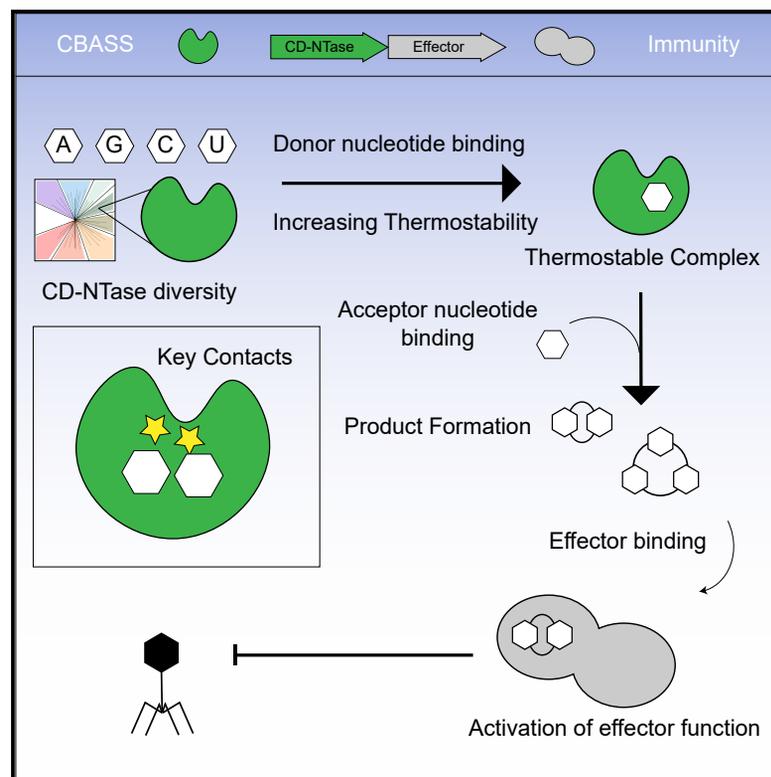


# Molecular basis of CD-NTase nucleotide selection in CBASS anti-phage defense

## Graphical abstract



## Authors

Apurva A. Govande,  
Brianna Duncan-Lowey,  
James B. Eaglesham, Aaron T. Whiteley,  
Philip J. Kranzusch

## Correspondence

philip\_kranzusch@dfci.harvard.edu

## In brief

CD-NTase enzymes synthesize nucleotide second messengers to control CBASS anti-phage defense. Govande et al. determine the structures of diverse CD-NTase enzymes, reveal key active-site contacts that control nucleotide selection, and establish rules to enable prediction of nucleotide signal specificity in diverse CBASS anti-phage systems.

## Highlights

- Structural analysis of the cAAG cyclic trinucleotide synthase *E. cloacae* CdnD
- Thermostable, pre-reactive complex formation controls CD-NTase nucleotide selection
- Nucleotide analogs reveal rules of substrate selection and product specificity
- Prediction of nucleotide signal specificity in diverse CBASS anti-phage systems



## Report

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

Apurva A. Govande,<sup>1,2</sup> Brianna Duncan-Lowey,<sup>1,2</sup> James B. Eaglesham,<sup>1,2</sup> Aaron T. Whiteley,<sup>3</sup> and Philip J. Kranzusch<sup>1,2,4,5,\*</sup>

<sup>1</sup>Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

<sup>2</sup>Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

<sup>3</sup>Department of Biochemistry, University of Colorado Boulder, Boulder, CO 80309, USA

<sup>4</sup>Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA 02115, USA

<sup>5</sup>Lead contact

\*Correspondence: [philip\\_kranzusch@dfci.harvard.edu](mailto:philip_kranzusch@dfci.harvard.edu)

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

## SUMMARY

cGAS/DncV-like nucleotidyltransferase (CD-NTase) enzymes are signaling proteins that initiate antiviral immunity in animal cells and cyclic-oligonucleotide-based anti-phage signaling system (CBASS) phage defense in bacteria. Upon phage recognition, bacterial CD-NTases catalyze synthesis of cyclic-oligonucleotide signals, which 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 *Enterobacter 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 the formation of a thermostable protein-nucleotide complex, and we 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

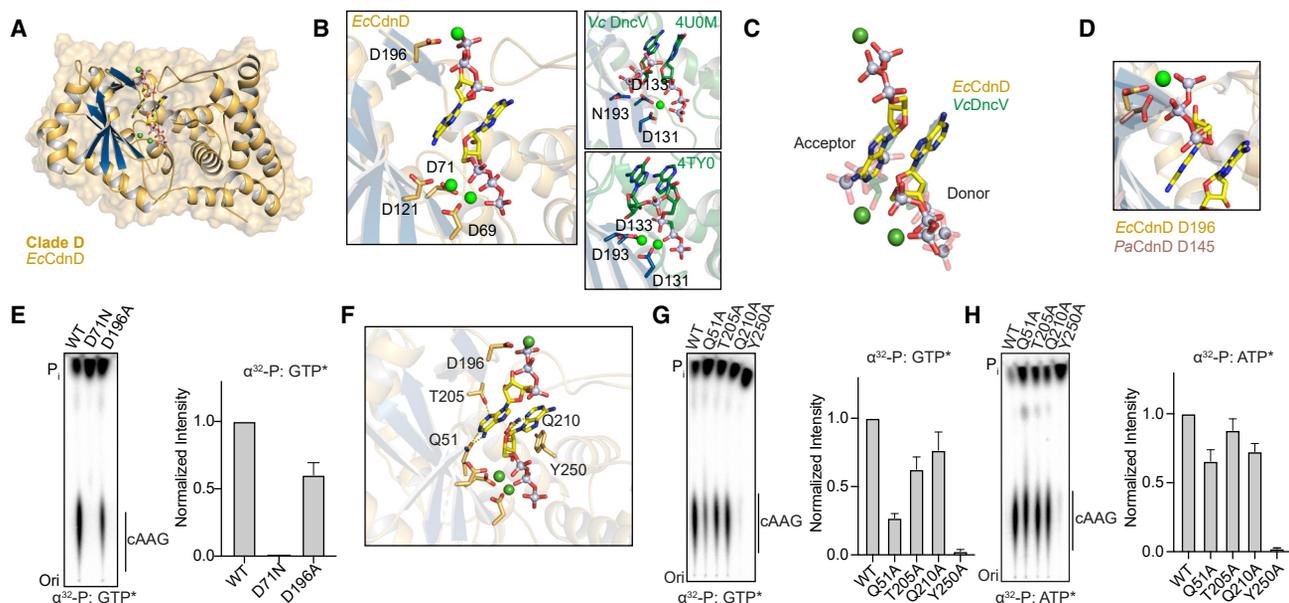
cGAS/DncV-like nucleotidyltransferase (CD-NTase) enzymes are signaling proteins in bacteria and animals, which sense viral infection and synthesize nucleotide second messengers to initiate antiviral defense (Kranzusch, 2019; Whiteley et al., 2019). In bacteria, CD-NTases function as part of defense operons, named cyclic-oligonucleotide-based anti-phage signaling systems (CBASS), which provide bacteriophage resistance (Cohen et al., 2019). After 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 (Cap) 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 cyclic GMP-AMP synthase and stimulator of interferon genes (cGAS-STING) immune pathway, which is 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 guanosine monophosphate (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, which necessitates CD-NTase enzymes that catalyze rapid synthesis of the correct nucleotide product to initiate antiviral defense (Lau et al., 2020; Lowey et al., 2020; Morehouse et al., 2020; Whiteley et al., 2019). Crystal structures of bacterial CD-NTases define enzyme features that direct nucleotide second-messenger formation (Kranzusch et al., 2014; Whiteley et al., 2019; Ye et al., 2020; Zhu et al., 2014), but current understanding is insufficient to predict product diversity across distantly related CD-NTases.

