Type III-A CRISPR-associated protein Csm6 degrades cyclic hexa-adenylate activator using both CARF and HEPN domains

Dalia Smalakyte1,†, Migle Kazlauskiene1,†, Jesper F. Havelund2, Audronė Rukšėnaitė1, Auguste Rimaite1, Giedre Tamulaitiene1, Nils J. Færgeman2, Gintautas Tamulaitis1,* and Virginijus Siksnys1,†

1Institute of Biotechnology, Vilnius University, Saulėtekio av. 7, LT-10257 Vilnius, Lithuania and 2Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark, Campusvej 55, DK-5230, Odense M, Denmark

Received May 22, 2020; Revised July 12, 2020; Editorial Decision July 14, 2020; Accepted August 06, 2020

ABSTRACT

The type III CRISPR–Cas systems provide immunity against invading nucleic acids through the coordinated transcription-dependent DNA targeting and cyclic adenylate (cAₙ)-activated RNA degradation. Here, we show that both these pathways contribute to the Streptococcus thermophilus (St) type III-A CRISPR–Cas immunity. HPLC-MS analysis revealed that in the heterologous Escherichia coli host the StCsm effector complex predominantly produces cA₅ and cA₆. cA₆ acts as a signaling molecule that binds to the CARF domain of StCsm6 to activate non-specific RNA degradation by the HEPN domain. By dissecting StCsm6 domains we demonstrate that both CARF and HEPN domains act as ring nucleases that degrade cAₙ molecules to switch signaling off. CARF ring nuclease converts cA₆ to linear A₆>p and to the final A₃>p product. HEPN domain, which typically degrades RNA, also shows ring nuclease activity and indiscriminately degrades cA₅ or other cAₙ molecules down to A₃>p. We propose that concerted action of both ring nucleases enables self-regulation of the RNase activity in the HEPN domain and eliminates all cAₙ secondary messengers in the cell when viral infection is combated by a coordinated action of StCsm effector and the cA₆-activated Csm6 ribonuclease.

INTRODUCTION

CRISPR–Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR-associated) adaptive immune systems protect many bacteria and archaea against invading nucleic acids (NAs) such as bacteriophages or plasmids (1). The type III CRISPR–Cas systems provide immunity through the coordinated transcription-dependent DNA targeting and cAₙ-dependent RNA degradation pathways. In response to infection, the Csm (type III-A and type-III D) or Cmr (type III-B and III-C) effector complex, guided by the crRNA, binds to the matching sequence in the invading RNA target and activates the Cas10 subunit (2). Activated Cas10 exhibits two different catalytic activities: ssDNase which degrades the target DNA that is being transcribed (3–9) and synthetase which produces cyclic oligoadenylates (cAₙ, n = 2–6) that act as secondary messengers (10–15). In addition to csm and cmr coding genes, csm6/csx1-like genes are frequently associated with type III CRISPR–Cas systems (16). Csm6 ribonuclease is preferentially associated with type III-A, while Csx1 ribonuclease is found with no clear link to a particular subtype (17,18). We and others have previously shown that cA₅ or cA₆ molecules bind to the CARF (CRISPR-associated Rossmann fold) domain of the Csm6/Csx1 RNases and activate their HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domain for RNA degradation (10–14). Activated Csm6/Csx1 RNases presumably degrade both cellular RNAs and phage transcripts at the later stages of phage infection, which can lead to cell death or dormancy (10,19,20). Some type III systems are predicted to have cA₅ binding CARF domains in conjunction with various effector domains such as putative transcription factors or DNases (17,21). The steady state concentration of cA₅ is in the cell and the cA₅-dependent Csm6/Csx1 RNAse activity depends on two factors: (i) cA₅ synthesis that is controlled through target RNA degradation by Csm3 (10,14) and (ii)
cA₆ degradation by specialized enzymes called ring nucleases. A family of ring nucleases composed of a sole CARF domain that degrade cA₆ has been first identified in *Sulfolobus solfataricus* and *Sulfolobus islandicus* (22,23). It has been shown later that the CARF domain of cA₄-dependent Csm6 RNases from *Thermus thermophilus* and *Thermococcus omniorum* also functions as a ring nuclease that slowly degrades cA₄ (24,25). Recently, a widespread new family of archaeal viral enzymes that efficiently degrade cA₄ has been identified, implying that these enzymes could function as type III anti-CRISPR proteins (26). However, the control and regulation mechanisms of cA₆-dependent CARF RNases remain to be elucidated.

