

Activation of CBASS Cap5 endonuclease immune effector by cyclic nucleotides

O. Rechkoblit, D. Kreitler

To be published in "Nature Structural & Molecular Biology"

February 2024

Photon Sciences

Brookhaven National Laboratory

U.S. Department of Energy

USDOE Office of Science (SC), Basic Energy Sciences (BES)

Notice: This manuscript has been authored by employees of Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy. The publisher by accepting the manuscript for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Activation of CBASS-Cap5 endonuclease immune effector by cyclic nucleotides: A view at high resolution

(Key words: cyclic dinucleotide, structure, bacterial immunity, CBASS, Cap5 effector endonuclease, viral defense).

Olga Rechkoblit^{1*}, Daniela Sciaky¹, Dale F. Kreitler², Angeliki Buku¹, Jithesh Kottur¹ and Aneel K. Aggarwal^{1*}

¹Department of Pharmacological Sciences
Icahn School of Medicine at Mount Sinai
New York, NY, USA

²Center for BioMolecular Structure
NSLS-II, Brookhaven National Laboratory
Upton, NY, USA

*Correspondence: olga.rechkoblit@mssm.edu (O.R.),
aneel.agarwal@mssm.edu (A.K.A.)

Summary

Many new defense systems have been discovered in bacteria recently, including the widespread cyclic oligonucleotide-based antiphage signaling system (CBASS) system with structural and functional similarities to the cGAS-STING system in humans. Both systems are defined by an enzyme that synthesizes a cyclic nucleotide upon viral infection and an effector protein that senses the second messenger to evoke the anti-viral response. The CBASS effectors do not activate transcription like STING, but instead kill the infected cell to prevent a viral epidemic within the bacterial cell population. Cap5, containing a SAVED domain coupled to an HNH DNA endonuclease domain, is the most abundant of CBASS effectors, yet the mechanism by which it becomes activated for cell killing is still not understood. We present here for the first-time high-resolution structures of full-length Cap5 from *Pseudomonas syringae* in complex with second messengers. We show that the key to PsCap5 activation is a dimer to tetramer transition. The binding of a second messenger to PsCap5 dimer triggers an open to closed transformation of the SAVED domains within each dimer, furnishing a surface for the assembly of the PsCap5 tetramer (dimer-of-dimers). This movement in the SAVED domains propagates to the HNH domains, juxtaposing and converting two of them into catalytically competent states for DNA destruction. PsCap5's preference for an asymmetrically linked second messenger (3'2'-cGAMP) is augmented by large asymmetric rotations of the SAVED domains (160° and 80°) compared to the unbound structure and thereby presenting different surfaces for interactions with the second messenger. Overall, the structures highlight the diversity of mechanisms for activating CBASS effectors and posit new therapeutic opportunities in controlling infections by pathogenic bacteria.

Antiviral signaling via cyclic nucleotides is a major defense system governing innate immunity in animals and antiphage defense in bacteria¹. Following infection, a nucleotide cyclase enzyme produces a cyclic nucleotide second messenger that binds and activates the stimulator of interferon genes (STING) in eukaryotes² or an effector protein in bacteria that induces cell death to prevent propagation of the virus (phage) to the rest of the bacterial population³⁻⁶. The most common such system in bacteria is CBASS defined by a cGAS/DncV-like nucleotidyltransferase cyclase (CD-NTase) enzyme and one or more CD-NTase-associated protein (Cap) effectors^{6,7}. Compared to STING in animals, the CBASS Cap effectors are much more diverse in function, ranging from degradation bacterial genomic DNA^{4,5,8,9}, depletion essential cellular metabolites^{10,11}, to compromise of the bacterial inner membrane^{12,13}, among other cell suicide mechanisms.

CBASS effectors can sense a cyclic nucleotide produced by a cognate CD-NTase via an animal-like STING domain or much more commonly by a structurally and functionally related SAVED (SMODS associated and fused to various effector domains) domain, among other means. The SAVED domain is a fusion of two CARF (CRISP-associated Rossman fold) modules co-opted from the type III CRISPR-Cas system^{5,7,9-11}. The largest group of these SAVED-domain containing proteins are Cap5 HNH-endonucleases, which once activated, degrade genomic DNA for abortive infection. In spite of their abundance, the mechanism by which the Cap5 effectors becomes activated by a second messenger for abortive infection is still not understood. Crystallographic studies on Cap5 from *Lactococcus lactis* (*Ll*) and *Asitccacaulis* sp.(*As*) have been reported recently⁹ but the all-important structure of a full-length Cap5 in complex with a second messenger has remained elusive. We present here high-resolution structures of full-length Cap5 from *Pseudomonas syringae* (*Ps*) in complex with cyclic nucleotides that elucidate for the first time how the binding of a second messenger induces oligomerization of Cap5 and how the HNH endonuclease domains become activated for DNA destruction.

Identification of *Ps*Cap5 activating ligand

Bacterial CBASS CD-NTases can use all four nucleotides for synthesizing cyclic di- and tri- nucleotide second messengers containing canonical 3'-3'- and non-canonical 2'-3' phosphodiester bonds^{5,14}. To identify a putative activating ligand for *Ps*Cap5, we screened a series of commercially available canonical bacterial 3'-5'/3'-5' (3'3') linked cyclic di-, tri- and tetra-nucleotides as well as mixed-linked 2'-5'/3'-5' (2'3') (or 3'-5'/2'-5' (3'2')) linked

depending on which base is named first) and 2'-5'/2'-5' (2'2') cyclic purine dinucleotides for their ability to activate *PsCap5* in plasmid DNA degradation assay (Fig. 1a,b and Extended Data Fig. 1). Strikingly, we observed plasmid digestion only with two of the tested ligands, 3'2'-cGAMP (3'2'GA) and 3'2'-c-diAMP (3'2'AA). Addition of 3'2'-cGAMP in particular leads to complete plasmid degradation with the digestion products being too short to be detected by ethidium bromide staining. Notably, the animal STING ligand 2'3'-cGAMP does not activate *PsCap5*. Likewise, we do not observe DNA cleavage with the symmetrically linked canonical bacterial 3'3'-cGAMP second messenger.

To better assess the concentrations of the activating ligands required for DNA digestion, we lowered *PsCap5* protein concentration to 50 nM and varied the ligand concentrations in the 10^{-2} – 10^4 nM range (Fig. 1c). We detected DNA cleavage with the strongly activating 3'2'-cGAMP ligand at a concentration of 1 nM, compared with an ~1000-fold higher concentration of 1 μ M with the weakly activating 3'2'-c-diAMP dinucleotide. This highlights the remarkable specificity of *PsCap5* for the 3'2'-cGAMP second messenger as compared to a closely related 3'2'-c-di-AMP. The only difference between 3'2'-cGAMP and 3'2'-diAMP is the substitution of guanine for adenine at one of the nucleobase positions. Our results are in line with DNA digestion studies performed with other *SAVED* domain containing effectors, which are similarly activated by low nM concentrations of cognate ligands. For example, several Cap4 proteins display robust DNA cleavage activities with 50 nM of cyclic tri-nucleotides⁵. Similarly, *B. pseudomaller* Cap5 is active with 50 nM of 3'3'-cGAMP⁵ and *AsCap5*. and *LICap5* proteins activated by 1-10 nM 3'2'-cGAMP⁹.

Oligomeric states of *PsCap5*

To assess the oligomeric states of *PsCap5* in absence and presence of the activating ligands we employed mass photometry. The characterization of apo *PsCap5* revealed that the protein exists predominantly as a dimer in solution, wherein 100% of the detected particles belonged to a single peak with the estimated molecular weight of 79 kDa, which is approximately a double of the theoretical mw of 42.7 kDa for the monomer (Fig. 1d, top panel). Upon the addition of 3'2'-cGAMP the mass distribution showed two peaks, corresponding to a dimer and a tetramer (Fig. 1d, middle panel). We did not detect any higher-order oligomers even with long 3h incubation times. Addition of the weaker activating dinucleotide 3'2'-c-diAMP also leads to the tetramer formation (Fig. 1d, bottom

panel). In contrast, with the non-activating ligands 3'3'-cGAMP and 2'3'-cGAMP, *PsCap5* remains a dimer. These data strongly indicate that the activated ligand-bound form of *PsCap5* is a tetramer.

