and secondary signaling

To expand our understanding of innate immune nucleotide based signaling, we sought to define novel receptors of 2'–5' OA and identified a metabolic enzyme, Decr1, that showed species-specific recognition of 2'–5' OA (Chapter 3). Linear 2'–5' OA is a unique signaling molecule that canonically signals downstream to effector RNase L, which leads to translational arrest and apoptosis, constituting a direct effector function (Hovanessian, Brown, and Kerr 1977; Silverman et al. 1988). A powerful aspect of second messenger dependent signaling is signal amplification, visible in the utility of both mammalian cGAS-STING and the bacterial CD-NTases. However, a small diffusible molecule in the cytoplasm of cells could potentially target other receptors as well, as a secondary arm of innate immune signaling in the event of viral inhibition of or in the absence of RNase L (Drappier and Michiels 2015; Kuusksalu et al. 1995).

Strikingly, the 2' phosphate and adenine base is a structural motif shared by NADPH, a common enzymatic cofactor that regulates many different enzymes. Our work showed that Decr1 can bind to 2'–5' OA even in the presence of NADPH (Chapter 3) and that 2'–5' OA could target other NADPH dependent enzymes in a similar manner, leading to broad anti-viral effects in a virally infected cell. A multi-faceted targeting of diverse downstream receptors by a single second messenger, reliant on a recognition of shared molecular contacts, could also help defend against a wide variety of viruses that rely on different mechanisms of replication and release and that have different metabolic requirements, and more broadly, intracellular bacteria as well.

Molecular contacts drive signaling and innate immune evolution

Overall, our work explored the role of molecular contacts within innate immune signaling, including the determinants of substrate nucleotide recognition, the evolved specificity in recognition of a signaling molecule driving a specific effector response within an anti-viral context, and the recognition of shared molecular patterns that enable more diverse signaling by diffusible second messengers. Protein–RNA contacts and hydrogen bonding interactions are at the heart of each of these signaling pathways, and each one can be reduced to an atomic level recognition between substrate and protein. First, the diversity of CD-

NTases leads to differential molecular contacts that drive substrate recognition and enable the formation of many different signals. Recognition by downstream receptors again relies on molecular contacts that drive distinct specificity, such as the specificity of bacterial STING for c-di-GMP over 3',3'-cGAMP and 2',3'-cGAMP. Lastly, we showed that shared molecular patterns can also drive recognition and binding, such as with Decr1, which evolved to bind NADPH but can also recognize a similar molecular motif within 2'–5' OA, enabling the evolution of adjacent signaling pathways. Altogether, our work broadly characterized and compared several different innate-immune signaling pathways and revealed both a shared pathway architecture and the importance of protein–RNA contacts that confer specificity and allow for specialization within each pathway. Further studies will only build on our understanding of these systems, the evolution of innate immunity, and give us a greater appreciation for the ancient and ongoing battle between organisms and their viral antagonists.

Supplementary Materials

Supplementary Table 3, associated with Chapter 1, can be downloaded from:

[https://www.cell.com/cell-reports/fulltext/S2211-1247\(21\)00555-](https://www.cell.com/cell-reports/fulltext/S2211-1247(21)00555-6?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2211124721005556%3Fshowall%3Dtrue#supplementaryMaterial)

[6?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2211124721005556%](https://www.cell.com/cell-reports/fulltext/S2211-1247(21)00555-6?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2211124721005556%3Fshowall%3Dtrue#supplementaryMaterial)

[3Fshowall%3Dtrue#supplementaryMaterial](https://www.cell.com/cell-reports/fulltext/S2211-1247(21)00555-6?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2211124721005556%3Fshowall%3Dtrue#supplementaryMaterial) listed as “Table S3. All clades and nucleotide-coordinating regions, related to Figure 4”

References

- Ablasser, Andrea, and Zhijian J. Chen. 2019. "CGAS in Action: Expanding Roles in Immunity and Inflammation." *Science* 363 (6431). <https://doi.org/10.1126/science.aat8657>.
- Ablasser, Andrea, Marion Goldeck, Taner Cavlar, Tobias Deimling, Gregor Witte, Ingo Röhl, Karl Peter Hopfner, Janos Ludwig, and Veit Hornung. 2013. "CGAS Produces a 2'-5'-Linked Cyclic Dinucleotide Second Messenger That Activates STING." *Nature* 498 (7454): 380–84. <https://doi.org/10.1038/nature12306>.
- Alphey, Magnus S, Wenhua Yu, Emma Byres, Ding Li, and William N Hunter. 2005. "Structure and Reactivity of Human Mitochondrial 2,4-Dienoyl-CoA Reductase: Enzyme-Ligand Interactions in a Distinctive Short-Chain Reductase Active Site." *The Journal of Biological Chemistry* 280 (4): 3068–77. <https://doi.org/10.1074/jbc.M411069200>.
- Berleman, James E., Benjamin M. Hasselbring, and Carl E. Bauer. 2004. "Hypercyst Mutants in *Rhodospirillum Centenum* Identify Regulatory Loci Involved in Cyst Cell Differentiation." *Journal of Bacteriology* 186 (17): 5834. <https://doi.org/10.1128/JB.186.17.5834-5841.2004>.
- Bernheim, Aude, and Rotem Sorek. n.d. "The Pan-Immune System of Bacteria: Antiviral Defence as a Community Resource." Accessed June 23, 2021. <https://doi.org/10.1038/s41579-019-0278-2>.
- Bisbal, C, C Martinand, M Silhol, B Lebleu, and T Salehzada. 1995. "Cloning and Characterization of a RNase L Inhibitor. A New Component of the Interferon-Regulated 2-5A Pathway." *The Journal of Biological Chemistry* 270 (22): 13308–17. <https://doi.org/10.1074/JBC.270.22.13308>.
- Bisbal, Catherine, and Robert H Silverman. 2007. "Diverse Functions of RNase L and Implications in Pathology." <https://doi.org/10.1016/j.biochi.2007.02.006>.
- Blomme, Arnaud, Catriona A. Ford, Ernest Mui, Rachana Patel, Chara Ntala, Lauren E. Jamieson, Mélanie Planque, et al. 2020. "2,4-Dienoyl-CoA Reductase Regulates Lipid Homeostasis in Treatment-Resistant Prostate Cancer." *Nature Communications* 2020 11:1 11 (1): 1–17. <https://doi.org/10.1038/s41467-020-16126-7>.
- Boehmer, Daniel F. R., Simone Formisano, Carina C. de Oliveira Mann, Stephan A. Mueller, Michael Kluge, Philipp Metzger, Meino Rohlf, et al. 2021. "OAS1/RNase L Executes RIG-I Ligand-Dependent Tumor Cell Apoptosis." *Science Immunology* 6 (61): eabe2550. <https://doi.org/10.1126/SCIIMMUNOL.ABE2550>.
- Breitenbach, Michael, Mark Rinnerthaler, Manuela Weber, Hannelore Breitenbach-Koller, Thomas Karl, Paul Cullen, Sukaniya Basu, Dana Haskova, and Jiri Hasek. 2018. "The Defense and Signaling Role of NADPH Oxidases in Eukaryotic Cells: Review." *Wiener Medizinische Wochenschrift (1946)* 168 (11): 286. <https://doi.org/10.1007/S10354-018-0640-4>.
- Burdette, Dara L., Kathryn M. Monroe, Katia Sotelo-Troha, Jeff S. Iwig, Barbara Eckert, Mamoru Hyodo, Yoshihiro Hayakawa, and Russell E. Vance. 2011. "STING Is a Direct Innate Immune Sensor of Cyclic Di-GMP." *Nature* 478 (7370): 515–18. <https://doi.org/10.1038/nature10429>.
- Burroughs, A. Maxwell, Dapeng Zhang, Daniel E. Schäffer, Lakshminarayan M. Iyer, and L. Aravind. 2015. "Comparative Genomic Analyses Reveal a Vast, Novel Network of Nucleotide-Centric Systems in Biological Conflicts, Immunity and Signaling." *Nucleic Acids Research* 43 (22): 10633–54. <https://doi.org/10.1093/nar/gkv1267>.
- Cavlar, Taner, Tobias Deimling, Andrea Ablasser, Karl Peter Hopfner, and Veit Hornung. 2013. "Species-Specific Detection of the Antiviral Small-Molecule Compound CMA by STING." *EMBO Journal* 32 (10): 1440–50. <https://doi.org/10.1038/emboj.2013.86>.