Here, we combine biochemical and structural biology approaches to characterize CD-NTase nucleotide second-messenger formation and define the determinants of nucleotide selectivity. We determine a crystal structure of the cyclic trinucleotide synthase *E. cloacae* CdnD, which enables a direct comparison with the cyclic dinucleotide synthase *V. cholerae* DncV.





**Figure 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 the 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 pppA[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, whereas 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 with 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 the third metal-ion-coordinating residue D196. D71N mutation abrogates synthesis of 3'3'3'-cAAG, whereas the D196A mutation does not prevent product formation but may show a slight defect.

(F) Detail of the 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.

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 before 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 side chains lining the top of the active site (Figure 1A) (Whiteley et al., 2019). Previous structures of CD-NTase-nucleotide complexes define “donor” and “acceptor” nucleotide pockets critical for substrate coordination and the sequential steps of product cyclization (Civril et al., 2013; Gao et al., 2013; Kranzusch et al., 2014, 2019; Whiteley et al., 2019; 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 1B, left). The *EcCdnD* donor ATP adopts a similar orientation to the donor nucleotide in structures of the 3'3'3'-cGAMP synthase *V. cholerae* DncV (*VcDncV*) (Figures 1C and S1A, inset) (Kato et al., 2015; Kranzusch et al., 2014; Zhu et al., 2014). In contrast, the ATP within the acceptor pocket of *EcCdnD* is

held in an unexpected, inverted orientation, which is rotated 120° relative to the VcDncV acceptor nucleotide (Figures 1B and 1C).

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 C $\alpha$  root-mean-square deviation [RMSD]) in complex with two molecules of guanosine triphosphate (GTP) (Table S1; Figure S1B). The SeCdnD-GTP structure reveals nearly identical nucleotide positioning with an inverted acceptor nucleotide (Figure S1B), 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 (Figures 1B and 1C). The inverted acceptor nucleotide is further coordinated by an auxiliary metal-binding site created between D196 and the nucleotide triphosphate (Figure 1D). EcCdnD and SeCdnD are most closely related to previous structures of *Pseudomonas aeruginosa* CdnD (PaCdnD, 3.1 Å C $\alpha$  RMSD) and *Escherichia coli* CdnC (2.7 Å C $\alpha$  RMSD), which 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 (Figures 1D and 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 S1D). 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 1E). In contrast, a D71N mutation of the main active site results in complete abrogation of EcCdnD enzymatic activity (Figure 1E).

