

# Crystal structure of the novel di-nucleotide cyclase from *Vibrio cholerae* (DncV) responsible for synthesizing a hybrid cyclic GMP-AMP

*Cell Research* (2014) 24:1270-1273. doi:10.1038/cr.2014.123; published online 23 September 2014

## Dear Editor,

Through production of cyclic di-nucleotides as second messengers, di-nucleotide cyclases in bacteria and more recently in mammalian cells have been found to play important roles in modulating a wide range of cellular processes [1]. During infection of *Vibrio cholerae* (*V. cholerae*), di-nucleotide cyclase Vibrio (DncV) is required for efficient intestinal colonization by *V. cholerae* and downregulates hyperinfectivity-associated *V. cholerae* chemotaxis through generation of a regulatory molecule cyclic GMP-AMP (cGAMP) [2].

In eukaryotes, cGAMP synthase (cGAS) was identified as a cytoplasmic DNA sensor that activates the type I interferon pathway by synthesizing the second messenger cGAMP [3, 4]. Structural, biochemical and cellular studies reveal that upon dsDNA binding, cGAS is activated through formation of a catalytically competent conformation [5, 6]. dsDNA binding can be disrupted by ablation of interactions with two or three successive phosphodiester linkages in either DNA strand, which results in impaired or completely abrogated cGAS function and in turn negatively impacts type I interferon production [6]. Similar to the formation of c-di-GMP, cGAMP is likely to be generated in a stepwise manner, involving an initial formation of G(2',5')pA linkage followed by A(3',5')pG generation. These sequential reactions lead to the formation of the pair of phosphodiester linkages with two molecules of pyrophosphate released as byproducts [5, 6].

Distinct from eukaryotic cGASs, DncV exhibits two unique biochemical features. One is its ability to specifically catalyze c(3',5')GAMP formation. This is supported by the fact that both DncV-produced and synthetic c(3',5')GAMP, but not cGAS-generated c[G(2',5')pA(3',5')p], were entirely blunted in their stimulatory activities in cells expressing a mutant STING (R231A) that only recognizes 2',5' linkage-containing cyclic di-nucleotides [7]. The other feature is its capability to produce three different cyclic di-nucleotides, c-di-GMP, c-di-AMP, and the dominant product cGAMP without an

activator DNA [2], though recent evidence showed that DncV may not produce c-di-GMP *in vivo* [8]. To gain a better understanding of the structural differences between these two classes of hybrid di-nucleotide cyclases, i.e., cGAS and DncV, we solved the crystal structure of the full-length DncV (residues V1 to G436) and investigated the catalytic as well as regulatory mechanisms of the NTase superfamily of di-nucleotide cyclases.