- Chan, Carmen, Ralf Paul, Dietrich Samoray, Nicolas C. Amiot, Bernd Giese, Urs Jenal, and Tilman Schirmer. 2004. "Structural Basis of Activity and Allosteric Control of Diguanylate Cyclase." *Proceedings of the National Academy of Sciences of the United States of America* 101 (49): 17084–89. <https://doi.org/10.1073/pnas.0406134101>.
- Chebaths, Judith, Philippe Benech, Ara Hovanessian, Julien Galabru, and Michel Revel. 1987. "THE JOURNAL OF BIOLOGICAL CHEMISTRY Four Different Forms of Interferon-Induced 2',5'-Oligo(A) Synthetase Identified by Immunoblotting in Human Cells*" 262 (8): 3852–57. <http://www.jbc.org/content/262/8/3852.full.pdf>.
- Chen, I. Min A., Ken Chu, Krishna Palaniappan, Manoj Pillay, Anna Ratner, Jinghua Huang, Marcel Huntemann, et al. 2019. "IMG/M v.5.0: An Integrated Data Management and Comparative Analysis System for Microbial Genomes and Microbiomes." *Nucleic Acids Research* 47 (D1): D666–77. <https://doi.org/10.1093/nar/gky901>.
- Chen, Vincent B., W. Bryan Arendall, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. 2010. "MolProbity: All-Atom Structure Validation for Macromolecular Crystallography." *Acta Crystallographica Section D: Biological Crystallography* 66 (1): 12–21. <https://doi.org/10.1107/S0907444909042073>.
- Civril, Filiz, Tobias Deimling, Carina C. De Oliveira Mann, Andrea Ablasser, Manuela Moldt, Gregor Witte, Veit Hornung, and Karl Peter Hopfner. 2013. "Structural Mechanism of Cytosolic DNA Sensing by CGAS." *Nature* 498 (7454): 332–37. <https://doi.org/10.1038/nature12305>.
- Cohen, Daniel, Sarah Melamed, Adi Millman, Gabriela Shulman, Yaara Oppenheimer-Shaanan, Assaf Kacen, Shany Doron, Gil Amitai, and Rotem Sorek. 2019. "Cyclic GMP–AMP Signalling Protects Bacteria against Viral Infection." *Nature* 574 (7780): 691–95. <https://doi.org/10.1038/s41586-019-1605-5>.
- Cong, Xiaoyan, Zenglin Yuan, Yijun Du, Bo Wu, Defen Lu, Xiangju Wu, Youjia Zhang, et al. 2019. "Crystal Structures of Porcine STING–CDN Complexes Reveal the Mechanism of Ligand Recognition and Discrimination of STING Proteins." *Journal of Biological Chemistry* 294 (30): 11420–32. <https://doi.org/10.1074/jbc.RA119.007367>.
- Cooper, Daphne A., Babal K. Jha, Robert H. Silverman, Jay R. Hesselberth, and David J. Barton. 2014. "Ribonuclease L and Metal-Ion-Independent Endoribonuclease Cleavage Sites in Host and Viral RNAs." *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gku118>.
- Danilchanka, Olga, and John J. Mekalanos. 2013. "Cyclic Dinucleotides and the Innate Immune Response." *Cell*. <https://doi.org/10.1016/j.cell.2013.08.014>.
- Dash, Chandranu, Timothy S. Fisher, Vinayaka R. Prasad, and Stuart F.J. Le Grice. 2006. "Examining Interactions of HIV-1 Reverse Transcriptase with Single-Stranded Template Nucleotides by Nucleoside Analog Interference." *Journal of Biological Chemistry* 281 (38): 27873–81. <https://doi.org/10.1074/jbc.M603970200>.
- Davies, Bryan W., Ryan W. Bogard, Travis S. Young, and John J. Mekalanos. 2012. "Coordinated Regulation of Accessory Genetic Elements Produces Cyclic Di-Nucleotides for *V. Cholerae* Virulence." *Cell* 149 (2): 358–70. <https://doi.org/10.1016/j.cell.2012.01.053>.
- Dey, Bappaditya, Ruchi Jain Dey, Laurene S. Cheung, Supriya Pokkali, Haidan Guo, Jong Hee Lee, and William R. Bishai. 2015. "A Bacterial Cyclic Dinucleotide Activates the Cytosolic Surveillance Pathway and Mediates Innate Resistance to Tuberculosis." *Nature Medicine* 21 (4): 401–8. <https://doi.org/10.1038/nm.3813>.