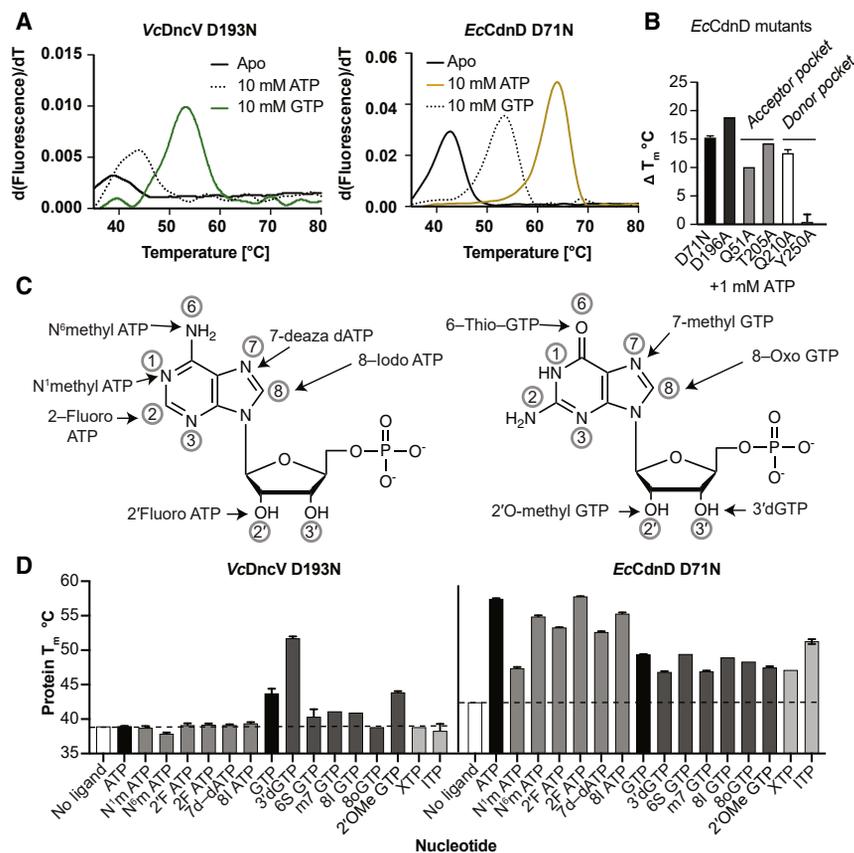
We next analyzed nucleotide contacts in the active site of EcCdnD (Figure 1F). In the donor pocket, Y250 stacks with the bound ATP nucleotide and stabilizes the base, whereas Q210 hydrogen bonds with the ATP N3 position. Mutation of Y250 leads to a complete disruption of 3'3'3'-cAAG synthesis, which is not the result of reduced enzyme stability (Figures 1G and S1E), whereas 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 between Q51 and the N1 and N6 positions of ATP (Figure 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 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 S1F). 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; 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 uses GTP in the donor nucleotide position to form a pppA[3'-5']pG linear intermediate (Kranzusch et al., 2014; Zhu et al., 2014). In the presence of the correct initiating donor nucleotide GTP, we observed significant stabilization with the melting temperature ( $T_m$ ) of VcDncV shifting ~13°C (Figure 2A). In contrast, the  $T_m$  of VcDncV shifts only 4°C in the presence of ATP (Figure 2A). We observed no synergistic effect when equimolar amounts of ATP and GTP are present, indicating enhanced thermostability is dependent on donor nucleotide interactions (Figure S1G). Nucleotides exhibit a similar ability to stabilize the  $T_m$  of EcCdnD, with ATP resulting in a ~21°C shift and GTP resulting in a 9°C shift (Figures 2A and S1G). 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 (Figures 2B and S1H). In contrast, mutations to the acceptor pocket and auxiliary-metal-binding site again demonstrate these interactions have a minor effect on thermo-stabilization (Figures 2B, S1G, and S1H). These results reveal that a key determinant of CD-NTase nucleotide selectivity is preferential interactions that lead to stabilization of an enzyme-substrate pre-reaction state.

Building on previous work with nucleotide-analog interference mapping (Dash et al., 2006; Launer-Felty and Strobel, 2018; Rausch et al., 2009; Wang et al., 2021), we next used a panel of 14 nucleotide analogs to determine how individual contacts affect CD-NTase substrate selectivity. The nucleotide panel included ATP and GTP analogs with nucleobase and ribose chemical modifications (Figures 2C, S1I, and S1J) and revealed major features controlling CD-NTase specificity. First, the data demonstrate that both VcDncV and EcCdnD are tolerant of a wide variety of nucleotide modifications (Figure 2D). Chemical modifications to the Watson-Crick and Hoogsteen edges do not disrupt binding and enhancement of CD-NTase thermostability (Figures 2D and S1K). Compared with previous studies with GGDEF- and DisA-family enzymes of cyclic dinucleotide syntheses (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 that hypothesis, we observed that the EcCdnD active site used for cyclic trinucleotide synthesis is more accommodating of modified nucleotides compared with the VcDncV active site used for cyclic dinucleotide synthesis (Figure 2D).

A second major feature of CD-NTase nucleotide specificity is preferential interaction with the nucleobase Watson-Crick



**Figure 2. 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 does. Right:  $T_m$  plot for catalytically inactive EcCdnD D71N, showing that ATP stabilizes to a greater extent than GTP does. Data are representative of three independent experiments.

(B)  $\Delta T_m$  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.

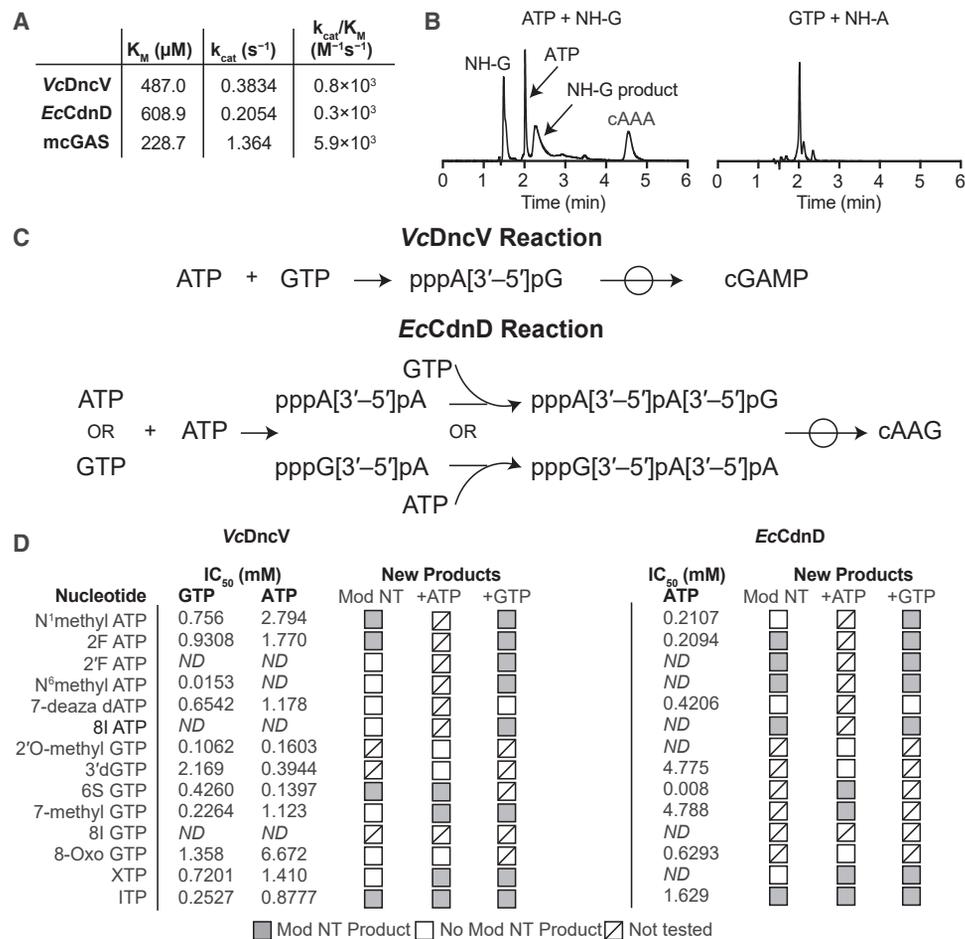
(D) Comparison of protein  $T_m$  for VcDncV D193N and EcCdnD D71N apo or bound to 1 mM ATP, GTP, ITP, XTP, or modified nucleotide as indicated. Data plotted are the mean of two independent experiments, and error bars denote the standard deviation.