Activated ligand-bound *PsCap5* tetramer

We solved and refined the structures of *PsCap5* with its favored ligand 3'2'-cGAMP, as well as the less preferred 3'2'-c-diAMP (Extended Data Table 1). The 3'2'-cGAMP liganded structure is refined to 1.87 Å resolution and has two tetrameric assemblies in the asymmetric unit (AU) (space group $P2_1$ with $a = 67.97$ Å, $b = 295.68$ Å, $c = 84.05$ Å, $\alpha = \gamma = 90.0^\circ$ and $\beta = 113.8^\circ$ unit cell). The 3'2'-c-diAMP liganded structure crystallized in two space groups: C2 with $a = 145.19$ Å, $b = 80.52$ Å, $c = 406.84$ Å, $\alpha = \gamma = 90.0^\circ$ and $\beta = 90.6^\circ$ unit cell; and $P3_22$ with the $a = b = 83.54$ Å, $c = 401.96$ Å, $\alpha = \beta = 90.0^\circ$ and $\gamma = 120^\circ$ unit cell. The C2 crystal form diffracted to 1.79 Å and has three tetramers in the AU, whereas the $P3_22$ crystal form diffracted to 2.10 Å and has a single tetramer in the AU. We first solved the 3'2'-c-diAMP liganded structure with a single tetramer complex by molecular replacement using an AlphaFold2 model¹⁵ derived by the Colabfold server¹⁶ and then used the refined model to solve the other structures.

Each *PsCap5* protomer consist of the N-terminal HNH domain (residues 1-96), the C-terminal SAVED domain (residues 127-388) and a mostly α -helical connector (residues 97-126) (Fig. 2a). The protomers assemble into dimer-of-dimers with both 3'2'-cGAMP and 3'2'-c-diAMP. The resulting tetramers (chains A, B, C and D) are similar in all of the ligand bound structures (Fig. 2a, Extended Data Fig. 2), and we note differences as we proceed.

PsCap5 dimers are arranged in a crisscross fashion to form the activated tetramer, wherein protomers A/B comprise one dimer and protomers C/D comprise the other dimer (Fig. 2a). The primary ligand binding pocket is formed mainly by the SAVED domain of protomer A, with the SAVED domain of protomer B forming a lid over the pocket (Fig. 2b). Thus, the two protomers of the A/B dimer interact asymmetrically with the ligand. The SAVED domains of protomers C and D of the C/D dimer bind the ligand in the same asymmetric manner (Fig. 2a). The electron densities for the entire 3'2'-cGAMP and 3'2'-c-diAMP ligands bound between the SAVED domains are extremely well-defined (shown

as polder maps at 4σ -level Fig. 2c,d) and allow for the unequivocal placement of 3'–2'- and 2'–3'- sugar phosphate linkages and the assignment of G versus A bases. There is a secondary ligand binding site on the outside of SAVED domains of protomers B and D (Fig. 2b and a) that exploit the same surface involved in ligand binding by the SAVED domains of A and C.

As expected, the SAVED domain of *PsCap5* is a fusion of two CARF modules (CARF1 and CARF2) with alternating β -strand and α -helical substructures (Fig 2b and Extended Data Fig. 3a,b). This fusion of CARF modules was first demonstrated for the SAVED domain of *Cap4*, wherein their pseudo-symmetrical arrangement provided a basis for the binding of the asymmetrical nucleotide second messenger 2'3'3'-c-AAA by *Cap4*⁵. Overall, the *PsCap5* SAVED shares a common architecture with the SAVED of *Cap4*⁵ and *AsCap5* and *LICap5*⁹ (Extended Data Fig. 3c-e), including a long bracing α -helix ($\alpha 6$ - $\alpha 7$ of CARF1 and $\alpha 10$ for CARF2 subdomains of *PsCap5* SAVED) and a central helix at CARF dimerization interface ($\alpha 8$ and $\alpha 11$).

Strikingly, HNH endonuclease domains of protomers A and C are in a catalytically active state, while the HNH domains of protomers B and D are in a catalytically inactive state (Fig. 2e and 3a,b). In particular, the secondary structure of the inactive protomers (Fig. 3b) deviates majorly from the characteristic $\beta\beta\alpha$ -topology of HNH endonucleases¹⁷. Thus, whereas residues 53-56 in the catalytically active HNH domains comprise the first β -strand ($\beta 3$) of the $\beta\beta\alpha$ substructure and carry the all-important catalytic His56, the same residues in the inactive domains form a loop and it is the ensuing residues (56-69, which we label as $\beta 11$) that form the first β -strand of the $\beta\beta\alpha$ substructure (Fig. 3a,b). The net effect of this structural rearrangement is that catalytic His56 is displaced by as much as ~ 7.5 Å ($C\alpha$ to $C\alpha$ distance) from the position in the catalytically active HNH domains (Fig. 3d,e), accompanied by other changes described below.

Activation of HNH endonuclease

The two SAVED domains in each dimer (A/C or B/D) are brought in close proximity by the binding of the ligand between them. This “closed” conformation of SAVED domains in turn mediates tetramerization via contacts predominately between protomers A and C (Extended Data Fig.4a-d). A net effect of this tetramer arrangement is that the HNH endonuclease domains of protomers A and C are positioned next to each other in

catalytically active states, whereas with HNH domains of protomers B and D, on the outside of the tetramer assembly, are in catalytically inactive states (Fig. 2e).

Members of HNH endonuclease superfamily share common DNA cleavage and binding motifs, but exhibit very limited sequence homology¹⁷. The active protomers A and C of *PsCap5* enable the canonical $\beta\beta\alpha$ -topology with short antiparallel β 3- and β 4-strands and α 4-helix (Fig. 3a and Extended Data Fig. 5). His91 and Asp95 of α 4-helix bind a catalytic Mg^{2+} ion with rigid coordination and solvation pattern (Fig. 3c) with the invariant catalytic His56 at the end of β 3 (fixed by hydrogen bond to Thr16) functioning as a general base to activate the water nucleophile for phosphodiester bond hydrolysis¹⁷. The β 3-strand (residues 53-56) of the active HNH domains is framed by stabilizing interactions that include hydrogen bonds with main chain carbonyl and amide group of Leu58 and the side chain of Asn82; interactions observed in some other HNH endonucleases¹⁷. One unique feature of the activated HNH domains in our structure is the bridging interactions between them, wherein a structural Mg^{2+} in each domain is coordinated by Asp92 and Glu53 to form an “ion clasp” (Fig. 3d and Extended Data Fig. 4c). Bridging interactions between the active HNH domains are further reinforced by two short antiparallel β -strands (β 1 and β 2) (Extended Data Fig. 4b).

The inactive HNH domains of protomers B and D lack not only the catalytic Mg^{2+} but also the structural Mg^{2+} (Fig. 3b and Extended Data Fig. 5). The α -helix equivalent to α 4 is shorter in the inactive state (α 4, residues 88-93), positioning the side chain of Asp95 ~7 Å away from the putative catalytic Mg^{2+} -coordinating position. Interestingly, in several inactive protomers of the ligand-bound structures reported here, the side chains of Asp95, and of the linker residues Asp97, Asp99, Glu103 and the main chain carbonyl of Tyr103 engage a different Mg^{2+} ion, thus helping to keep Asp95 in the inactive state (Extended Data Fig. 6). The absence of a structural Mg^{2+} in the HNH domains of protomers B and D is consistent with their far apart positions on the outside the *PsCap5* tetramer assembly, precluding the necessity of an “ion clasp” to juxtapose them. Interestingly, β -strands β 1 and β 2 of the active HNH domains that also partake in their bridging interactions are remodeled into a short α -helix followed by loop in the inactive state and participate in interactions with the neighboring protomer (B with A, and D with C).

Mechanism of DNA cleavage

HNH endonucleases can function as monomers and cut one strand of a DNA duplex or as dimers to nick both DNA strands at the same time¹⁷. The activated HNH domains of protomers A and C of the ligand bound *PsCap5* tetramer form a crescent-shaped dimeric active site with the α -helices of the $\beta\beta\alpha$ substructures collocated as in other dimeric HNH endonucleases (Fig. 3d,e and Extended Data Fig. 7). The conserved $\beta\beta\alpha$ -elements of the dimer (Fig. 3d) are expected to insert into the minor groove of a DNA duplex (Fig. 3f), as observed with DNA complexes of dimeric HNH endonucleases such as I-Ppol¹⁸, Hpy99I¹⁹ and T4 Endo VII²⁰ (Extended Data Fig. 7). The distance between the two catalytic Mg^{2+} ions (19.3 Å) is similar to that observed in the I-Ppol and Hpy99I DNA complexes (18.7 Å and 22.3 Å, respectively) and consistent with the diameter of B-DNA diameter (~20 Å) for cleavage of opposite strands. In I-Ppol, Hpy99I and T4 Endo VII the dimeric active site is positively charged as would be expected for interactions with negatively charged sugar-phosphate DNA backbone. Surprisingly, by contrast, the surface area between the catalytic ions in *PsCap5* is negatively charged in order to coordinate the structural Mg^{2+} ions of the ion clasp. These ions may directly contact the DNA backbone and appear to be surrogates for the Lys and Arg residues in other HNH endonucleases. Although the HNH domains of protomers B and D are catalytically inactive and devoid of Mg^{2+} ions, substantial portion of their surface is positively charged and may contribute to DNA binding (Extended Data Fig. 7).