- Dong, B, and R H Silverman. 1995. "2-5A-Dependent RNase Molecules Dimerize during Activation by 2-5A." *The Journal of Biological Chemistry* 270 (8): 4133–37. <https://doi.org/10.1074/JBC.270.8.4133>.
- Donovan, Jesse, Matthew Dufner, and Alexei Korennykh. 2013. "Structural Basis for Cytosolic Double-Stranded RNA Surveillance by Human Oligoadenylate Synthetase 1." *Proceedings of the National Academy of Sciences of the United States of America* 110 (5): 1652–57. <https://doi.org/10.1073/pnas.1218528110>.
- Donovan, Jesse, Gena Whitney, Sneha Rath, and Alexei Korennykh. 2015. "Structural Mechanism of Sensing Long DsRNA via a Noncatalytic Domain in Human Oligoadenylate Synthetase 3." *Proceedings of the National Academy of Sciences* 112 (13). <https://doi.org/10.1073/pnas.1419409112>.
- Drappier, Melissa, and Thomas Michiels. 2015. "Inhibition of the OAS/RNase L Pathway by Viruses." *Current Opinion in Virology* 15: 19–26. <https://doi.org/10.1016/j.coviro.2015.07.002>.
- Eaglesham, James B., Youdong Pan, Thomas S. Kupper, and Philip J. Kranzusch. 2019. "Viral and Metazoan Poxins Are CGAMP-Specific Nucleases That Restrict CGAS–STING Signalling." *Nature* 566 (7743): 259–63. <https://doi.org/10.1038/s41586-019-0928-6>.
- Elkhateeb, Enas, Hassan T. Tag-El-Din-Hassan, Nobuya Sasaki, Daisuke Torigoe, Masami Morimatsu, and Takashi Agui. 2016. "The Role of Mouse 2',5'-Oligoadenylate Synthetase 1 Paralogs." *Infection, Genetics and Evolution* 45 (November): 393–401. <https://doi.org/10.1016/J.MEEGID.2016.09.018>.
- Emsley, Paul, and Kevin Cowtan. 2004. "Coot: Model-Building Tools for Molecular Graphics." *Acta Crystallographica Section D: Biological Crystallography* 60 (12 I): 2126–32. <https://doi.org/10.1107/S0907444904019158>.
- Ergun, Sabrina L., Daniel Fernandez, Thomas M. Weiss, and Lingyin Li. 2019. "STING Polymer Structure Reveals Mechanisms for Activation, Hyperactivation, and Inhibition." *Cell* 178 (2): 290-301.e10. <https://doi.org/10.1016/j.cell.2019.05.036>.
- Essuman, Kow, Daniel W Summers, Yo Sasaki, Xianrong Mao, Aldrin Kay, Yuen Yim, Aaron Diantonio, and Jeffrey Milbrandt. 2018. "TIR Domain Proteins Are an Ancient Family of NAD + -Consuming Enzymes." *Current Biology* 28: 421–30. <https://doi.org/10.1016/j.cub.2017.12.024>.
- Faozia, Sabrina, Tazin Fahmi, Gary C. Port, and Kyu Hong Cho. 2021. "C-Di-AMP-Regulated K⁺ Importer KtrAB Affects Biofilm Formation, Stress Response, and SpeB Expression in *Streptococcus Pyogenes*." *Infection and Immunity* 89 (4). <https://doi.org/10.1128/IAI.00317-20>.
- Fillgrove, Kerry L, and Vernon E Anderson. n.d. "The Mechanism of Dienoyl-CoA Reduction by 2,4-Dienoyl-CoA Reductase Is Stepwise: Observation of a Dienolate Intermediate." Accessed June 13, 2018. <https://doi.org/10.1021/bi0111606>.
- Freudenthal, Bret D., William A. Beard, David D. Shock, and Samuel H. Wilson. 2013. "Observing a DNA Polymerase Choose Right from Wrong." *Cell* 154 (1): 157–68. <https://doi.org/10.1016/J.CELL.2013.05.048>.
- Friebe, Andreas, Peter Sandner, and Achim Schmidtke. 2020. "CGMP: A Unique 2nd Messenger Molecule – Recent Developments in CGMP Research and Development." *Naunyn-Schmiedeberg's Archives of Pharmacology*. Springer. <https://doi.org/10.1007/s00210-019-01779-z>.
- Gao, Pu, Manuel Ascano, Yang Wu, Winfried Barchet, Barbara L Gaffney, Thomas Zillinger, Artem A Serganov, et al. 2013. "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.

<https://doi.org/10.1016/j.cell.2013.04.046>.

- Gao, Pu, Manuel Ascano, Thomas Zillinger, Weiyi Wang, Peihong Dai, Artem A. Serganov, Barbara L. Gaffney, et al. 2013. "XStructure-Function Analysis of STING Activation by c[G(2',5') PA(3',5')p] and Targeting by Antiviral DMXAA." *Cell* 154 (4): 748–62. <https://doi.org/10.1016/j.cell.2013.07.023>.
- Ghosh, Arundhati, Saumendra N Sarkar, and Ganes C Sen. n.d. "Cell Growth Regulatory and Antiviral Effects of the P69 Isozyme of 2X5 (A) Synthetase." Accessed November 11, 2017. <https://doi.org/10.1006/viro.1999.0085>.
- Ghosh, Joydeep, Patricia J. Anderson, Sukantha Chandrasekaran, and Michael G. Caparon. 2010. "Characterization of *Streptococcus Pyogenes* β -NAD⁺ Glycohydrolase: Re-Evaluation of Enzymatic Properties Associated with Pathogenesis." *Journal of Biological Chemistry* 285 (8): 5683–94. <https://doi.org/10.1074/jbc.M109.070300>.
- Gui, Xiang, Hui Yang, Tuo Li, Xiaojun Tan, Peiqing Shi, Minghao Li, Fenghe Du, and Zhijian J. Chen. 2019. "Autophagy Induction via STING Trafficking Is a Primordial Function of the CGAS Pathway." *Nature* 567 (7747): 262–66. <https://doi.org/10.1038/s41586-019-1006-9>.
- Gundlach, Jan, Felix M.P. Mehne, Christina Herzberg, Jan Kampf, Oliver Valerius, Volkhard Kaefer, and Jörg Stülke. 2015. "An Essential Poison: Synthesis and Degradation of Cyclic Di-AMP in *Bacillus Subtilis*." *Journal of Bacteriology* 197 (20): 3265–74. <https://doi.org/10.1128/JB.00564-15>.
- Han, Yuchen, Jesse Donovan, Sneha Rath, Gena Whitney, Alisha Chitrakar, and Alexei Korennykh. 2014. "Structure of Human RNase L Reveals the Basis for Regulated RNA Decay in the IFN Response." *Science* 343 (6176): 1244–48. <https://doi.org/10.1126/SCIENCE.1249845>.
- Han, Yuchen, Gena Whitney, Jesse Donovan, and Alexei Korennykh. 2012. "Innate Immune Messenger 2-5A Tethers Human RNase L into Active High-Order Complexes." *Cell Reports* 2 (4): 902–13. <https://doi.org/10.1016/J.CELREP.2012.09.004>.
- Hancks, Dustin C., Melissa K. Hartley, Celia Hagan, Nathan L. Clark, and Nels C. Elde. 2015. "Overlapping Patterns of Rapid Evolution in the Nucleic Acid Sensors CGAS and OAS1 Suggest a Common Mechanism of Pathogen Antagonism and Escape." Edited by Jianzhi Zhang. *PLOS Genetics* 11 (5): e1005203. <https://doi.org/10.1371/journal.pgen.1005203>.
- Hartmann, R, H S Olsen, S Widder, R Jorgensen, and J Justesen. 1998. "P59OASL, a 2'-5' Oligoadenylate Synthetase like Protein: A Novel Human Gene Related to the 2'-5' Oligoadenylate Synthetase Family." *Nucleic Acids Research* 26 (18): 4121–28. <http://www.ncbi.nlm.nih.gov/pubmed/9722630>.
- Heaton, Nicholas S, and Glenn Randall. 2011. "Microbes and Metabolism Multifaceted Roles for Lipids in Viral Infection." <https://doi.org/10.1016/j.tim.2011.03.007>.
- Helander, Heli M, Kari T Koivuranta, Nina Horelli-Kuitunen, Jorma J Palvimo, Aarno Palotie, and J Kalervo Hiltunen. 1997. "Molecular Cloning and Characterization of the Human Mitochondrial 2,4-Dienoyl-CoA Reductase Gene (DECR)." *GENOMICS* 46: 112–19. https://ac-els-cdn-com.ezp-prod1.hul.harvard.edu/S0888754397950047/1-s2.0-S0888754397950047-main.pdf?_tid=eeb4d750-da89-11e7-9fa1-00000aab0f02&acdnat=1512567341_d77f2081c8ef9307ba1d6125f3f183a1.
- Horsefield, Shane, Hayden Burdett, Xiaoxiao Zhang, Mohammad K. Manik, Yun Shi, Jian Chen, Tiancong Qi, et al. 2019. "NAD⁺ Cleavage Activity by Animal and Plant TIR Domains in Cell Death Pathways." *Science* 365 (6455): 793–99. <https://doi.org/10.1126/science.aax1911>.
- Hovanessian, Ara G., Ronald E. Brown, and Ian M. Kerr. 1977. "Synthesis of Low Molecular Weight