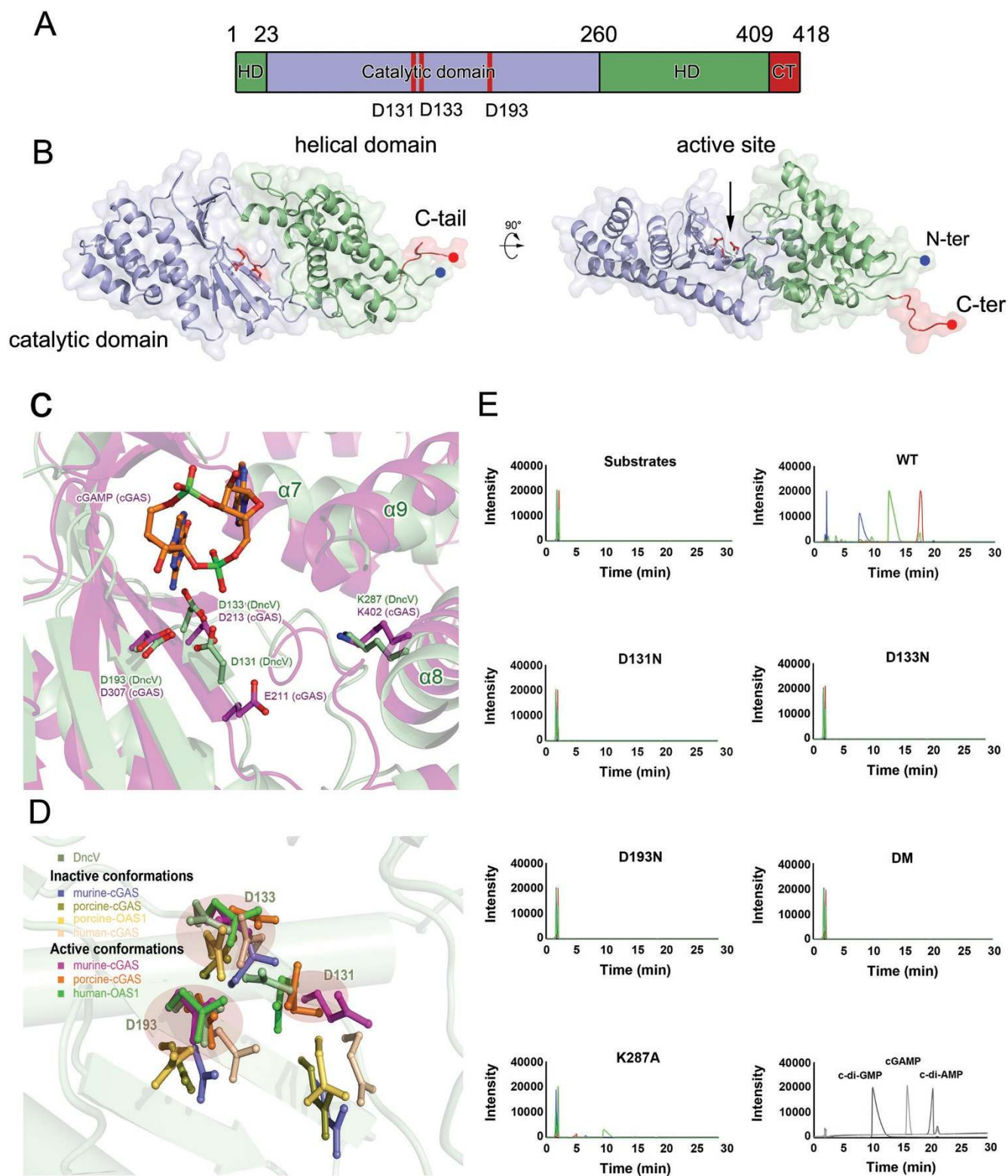
Initially we were unable to crystallize wide-type DncV protein (wt-DncV). Based on previous findings that two conserved acidic residues, D131 and D133, were essential for DncV function [2], we generated a double-mutant protein (dm-DncV) with the two predicted catalytic residues replaced by asparagine residues. The dm-DncV protein showed good properties for crystallization and structure determination. Eventually the structure of dm-DncV was determined by the selenium-based single-wavelength anomalous dispersion (SAD) method. After numerous trials, we also succeeded in determination of the crystal structure of wt-DncV at 3.3 Å resolution. The crystal structures of wt-DncV and dm-DncV adopt a similar overall fold, with 1.1 Å root-mean-square deviation (*r.m.s.d*) for 343 equivalent Ca atoms. Thus, we focus on the wt-DncV structure, and refer to it as DncV for simplicity in the rest of the manuscript.

Comparison of DncV structure with other structures in the Protein Data Bank (PDB) using the structure comparison service DALI [9] yielded several structural homologues. The most significant matches, based on the highest Z-scores, include the structures of human cGAS [10] (PDB code: 4KM5, Z-score: 18.7), yeast polyadenylate polymerase (PAP) [11] (PDB code: 1FA0, Z-score: 17.1), and porcine OAS1 [12] (PDB code: 1PX5, Z-score: 13.8). These hits suggest that DncV adopts a conserved NTase fold, which is consistent with the previous knowledge [2].