Recognition of 3'2'-c-GAMP and 3'2'-c-diAMP second messengers

The 3'2'-c-GAMP and 3'2'-c-diAMP are asymmetrically linked cyclic dinucleotides that bind between the SAVED domains of the crisscross dimers A/B and C/D of the *PsCap5* tetramer (Fig. 2a,b). The SAVED of protomer A (and C) comprises most of the ligand binding pocket with the SAVED of protomer B (and D) using a different interface to form a smaller lid to fully enclose the ligand (Fig. 4a,b and Extended Data Fig. 3a,b and 8). For the recognition of 3'2'-c-GAMP, SAVED of *PsCap5* protomers A and B make hydrophobic and polar contacts that together provide specificity for the G and A nucleobases and the asymmetric sugar-phosphate linkages between them (Fig. 4a). Interactions from SAVED of protomer A involve predominantly residues that stem from loops that precede several α -helices ($\alpha 5$, $\alpha 8$, $\alpha 9$, and $\alpha 11$) and β -strands ($\beta 9$) that coalesce at the interface between the two CARF modules. There is an interesting division of labor, wherein residues from CARF1 (His138, Phe139, Leu216, Ala217, Ile219, Phe240 and Arg242) cluster around the A nucleobase and residues from CARF2 (Ala278, Val280, Pro281 and Phe374)

surround the G nucleobase. However, it is primarily residues from CARF2 (Ser277, Ala339, Ala340, Ile341, Pro342, Ala343, Asp365) that encompass the sugar-phosphate linkages between the nucleobases. Residues from SAVED of protomer B (Arg232, Asp231, Gln356 and His357) add to the interactions with 3'2'-c-GAMP; for example, His357 stacking against the A nucleobase.

How to explain the specificity of *PsCap5* for 3'2'-c-GAMP? From the structure of the complex, specificity for the G nucleobase is achieved via interactions with residues from the SAVED of protomer A only (Fig. 4a). The side chain of Arg366 recognizes the Hoogsteen edge of G via contacts with the N7 and O⁶ atoms and the phenolic OH-group of Tyr304 contacts the N² amino group of the base. The Watson-Crick edge of the G is hydrated with three water molecules and the water bridge extends to the main chain carbonyl of Gln279. Recognition of the A nucleobase is achieved via residues from both A and B protomers. Residues Asp231 and Arg232 of protomer B make direct hydrogen bonds with the Hoogsteen edge of A, whereas Phe240 and Ala343 of protomer A make water mediated contacts with A. Specificity for the 3'2' (3'-5'/2'-5') linkage is derived through a combination of direct and water-mediated hydrogen bonds from residues Arg178, Ile219, Arg242, Ser277, Tyr304, Ile341, Ala343 and Asp365 of protomer A and residues Gln356 and His357 of protomer B. In comparison to the *PsCap5*-3'2'-c-GAMP complex, the nucleobase A in place of G in the *PsCap5*-3'2'-c-diAMP complex loses two critical hydrogen bonds: one from the side chain of Arg366 (to O⁶ of G) and another from the side chain of Tyr309 (to N² of G) (Fig. 4b). The retained hydrogen bonds are from Arg366 to the N7 atom and the water bridge to Gln279. The loss of these hydrogen bonds helps to explain the lower preference of *PsCap5* for 3'-2'-c-diAMP as the second messenger. Contacts to the 3'-5' and 2'-5' sugar-phosphate linkages are nearly identical in the *PsCap5*-3'2'-c-GAMP and *PsCap5*-3'2'-c-diAMP complexes (Fig. 4a,b). These contacts are extensive, and their geometry is such that it will not easily accommodate deviations from the 3'2' linkage and provides a basis for understanding why ligands closely related to 3'2'-c-GAMP such as 3'3'-cGAMP and 2'3'-cGAMP are unable to activate *PsCap5*.

Interestingly, in the *PsCap5*-3'2'-c-GAMP and *PsCap5*-3'2'-c-diAMP complexes, the 3'-5' linked nucleobase (G or A) is in the *syn* conformation about the N-glycosidic bond,

whereas 2'–5' linked nucleobase (A) is in the *anti* conformation. However, the electron density for the 3'–5' linked A nucleobase of 3'2'-c-diAMP in the secondary ligand binding site on the outside of SAVED of protomer B (and equivalent protomers in other structures) suggests a *syn-anti* equilibria, whereas the 3'–5' linked G of 3'2'-cGAMP is always in the *syn* conformation (Extended Data Fig. 9). The propensity of the nucleobase G for the *syn* conformation may be additional factor favoring the binding of 3'2'-c-GAMP over 3'2'-c-diAMP in the primary binding site.

Structure of *PsCap5* in the absence of a ligand

We succeeded in crystallizing full-length *PsCap5* in the absence of an activating ligand. The apo crystal diffract to medium 3.15 Å resolution and contain *PsCap5* as a dimer rather than the tetramer with an activating ligand. This is consistent with our mass photometry data showing that *PsCap5* exists as a dimer in the absence of an activating ligand and that it transitions to a tetramer only in the presence of an activating ligand. The apo dimer is arranged in a crisscross manner with protomers A and B related by twofold symmetry and assuming almost identical conformations (rmsd of 1.57 Å for 354 C α s) (Fig. 5, top panel). The SAVED domains are relatively far apart with most of the contacts between them limited to loop residues 201-208 (Extended Data Fig. 10). We refer to the SAVED domains in this apo state as in an “open” configuration, as opposed to the closed configuration with a ligand. The interfaces between the HNH domains and the connector helices are also different from the ligand bound state (Extended Data Figs. 4, 10-12). The HNH domains lack a catalytic Mg²⁺ and the general base His56 is disordered (Extended Data Fig. 13) Together, these changes underscore the extraordinary remodeling of *PsCap5* that takes place upon ligand binding (Figure 5).

Transition of *PsCap5* from apo to ligand bound state

Upon ligand binding the SAVED domains close and rotate relative to the HNH-endonuclease domains and the connector helices. Compared to the apo state, the SAVED of protomer A rotates by 160° and presents its opposite surface to the SAVED of protomer B. For example, whereas the β 9- α 10 structural element (selected for visual reference) in the apo protomer A faces outward, it faces inward in the ligand bound state (Fig 5, top and middle panels). Protomer B on the other hand rotates by lesser 82° with the β 9- α 10 structural element facing outward. This unequal or asymmetric rotation of the two SAVED domains (relative to the apo state) augments the recognition of asymmetrical ligands by

presenting two different surfaces of the SAVED domains. Approximately 1540 Å² of surface area is buried between the two SAVED domains, as compared to only 600 Å² in the apo state. Importantly, the two ligand-bound crisscross dimers in turn expose a surface that lends to formation of the active tetramer with the HNH domains of protomers A and C activated through secondary structure remodeling and juxtaposed for DNA cleavage (Fig 5, bottom panel). By contrast, the HNH domains of protomers B and D are positioned on the outside of the tetramer assembly in catalytically inactive states, which resemble the structure of HNH domains in the apo state in lacking a catalytic Mg²⁺ (Extended Data Fig.13), among other changes. Overall, the tetramer is held together by extensive protein-protein interactions, wherein each protomer makes contacts with the all three of the other protomers (Extended Data Figs. 4, 11-12).

Conclusions

We present here for the first-time high-resolution structures of a full-length Cap5 from *Pseudomonas syringae* in complex with second messengers. The structures show how *PsCap5* senses a second messenger and becomes allosterically activated for DNA destruction. We show here that the key to *PsCap5* activation is a dimer to tetramer transition. The binding of a second messenger to *PsCap5* dimer triggers an open to closed transformation of the SAVED domains within each dimer and, in turn, furnishing a surface for the assembly of the *PsCap5* tetramer (dimer-of-dimers). This configurational change in the SAVED domains propagates to the HNH endonuclease domains, converting two of them into catalytically competent states and bringing them in close proximity for DNA degradation. Change in oligomerization is emerging as a common theme in CBASS effectors, though the nature of the oligomer varies. From electron microscopy analysis, the CBASS effectors Cap4 (restriction endonuclease (RE)-SAVED fusion)⁵, Cap12 (NAD⁺ cleaving TIR-STING fusion)²¹ and TIR-SAVED¹¹ form filaments in the presence of a second messenger, and from crystallographic analysis the CBASS effector NucC (RE-like fold) transitions from a trimer to a hexamer⁸. Interestingly, NucC lacks a domain (like SAVED or STING) dedicated for nucleotide binding. The second messenger binds instead at the periphery of the NucC RE-fold and the conformational changes induced are much subtler than the gross domain movements observed for *PsCap5*. Overall, a dimer to tetramer transition observed with *PsCap5* highlights the diversity of multimerization mechanisms for activating CBASS effectors. The “closed symmetry” of *PsCap5* tetramer

is also counter to the belief that CBASS (and other innate immune) effectors containing a SAVED domain are activated by “open symmetry” multimerization leading to filaments.