- Inhibitor of Protein Synthesis with Enzyme from Interferon-Treated Cells." *Nature* 268 (5620): 537–40. <https://doi.org/10.1038/268537a0>.
- HOVANESSIAN, Ara G., and Ian M. KERR. 1978. "Synthesis of an Oligonucleotide Inhibitor of Protein Synthesis in Rabbit Reticulocyte Lysates Analogous to That Formed in Extracts from Interferon-Treated Cells." *European Journal of Biochemistry* 84 (1): 149–59. <https://doi.org/10.1111/j.1432-1033.1978.tb12151.x>.
- Huang, Hao, Elton Zeqiraj, Beihua Dong, Babal Kant Jha, Nicole M. Duffy, Stephen Orlicky, Neroshan Thevakumaran, et al. 2014. "Dimeric Structure of Pseudokinase RNase L Bound to 2-5A Reveals a Basis for Interferon-Induced Antiviral Activity." *Molecular Cell* 53 (2): 221–34. <https://doi.org/10.1016/J.MOLCEL.2013.12.025>.
- Huang, Yi He, Xiang Yu Liu, Xiao Xia Du, Zheng Fan Jiang, and Xiao Dong Su. 2012. "The Structural Basis for the Sensing and Binding of Cyclic Di-GMP by STING." *Nature Structural and Molecular Biology* 19 (7): 728–30. <https://doi.org/10.1038/nsmb.2333>.
- Ibsen, Mikkel Søres, Hans Henrik Gad, Line Lykke Andersen, Veit Hornung, Ilkka Julkunen, Saumendra N Sarkar, and Rune Hartmann. 2015. "Structural and Functional Analysis Reveals That Human OASL Binds DsRNA to Enhance RIG-I Signaling." *Nucleic Acids Research* 43 (10): 5236–48. <https://doi.org/10.1093/nar/gkv389>.
- Ibsen, Mikkel Søres, Hans Henrik Gad, Karthiga Thavachelvam, Thomas Boesen, Philippe Desprès, and Rune Hartmann. 2014. "The 2'-5'-Oligoadenylate Synthetase 3 Enzyme Potently Synthesizes the 2'-5'-Oligoadenylates Required for RNase L Activation." *Journal of Virology* 88 (24): 14222. <https://doi.org/10.1128/JVI.01763-14>.
- Ishikawa, Hiroki, and Glen N. Barber. 2008. "STING Is an Endoplasmic Reticulum Adaptor That Facilitates Innate Immune Signalling." *Nature* 455 (7213): 674–78. <https://doi.org/10.1038/nature07317>.
- Jenal, Urs, Alberto Reinders, and Christian Lori. 2017. "Cyclic Di-GMP: Second Messenger Extraordinaire." *Nature Reviews Microbiology*. Nature Publishing Group. <https://doi.org/10.1038/nrmicro.2016.190>.
- Ju, Huai-Qiang, Jin-Fei Lin, Tian Tian, Dan Xie, and Rui-Hua Xu. 2020. "NADPH Homeostasis in Cancer: Functions, Mechanisms and Therapeutic Implications." *Signal Transduction and Targeted Therapy* 2020 5:1 5 (1): 1–12. <https://doi.org/10.1038/s41392-020-00326-0>.
- Kakuta, Shigeru, Shinwa Shibata, and Yoichiro Iwakura. 2004. "Genomic Structure of the Mouse 2',5'-Oligoadenylate Synthetase Gene Family." *https://Home.Liebertpub.Com/Jir* 22 (9): 981–93. <https://doi.org/10.1089/10799900260286696>.
- Kalia, Dimpy, Gökçe Merey, Shizuka Nakayama, Yue Zheng, Jie Zhou, Yiling Luo, Min Guo, Benjamin T. Roembke, and Herman O. Sintim. 2013. "Nucleotide, c-Di-GMP, c-Di-AMP, CGMP, CAMP, (p)PpGpp Signaling in Bacteria and Implications in Pathogenesis." *Chemical Society Reviews* 42 (1): 305–41. <https://doi.org/10.1039/c2cs35206k>.
- Karplus, P. Andrew, and Kay Diederichs. 2012. "Linking Crystallographic Model and Data Quality." *Science* 336 (6084): 1030–33. <https://doi.org/10.1126/science.1218231>.
- Kato, Kazuki, Ryohei Ishii, Seiichi Hirano, Ryuichiro Ishitani, and Osamu Nureki. 2015. "Structural Basis for the Catalytic Mechanism of DncV, Bacterial Homolog of Cyclic GMP-AMP Synthase." *Structure* 23 (5): 843–50. <https://doi.org/10.1016/j.str.2015.01.023>.

- Katoh, Kazutaka, John Rozewicki, and Kazunori D. Yamada. 2018. "MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization." *Briefings in Bioinformatics* 20 (4): 1160–66. <https://doi.org/10.1093/bib/bbx108>.
- Kerr, Ian M., Ronald E. Brown, and Ara G. Hovanessian. 1977. "Nature of Inhibitor of Cell-Free Protein Synthesis Formed in Response to Interferon and Double-Stranded RNA." *Nature* 268 (5620): 540–42. <https://doi.org/10.1038/268540a0>.
- Kerr, Ian M, and Ronald E Brown. 1978. "Biochemistry PppA2'p5'A2'p5'A: An Inhibitor of Protein Synthesis Synthesized with an Enzyme Fraction from Interferon-Treated Cells (Cell-Free Systems/Double-Stranded RNA/Low Molecular Weight Inhibitor/Oligomeric Series/Protein Kinase)" 75 (1): 256–60. <http://www.pnas.org/content/75/1/256.full.pdf>.
- Kimata, Keiko, Hideyuki Takahashi, Toshifumi Inada, Pieter Postma, and Hiroji Aiba. 1997. "CAMP Receptor Protein–CAMP Plays a Crucial Role in Glucose–Lactose Diauxie by Activating the Major Glucose Transporter Gene in Escherichia Coli." *Proceedings of the National Academy of Sciences* 94 (24): 12914–19. <https://doi.org/10.1073/PNAS.94.24.12914>.
- Kjaer, Karina Hansen, Jesper Buchhave Poulsen, Tõnu Reintamm, Emilie Saby, Pia Moeller Martensen, Merike Kelve, and Just Justesen. 2009. "Evolution of the 2'-5'-Oligoadenylate Synthetase Family in Eukaryotes and Bacteria." *Journal of Molecular Evolution* 69 (6): 612–24. <https://doi.org/10.1007/s00239-009-9299-1>.
- Kottur, Jithesh, and Deepak T Nair. 2018. "Pyrophosphate Hydrolysis Is an Intrinsic and Critical Step of the DNA Synthesis Reaction." *Nucleic Acids Research* 46 (12): 5875–85. <https://doi.org/10.1093/nar/gky402>.
- Kranzusch, Philip J., Amy S.Y. Lee, Stephen C. Wilson, Mikhail S. Solovykh, Russell E. Vance, James M. Berger, and Jennifer A. Doudna. 2014. "Structure-Guided Reprogramming of Human Cgas Dinucleotide Linkage Specificity." *Cell* 158 (5): 1011–21. <https://doi.org/10.1016/j.cell.2014.07.028>.
- Kranzusch, Philip J., Stephen C. Wilson, Amy S.Y. Lee, James M. Berger, Jennifer A. Doudna, and Russell E. Vance. 2015. "Ancient Origin of CGAS-STING Reveals Mechanism of Universal 2',3' CGAMP Signaling." *Molecular Cell* 59 (6): 891–903. <https://doi.org/10.1016/j.molcel.2015.07.022>.
- Kranzusch, Philip J. 2019. "CGAS and CD-NTase Enzymes: Structure, Mechanism, and Evolution This Review Comes from a Themed Issue on Protein Nucleic Acid Interactions." *Current Opinion in Structural Biology* 59: 178–87. <https://doi.org/10.1016/j.sbi.2019.08.003>.
- Krissinel, E., and K. Henrick. 2004. "Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions." *Acta Crystallographica Section D: Biological Crystallography* 60 (12 I): 2256–68. <https://doi.org/10.1107/S0907444904026460>.
- Kulesekara, Hemantha, Vincent Lee, Anja Brencic, Nicole Liberati, Jonathan Urbach, Sachiko Miyata, Daniel G. Lee, et al. 2006. "Analysis of Pseudomonas Aeruginosa Diguanylate Cyclases and Phosphodiesterases Reveals a Role for Bis-(3'-5')-Cyclic-GMP in Virulence." *Proceedings of the National Academy of Sciences of the United States of America* 103 (8): 2839–44. <https://doi.org/10.1073/pnas.0511090103>.
- Kuusksalu, Anne, Arno Pihlak, Werner E.G. Müller, and Merike Kelve. 1995. "The (2'-5')Oligoadenylate Synthetase Is Present in the Lowest Multicellular Organisms, the Marine Sponges: Demonstration of the Existence and Identification of Its Reaction Products." *European Journal of Biochemistry* 232 (2): 351–57. <https://doi.org/10.1111/j.1432-1033.1995.351zz.x>.
- Kwon, Young-Chan, Ju-Il Kang, Soon B. Hwang, and Byung-Yoon Ahn. 2013. "The Ribonuclease L-