Similar to the overall appearance of cGAS and OAS1 (Supplementary information, Figure S1A), the structure of DncV comprises two distinct domains: a helical domain (HD) containing residues V1 to V23 and P261 to

L409, and a catalytic domain (CD) consisting of residues L24 to D260, which covers the three predicted catalytic residues D131, D133 and D193 (Figure 1A and 1B). The N-terminal extension, together with helix  $\alpha 7$  of the HD,

is responsible for tethering of the HD to the CD (Supplementary information, Figure S1A). Fundamentally distinct from the corresponding structural elements in OAS1 and cGASs (Supplementary information, Figure S1A),



**Figure 1** Structure of DncV from *V. cholerae* responsible for synthesizing a hybrid cGAMP molecule. **(A)** Schematic diagram of domain organization of DncV in primary sequence. The catalytic domain, in which the three catalytic residues are highlighted, and the helical domain (HD) followed by the C-terminal tail (CT) are colored as blue, green and red, respectively. The numbers for the residues are labeled in the upper panel. **(B)** The overall structure of DncV. The molecule for DncV is shown in a cartoon representation. The catalytic domain, helical domain and the C-terminal tail are colored in blue, green and red, respectively. Three key residues expected for catalytic reaction are shown as colored sticks. **(C)** Active site of DncV. Comparison of the active sites of DncV and active murine cGAS (PDB code: 4K9B) [6]. The polypeptides of DncV and active cGAS were shown as pale green and purple cartoons. The catalytic acidic residues from two structures were represented as colored sticks. A cGAMP molecule, which was modeled from the complex structure of cGAS, was shown as a colored stick diagram. Three secondary structures,  $\alpha 7$ ,  $\alpha 8$  and  $\alpha 9$ , of DncV were labeled out. **(D)** Comparison of the key residues of active sites from DncV, inactive (PDB code: 4K8V) and active (PDB code: 4K9B) murine cGAS [6], inactive (PDB code: 4JLX) and active porcine cGAS (PDB code: 4KB6) [5], inactive porcine (PDB code: 1PX5) [12] and active human OAS1 (PDB code: 4IG8) [15], and inactive human cGAS (PDB code: 4KM5) [10]. **(E)** Key residues responsible for DncV activity. DncV-catalyzed reaction products and commercial standards of 3 cyclic di-nucleotides were fractionated by RP-HPLC. Wild-type (240  $\mu$ M) and mutant DncVs (240  $\mu$ M) were incubated with ATP (1 mM), GTP (1 mM), or ATP+GTP (1 mM for each) at 37 °C for 60 min and the products were fractionated by RP-HPLC. Red, blue and green denote products derived from reactions using ATP, GTP, and ATP+GTP as substrates, respectively. The elution peaks for chemical standards c-di-GMP, c-di-AMP and cGAMP are shown in the last panel. x axis represents the retention time in the C18 column. Please note that wt-DncV-catalyzed reaction products were uniformly eluted a bit earlier than the corresponding chemical standards, which might be caused by the heating treatment during the preparation of the product samples.

the N-terminal extension in DncV is mostly integral to the HD, resulting in a continuous hydrophobic core of the HD. This result suggests that the N-terminal extension of DncV plays an essential role in the folding of the HD.

The active site of the NTase superfamily of enzymes is known to be characterized by two conserved structural features, i.e., the three conserved acidic residues, and the G[G/S]xx motif that precedes the first triad residue and is expected to bind to the  $\gamma$ -phosphate of the incoming nucleotide [12]. DncV indeed possesses these two structural features; the corresponding catalytic residues are D131 and D133 in strand  $\beta 1$ , and D193 in strand  $\beta 5$  (Figure 1A-1C), and the G[G/S]xx motif is defined by residues G113-S114-F115-Q116 in a well-defined loop linking helix  $\alpha 5$  and strand  $\beta 1$ . All these structural elements of DncV are positioned on a pocket formed between the CD and the HD (Supplementary information, Figure S1B). The critical role of the pocket in DncV-mediated cyclic di-nucleotide formation is supported by the fact that the homology of prokaryotic DncV orthologs mainly maps to the pocket and its immediate surrounding regions, with the predicted catalytic acidic residues and residues of the L7 loop that make up the signature motif at the bottom of the pocket showing the highest level of conservation (Supplementary information, Figure S1B and S1C).