We show here that the primary binding site for the second messenger is at the interface between the two SAVED domains within each *PsCap5* dimer of the activated tetramer. The second messenger acts effectively as a glue to close the SAVED domains within each dimer for subsequent tetramerization. A packing of the SAVED domains on ligand binding has previously been postulated for Cap5 but not structurally observed⁹. Unexpectedly, we find that one SAVED domain makes the majority of interactions with the second messenger while the other forms a lid over it to enclose it. This asymmetry in interactions occurs because the two SAVED domains undergo strikingly different rotations (160° and 82°) compared to the unbound structure and thereby present different surfaces for interactions with the second messenger. As first shown for Cap4⁵, the SAVED domain is a fusion of two pseudo-symmetrically related CARF modules that obviates the need for two-fold symmetric second messengers. The fact that *PsCap5* employs two SAVED domains (or four CARF modules) to recognize a cyclic dinucleotide further expands the repertoire of second messenger signals that can be recognized by Cap5 and other CBASS effectors. Also, because the two SAVED domains undergo drastically different rotations on nucleotide binding it further averts the requirement for internally two-fold symmetric second messengers. This is important as some of the key signaling molecules for normal cellular processes in bacteria are two-fold symmetric cyclic dinucleotides such as 3'3'-c-diGMP and 3'3'-c-diAMP. We show that neither of these symmetric cyclic dinucleotides can activate *PsCap5*, guarding the effector from inadvertent activation by signals generated during normal cellular processes.

The CBASS system is common in environmental and pathogenic bacteria. An outstanding question in future work is whether the cell killing activity of CBASS effectors such as Cap5 can be leveraged or harnessed to limit infection by pathogenic bacteria. The recent identity of CBASS effectors and their activating cyclic nucleotides (or more membrane penetrable versions) may offer new therapeutic opportunities in controlling infections by pathogenic bacteria in humans.

METHODS

Protein expression and purification

The DNA sequence coding for full-length (residues 1-388) *Pseudomonas syringae* (*Ps*) Cap5 (TRN83959.1) was codon optimized for protein expression in *E. Coli*, synthesized and inserted between BamHI and HindIII cloning sites of a modified pET28b(+) vector (Novagen) to produce His6-SUMO (small ubiquitin-like modifier) N-terminally tagged fusion protein (GenScript). The plasmid was transformed to *Escherichia coli* BL21 (DE3) competent cells (Agilent). The cells were grown in LB medium with 30 µg/mL kanamycin at 37°C till OD₆₀₀ reached ~0.6, the temperature was then decreased to 18°C and protein expression was induced by the addition of 0.4 mM isopropyl-β-d-1-thiogalactopyranoside (IPTG). After an overnight growth, the bacterial cells were pelleted by centrifugation, resuspended in 25 mM TRIS-HCl pH 8.0, 20 mM Imidazole, 500 mM NaCl, 10% glycerol, 2 mM β-mercaptoethanol (BME) with cOmplete™ EDTA-free protease inhibitor cocktail tablet (Roche) and lysed by sonication. The lysate was centrifuged at 18,000 rpm for 1h and the protein was purified from the soluble fraction by a nickel chelating His60 Ni Superflow 5 mL column (Takara Bio Inc.) connected to an ÄKTA FPLC system (GE Healthcare Life Sciences). The protein was eluted from the column with 25 mM TRIS-HCl pH 8.0, 300 mM Imidazole, 300 mM NaCl, 10% glycerol, 2 mM BME buffer. The His6-SUMO tag was cleaved with Ulp1 protease during an overnight dialysis into 25 mM TRIS-HCl pH 8.0, 300 mM NaCl, 2 mM BME buffer. The protein was passed through the His60 Ni column to bind the cleaved His6-SUMO tag and His6-tagged Ulp1. *Ps*Cap5 protein was then purified by gel-filtration on Superdex 75 column (GE Healthcare) and concentrated in 25 mM TRIS-HCl pH 8.0, 250 mM NaCl, and 2 mM tris(2-carboxyethyl) phosphate (TCEP) to ~32 mg/ml, and stored in aliquots at -80 °C.

DNA plasmid digestion assay

For DNA cleavage assay we have used pET-29b(+) plasmid with human *FEN1* gene residues 2-366 inserted between NdeI and XhoI sites for a total of 6.4 kb (kilo base pairs). *Ps*Cap5 protein was incubated with (or without) the cyclic nucleotide ligands 15 min on ice in 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 10 mM MgCl₂ and 1 mM TCEP. The reactions were initiated by addition of 500 ng of the DNA plasmid. The total reaction volume was 25 µL. For the identification of the putative activating ligands, we have used 500 nM protein and 500 nM ligands concentrations; 3'-5'/3'-5' (3'3') linked cyclic di-, tri- and tetra-nucleotides as well as mixed-linked 2'-5'/3'-5' (2'3') and 2'-5'/2'-5' (2'2') cyclic purine

dinucleotides ligands were obtained from BIOLOG Life Science Institute (Germany). For the evaluation of the activating ligand concentration the protein concentration was reduced to 50 nM and the ligand concentrations was varied in in the 10^{-2} – 10^4 nM range as indicated on the corresponding figures. The reactions were incubated 20 min at 37°C and stopped by addition of 6x loading buffer (60 mM EDTA pH 8.0, 20 mM Tris-HCl pH 7.5, 30% glycerol, 0.1% SDS, 0.012% xylene cyanol). The DNA cleavage products were separated by electrophoresis on 1.5% agarose gel with TAE buffer, 1 kb DNA ladder from New England Biolabs was used as DNA mobility weight marker. After running the gel for 30 min at 100 V, the gels were stained with 10 µg/mL ethidium bromide in water for 30 min and then de-stained in water for 10 min. The gels were photographed with GelDoc-It2 imaging system with VisionWorks software. The black and white colors of the original gels were inverted for figure clarity.

Mass photometry

Mass photometry experiments were performed on a Refeyn OneMP mass photometer (Refeyn Ltd, Oxford, UK), the datasets were recorded with the AcquireMP and analyzed with DiscoverMP software. To form a sample chamber, self-adhesive silicon gasket (4-6 wells, ~15 µL volume each) was placed on a cleaned and dried borosilicate glass microscope cover slide. To conduct a measurement, the glass slide/silicone gasket assembly was placed on the instrument's objective and centered on a single well, 12 µL of sample buffer was then added to the well and the focal position of the glass surface was determined and held constant using an autofocus system. 1 µL of protein sample was then added to the well (13-fold dilution) and mixed with the buffer. A 90 sec movie was recorded immediately after the dilution. A fresh well was used for each measurement and the measurement was repeated at least three times for each sample. An aliquot of *PsCap5* was de-frozen on ice and serially diluted in the sample buffer (25 mM Tris-HCl pH 8.0, 250 mM NaCl, 2 mM TCEP and 5 mM MgCl₂) so that the number of the detected events (particle counts) during a measurement was ~3000 to 5000 per 90 sec movie for an optimum data acquisition and processing. Contrast-to-mass linear calibration curve was created using BSA (monomer and dimer) and apoferritin protein standards diluted in the sample buffer. To calculate the molecular weight of the major species observed on the particle counts versus molecular mass distribution histograms we have used Gaussian function in DiscoverMP software. The final concentration of *PsCap5* in a sample well was 5.8 nM and the ligand concentration (if present) was 25 µM. We noticed that the fraction

of the tetramer peak in the presence of the activating ligands 3'2'-cGAMP or 3'2'-c-diAMP increased if the protein was incubated with the ligand in a test tube prior to the mass photometry measurement. The fraction of the tetramers reached its maximum ~15 min after the sample mixing. Further incubation (up to 3h, on ice or at a room temperature) did not increase the proportion of the tetramers. We did not observe any higher order complexes even with 3h incubation times. Decreasing NaCl salt concentration in the buffer to 100 mM did not affect mass distribution.