- Dependent Antiviral Roles of Human 2',5'-Oligoadenylate Synthetase Family Members against Hepatitis C Virus." *FEBS Letters* 587 (2): 156–64. <https://doi.org/10.1016/J.FEBSLET.2012.11.010>.
- Lau, Rebecca K., Qiaozhen Ye, Erica A. Birkholz, Kyle R. Berg, Lucas Patel, Ian T. Mathews, Jeramie D. Watrous, et al. 2020. "Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity." *Molecular Cell* 77 (4): 723-733.e6. <https://doi.org/10.1016/j.molcel.2019.12.010>.
- Launer-Felty, Katherine D, and Scott A Strobel. 2018. "Enzymatic Synthesis of Cyclic Dinucleotide Analogs by a Promiscuous Cyclic-AMP-GMP Synthetase and Analysis of Cyclic Dinucleotide Responsive Riboswitches." *Nucleic Acids Research* 46 (6): 2765–76. <https://doi.org/10.1093/nar/gky137>.
- Letunic, Ivica, and Peer Bork. 2019. "Interactive Tree of Life (ITOL) v4: Recent Updates and New Developments." *Nucleic Acids Research* 47 (W1). <https://doi.org/10.1093/nar/gkz239>.
- Leu, Guang-Zhou, Tiao-Yin Lin, and John T.A Hsu. 2004. "Anti-HCV Activities of Selective Polyunsaturated Fatty Acids." *Biochemical and Biophysical Research Communications* 318 (1): 275–80. <https://doi.org/10.1016/j.bbrc.2004.04.019>.
- Li, Geqiang, Ying Xiang, Kanaga Sabapathy, and Robert H. Silverman. 2004. "An Apoptotic Signaling Pathway in the Interferon Antiviral Response Mediated by RNase L and C-Jun NH₂-Terminal Kinase." *Journal of Biological Chemistry* 279 (2): 1123–31. <https://doi.org/10.1074/jbc.M305893200>.
- Li, Geqiang, Ying Xiang, Kanaga Sabapathy, and Robert H Silverman. 2003. "An Apoptotic Signaling Pathway in the Interferon Antiviral Response Mediated by RNase L and C-Jun NH₂-Terminal Kinase*." <https://doi.org/10.1074/jbc.M305893200>.
- Li, Yize, Shuvajit Banerjee, Yuyan Wang, Stephen A Goldstein, Beihua Dong, Christina Gaughan, Robert H Silverman, and Susan R Weiss. 2016. "Activation of RNase L Is Dependent on OAS3 Expression during Infection with Diverse Human Viruses." *Proceedings of the National Academy of Sciences of the United States of America* 113 (8): 2241–46. <https://doi.org/10.1073/pnas.1519657113>.
- Liebschner, Dorothee, Pavel V. Afonine, Matthew L. Baker, Gábor Bunkóczi, Vincent B. Chen, Tristan I. Croll, Bradley Hintze, et al. 2019. "Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix." *Acta Crystallographica. Section D, Structural Biology* 75 (Pt 10): 861. <https://doi.org/10.1107/S2059798319011471>.
- Lohöfener, Jan, Nicola Steinke, Penelope Kay-Fedorov, Petra Baruch, Alexey Nikulin, Svetlana Tishchenko, Dietmar J. Manstein, and Roman Fedorov. 2015. "The Activation Mechanism of 2'-5'-Oligoadenylate Synthetase Gives New Insights Into OAS/CGAS Triggers of Innate Immunity." *Structure* 23 (5): 851–62. <https://doi.org/10.1016/j.str.2015.03.012>.
- Lopatina, Anna, Nitzan Tal, and Rotem Sorek. 2020. "Annual Review of Virology Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy." <https://doi.org/10.1146/annurev-virology-011620>.
- Loring, Heather S., Janneke D. Ico, Venkatesh V. Nemmara, and Paul R. Thompson. 2020. "Initial Kinetic Characterization of Sterile Alpha and Toll/Interleukin Receptor Motif-Containing Protein 1." *Biochemistry* 59 (8): 933–42. <https://doi.org/10.1021/acs.biochem.9b01078>.
- Lowey, Brianna, Aaron T. Whiteley, Alexander F.A. Keszei, Benjamin R. Morehouse, Ian T. Mathews, Sadie P. Antine, Victor J. Cabrera, 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 (1): 38-49.e17. <https://doi.org/10.1016/j.cell.2020.05.019>.