To evaluate the functional relevance of these structural elements, we performed *in vitro* di-nucleotide cyclase assays using reverse phase high performance liquid chromatography (RP-HPLC). RP-HPLC revealed that DncV alone is capable of making cGAMP from GTP and ATP, c-di-AMP from ATP, and c-di-GMP from GTP (Figure 1E). We further characterized DncV-generated chemical products by MALDI-TOF-MS and nuclear magnetic resonance (NMR) analyses (Supplementary informa-

tion, Figure S1D). MALDI-TOF-MS analysis revealed that the molecular mass of these three products is 673.0, 689.1 and 657.1 Da, matching the molecular mass of cGAMP, c-di-GMP and c-di-AMP, respectively. Notably, RP-HPLC and NMR analysis revealed that a mixture of GTP and ATP triggers the synthesis of cGAMP as the dominant product, with little c-di-AMP and c-di-GMP generated (Figure 1E and Supplementary information, Figure S1D), reminiscent of a previous report that DncV preferentially generates the hybrid cGAMP molecule [2]. Our NMR analyses provide strong evidence to demonstrate that DncV produces 3',5'-cGAMP, 3',5'-c-di-AMP or 3',5'-c-di-GMP *in vitro* (Supplementary information, Figure S1D). We also assessed the effects of mutating the three conserved acidic residues (D131, D133 and D193) in the DncV active site using RP-HPLC. Our results showed that neither dm-DncV nor any of the three single mutant (D131N, D133N or D193N) could maintain their abilities to generate detectable cyclic di-nucleotide products (Figure 1E). In contrast, mutating one of the invariant residues, K287, which corresponds to a residue that participates in the interaction with pyrophosphate [13], impaired but did not completely eliminate its ability to catalyze c-di-GMP formation (Figure 1E). These observations, together with other reported evidence [2], suggest that all three acidic residues in the active site are indispensable for the catalytic activity of DncV.

As shown by previous studies [5, 10, 14], cGASs adopt a bilobal scaffold with mixed  $\alpha/\beta$  topology, which represents a structural characteristic of members of the NTase superfamily. However, cGASs cannot execute di-nucleotide cyclization function in the absence of an activator dsDNA. In contrast, DncV in its free form appears to adopt an active conformation, as judging from the architecture of its three catalytic acidic residues and

the widened entrance to the catalytic pocket (Supplementary information, Figure S1E).

One of the most striking features revealed by the structures of free and dsDNA-bound cGAS is the drastic conformational change of the active site upon dsDNA binding (Supplementary information, Figure S1E). In the case of murine cGAS, all dsDNA-bound cGAS structures show identical localizations of the catalytic E211, D213 and D307 residues on the well-formed active site, while these residues in the structure of free cGAS are mis-aligned, especially for E211, which is at least 5.2 Å away from its position in the active structure (Figure 1D). Consistent with these observations, the catalytic triad residues of dsDNA-bound porcine cGAS, E200, D202 and D296, are located in positions similar to those of active murine cGAS, whereas free-state human and porcine cGASs position their catalytic triads in the same way as inactive murine cGAS (Figure 1D). Surprisingly, the architecture of the three catalytic residues of free DncV is identical to that of the active form of cGASs, suggesting that the active site of DncV is well formed before substrate binding and/or activator induction.

Given that di-nucleotide cyclases and other NTase superfamily members are conserved in the two-domain architecture and the catalytic acidic residues, we reasoned that other DncV homologs of the NTase superfamily might adopt a similar architecture of the catalytic residues upon activation. Supporting our hypothesis, the conformation of the catalytic residues of dsRNA-bound human OAS1 is quite different from that of porcine OAS1 in its free state, with the former resembling the active form while the latter resembling the inactive form of cGAS proteins (Figure 1D).

Although DncV exhibits an active conformation in the absence of DNA binding, we wonder whether DncV activity can be regulated by other factors. By comparing the structures of DncV and dsDNA-activated cGAS, we found that dsDNA could also fit very well on backside of the active site of DncV. As long-rang interactions are formed primarily in complementary electrostatic nature, we performed surface electrostatic potential analysis to find whether these cyclases share any similar electrostatic surfaces. Interestingly, both DncV and cGAS possess several surface patches with pronounced positive electrostatic potential on their backsides (Supplementary information, Figure S1F). These linear positive-nature patches are completely solvent-exposed and poised for binding to other cofactors with complementary electrostatic potential, and thereby may serve as an interaction motif of these di-nucleotide cyclases. As this potential interaction motif is also found in the structures of human OAS1 and yeast PAP (Supplementary information, Figure S1F), it

is likely involved in the regulation of functions of the NTase family proteins.