edge. Both VcDncV and EcCdnD maintain specific interactions with the ATP N<sup>1</sup> position and the GTP N<sup>1</sup> and N<sup>2</sup> positions. VcDncV in complex with inosine triphosphate (ITP), which shares the N<sup>1</sup> and O<sup>6</sup> positions with GTP but lacks the amine group at position 2, has a similar  $T_m$  to that of the VcDncV-ATP complex (Figure 2D). The enhanced VcDncV thermostability upon interaction with GTP is, therefore, dependent on hydrogen-bond contacts with the N<sup>2</sup> 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 2D). Xanthine triphosphate (XTP), which adds an additional oxygen atom to the N<sup>2</sup> position and includes a hydrogen bond donor at both positions 1 and 3, destabilizes VcDncV and slightly increases the stability of EcCdnD (Figures 2D and S1K). For the ATP Watson-Crick edge, modifications to the N<sup>1</sup> position destabilize the interaction with both VcDncV and EcCdnD, whereas modifications to the N<sup>6</sup> position destabilize the interaction with VcDncV but only have a minimal effect on EcCdnD stability (Figures 2D and S1K). Strikingly, modification of the N<sup>2</sup> position of ATP with 2-fluoro-ATP does not impair the ability of ATP to stabilize either enzyme (Figures 2D and S1K). Overall, these interactions with modified nucleotides agree with contacts observed in the crystal structures of EcCdnD and VcDncV intermediate states (Figure 1) (Kranzusch et al., 2014; Zhu et al., 2014) and demonstrate specific CD-NTase donor and acceptor pocket interactions controlling nucleotide selection.

Lastly, nucleotide analogs provide insight into the importance of the nucleobase positions that do not form clear contacts in existing CD-NTase structures. Substituting the GTP O<sup>6</sup> hydrogen-bond acceptor position with 6-thio-GTP results in a decrease in the stability of VcDncV but not EcCdnD (Figure 2D). Masking the N<sup>7</sup> purine ring position destabilizes the interaction with both CD-NTases, suggesting transient recognition of this position could have a role in discriminating against pyrimidine nucleotides (Figures 2D and S1K). Notably, chemical modifications at nucleobase position 8 reveal additional differences between VcDncV and EcCdnD with 8-iodo-GTP and 8-oxo-GTP stabilizing EcCdnD, whereas only 8-iodo-GTP stabilizes VcDncV and 8-oxo-GTP has a specific destabilizing effect (Figure 2D). Likewise, VcDncV tolerates ribose 2' or 3' OH modifications, whereas EcCdnD is destabilized (Figure 2D). 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  $\sim 10$  times slower than DNA-activated mouse



**Figure 3. 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 ( $k_{\text{cat}}/K_M$ ) of the bacterial CD-NTases under these conditions is  $\sim 10$  times less than that of 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%  $\text{NaH}_2\text{PO}_4$  and 6% acetonitrile. Data are representative of three independent experiments.

(C) Schematic comparison of reaction order for VcDncV 3'3'3'-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,  $\text{IC}_{50}$  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.

cGAS or human cGAS (Figures 3A, S2A, and S2B) (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 3B; Table S2). Furthermore, in the presence of ATP and Apcpp, EcCdnD forms a linear reaction intermediate, whereas GTP and Gpcpp are not sufficient for EcCdnD product formation (Figure S2C). These results confirm that EcCdnD uses ATP as the initial donor nucleotide, followed by selection of GTP, and reveal that cyclic trinucle-

otide synthesizing CD-NTases also use a defined reaction order (Figure 3C) (Kranzusch et al., 2014; Whiteley et al., 2019).

As previously demonstrated for VcDncV (Laurer-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 (Figures 3D and S3; Table S2). For each nucleotide analog, we calculated the  $\text{IC}_{50}$  as the concentration of nucleotide required to decrease product formation by 50% (Figure 3D; Data S1). In each case, nucleotide-analog incorporation was more efficient in the presence of the correct ATP or GTP partner nucleotide, indicating maintenance of the preference for synthesizing mixed-base products. The reactions with nucleotide analogs largely fall into

three groups: (1) 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*; (2) nucleotides capable of binding but not reacting (no modified product formation,  $IC_{50} < 0.5$  mM), including 2'O-methyl GTP for *VcDncV* and N<sup>1</sup> methyl ATP for *EcCdnD*; and (3) 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 3D; Data S1G). Together, these results confirm our thermostability findings and further explain the donor- and acceptor-pocket interactions controlling CD-NTase product specificity.

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

Bacterial CD-NTase enzymes are remarkably diverse and cluster into eight major clades, designated A–H (Millman et al., 2020; Whiteley et al., 2019). We used the *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 4B; Table S1). 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 the 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 4A). Lid residues correlate with major product synthesis (Figures 4B and 4C; Table S3), 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 4C), (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 that helix are a critical determinant for nucleobase incorporation (Figure 4C; Table S3).