- Madden, J. C., Jr., Dan Cui, and M. A. Brinton. 2019. "RNase L Antiviral Activity Is Not a Critical Component of the Oas1b-Mediated Flavivirus Resistance Phenotype." *Journal of Virology* 93 (22). <https://doi.org/10.1128/JVI.00946-19>.
- Manivannan, Praveen, Mohammad Adnan Siddiqui, and Krishnamurthy Malathi. 2020. "RNase L Amplifies Interferon Signaling by Inducing Protein Kinase R-Mediated Antiviral Stress Granules." *Journal of Virology* 94 (13). <https://doi.org/10.1128/JVI.00205-20>.
- Margolis, Shally R., Stephen C. Wilson, and Russell E. Vance. 2017. "Evolutionary Origins of CGAS-STING Signaling." *Trends in Immunology*. Elsevier Ltd. <https://doi.org/10.1016/j.it.2017.03.004>.
- Martinand, C, T Salehzada, M Silhol, B Lebleu, and C Bisbal. 1998. "The RNase L Inhibitor (RLI) Is Induced by Double-Stranded RNA." *JOURNAL OF INTERFERON AND CYTOKINE RESEARCH* 18: 1031–38. <http://online.liebertpub.com.ezp-prod1.hul.harvard.edu/doi/pdf/10.1089/jir.1998.18.1031>.
- McFarland, Adelle P, Shukun Luo, Fariha Ahmed-Qadri, Meghan Zuck, Elizabeth F Thayer, Young Ah Goo, Kevin Hybiske, Liang Tong, and Joshua J Woodward. 2017. "Sensing of Bacterial Cyclic Dinucleotides by the Oxidoreductase RECON Promotes NF-KB Activation and Shapes a Proinflammatory Antibacterial State." *Immunity* 46 (3): 433–45. <https://doi.org/10.1016/j.immuni.2017.02.014>.
- Millman, Adi, Sarah Melamed, Gil Amitai, and Rotem Sorek. 2020. "Diversity and Classification of Cyclic-Oligonucleotide-Based Anti-Phage Signalling Systems." *Nature Microbiology* 5 (12): 1608–15. <https://doi.org/10.1038/s41564-020-0777-y>.
- Morehouse, Benjamin R., Apurva A. Govande, Adi Millman, Alexander F.A. Keszei, Brianna Lowey, Gal Ofir, Sichen Shao, Rotem Sorek, and Philip J. Kranzusch. 2020. "STING Cyclic Dinucleotide Sensing Originated in Bacteria." *Nature*, September, 1–5. <https://doi.org/10.1038/s41586-020-2719-5>.
- Morin, B, N Rabah, J Boretto-Soler, H Tolou, K Alvarez, and B Canard. 2010. "High Yield Synthesis, Purification and Characterisation of the RNase L Activators 5-Triphosphate 2-5 Oligoadenylates." *Antiviral Research* 87: 345–52. <https://doi.org/10.1016/j.antiviral.2010.06.003>.
- Nassar, Zeyad D, Chui Yan Mah, Jonas Dehairs, Ingrid JG Burvenich, Swati Irani, Margaret M Centenera, Madison Helm, et al. 2020. "Human DECR1 Is an Androgen-Repressed Survival Factor That Regulates PUFA Oxidation to Protect Prostate Tumor Cells from Ferroptosis." *ELife* 9 (July): 1–34. <https://doi.org/10.7554/ELIFE.54166>.
- Nobeli, Irene, Roman A Laskowski, William S J Valdar, and Janet M Thornton. 2001. "On the Molecular Discrimination between Adenine and Guanine by Proteins." *Nucleic Acids Research*. Vol. 29. <http://www.daylight.com/dayhtml/smiles/>.
- Ofir, Gal, Sarah Melamed, Hila Sberro, Zohar Mukamel, Shahar Silverman, Gilad Yaakov, Shany Doron, and Rotem Sorek. 2018. "DISARM Is a Widespread Bacterial Defence System with Broad Anti-Phage Activities." *Nature Microbiology* 3 (1): 90–98. <https://doi.org/10.1038/s41564-017-0051-0>.
- Oliveira Mann, Carina C. de, Megan H. Orzalli, David S. King, Jonathan C. Kagan, Amy S.Y. Lee, and Philip J. Kranzusch. 2019. "Modular Architecture of the STING C-Terminal Tail Allows Interferon and NF-KB Signaling Adaptation." *Cell Reports* 27 (4): 1165-1175.e5. <https://doi.org/10.1016/j.celrep.2019.03.098>.
- Ouyang, Songying, Xianqiang Song, Yaya Wang, Heng Ru, Neil Shaw, Yan Jiang, Fengfeng Niu, et al. 2012. "Structural Analysis of the STING Adaptor Protein Reveals a Hydrophobic Dimer Interface

- and Mode of Cyclic Di-GMP Binding." *Immunity* 36 (6): 1073–86.
<https://doi.org/10.1016/j.immuni.2012.03.019>.
- Pei, Jimin, and Nick V. Grishin. 2014. "PROMALS3D: Multiple Protein Sequence Alignment Enhanced with Evolutionary and Three-Dimensional Structural Information." *Methods in Molecular Biology* 1079: 263–71. https://doi.org/10.1007/978-1-62703-646-7_17.
- Perelygin, Andrey A, Svetlana V Scherbik, Igor B Zhulin, Bronislava M Stockman, Yan Li, and Margo A Brinton. 2002. "Positional Cloning of the Murine Flavivirus Resistance Gene." *PNAS July*. Vol. 99. www.pnas.org/cgi/doi/10.1073/pnas.142287799.
- Peterson, Elizabeth, Emma Shippee, Margo A. Brinton, and Parjit Kaur. 2019. "Biochemical Characterization of the Mouse ABCF3 Protein, a Partner of the Flavivirus-Resistance Protein OAS1B." *The Journal of Biological Chemistry* 294 (41): 14937.
<https://doi.org/10.1074/JBC.RA119.008477>.
- Pollak, Nadine, Marc Niere, and Mathias Ziegler. 2007. "NAD Kinase Levels Control the NADPH Concentration in Human Cells." *The Journal of Biological Chemistry* 282 (46): 33562–71.
<https://doi.org/10.1074/jbc.M704442200>.
- Poulsen, Jesper Buchhave, Karina Hansen Kjær, Just Justesen, and Pia Møller Martensen. 2015. "Enzyme Assays for Synthesis and Degradation of 2-5As and Other 2'-5' Oligonucleotides." *BMC Biochemistry* 16 (June): 15. <https://doi.org/10.1186/s12858-015-0043-8>.
- RALL, T. W., and E. W. SUTHERLAND. 1958. "FORMATION OF A CYCLIC ADENINE RIBONUCLEOTIDE BY TISSUE PARTICLES." *Journal of Biological Chemistry* 232 (2): 1065–76.
[https://doi.org/10.1016/S0021-9258\(19\)77422-5](https://doi.org/10.1016/S0021-9258(19)77422-5).
- Rausch, Jason W., Linda Chelico, Myron F. Goodman, and Stuart F.J. Le Grice. 2009. "Dissecting APOBEC3G Substrate Specificity by Nucleoside Analog Interference." *Journal of Biological Chemistry* 284 (11): 7047–58. <https://doi.org/10.1074/jbc.M807258200>.
- Ravnskjaer, Kim, Anila Madiraju, and Marc Montminy. 2015. "Role of the CAMP Pathway in Glucose and Lipid Metabolism." In , 29–49. Springer, Cham. https://doi.org/10.1007/164_2015_32.
- Rebouillat, Dominique, Alain Hovnanian, Isabelle Marié, and Ara G Hovanessian. n.d. "The 100-KDa 2,5'-Oligoadenylate Synthetase Catalyzing Preferentially the Synthesis of Dimeric PppA2p5A Molecules Is Composed of Three Homologous Domains*." Accessed October 31, 2017.
<http://www.jbc.org/content/274/3/1557.full.pdf>.
- Reintamm, Tõnu, Anne Kuusksalu, Madis Metsis, Mailis Päre, Kerli Vallmann, Annika Lopp, Just Justesen, and Merike Kelve. 2008. "Sponge OAS Has a Distinct Genomic Structure within the 2-5A Synthetase Family." *Molecular Genetics and Genomics* 280 (5): 453–66.
<https://doi.org/10.1007/s00438-008-0379-5>.
- Rice, Andrew P, Walden K Roberts, and Ian M Kerr. 1984. "2-5A Accumulates to High Levels in Interferon-Treated, Vaccinia Virus-Infected Cells in the Absence of Any Inhibition of Virus Replication." *JOURNAL OF VIROLOGY* 50 (1): 220–28. <http://jvi.asm.org/content/50/1/220.full.pdf>.
- RJ, Lin, Yu HP, Chang BL, Tang WC, Liao CL, and Lin YL. 2009. "Distinct Antiviral Roles for Human 2',5'-Oligoadenylate Synthetase Family Members against Dengue Virus Infection." *Journal of Immunology (Baltimore, Md. : 1950)* 183 (12): 8035–43.
<https://doi.org/10.4049/JIMMUNOL.0902728>.
- Rosenberg, Jonathan, Achim Dickmanns, Piotr Neumann, Katrin Gunka, Johannes Arens, Volkhard