Crystallization, data collection, and structure determination

Initial crystallization conditions for apo *PsCap5* were determined with matrix screens (Hampton Research, Molecular Dimensions and Qiagen) set up with OryxNano crystallization robot (Douglas Instruments) using 10 mg/mL protein in buffer containing 25 mM Tris-HCl pH 8.0, 250 mM NaCl, 2 mM TCEP and 5 mM MgCl₂ at 20°C. Small crystals appeared after 3 days of incubation in 1M trisodium citrate with 0.1M Imidazole pH 8.0. Large rectangular bipyramid crystals up to ~0.4x0.2x0.2 mm in size were optimized to grow in 1M trisodium citrate with 0.1M TRIS-HCl pH 8.0 in hanging drop vapor diffusion experiments; large triangular prism shaped crystals were also observed in different drops under the same conditions. The crystals cryoprotected in mother liquor-glycerol mixture (1.26M trisodium citrate, 0.1M TRIS-HCl pH 8.0 and 10% glycerol) diffracted poorly (~12-15 Å resolution) with synchrotron radiation under cryogenic conditions (NSLS-II 17-ID-1 and 17-ID-2 beamlines at the Brookhaven National Laboratory and NE-CAT 24-ID-E beamline at Argonne National Laboratory). To improve the diffraction, we have screened various cryoprotective mixtures. The use of 1.4M trisodium citrate with 0.375 M sodium formate and 50mM TRIS-HCl pH 8.0 as a cryoprotectant resulted in consistent resolution improvement to ~5-7 Å, and ~2h soaking in the cryoprotectant further improved the resolution to ~4-6 Å. About 325 crystals have been screened in total and the best crystal diffracted to 3.16 Å at 24-ID-E beamline. The apo *PsCap5* crystal belongs to P4₁2₁2 space group with a= b=159.4 Å, c=433.9 Å $\alpha=\beta=\gamma= 90.0^\circ$ unit cell. We have used relaxed AlphaFold2¹⁵ models for *PsCap5* monomer and dimer derived by the Colabfold server¹⁶ for molecular replacement (MR) using Phaser²² module in Phenix²³. The models were processed with phenix.process prior to MR search. The best search results were obtained by using a SAVED domain and a predicted HNH endonuclease domain dimer (with some regions manually trimmed to eliminated clashes) as the search models. Phaser was able to place six copies of the HNH endonuclease domain dimers and 10 copies of SAVED.

We have docked remaining two copies of SAVED into density modified (DM) map generated in Phenix for a total of 12 copies of *PsCap5* protein (6 dimers) in the AU. For refinement in Phenix-refine, we have used a *PsCap5* monomer from a high-resolution structure with 3'2'-c-diAMP (see below) obtained at that time as a reference model. We were able to refine the apo *PsCap5* structure to $R_{\text{work}}/R_{\text{free}}$ of $\sim 24.0/29.5$ but with pronounced number of geometry outliers likely due to differences in conformations in many regions of the apo structure compared to the ligand-bound tetramer. In addition, several regions of the structure had poor electron density. To improve the fit and geometry of the apo *PsCap5* model, we remodeled each copy of the apo dimer by removing the poorly defined regions and using the good regions as a template for new AlfaFold2 predictions as described in ref²⁴, then placing the predicted rebuilt dimers back into the model and re-refining with rigid body and ADP in Phenix. This allowed to increase the fraction of residues in Ramachandran favored regions from $\sim 88\%$ to $\sim 96\%$. The rebuilt model was then relaxed into the sharpened experimental electron density map with the molecular dynamics refinement tool ISOLDE²⁵. The refinement was completed with Phenix-refine and the final model was refined to 3.16 Å resolution with R_{work} and R_{free} values of xx and xx%, respectively, and displayed good stereochemistry with % of residues in Ramachandran favored regions (Supplementary Data Table 1).

To obtain crystals of activated *PsCap5* we have re-screened *PsCap5* in the presence of 1 mM 3'2'-cGAMP and 3'2'-c-diAMP cyclic dinucleotides with the matrix screens. The 3'2'-cGAMP-liganded crystals were optimized to grow as thick large rectangular plates in 0.1M Bis-Tris pH 5.5-6.0 with 0.2M MgCl_2 and PEG3350 19-20% with several rounds of seeding. The 3'2'-c-diAMP-liganded crystals grew under similar conditions as hexagonal plates. The crystals were cryoprotected in mother liquor solution with 24% PEG3350 and stepwise increase of glycerol content to 20%. Most of these crystals diffracted to high resolution at NSLS-II 17-ID-1 and 17-ID-2 beamlines at the Brookhaven National Laboratory. The 3'2'-cGAMP liganded structure is refined to 1.87 Å resolution and has two tetrameric assemblies in the asymmetric unit (AU) (space group $P2_1$ with $a = 67.97$ Å, $b = 295.68$ Å, $c = 84.05$ Å, $\alpha = \gamma = 90.0^\circ$ and $\beta = 113.8^\circ$ unit cell). The 3'2'-c-diAMP liganded structure crystallized in two space groups: $C2$ with $a = 145.19$ Å, $b = 80.52$ Å, $c = 406.84$ Å, $\alpha = \gamma = 90.0^\circ$ and $\beta = 90.6^\circ$ unit cell; and $P3_22$ with the $a = b = 83.54$ Å, $c = 401.96$ Å, $\alpha = \beta = 90.0^\circ$ and $\gamma = 120^\circ$ unit cell. The $C2$ crystal form diffracted to 1.79 Å and has three tetramers in the AU, whereas the $P3_22$ crystal form diffracted to 2.10 Å and has a single

tetramer in the AU. We first solved the 3'2'-c-diAMP liganded structure with a single tetramer complex by molecular replacement using the AlphaFold2 models as described above for the apo structure and then used the refined model to solve the other structures. The structures have been refined in Phenix-refine with manual rebuilding in Coot. The 3'2'-c-diAMP-liganded C2 crystal form has been refined to 1.79 Å with R_{work} and R_{free} values of 17.9 and 22.5%, respectively, and the P3₂2 to 2.10 Å with $R_{\text{work}}/R_{\text{free}}$ xx/xx %. The 1.87 Å resolution 3'2'-cGAMP liganded structure has R_{work} and R_{free} values of xx and xx%, respectively. The data collection and refinement statistics are summarized in Supplementary Data Table 1 for all the structures. All molecular graphic figures were prepared using PyMOL (Schrödinger LLC). Surface areas were calculated using PDBePISA (the protein interfaces, surfaces and assemblies service (PISA) at the European Bioinformatics Institute; http://www.ebi.ac.uk/pdbe/prot_int/pistart.html)²⁶. For secondary structure analysis and comparison we have used DSSP method²⁷ at <https://2struccompare.cryst.bbk.ac.uk/index.php>²⁸.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary submitted with this manuscript.

Data availability

Associated data are provided in Extended Data Figs. 1–13 and Supplementary Table 1. Structures of *PsCap5* with 3'2'-cGAMP, 3'2'-c-diAMP (with three and one tetramer assemblies in the AU) and without a ligand are available in the Protein Data Bank under accession codes XXX, XXX, XXX, and XXX, respectively.

References

- 1 Wein, T. & Sorek, R. Bacterial origins of human cell-autonomous innate immune mechanisms. *Nat Rev Immunol* (2022). <https://doi.org/10.1038/s41577-022-00705-4>
- 2 Domizio, J. D. *et al.* The cGAS-STING pathway drives type I IFN immunopathology in COVID-19. *Nature* **603**, 145-151 (2022). <https://doi.org/10.1038/s41586-022-04421-w>
- 3 Cohen, D. *et al.* Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* **574**, 691-695 (2019). <https://doi.org/10.1038/s41586-019-1605-5>

- 4 Ye, Q. *et al.* HORMA Domain Proteins and a Trip13-like ATPase Regulate Bacterial cGAS-like Enzymes to Mediate Bacteriophage Immunity. *Mol Cell* **77**, 709-722 e707 (2020). <https://doi.org:10.1016/j.molcel.2019.12.009>
- 5 Lowey, B. *et al.* 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 (2020). <https://doi.org:10.1016/j.cell.2020.05.019>
- 6 Duncan-Lowey, B. & Kranzusch, P. J. CBASS phage defense and evolution of antiviral nucleotide signaling. *Curr Opin Immunol* **74**, 156-163 (2022). <https://doi.org:10.1016/j.coi.2022.01.002>
- 7 Burroughs, A. M., Zhang, D., Schaffer, D. E., Iyer, L. M. & Aravind, L. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res* **43**, 10633-10654 (2015). <https://doi.org:10.1093/nar/gkv1267>
- 8 Lau, R. K. *et al.* Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity. *Mol Cell* **77**, 723-733 e726 (2020). <https://doi.org:10.1016/j.molcel.2019.12.010>
- 9 Fatma, S., Chakravarti, A., Zeng, X. & Huang, R. H. Molecular mechanisms of the CdnG-Cap5 antiphage defense system employing 3',2'-cGAMP as the second messenger. *Nat Commun* **12**, 6381 (2021). <https://doi.org:10.1038/s41467-021-26738-2>
- 10 Morehouse, B. R. *et al.* Cryo-EM structure of an active bacterial TIR-STING filament complex. *Nature* (2022). <https://doi.org:10.1038/s41586-022-04999-1>
- 11 Hogrel, G. *et al.* Cyclic nucleotide-induced helical structure activates a TIR immune effector. *Nature* (2022). <https://doi.org:10.1038/s41586-022-05070-9>
- 12 Duncan-Lowey, B., McNamara-Bordewick, N. K., Tal, N., Sorek, R. & Kranzusch, P. J. Effector-mediated membrane disruption controls cell death in CBASS antiphage defense. *Mol Cell* **81**, 5039-5051 e5035 (2021). <https://doi.org:10.1016/j.molcel.2021.10.020>
- 13 Severin, G. B. *et al.* Direct activation of a phospholipase by cyclic GMP-AMP in *El Tor* *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **115**, E6048-E6055 (2018). <https://doi.org:10.1073/pnas.1801233115>
- 14 Whiteley, A. T. *et al.* Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194-199 (2019). <https://doi.org:10.1038/s41586-019-0953-5>
- 15 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). <https://doi.org:10.1038/s41586-021-03819-2>
- 16 Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679-682 (2022). <https://doi.org:10.1038/s41592-022-01488-1>
- 17 Wu, C. C., Lin, J. L. J. & Yuan, H. S. Structures, Mechanisms, and Functions of His-Me Finger Nucleases. *Trends Biochem Sci* **45**, 935-946 (2020). <https://doi.org:10.1016/j.tibs.2020.07.002>
- 18 Flick, K. E., Jurica, M. S., Monnat, R. J., Jr. & Stoddard, B. L. DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI. *Nature* **394**, 96-101 (1998). <https://doi.org:10.1038/27952>