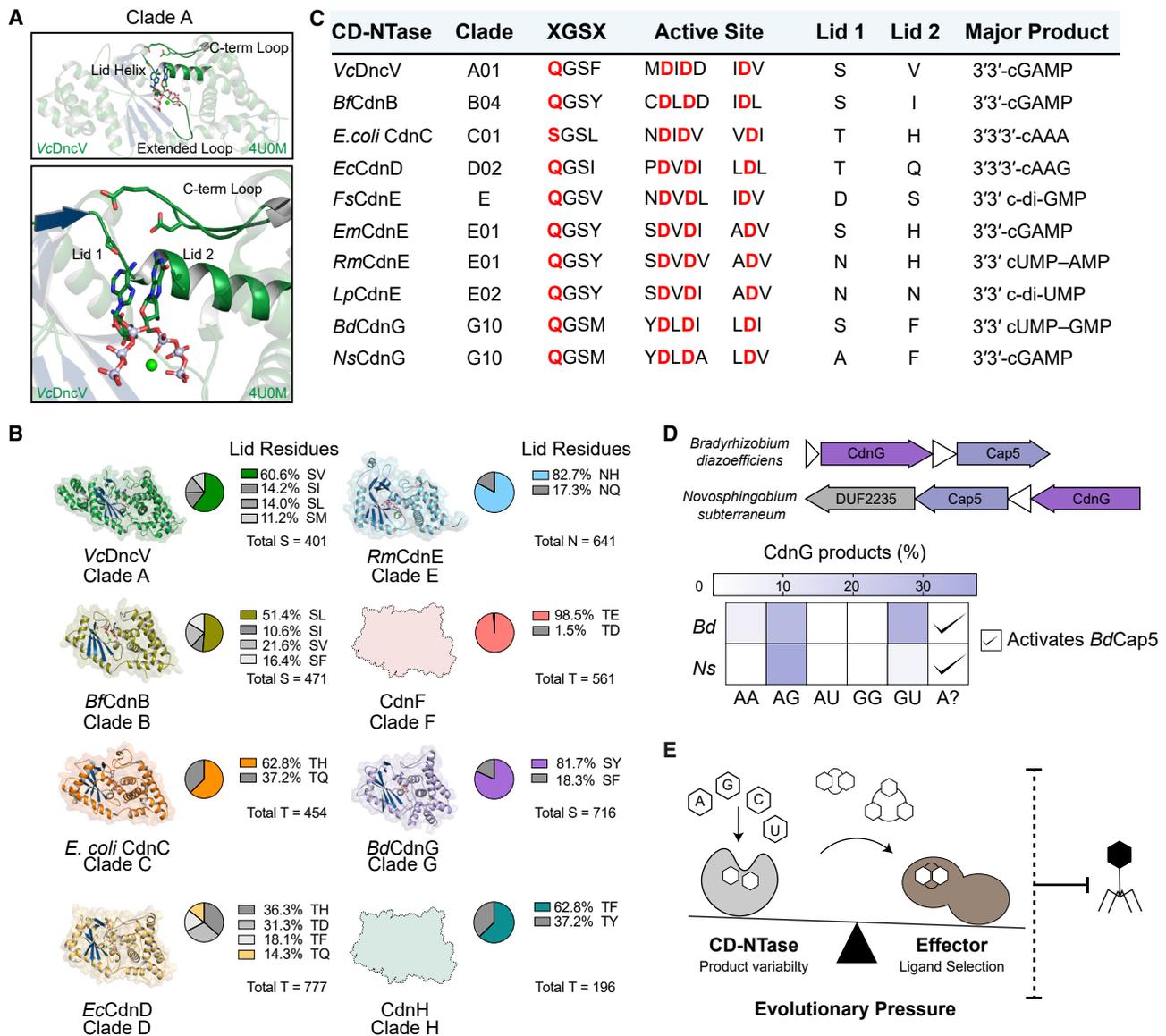
Previous large-scale analysis of nucleic-acid-binding proteins demonstrates a close relationship between side-chain identity and nucleobase interactions (Nobeli et al., 2001). Using experimentally determined CD-NTase product specificity and known protein-nucleobase interaction patterns as a guide, we next analyzed all CD-NTase enzyme sequences and developed rules to enable prediction of CD-NTase product specificity (Table S3; STAR 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 S4A). *BdCdnG* encodes a serine at lid helix position 1 similar to *VcDncV*, *BfCdnB*, and *EmCdnE* (Figure 4C),

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 S4B). 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, whereas *NsCdnG* synthesizes primarily 3'3'-cGAMP (Figures 4C and S4C–S4E).

The anti-phage activity of CBASS operons depends on 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 the 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 (Figures 4D and S4F–S4H). 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 (Figures 4D, S4G, and S4H), as well as *NsCdnG* reactions, suggesting that *BdCap5* recognizes a minor product that is conserved between closely related CD-NTase enzymes (Figure S4H). *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.

### 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 for 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 a 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, which result in the 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 4), supporting the idea 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 the primary CD-NTase protein sequence alone.



**Figure 4. Structural and sequence-based comparison of divergent bacterial CD-NTases reveals correlates of nucleotide selection**

(A) VcDncV structure PDB: 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 (i.e., Q51 in EcCdnD), the 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 heatmap 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 because effector activation is critical for an anti-phage response and survival of the bacterial colony.

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) (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 4E). Overall, our work further establishes the determinants of CD-NTase specificity that controls nucleotide second-messenger formation and demonstrates how antiviral signaling systems are balanced for both functional defense and adaptability.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Protein expression and purification
  - Crystallization and structure determination
  - Thin-layer chromatography
  - Thermal denaturation assay
  - HPLC enzymatic reaction analysis
  - Mass spectrometry analysis
  - CD-NTase enzyme kinetics
  - BdCap5 nuclease assay
  - Sequence-based determination of CD-NTase product specificity
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2021.109206>.