## Acknowledgments

The authors are grateful to the staffs at PF and SSRF, and Dr Haitao Li and Dr Yuanyuan Li at Tsinghua University for their assistance with X-ray diffraction data collection. This work was supported by the National Basic Research Program of China (973 Program; 2013CB911103 and 2012CB518904), and the National Natural Science Foundation of China (81322023, 31170678, 31170158, 31370733 and 31000332).

Zhenhua Ming<sup>1,\*</sup>, Wei Wang<sup>1,4,\*</sup>, Yuchen Xie<sup>1</sup>, Pengfei Ding<sup>5</sup>, Yuchao Chen<sup>1,4</sup>, Dazhi Jin<sup>6</sup>, Yuna Sun<sup>3</sup>, Bin Xia<sup>5</sup>, Liming Yan<sup>2</sup>, Zhiyong Lou<sup>1,7</sup>

<sup>1</sup>Structural Biology Laboratory and MOE laboratory of Protein Science, School of Medicine and Life Science, Tsinghua University, Beijing 100084, China; <sup>2</sup>Tsinghua-Peking Center for Life Sciences, Laboratory of Structural Biology, School of Medicine, Tsinghua University, Beijing, 100084, China; <sup>3</sup>National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, China; <sup>4</sup>Institute of Plant Stress Biology, State Key Laboratory of Cotton Biology, School of Life Sciences, Henan University, Kaifeng, Henan 475004, China; <sup>5</sup>Beijing Nuclear Magnetic Resonance Center, College of Chemistry and Molecular Engineering, and School of Life Sciences, Peking University, Beijing 100871, China; <sup>6</sup>Zhejiang Provincial Center for Disease Control and Prevention, Zhejiang 310051, China; <sup>7</sup>Collaborative Innovation Center for Biotherapy, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, West China Medical School, Sichuan University, Chengdu, Sichuan 610207China

\*These two authors contributed equally to this paper.

Correspondence: Zhiyong Lou<sup>a</sup>, Liming Yan<sup>b</sup>

<sup>a</sup>E-mail: louzy@xtal.tsinghua.edu.cn

<sup>b</sup>E-mail: yanlm@xtal.tsinghua.edu.cn

## References

- 1 Danilchanka O, Mekalanos JJ. *Cell* 2013; **154**:962-970.
- 2 Davies BW, Bogard RW, Young TS, et al. *Cell* 2012; **149**:358-370.
- 3 Sun L, Wu J, Du F, et al. *Science* 2012; **339**:786-791.
- 4 Wu J, Sun L, Chen X, et al. *Science* 2013; **339**:826-830.
- 5 Civril F, Deimling T, de Oliveira Mann CC, et al. *Nature* 2013; **498**:332-337.
- 6 Gao P, Ascano M, Wu Y, et al. *Cell* 2013; **153**:1094-1107.
- 7 Ablasser A, Goldeck M, Cavlar T, et al. *Nature* 2013; **498**:380-384.
- 8 Kellenberger CA, Wilson SC, Sales-Lee J, et al. *J Am Chem Soc* 2013; **135**:4906-4909.
- 9 Holm L, Rosenstrom P. *Nucleic Acids Res* 2010; **38**:W545-W549.
- 10 Kranzusch PJ, Lee AS, Berger JM, et al. *Cell Rep* 2013; **3**:1362-1368.
- 11 Bard J, Zhelkovsky AM, Helmling S, et al. *Science* 2000; **289**:1346-1349.
- 12 Hartmann R, Justesen J, Sarkar SN, et al. *Mol Cell* 2003; **12**:1173-1185.
- 13 Martin G, Keller W, Doublet S. *EMBO J* 2000; **19**:4193-4203.
- 14 Wu J, Sun L, Chen X, et al. *Science* 2012; **339**:826-830.
- 15 Donovan J, Dufner M, Korennykh A. *Proc Natl Acad Sci USA* 2013; **110**:1652-1657.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)