- Kaefer, Jörg Stülke, Ralf Ficner, and Fabian M Commichau. 2015. "Structural and Biochemical Analysis of the Essential Diadenylate Cyclase CdaA from *Listeria Monocytogenes*." *The Journal of Biological Chemistry* 290 (10): 6596–6606. <https://doi.org/10.1074/jbc.M114.630418>.
- Scherbik, Svetlana V., Kerri Kluetzman, Andrey A. Perelygin, and Margo A. Brinton. 2007. "Knock-in of the Oas1br Allele into a Flavivirus-Induced Disease Susceptible Mouse Generates the Resistant Phenotype." *Virology* 368 (2): 232–37. <https://doi.org/10.1016/J.VIROL.2007.08.017>.
- Serezani, Carlos H., Megan N. Ballinger, David M. Aronoff, and Marc Peters-Golden. 2008. "Cyclic AMP: Master Regulator of Innate Immune Cell Function." *American Journal of Respiratory Cell and Molecular Biology* 39 (2): 127. <https://doi.org/10.1165/RCMB.2008-0091TR>.
- Severin, Geoffrey B, Miriam S Ramliden, Lisa A Hawver, Kun Wang, Macy E Pell, Ann-Katrin Kieninger, Atul Khataokar, et al. 2018. "Direct Activation of a Phospholipase by Cyclic GMP-AMP in *El Tor* *Vibrio Cholerae*." *Proceedings of the National Academy of Sciences of the United States of America* 115 (26): E6048–55. <https://doi.org/10.1073/pnas.1801233115>.
- Shang, Guijun, Conggang Zhang, Zhijian J. Chen, Xiao chen Bai, and Xuwu Zhang. 2019. "Cryo-EM Structures of STING Reveal Its Mechanism of Activation by Cyclic GMP-AMP." *Nature* 567 (7748): 389–93. <https://doi.org/10.1038/s41586-019-0998-5>.
- Shang, Guijun, Deyu Zhu, Ning Li, Junbing Zhang, Chunyuan Zhu, Defen Lu, Cuilan Liu, et al. 2012. "Crystal Structures of STING Protein Reveal Basis for Recognition of Cyclic Di-GMP." *Nature Structural and Molecular Biology* 19 (7): 725–27. <https://doi.org/10.1038/nsmb.2332>.
- Shu, Chang, Guanghui Yi, Tylan Watts, C. Cheng Kao, and Pingwei Li. 2012. "Structure of STING Bound to Cyclic Di-GMP Reveals the Mechanism of Cyclic Dinucleotide Recognition by the Immune System." *Nature Structural and Molecular Biology* 19 (7): 722–24. <https://doi.org/10.1038/nsmb.2331>.
- Silverman, Robert H, David D Jung, Nancy L Nolan-Sorden, Carl W Dieffenbach, Vishram P Kedar, and Dibyendu N Sengupta. 1988. "Purification and Analysis of Murine 2-5A-Dependent RNase*." *THE JOURNAL OF BIOLOGICAL CHEMISTRY* Vol 263 (15): 7336–41. <http://www.jbc.org/content/263/15/7336.full.pdf>.
- Sixt, Barbara S., Robert J. Bastidas, Ryan Finethy, Ryan M. Baxter, Victoria K. Carpenter, Guido Kroemer, Jörn Coers, and Raphael H. Valdivia. 2017. "The Chlamydia Trachomatis Inclusion Membrane Protein CpoS Counteracts STING-Mediated Cellular Surveillance and Suicide Programs." *Cell Host and Microbe* 21 (1): 113–21. <https://doi.org/10.1016/j.chom.2016.12.002>.
- Soveg, Frank W, Johannes Schwerk, Nandan S Gokhale, Karen Cerosaletti, Julian R Smith, Erola Pairo-Castineira, Alison M Kell, et al. 2021. "Endomembrane Targeting of Human OAS1 P46 Augments Antiviral Activity." *ELife* 10 (August). <https://doi.org/10.7554/ELIFE.71047>.
- Sun, Lijun, Jiayi Wu, Fenghe Du, Xiang Chen, and Zhijian J. Chen. 2013a. "Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway." *Science* 339 (6121): 786–91. <https://doi.org/10.1126/science.1232458>.
- Sun, Lijun, Jiayi Wu, Fenghe Du, Xiang Chen, and Zhijian J Chen. 2013b. "Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway." *Science (New York, N.Y.)* 339 (6121): 786–91. <https://doi.org/10.1126/science.1232458>.
- SUTHERLAND, E. W., and T. W. RALL. 1958. "FRACTIONATION AND CHARACTERIZATION OF A CYCLIC ADENINE RIBONUCLEOTIDE FORMED BY TISSUE PARTICLES." *Journal of Biological Chemistry* 232 (2): 1077–91. [https://doi.org/10.1016/S0021-9258\(19\)77423-7](https://doi.org/10.1016/S0021-9258(19)77423-7).

- Tak, Uday, Jiri Vlach, Acely Garza-Garcia, Doreen William, Olga Danilchanka, Luiz Pedro Sório de Carvalho, Jamil S. Saad, and Michael Niederweis. 2019. "The Tuberculosis Necrotizing Toxin Is an NAD and NADP Glycohydrolase with Distinct Enzymatic Properties." *Journal of Biological Chemistry* 294 (9): 3024–36. <https://doi.org/10.1074/jbc.RA118.005832>.
- Tang, Guang, Liwei Peng, Philip R. Baldwin, Deepinder S. Mann, Wen Jiang, Ian Rees, and Steven J. Ludtke. 2007. "EMAN2: An Extensible Image Processing Suite for Electron Microscopy." *Journal of Structural Biology* 157 (1): 38–46. <https://doi.org/10.1016/j.jsb.2006.05.009>.
- Toshchakov, Vladimir Y., and Andrew F. Neuwald. 2020. "A Survey of TIR Domain Sequence and Structure Divergence." *Immunogenetics* 72 (3): 181–203. <https://doi.org/10.1007/s00251-020-01157-7>.
- VanLinden, Magali R., Renate Hvidsten Skoge, and Mathias Ziegler. 2015. "Discovery, Metabolism and Functions of NAD and NADP." *The Biochemist* 37 (1): 9–13. <https://doi.org/10.1042/BIO03701009>.
- Wan, Li, Kow Essuman, Ryan G. Anderson, Yo Sasaki, Freddy Monteiro, Eui Hwan Chung, Erin Osborne Nishimura, et al. 2019. "TIR Domains of Plant Immune Receptors Are NAD⁺-Cleaving Enzymes That Promote Cell Death." *Science* 365 (6455): 799–803. <https://doi.org/10.1126/science.aax1771>.
- Wang, Jun, Guofan Zhang, Xiaodong Fang, Ximing Guo, Li Li, Ruibang Luo, Fei Xu, et al. 2012. "The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation." *Nature* 490 (7418): 49–54. <https://doi.org/10.1038/nature11413>.
- Wang, Zhenghua, Cancan Zhao, Chuanlin Wang, Hang Zhang, Dejun Ma, Qiangzhe Zhang, Xin Wen, Luyuan Li, and Zhen Xi. 2021. "Synthesis and Biological Evaluation of All Possible Inosine-Mixed Cyclic Dinucleotides That Activate Different HSTING Variants." *Bioorganic and Medicinal Chemistry* 29 (January). <https://doi.org/10.1016/j.bmc.2020.115899>.
- Waterhouse, Andrew M., James B. Procter, David M.A. Martin, Michèle Clamp, and Geoffrey J. Barton. 2009. "Jalview Version 2-A Multiple Sequence Alignment Editor and Analysis Workbench." *Bioinformatics* 25 (9): 1189–91. <https://doi.org/10.1093/bioinformatics/btp033>.
- Weiss, Manfred S. 2001. "Global Indicators of X-Ray Data Quality." *Journal of Applied Crystallography* 34 (2): 130–35. <https://doi.org/10.1107/S0021889800018227>.
- Wexler, Aaron G., and Andrew L. Goodman. 2017. "An Insider's Perspective: Bacteroides as a Window into the Microbiome." *Nature Microbiology*. Nature Publishing Group. <https://doi.org/10.1038/nmicrobiol.2017.26>.
- Whiteley, Aaron T., James B. Eaglesham, Carina C. de Oliveira Mann, Benjamin R. Morehouse, Brianna Lowey, Eric A. Nieminen, Olga Danilchanka, et al. 2019. "Bacterial CGAS-like Enzymes Synthesize Diverse Nucleotide Signals." *Nature* 567 (7747): 194–99. <https://doi.org/10.1038/s41586-019-0953-5>.
- Witte, Gregor, Sophia Hartung, Katharina Büttner, and Karl Peter Hopfner. 2008. "Structural Biochemistry of a Bacterial Checkpoint Protein Reveals Diadenylate Cyclase Activity Regulated by DNA Recombination Intermediates." *Molecular Cell* 30 (2): 167–78. <https://doi.org/10.1016/j.molcel.2008.02.020>.
- Wood, Edgar R, Randy Bledsoe, Jing Chai, Philias Daka, Hongfeng Deng, Yun Ding, Sarah Harris-Gurley, et al. 2015. "The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors." *The Journal of Biological Chemistry* 290 (32): 19681–96. <https://doi.org/10.1074/jbc.M115.653113>.