## ACKNOWLEDGMENTS

The authors are grateful to A. Lee, K. Chat, and members of the Kranzusch laboratory for helpful discussion; and A. Engelman, D. Pearson, M. Danielsen, and D. Malhihero 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 grad-

uate research fellowship, and 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

Experiments were designed by A.A.G. and P.J.K. All enzyme activity experiments, nucleotide-binding studies, and data analysis were conducted by A.A.G. with assistance from B.L. X-ray crystallography data were collected by A.A.G., B.L., J.B.E., and P.J.K., and molecular modeling was completed by A.A.G. and B.L. A.T.W. assisted with initial enzyme-activity experiments and data analysis. The manuscript was written by A.A.G. and P.J.K., and all authors contributed to editing the manuscript and support the conclusions.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 29, 2021

Revised: April 9, 2021

Accepted: May 11, 2021

Published: June 1, 2021

## REFERENCES

- Burroughs, A.M., Zhang, D., Schäffer, D.E., Iyer, L.M., and Aravind, L. (2015). Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* *43*, 10633–10654.
- Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U., and Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. *Proc. Natl. Acad. Sci. USA* *101*, 17084–17089.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 12–21.
- Civril, F., Deimling, T., de Oliveira Mann, C.C., Ablasser, A., Moldt, M., Witte, G., Hornung, V., and Hopfner, K.P. (2013). Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* *498*, 332–337.
- Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* *574*, 691–695.
- Dash, C., Fisher, T.S., Prasad, V.R., and Le Grice, S.F. (2006). Examining interactions of HIV-1 reverse transcriptase with single-stranded template nucleotides by nucleoside analog interference. *J. Biol. Chem.* *281*, 27873–27881.
- Davies, B.W., Bogard, R.W., Young, T.S., and Mekalanos, J.J. (2012). Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell* *149*, 358–370.
- Eaglesham, J.B., Pan, Y., Kupper, T.S., and Kranzusch, P.J. (2019). Viral and metazoan poxins are cGAMP-specific nucleases that restrict cGAS-STING signalling. *Nature* *566*, 259–263.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* *60*, 2126–2132.
- Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B.L., Zillinger, T., Serganov, A.A., Liu, Y., Jones, R.A., Hartmann, G., et al. (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.
- Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. *Science* *336*, 1030–1033.
- Kato, K., Ishii, R., Hirano, S., Ishitani, R., and Nureki, O. (2015). Structural basis for the catalytic mechanism of DncV, bacterial homolog of cyclic GMP-AMP synthase. *Structure* *23*, 843–850.

- Kranzusch, P.J. (2019). cGAS and CD-NTase enzymes: structure, mechanism, and evolution. *Curr. Opin. Struct. Biol.* *59*, 178–187.
- Kranzusch, P.J., Lee, A.S.Y., Wilson, S.C., Solovykh, M.S., Vance, R.E., Berger, J.M., and Doudna, J.A. (2014). Structure-guided reprogramming of human cGAS dinucleotide linkage specificity. *Cell* *158*, 1011–1021.
- Lau, R.K., Ye, Q., Birkholz, E.A., Berg, K.R., Patel, L., Mathews, I.T., Watrous, J.D., Ego, K., Whiteley, A.T., Lowey, B., et al. (2020). Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity. *Mol. Cell* *77*, 723–733.e6.
- Lauer-Felty, K.D., and Strobel, S.A. (2018). Enzymatic synthesis of cyclic dinucleotide analogs by a promiscuous cyclic-AMP-GMP synthetase and analysis of cyclic dinucleotide responsive riboswitches. *Nucleic Acids Res.* *46*, 2765–2776.
- Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczy, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* *75*, 861–877.
- Lowey, B., Whiteley, A.T., Keszei, A.F.A., Morehouse, B.R., Mathews, I.T., Antine, S.P., Cabrera, V.J., Kashin, D., Niemann, P., Jain, M., et al. (2020). CBASS immunity uses CARF-related effectors to sense 3′-5′- and 2′-5′-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* *182*, 38–49.e17.
- Millman, A., Melamed, S., Amitai, G., and Sorek, R. (2020). Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. *Nat. Microbiol.* *5*, 1608–1615.
- Morehouse, B.R., Govande, A.A., Millman, A., Keszei, A.F.A., Lowey, B., Ofir, G., Shao, S., Sorek, R., and Kranzusch, P.J. (2020). STING cyclic dinucleotide sensing originated in bacteria. *Nature* *586*, 429–433.
- Nobell, I., Laskowski, R.A., Valdar, W.S., and Thornton, J.M. (2001). On the molecular discrimination between adenine and guanine by proteins. *Nucleic Acids Res.* *29*, 4294–4309.
- Rausch, J.W., Chelico, L., Goodman, M.F., and Le Grice, S.F. (2009). Dissecting APOBEC3G substrate specificity by nucleoside analog interference. *J. Biol. Chem.* *284*, 7047–7058.
- Rosenberg, J., Dickmanns, A., Neumann, P., Gunka, K., Arens, J., Kaever, V., Stülke, J., Ficner, R., and Commichau, F.M. (2015). Structural and biochemical analysis of the essential diadenylate cyclase CdaA from *Listeria monocytogenes*. *J. Biol. Chem.* *290*, 6596–6606.
- Wang, Z., Zhao, C., Wang, C., Zhang, H., Ma, D., Zhang, Q., Wen, X., Li, L., and Xi, Z. (2021). Synthesis and biological evaluation of all possible inosine-mixed cyclic dinucleotides that activate different hSTING variants. *Bioorg. Med. Chem.* *29*, 115899.
- Weiss, M.S. (2001). Global indicators of X-ray data quality. *J. Appl. Crystallogr.* *34*, 130–135.
- Whiteley, A.T., Eaglesham, J.B., de Oliveira Mann, C.C., Morehouse, B.R., Lowey, B., Nieminen, E.A., Danilchanka, O., King, D.S., Lee, A.S.Y., Mekalanos, J.J., and Kranzusch, P.J. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* *567*, 194–199.
- Witte, G., Hartung, S., Büttner, K., and Hopfner, K.P. (2008). Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol. Cell* *30*, 167–178.
- Ye, Q., Lau, R.K., Mathews, I.T., Birkholz, E.A., Watrous, J.D., Azimi, C.S., Pogliano, J., Jain, M., and Corbett, K.D. (2020). HORMA domain proteins and a Trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity. *Mol. Cell* *77*, 709–722.e7.
- Zhou, W., Whiteley, A.T., de Oliveira Mann, C.C., Morehouse, B.R., Nowak, R.P., Fischer, E.S., Gray, N.S., Mekalanos, J.J., and Kranzusch, P.J. (2018). Structure of the human cGAS-DNA complex reveals enhanced control of immune surveillance. *Cell* *174*, 300–311.e11.
- Zhu, D., Wang, L., Shang, G., Liu, X., Zhu, J., Lu, D., Wang, L., Kan, B., Zhang, J.R., and Xiang, Y. (2014). Structural biochemistry of a *Vibrio cholerae* dinucleotide cyclase reveals cyclase activity regulation by folates. *Mol. Cell* *55*, 931–937.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
<i>E. coli</i> BL21-DE3 RIL	Agilent	230245
<i>E. coli</i> MG1655	Coli Genetic Stock Center	CGSC6300
<b>Chemicals, peptides, and recombinant proteins</b>		
Ni-NTA Agarose	QIAGEN	30250
HiLoad 16/600 Superdex 75 PG	GE Healthcare	28989333
HiLoad 16/600 Superdex 200 PG	GE Healthcare	28989336
Zorbax Bonus-RP	Agilent	863668-901
[ $\alpha$ - <sup>32</sup> P] ATP	Perkin Elmer	BLU003H250UC
[ $\alpha$ - <sup>32</sup> P] GTP	Perkin Elmer	BLU006H250UC
PEI-Cellulose F TLC plate	EMD Biosciences	EM1.05579.0001
Alkaline Phosphatase, Calf Intestinal (CIP)	New England Biolabs	M0290S
ATP, GTP, UTP, CTP	New England Biolabs	N0450S
HEPES	VWR	97061-824
Tris base	VWR	97062-416
PEG-400	Sigma-Aldrich	202398
PEG-3350	Sigma-Aldrich	202444
Imidazole	VWR	97065-016
Tris[2-carboxyethyl] phosphine hydrochloride (TCEP)	GoldBio	TCEP50
Ethylene glycol	VWR	97061-964
3'3'3'-cAAA	Biolog Life Science Institute	362
3'3'3'-cAAG	Biolog Life Science Institute	361
3'3'-cGAMP	Biolog Life Science Institute	117
3'3'-cAMP-UMP	Biolog Life Science Institute	357
3'3'-cGMP-UMP	Biolog Life Science Institute	371
c-di-AMP	Biolog Life Science Institute	088
<i>EcCdnD</i> recombinant protein, and mutants as described	This paper	N/A
<i>VcDncV</i> recombinant protein, and D193N mutant	This paper	N/A
<i>BdCdnG</i> recombinant protein	This paper	N/A
<i>BdCap5</i> recombinant protein	<a href="#">Lowey et al., 2020</a>	N/A
<i>NsCdnG</i> recombinant protein	This paper	N/A
<i>SeCdnD</i> recombinant protein	This paper	N/A
<i>BfCdnB</i> recombinant protein	This paper	N/A
<b>Critical commercial assays</b>		
DNA Extraction Kit	QIAGEN	69504
<b>Deposited data</b>		
<i>EcCdnD</i> -2 × ATP	This paper	PDB: 7LJL
<i>SeCdnD</i> -2 × GTP	This paper	PDB: 7LJM
<i>BfCdnB</i> -ADP	This paper	PDB: 7LJO
<i>BdCdnG</i> -GTP	This paper	PDB: 7LJN
<i>Vibrio cholerae</i> di-nucleotide cyclase (DncV) mutant D193N in complex with ATP, GTP and 5MTHFGLU2	<a href="#">Zhu et al., 2014</a>	PDB:4U0M
<i>Vibrio cholerae</i> DncV cyclic AMP-GMP synthase in complex with linear intermediate 5' pppA(3',5')pG	<a href="#">Kranzusch et al., 2014</a>	PDB:4TY0