- 19 Sokolowska, M., Czapinska, H. & Bochtler, M. Crystal structure of the beta beta alpha-Me type II restriction endonuclease Hpy99I with target DNA. *Nucleic Acids Res* **37**, 3799-3810 (2009). <https://doi.org:10.1093/nar/gkp228>
- 20 Biertumpfel, C., Yang, W. & Suck, D. Crystal structure of T4 endonuclease VII resolving a Holliday junction. *Nature* **449**, 616-620 (2007). <https://doi.org:10.1038/nature06152>
- 21 Morehouse, B. R. *et al.* STING cyclic dinucleotide sensing originated in bacteria. *Nature* **586**, 429-433 (2020). <https://doi.org:10.1038/s41586-020-2719-5>
- 22 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007). <https://doi.org:10.1107/S0021889807021206>
- 23 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). <https://doi.org:10.1107/S09074444909052925>
- 24 Terwilliger, T. C. *et al.* Improved AlphaFold modeling with implicit experimental information. *Nat Methods* **19**, 1376-1382 (2022). <https://doi.org:10.1038/s41592-022-01645-6>
- 25 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018). <https://doi.org:10.1107/S2059798318002425>
- 26 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797 (2007). <https://doi.org:10.1016/j.jmb.2007.05.022>
- 27 Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983). <https://doi.org:10.1002/bip.360221211>
- 28 Drew, E. D. & Janes, R. W. 2StrucCompare: a webserver for visualizing small but noteworthy differences between protein tertiary structures through interrogation of the secondary structure content. *Nucleic Acids Res* **47**, W477-W481 (2019). <https://doi.org:10.1093/nar/gkz456>
- 29 Mate, M. J. & Kleanthous, C. Structure-based analysis of the metal-dependent mechanism of H-N-H endonucleases. *J Biol Chem* **279**, 34763-34769 (2004). <https://doi.org:10.1074/jbc.M403719200>

Acknowledgements

We thank the staff at the National Synchrotron Light Source II (NSLS-II) beamlines 17-ID-1 and 17-ID-2 for facilitating X-ray data collection. This research used the AMX (17-ID-1) and FMX (17-ID-2) beamlines of the National Synchrotron Light Source II, a United States Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0012704. The Life Science Biomedical Technology Research resource, which supports AMX and FMX, is primarily supported by the NIH (NIGMS) through a Biomedical Technology Research Resource P41 grant (P41GM111244), and by the DOE Office of Biological and

Environmental Research (KP1605010). Most of the datasets on apo *PsCap5* crystals have been collected at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on the 24-ID-E beam line is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We thank NE-CAT beamline staff for their assistance with data collection.

Mass photometry experiments were performed on a Refeyn OneMP at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (SF349247), NYSTAR, and the NIH National Institute of General Medical Sciences (GM103310). We thank Dr. William (Chase) Budell for the assistance with the mass photometry experiments.

Contributions

O.R. and A.K.A. designed the experiments. O.R. performed mass photometry, crystallization, structure solution and refinement for all the structures. D.S. performed DNA plasmid digestion assays. D.F.K. contributed to molecular replacement phasing of the apo *PsCap5* structure and performed advanced refinement for the structure (Alfalfold2 and ISOLDE) and assisted with refinement for the other structures. J.K. assisted in X-ray data collection and mass photometry experiments. A.B. conducted protein expression and purification. A.K.A. guided the project. O.R. and A.K.A. wrote the manuscript.

Corresponding authors

Correspondence to Olga Rechkoblit at olga.rechkoblit@mssm.edu or Aneel K. Aggarwal at aneel.aggarwal@mssm.edu.

Ethics declarations

Competing interests

The authors declare no competing interests.

Figure Legends

Fig.1. Identification of activating ligands and oligomerization states of *PsCap5*. **a**, Chemical structures of 2'3'-mixed linked (or 3'2'-mixed linked depending which nucleobase is named first) 3'2'-cGAMP, 3'2'-c-diAMP and eukaryotic STING ligand 2'3'-cGAMP, and bacterial canonically 3'3'-linked 3'3'-cGAMP cyclic purine dinucleotides. **b**, DNA plasmid digestion by *PsCap5* in the presence of cyclic purine dinucleotides with 3'3', 2'3' and 2'2' linkages under 500 nM protein and ligand concentrations. Plasmid digestion is observed only with 3'2'-cGAMP (3'2'GA) and 3'2'-c-diAMP (3'2'AA). DNA degradation in the presence of other di, tri- and tetra- nucleotides is shown on Extended Data Fig. 1. For gel source data, see Fig. S1. **c**, Determination of the *PsCap5* activating concentrations for 3'2'-cGAMP and 3'2'-c-diAMP. Concentration of *PsCap5* is 50 nM and the ligand concentrations are varied in the 10^{-2} – 10^4 nM range as indicated. The digestion products are analyzed by agarose gel electrophoresis followed by staining with ethidium bromide. The black and white colors are inverted for clarity. Plasmid DNA includes closed and open circle species. DNA ladder is marked in kilo base pair (kb). Each gel is a representative of at least three independent experiments. For gel source data, see Fig. S1. **d**, Mass photometry analysis of apo *PsCap5* and in presence of 3'2'-cGAMP and 3'2'-c-diAMP. Solid lines represent Gaussian function fit to major species observed on the particle counts versus molecular mass distribution histograms. The expected theoretical masses of apo *PsCap5* monomer, dimer and tetramer are 42.7, 85.4 and 170.8 kDa, respectively.

Fig.2. Structure of the activated ligand-bound *PsCap5* tetramer. **a**, The overall dimer-of-dimers architecture of the 3'2'-c-diAMP bound *PsCap5* tetramer. Two crisscross dimers (A/B and C/D), formed by *PsCap5* protomers A and B (represented as cyan and beige cartoons, respectively) and C and D (colored dark blue and pink, respectively) produce an activated tetramer. Each *PsCap5* protomer consist of the N-terminal HNH endonuclease domain (residues 1-96), the C-terminal SAVED domain (residues 127-388) and a mostly α -helical connector (residues 97-126). HNH endonuclease domain of protomers A and C each carries a catalytic and a structural Mg^{2+} ions (Mg) shown as green spheres and HNH endonuclease of each protomer has a structural Zn^{2+} ion (Zn) shown as a grey sphere. There are two 3'2'-c-diAMP ligands (shown as sticks) bound to the primary sites between the SAVED domains of dimers A/B and C/D, respectively, and to the secondary sites on

the outside of the SAVED of B and D, respectively. **b**, The crisscross dimer A/B with secondary structure elements labeled. The SAVED domains of A and B present different surfaces for interactions with the ligand, see Extended Data Fig. 3 for details. **c** and **d**, Polder electron density maps contoured at 4σ -level are shown as blue mesh at 1.87 Å resolution for 3'2'-cGAMP and at 1.79 Å resolution for 3'2'-c-diAMP ligands, respectively. **e**, HNH endonuclease domains arrangement in the tetramer. The HNH endonucleases of protomers A and C are positioned next to each other in a catalytically active state, whereas the HNH domains of protomers B and D are on the outside of the assembly in a catalytically inactive state.