- Woodward, Joshua J., Anthony T. Lavarone, and Daniel A. Portnoy. 2010. "C-Di-AMP Secreted by Intracellular *Listeria Monocytogenes* Activates a Host Type I Interferon Response." *Science* 328 (5986): 1703–5. <https://doi.org/10.1126/science.1189801>.
- Yang, Wei, Jae Young Lee, and Marcin Nowotny. 2006. "Making and Breaking Nucleic Acids: Two-Mg²⁺-Ion Catalysis and Substrate Specificity." *Molecular Cell*. <https://doi.org/10.1016/j.molcel.2006.03.013>.
- Ye, Qiaozhen, Rebecca K. Lau, Ian T. Mathews, Erica A. Birkholz, Jeramie D. Watrous, Camillia S. Azimi, Joe Pogliano, Mohit Jain, and Kevin D. Corbett. 2020. "HORMA Domain Proteins and a Trip13-like ATPase Regulate Bacterial CGAS-like Enzymes to Mediate Bacteriophage Immunity." *Molecular Cell* 77 (4): 709–722.e7. <https://doi.org/10.1016/j.molcel.2019.12.009>.
- Yin, Qian, Yuan Tian, Venkataraman Kabaleeswaran, Xiaomo Jiang, Daqi Tu, Michael J. Eck, Zhijian J. Chen, and Hao Wu. 2012. "Cyclic Di-GMP Sensing via the Innate Immune Signaling Protein STING." *Molecular Cell* 46 (6): 735–45. <https://doi.org/10.1016/j.molcel.2012.05.029>.
- Zhang, Conggang, Guijun Shang, Xiang Gui, Xuewu Zhang, Xiao chen Bai, and Zhijian J. Chen. 2019. "Structural Basis of STING Binding with and Phosphorylation by TBK1." *Nature* 567 (7748): 394–98. <https://doi.org/10.1038/s41586-019-1000-2>.
- Zhang, Heng, Min Jie Han, Jianli Tao, Zhao Yang Ye, Xiao Xia Du, Ming Jing Deng, Xiao Yan Zhang, Lan Fen Li, Zheng Fan Jiang, and Xiao Dong Su. 2015. "Rat and Human STINGs Profile Similarly towards Anticancer/Antiviral Compounds." *Scientific Reports* 5 (December). <https://doi.org/10.1038/srep18035>.
- Zhang, Kai. 2016. "Gctf: Real-Time CTF Determination and Correction." *Journal of Structural Biology* 193 (1): 1–12. <https://doi.org/10.1016/j.jsb.2015.11.003>.
- Zhang, Xu, Heping Shi, Jiayi Wu, Xuewu Zhang, Lijun Sun, Chuo Chen, and Zhijian J. Chen. 2013. "Cyclic GMP-AMP Containing Mixed Phosphodiester Linkages Is an Endogenous High-Affinity Ligand for STING." *Molecular Cell* 51 (2): 226–35. <https://doi.org/10.1016/j.molcel.2013.05.022>.
- Zhao, Baoyu, Fenglei Du, Pengbiao Xu, Chang Shu, Banumathi Sankaran, Samantha L. Bell, Mengmeng Liu, et al. 2019. "A Conserved PLPLRT/SD Motif of STING Mediates the Recruitment and Activation of TBK1." *Nature* 569 (7758): 718–22. <https://doi.org/10.1038/s41586-019-1228-x>.
- Zhong, Bo, Yan Yang, Shu Li, Yan Yi Wang, Ying Li, Feici Diao, Caoqi Lei, et al. 2008. "The Adaptor Protein MITA Links Virus-Sensing Receptors to IRF3 Transcription Factor Activation." *Immunity* 29 (4): 538–50. <https://doi.org/10.1016/j.immuni.2008.09.003>.
- Zhou, Wen, Aaron T. Whiteley, Carina C. de Oliveira Mann, Benjamin R. Morehouse, Radosław P. Nowak, Eric S. Fischer, Nathanael S. Gray, John J. Mekalanos, and Philip J. Kranzusch. 2018. "Structure of the Human CGAS–DNA Complex Reveals Enhanced Control of Immune Surveillance." *Cell* 174 (2): 300–311.e11. <https://doi.org/10.1016/j.cell.2018.06.026>.
- Zhu, Deyu, Lijun Wang, Guijun Shang, Xue Liu, Jing Zhu, Defen Lu, Lei Wang, Biao Kan, Jing-ren Zhang, and Ye Xiang. 2014. "Structural Biochemistry of a *Vibrio Cholerae* Dinucleotide Cyclase Reveals Cyclase Activity Regulation by Folates." *Molecular Cell* 55 (6): 931–37. <https://doi.org/10.1016/j.molcel.2014.08.001>.
- Zhu, Jianzhong, Yugen Zhang, Arundhati Ghosh, Rolando A. Cuevas, Adriana Forero, Jayeeta Dhar, Mikkel Sørensen, et al. 2014. "Antiviral Activity of Human OASL Protein Is Mediated by Enhancing Signaling of the RIG-I RNA Sensor." *Immunity* 40 (6): 936–48. <https://doi.org/10.1016/J.IMMUNI.2014.05.007>.

Zilberstein, A, A Kimchi, A Schmidt, and M Revel. 1978. "Isolation of Two Interferon-Induced Translational Inhibitors: A Protein Kinase and an Oligo-Isadenylate Synthetase (Translation Control/Interferon/L Cells)." *Biochemistry*. Vol. 75.