(Continued on next page)

<b>Continued</b>		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>E. coli</i> MS115-1 CdnC + ATP	Ye et al., 2020	PDB: 6P80
<i>P. aeruginosa</i> ATCC27853 CdnD, Apo form 1	Ye et al., 2020	PDB: 6P82
<i>Rhodothermus marinus</i> CdnE c-UMP-AMP synthase with Apcpp and Upnpp	Whiteley et al., 2019	PDB: 6E0L
<b>Oligonucleotides</b>		
DNA45 sense 5'-TACAGATCTACTAGTGATCTATGACTG ATCTGTACATGATCTACA-3'	Integrated DNA Technologies	N/A
DNA45 antisense 5'-TGTAGATCATGTACAGATCAGTCATAG ATCACTAGTAGATCTGTA-3'	Integrated DNA Technologies	N/A
<b>Recombinant DNA</b>		
pET16SUMO2- <i>Ec</i> CdnD	Lowey et al., 2020	N/A
pET16SUMO2- <i>Se</i> CdnD	This paper	N/A
pET16SUMO2- <i>Vc</i> DncV	Whiteley et al., 2019	N/A
pET16SUMO2- <i>Bd</i> CdnG	Whiteley et al., 2019	N/A
pET16SUMO2- <i>Bf</i> CdnB	Whiteley et al., 2019	N/A
pET16SUMO2- <i>Ns</i> CdnG	This paper	N/A
pET16SUMO2- <i>Bd</i> Cap5	Lowey et al., 2020	N/A
<b>Software and algorithms</b>		
Phenix 1.18.2-3874	Liebschner et al., 2019	<a href="https://www.phenix-online.org/">https://www.phenix-online.org/</a>
Coot 0.8.9.3-pre	Emsley and Cowtan, 2004	<a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a>
Pymol v2.4.0	Schrödinger, LLC	<a href="https://pymol.org/2/">https://pymol.org/2/</a>
Prism 9.0d	GraphPad software	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Philip Kranzusch ([philip\\_kranzusch@dfci.harvard.edu](mailto:philip_kranzusch@dfci.harvard.edu)).