Fig. 3. Active and inactive HNH endonuclease domains of *PsCap5* tetramer and the catalytically competent dimeric active site. **a**, Topology and structure of the active conformation of the HNH endonuclease (protomers A and C) with the characteristic $\beta\beta\alpha$ -topology of HNH endonucleases colored in green. Catalytic His56 is highlighted in red. **b**, Topology and structure of the inactive conformation of the HNH endonuclease (protomers B and D). **c**, Polder electron density map for catalytic Mg^{2+} ion of protomer A at 1.79 Å resolution is contoured at 5σ -level and is shown as blue mesh. **d**, Structure of the dimeric active site formed by the HNH endonucleases of protomers A and C of the tetramer. **e**, Electrostatic surface representation of the dimeric active site with views from the top and the side. **f**, Superposition of the $\beta\beta\alpha$ module of Colicin E9 HNH endonuclease monomer with a scissile DNA strand (shown in as a light grey cartoon) (PDB ID: 1V15²⁹) to the $\beta\beta\alpha$ modules of the HNH endonucleases of protomers A and C. The positioning of the DNA duplex model resulting from the superposition relative to the dimeric A/C active site of the tetramer suggests that *PsCap5* produces cleavage products with a single base overhang.

Fig. 4. Recognition of the 3'2'-cGAMP and 3'2'-c-diAMP ligands by *PsCap5* and interaction diagram. **a** and **b**, Interactions of protomer A and B with 3'2'-cGAMP and 3'2'-c-diAMP ligands, respectively. The ligands, shown as sticks with carbon atoms colored yellow, are bound between the SAVED domains of protomer A and B (and C and D) in the *PsCap5* activated tetramer. The other atoms are colored by the elements. The protein residues involved in ligand recognition are shown as sticks with the carbon atoms of the residues colored in cyan for protomer A and in beige for protomer B. The interaction diagram was produced with the Maestro module in the Schrodinger suite of programs (Schrodinger LLC).

Fig. 5. Structure of apo *PsCap5* dimer and transition to the ligand-bound activated tetramer. The structures of the apo dimer, ligand-bound dimer (one of the two dimers in the tetramer) and the dimer-of-dimers tetramer are shown in the same view with the apo and tetramer structures superimposed by the HNH endonuclease domains of protomers A and B. Top panel, the structure of apo *PsCap5* crisscross dimer with the SAVED domains in the “open” conformation. Middle panel, the binding of the activating ligand triggers the SAVED domains closure and rotation relative to the HNH-endonuclease domains and the connector helices. The SAVED of protomer A rotates by 160° and presents the opposite surface to the SAVED of protomer B. For example, whereas the β 9- α 10 structural element in the apo protomer A faces outward, it faces inward in the ligand bound state. Protomer B on the other hand rotates by 82° with the β 9- α 10 structural element facing outward. Bottom panel, the two ligand-bound dimers form the tetramer with the HNH endonuclease domains of protomers A and C activated for DNA cleavage. To produce the model of longer DNA bound to the *PsCap5* tetramer, the short DNA duplex described in Fig. 3f is inserted in the middle of the ideal B-DNA duplex.

Extended Data Figures and Tables

Activation of CBASS-Cap5 endonuclease immune effector by cyclic nucleotides: A view at high resolution

Extended Data Figures 1-12

Extended Data Table 1

Extended Data Fig. 1. DNA plasmid digestion assay with di-, tri- and tetra- cyclic nucleotides with 3'3', 2'3' and 2'2' linkages. Related to Fig. 1a that shows only purine dinucleotide ligands. Plasmid digestion is observed only with 3'2'-cGAMP (3'2'GA) and 3'2'-c-diAMP (3'2'AA). Concentrations of *PsCap5* and the ligands are 500 nM. The DNA cleavage products are analyzed by agarose gel electrophoresis followed by staining with ethidium bromide. The black and white colors are inverted for clarity. Plasmid DNA includes closed and open circle species. Each gel is a representative of at least three independent experiments. For gel source data, see Fig. S2.

Extended Data Fig. 2. Composition of the asymmetric unit (AU) in the crystal structures of the activated *PsCap5* tetramer with 3'2'-cGAMP and 3'2'-c-diAMP ligands. **a**, The 3'2'-c-diAMP liganded structure in C2 space group with $a = 145.19 \text{ \AA}$, $b = 80.52 \text{ \AA}$, $c = 406.84 \text{ \AA}$, $\alpha = \gamma = 90.0^\circ$ and $\beta = 90.6^\circ$ unit cell is refined to 1.79 \AA resolution and has three tetramer assemblies in the AU. **b**, The 3'2'-cGAMP containing structure in P2₁ space group and with $a = 67.97 \text{ \AA}$, $b = 295.68 \text{ \AA}$, $c = 84.05 \text{ \AA}$, $\alpha = \gamma = 90.0^\circ$ and $\beta = 113.8^\circ$ unit cell is refined to 1.87 \AA resolution and has two tetramer complexes in the AU. **c**, The 3'2'-c-diAMP liganded structure in P3₂2 space group with the $a = b = 83.54 \text{ \AA}$, $c = 401.96 \text{ \AA}$, $\alpha = \beta = 90.0^\circ$ and $\gamma = 120^\circ$ unit cell is refined to 2.10 \AA and has a single tetramer assembly in the AU.

Extended Data Fig. 3. Two CARF-like unit architecture of the SAVED domain of *PsCap5* and ligand recognition. **a**. Topology diagram of the SAVED domain of *PsCap5* highlights that the SAVED domain has a pseudo-symmetry and is a fusion of two different in sequence CRISP-associated Rossmann fold (CARF) domains with alternating β -strands and α -helical segments. CARF1 module of protomers A and B shown in light cyan and yellow, respectively, and CARF2 module of protomers A and B are colored in cyan and light orange, respectively. In the *PsCap5* tetramer, the activating ligand is bound between the SAVED domains of the crisscross dimer A/B (and C/D). The protomers A and B use different residues of CARF1 and CARF2 modules for ligand recognition (labeled as red triangles). **b**. Structures of the SAVED domains of dimer A/B with CARF1/CARF2 modules colored as described above bound to the 3'2'-cGAMP ligand. **c**, *Acinetobacter baumannii* (*Ab*) Cap4 SAVED domain with 2'3'3'-c-triAAA (PDB ID: 6VM6)⁵ with CARF1/CARF2 modules colored light cyan and cyan, respectively. **d**, *L.Lactis* Cap5 SAVED domain with CARF1/CARF2 modules colored as above bound to 3'2'-cGAMP (PDB ID: 7RWS)⁹. **e**,

Asticcacaulis sp Cap5 SAVED with CARF1/CARF2 modules colored as in **c** without a ligand (PDB ID: 7RWK)⁹.

Extended Data Fig. 4. Interface between protomers A and C in the ligand-bound *PsCap5* tetramer. **a**, Overall structure of protomers A (cyan) and C (dark blue) in the tetramer shown in Fig. 2a. Protomers A and C belong to different crisscross dimers (A/B and C/D, respectively) that constitute the tetramer. **b**, The interface between the catalytically active HNH endonuclease domains of protomers A and C. In addition to protein-protein contacts bridging the two endonucleases, a structural Mg^{2+} in each domain is coordinated by Asp92 and Glu53 to form an “ion clasp”. **c**, Flexibility in coordination of the structural Mg^{2+} ions and the “ion clasp”. The two structural Mg^{2+} ions are coordinated slightly different in the three tetramer assemblies of the AU of the 1.79 Å resolution 3′2′-c-diAMP containing structure. For instance, in the A/B-C/D tetramer (left panel), Mg^{2+} of protomer A (MgA) forms direct and water-mediated contacts with Glu53 and Asp92 of protomer A and Glu53 of protomer B, whereas MgC interacts with Glu53 and Asp92 of protomer B only. However, in the E/F-G/H tetramer (middle panel), MgE is contacted by Glu53 and Asp92 of protomer E, whereas MgG interacts with Glu53 and Asp92 of protomer G and Glu53 of protomer E. In K/L-M/N tetramer (right panel), the Mg^{2+} ion’s coordination by Glu53 and Asp92 is as the one observed in the E/F-G/H tetramer, but the ion’s hydration patterns are different. **d**, Interactions between the SAVED domains of protomers A and C.

Extended Data Fig. 5. Sequence and secondary structures of catalytically active and inactive conformations of the HNH endonuclease domains in the tetramer. HNH endonucleases of protomers A and C have active conformation with the characteristic $\beta\beta\alpha$ -topology of HNH endonucleases colored in green. Catalytic His56 is highlighted in red. Residues that are involved in coordination of the catalytic and structural Mg^{2+} are indicated by green stars. Cys32, Cys35, Cys87 and Cys90 coordinate a structural Zn^{2+} ion and are indicated by the grey dot.

Extended Data Fig. 6. Conformations of the linker residues immediately adjacent to the C-terminus of the inactive HNH endonuclease domains in the activated tetramers. **a**, Structure of the residues 95-103 of the protomer B (colored in beige) with the inactive HNH endonuclease domain in the 3′2′-c-diAMP liganded structure with three tetramer assemblies. Residues 95-103 in the other catalytically inactive protomers (D, F, H, L and

N) in this structure have similar conformation. The α -helix of the characteristic for HNH endonucleases $\beta\beta\alpha$ module is shorter in the inactive form (residues 88-93) versus the active form (residues 88-96) and excludes Asp95 that is involved in coordination of the catalytic Mg^{2+} ion in the active form. This positions the side chain of Asp95 ~ 7 Å away from the putative catalytic Mg^{2+} -coordinating position. **b**, Protomer B (top panel) in the two-tetramer assembly 3'2'-c-GAMP ligated structure has a Mg^{2+} ion coordinated by the side chains of Asp95, and of the linker residues Asp97, Asp99, Glu103 and the main chain carbonyl of Tyr103 that helps to keep Asp95 in the inactive state. Bottom panel, residues 95-103 of the protomers D, F and H take conformation that is similar to the one in **a** and do not carry a Mg^{2+} ion. **c**, Residues 95-103 of both protomers (B and D) of the one-tetramer assembly 3'2'-c-diAMP liganded structure also coordinate the Mg^{2+} ion.

Extended Data Fig. 7. Structures of dimeric HNH endonuclease complexes with DNA and their comparison with the dimeric HNH active site of *PsCap5*. HNH endonuclease I-PpoI¹⁸ (PDB ID: 1A74, with the catalytic ions modeled based on the complex with cleaved DNA PDB ID: 1A73) and Hpy99I¹⁹ (PDB ID: 3GOX). In the displayed Hpy99I structure, the N-terminal β -barrel (residues 1–53), a linker (residues 54–64) and the residues of the HNH domain which approach the major groove of the DNA for sequence-specific interactions (residues 79-98) have been omitted for clarity. The phage T4 Endo VII²⁰ (PDB ID: 2QNC) cleaves two symmetrical strands of a four-way junction DNA. For all proteins, a front and a side views of a DNA-binding surface colored by electrostatic potential are shown; the conserved $\beta\beta\alpha$ structural elements of the dimer approaching duplex DNA for cleavage from the minor groove side is colored in green, the catalytic ions are shown as green spheres and the catalytic His residues shown as red sticks. A front view of a DNA binding surface of dimeric HNH endonuclease of *PsCap5* tetramer, formed by the active conformation protomers A and C. The view also includes the inactive protomers B and D. The side view of HNH endonuclease domains A/C only and of the four endonuclease domains of the tetramer. The two structural Mg^{2+} ions are positioned in-between of the catalytic ions. A model of HNH endonuclease domains A/C interactions with DNA is based on the structure of Colicin E9 HNH endonuclease monomer with DNA (PDB ID: 1V15²⁹). The $\beta\beta\alpha$ module of Colicin E9 HNH endonuclease monomer with a scissile DNA strand was superimposed to the $\beta\beta\alpha$ modules of the HNH endonucleases of protomers A and C, as shown in Fig. 3f. This produced a 5-base pair (bp) long DNA duplex with slightly wider minor groove than it is in the ideal B-DNA. To generate a longer DNA model, a 23 bp ideal

B-DNA generated in Coot was superimposed to the 5 bp DNA duplex, the 5 bp fragment was then inserted into the ideal B-DNA.

Extended Data Fig. 8. Sequence and secondary structures of the SAVED domains of dimer A/B (and C/D) in the tetramer. The SAVED domain of protomer A (and C) forms most of the ligand binding pocket (secondary structure elements are colored in cyan). The SAVED domain of protomer B forms a “lid” of the pocket (secondary structure elements are colored in beige). The protein residues interacting with the 3′2′-c-dGAMP ligand bound between the SAVED domains of dimer A/B (and C/D) are indicated by the red triangles.

Extended Data Fig. 9. Electron density maps for 3′2′-c-GAMP and 3′2′-c-diAMP ligands in the primary binding site between the SAVED domains A/B (and C/D) and in the secondary binding sites on the outside of SAVED B (and D). Polder electron density maps contoured at 4σ -level are shown as blue mesh at 1.87 Å resolution for 3′2′-cGAMP and at 1.79 Å resolution for 3′2′-c-diAMP ligands, respectively. 2Fo-Fc contoured at 2σ -level are shown in green mesh. For the 3′2′-c-diAMP ligand in the outside site, 2Fo-Fc map is also shown contoured at 1σ -level in grey mesh. At the primary ligand binding sites between the SAVED domains in the *PsCap5*-3′2′-c-GAMP and *PsCap5*-3′2′-c-diAMP complexes, the 3′–5′ linked nucleobase (G or A) is the *syn* conformation about the N-glycosidic bond, whereas 2′–5′ linked nucleobase (A) is in the *anti* conformation. However, the electron density for the 3′–5′ linked A nucleobase of 3′2′-c-diAMP in the secondary ligand binding site on the outside of SAVED of protomer B (and equivalent protomers L and N) indicates predominantly the *anti* conformation for the base. At the other secondary binding sites of the three-tetramer assembly structure (protomers D, F and H), the 3′–5′ linked A base remains mainly in the *syn* conformation. These observations suggest a *syn-anti* equilibria of the 3′–5′ linked A of 3′2′-c-diAMP, whereas the 3′–5′ linked G of 3′2′-cGAMP is always in the *syn* conformation.

Extended Data Fig. 10. Interface between protomers A and B in the apo *PsCap5* dimer. Overall structure of protomers A (cyan) and B (beige) in the apo dimer with the interaction areas between the protomers indicated by black rectangles. The inserts show the details of interactions between the HNH endonuclease domains, the connector helices-SAVED domains, HNH endonuclease-SAVED domains and between the SAVED domains.

Extended Data Fig. 11. Interface between protomers A and B of the dimer A/B in the ligand bound activated *PsCap5* tetramer. Overall structure of protomers A (cyan) and B (beige) in the dimer with the interaction areas between the protomers indicated by black rectangles. The inserts show the details of interactions between the HNH endonuclease domains, the connector helices-SAVED domains, and between the SAVED domains. The interactions within the second dimer C/D of the tetramer are similar to the ones within dimer A/B.

Extended Data Fig. 12. Interface between protomers A and D and between protomers B and D in the ligand-bound activated *PsCap5* tetramer. **a**, Overall structure of protomers A (cyan) and D (pink) in the tetramer with the interaction areas between the protomers indicated by black rectangles. The inserts show the details of interactions between the HNH endonuclease domains and the SAVED domains. The protomers A and D belong to the different dimers in the tetramer (A/B and C/D, respectively). **b**, Overall structure of protomers B (beige) and D (pink) in the tetramer with the interaction area between the protomers indicated by black rectangle. The interactions between protomers B and D involve SAVED domain $\alpha 7$ residues 199-201 and adjacent loop residues 202-205 of both protomers. Interestingly, these interactions occur almost exclusively via weaker more transient C-H---O hydrogen bonds (pink dash lines) rather than conventional hydrogen bonds (black dash lines). Notably, in the apo dimer, the same residues remodel into a different conformation and provide the interface between SAVED A and B of the apo dimer via conventional hydrogen bonds (Extended Data Fig 10).

Extended Data Fig. 13. Inactive conformation of the HNH endonuclease domain in the apo *PsCap5* dimer and its comparison with the inactive and active conformations of the HNH endonucleases in the ligand-bound tetramer. **a**, Sequence and secondary structures of catalytically inactive endonucleases in the apo dimer A/B (A is similar to B and colored light blue), and inactive and active conformations of the HNH endonuclease domains in the tetramer (beige and cyan). In the tetramer, HNH endonucleases of protomers A (and C) have active conformation with the characteristic $\beta\beta\alpha$ -topology of HNH endonucleases colored in green. Catalytic His56 is highlighted in red. Residues that are involved in coordination of the catalytic and structural Mg^{2+} in the catalytically active conformation are indicated by green stars. Cys32, Cys35, Cys87 and Cys90 coordinate a structural Zn^{2+} ion

and are indicated by the grey dot. **b**, Structures of the HNH endonuclease domain in the inactive conformations in the apo and ligand-bound tetramer and the catalytically active conformation in the tetramer. In the apo structure, the side chain of the catalytic His56 is disordered in both protomers in all six copies of the dimer assembly of the AU; the catalytic and structural Mg²⁺ are lacking. Furthermore, the α helix of the $\beta\beta\alpha$ -topology is shorter than in the active conformation (α 14, residues 88-93 vs α 4, residues 88-96), this positions the side chain of Asp95 ~7 Å away from the putative catalytic Mg²⁺-coordinating position. See the main text and Fig. 3 for description of the inactive and active conformations of HNH endonucleases in the tetramer.

Supplementary information

Activation of CBASS-Cap5 endonuclease immune effector by cyclic nucleotides: A view at high resolution

Guide to Supplementary Information

Supplementary Figure 1

Unprocessed gels associated with the data in Figure 1b and c.

Supplementary Figure 2

Unprocessed gels associated with the data in Extended Data Figure 1.