### 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 Na<sub>2</sub>PO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 0.25% aspartic acid, 100 mg mL<sup>-1</sup> ampicillin, 34 mg mL<sup>-1</sup> chloramphenicol, and trace metals) and used to inoculate 1 L M9ZB media cultures (0.5% glycerol, 1% CAS amino acids, 47.8 mM Na<sub>2</sub>PO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 18.7 mM NH<sub>4</sub>Cl, 85.6 mM NaCl, 2 mM MgSO<sub>4</sub>, 100 mg mL<sup>-1</sup> ampicillin, 34 mg mL<sup>-1</sup> 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 OD<sub>600</sub> ~0.05. Cultures were grown at 37°C, 230 RPM until OD<sub>600</sub> reached ~2.2, chilled on ice for 10 min, induced with 0.5 mM IPTG, and incubated at 16°C, 230 RPM for ~16 h before harvest. Cultures for production of selenomethionine-substituted (SeMet) protein were grown as previously described (Eaglesham et al., 2019). Briefly, SeMet protein was grown in overnight MDG starter cultures and used to inoculate 1 L induction cultures of M9 medium (47.8 mM Na<sub>2</sub>PO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 18.7 mM NH<sub>4</sub>Cl, 85.6 mM NaCl, 2 mM MgSO<sub>4</sub>, 100 mg mL<sup>-1</sup> ampicillin, 34 mg mL<sup>-1</sup> chloramphenicol, 1 μg mL<sup>-1</sup> 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<sup>-1</sup> leucine, isoleucine, and valine (VWR), 100 mg L<sup>-1</sup> phenylalanine, lysine, and threonine (VWR), and 75 mg L<sup>-1</sup> 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<sup>-1</sup>, 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<sup>-1</sup>, 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<sup>-1</sup> with 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP, supplemented with 10.5 mM MgCl<sub>2</sub> and 5 mM ATP, mixed 1:1 with reservoir solution (100 mM MES pH 6.5, 19%–21% PEG-1000) and allowed to grow for 3 days. Crystals were further optimized with micro-seeding, grown for 2 days and then harvested by cryoprotecting with reservoir solution supplemented with 17.5% ethylene glycol MgCl<sub>2</sub> and ATP prior to freezing in liquid nitrogen; 2) *SeCdnD* + GTP: Native protein was diluted to 8.25 mg mL<sup>-1</sup>, supplemented with 10.5 mM MgCl<sub>2</sub> 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 mL<sup>-1</sup> (Agilent), supplemented with 10 mM MgCl<sub>2</sub> 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 mL<sup>-1</sup> (Agilent) with 20 mM HEPES-KOH, 10 mM KCl, 1 mM TCEP, supplemented with 10.5 mM MgCl<sub>2</sub>, 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](#) (Chen et al., 2010; Karplus and Diederichs, 2012; Weiss, 2001).

### Thin-layer chromatography

CD-NTase reactions for nucleotide second messenger product characterization were performed in 20 μL reactions with 50 μM enzyme, 12.5 μM NTP and trace α<sup>32</sup>P-ATP or α<sup>32</sup>P-GTP, as indicated. *VcDncV* reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM MgCl<sub>2</sub>,

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  $IC_{50}$  calculations were carried out for 15 min at 37°C before heat inactivation at 95°C for 2 min. Prior to thin-layer chromatography analysis, all reactions were treated with 0.5  $\mu$ L of Quick CIP (NEB) at 37°C for 30 min. Reactions were analyzed by thin-layer chromatography by spotting 1  $\mu$ L of each phosphatase-treated reaction on a PEI-cellulose plate (Millipore) and developed in 1.5 M  $KH_2PO_4$  pH 3.8 until buffer had reached  $\sim$ 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 wild-type enzyme where indicated. For  $IC_{50}$  calculation, pixel intensity corresponding to radiolabeled c-di-AMP, c-di-GMP (*VcDncV*), or 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  $\mu$ M) and normalized to reactions containing no modified nucleotide. Inhibition data were fit using a non-linear regression curve ([Inhibitor] versus 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  $\mu$ L reactions containing 10  $\mu$ M protein and no nucleotide, or a range of nucleotide concentrations as indicated, 0.001 – 20  $\mu$ M. *VcDncV* reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT. Protein and nucleotide solutions were incubated on ice for at least 45 min until measuring melting temperature of 10  $\mu$ L of reaction between 35–95°C over 3 min.

### HPLC enzymatic reaction analysis

For HPLC analysis, CD-NTase reactions were performed in 20  $\mu$ L reactions with 50  $\mu$ M enzyme and 500  $\mu$ M NTPs. *VcDncV* reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT. Reactions were incubated at 37°C overnight for  $\sim$ 16 h, heat inactivated at 95°C for 2 min, diluted to 200  $\mu$ 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- $\mu$ m) and Agilent 1200 Infinity Series LC system. 20  $\mu$ 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  $NaH_2PO_4$  pH 6.8 supplemented with 3% HPLC grade acetonitrile for *VcDncV* reactions and 6% acetonitrile for *EcCdnD* reactions. For quick-CIP treated reactions, 20  $\mu$ L of heat inactivated reaction was treated with 1  $\mu$ 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  $\mu$ 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  $\mu$ M protein (*VcDncV*, *EcCdnD*, or *mcGAS*) and nucleotide concentrations 125, 250, 500, 750  $\mu$ M, or 1 mM ATP and GTP. *VcDncV* and *mcGAS* reactions contained 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT and *EcCdnD* reactions contained 50 mM Tris-HCl pH 9.0, 10 mM NaCl, 5 mM  $MgCl_2$ , 1 mM DTT. For each nucleotide concentration, 100  $\mu$ L reactions were incubated at 37°C and 20  $\mu$ 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  $\mu$ L aliquots were diluted to 200  $\mu$ 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  $\mu$ L of each sample was then injected and reactions were analyzed by HPLC as detailed above. Absorbance units were converted to  $\mu$ M by comparing to a standard curve from 10–750  $\mu$ 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  $\mu$ L<sup>-1</sup> pGEM9z plasmid (Promega) or 10 ng  $\mu$ L<sup>-1</sup> pcDNA4 plasmid in a 10  $\mu$ L reaction for 30 min at 37°C containing 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM  $MgCl_2$ , and 1 mM TCEP. Nuclease reactions were supplemented with filtered CD-NTase product reactions or chemically synthesized nucleotide second messengers at 1 pM, 10 pM, 100 pM, 1 nM, 10 nM or 100 nM as indicated. Reactions were separated on a 2% TAE agarose gel containing 10 mg mL<sup>-1</sup> 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 “[Table S3](#), 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. \(2019\)](#) and [Table S2](#). 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 (XDXX, 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 [Table S3](#). 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, [Table S3](#) 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 sub-clade 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. \(2020\)](#); and [Whiteley et al. \(2019\)](#).

### QUANTIFICATION AND STATISTICAL ANALYSIS

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