To address this question we have focused on the well-characterized *Streptococcus thermophilus* type III-A CRISPR–Cas system (8,10,27,28). We have previously shown that in vitro StCsm complex produces cA₃, cA₄, cA₅ and cA₆ in decreasing order of abundance and traces of cA₂ (10). Although cA₃ was the major reaction product in vitro, the least abundant cA₆ acted as the activator of StCsm6 and StCsm6 ′ RNases (10), raising a question whether a similar or different set of cA₆s is produced in vivo. Using a targeted HPLC-MS analysis of cell metabolites we show here that StCsm complex in the heterologous *E. coli* host produces a range of different cA₆s, with the equilibrium shifted towards cA₅ and cA₆ species. We further show that cells expressing wild type (WT) StCsm6 and StCsm6 ′ exhibit dramatically lower cA₆ levels. Finally, we demonstrate that both CARF and HEPN domains of StCsm6 RNases function as ring nucleases that degrade the cA₆ and other cA₆s to auto-regulate the level of the signaling molecules and to limit the degree of RNA degradation in the cell after phage is eliminated.

**MATERIALS AND METHODS**

**Construction of plasmids**

pCas/Csm plasmid carrying genes *cas6*, *cas10*, *csm2*, *csm3*, *csm4*, *csm5*, *csm6* ′ and *csm6* from *S. thermophilus* DGCC8004 CRISPR2 region (27) and its derivatives was used as a template to generate the following mutant variants: pCas/Csm_dCsm3_dHDCas10, pCas/Csm_dCsm3_dHDCas10 ΔCsm6 ΔCsm6 ′, pCas/Csm_dCsm3_dHDCas10_dPalmCas10 ΔCsm6 ′ ΔCsm6 (see Supplementary Table S1).

To obtain pCas/Csm_dCsm3_dHDCas10, a Cas10(D16A) mutation-containing fragment was PCR amplified from pBAD24_Cas10(D16A)-C-HSH (8) and cloned into pCas/Csm_dCsm3(D33A) (27) via SacI and NdeI sites.

To obtain pCas/Csm_dCsm3_dHDCas10 ΔCsm6 ′ ΔCsm6, the constructed pCas/Csm_dCsm3_dHDCas10 was cleaved with PpiI and XmaI. Sticky ends of the resulting DNA fragment were converted to blunt ends and subjected to ligation. This resulted in the Csm6 ′ ORF truncation from 386 to 181 codons and elimination of Csm6 ORF.

To obtain pCas/Csm_dCsm3_dHDCas10_dPalmCas10 ΔCsm6 ′ ΔCsm6, a fragment containing Cas10(D575A,D576A) was produced from pBAD24_Cas10(D575A,576A) (8) by PCR. pCas/Csm_dCsm3_dHDCas10 ΔCsm6 ′ ΔCsm6 was digested with AjiI. The resultant vector and PCR fragments were assembled using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs).

For targeting TcR transcript, artificial CRISPR locus was constructed. A synthetic 445 nt CRISPR locus containing four identical 36 nt spacers complementary to the targeted region of the TcR RNA separated by four 36 nt repeats and flanked by the leader sequence and the terminal repeat of the *S. thermophilus* DGCC8004 CRISPR2 system was cloned into the pACYC-Duet-1 (Novagen) vector, generating the pCRISPR_Tc plasmid.

pCRISPR_MS2 containing MS2⁰⁰ spacer flanked by the leader sequence, repeat sequence and a terminal repeat from the *S. thermophilus* DGCC8004 CRISPR2 system was constructed from the pCRISPR_S3 plasmid (27). pCRISPR_S3 was cleaved with XhoI and a fragment containing part of CRISPR locus and two SapI sites was ligated to produce an intermediate plasmid. Further, a fragment containing MS2⁰⁰ spacer was cloned via SapI site into the intermediate plasmid generating a pCRISPR_MS2 plasmid.

pTarget_TcΔhandle mutant plasmid, a pBR322 derivative having the 3′-flanking region of the Tc target in the TcR gene mutated to 5′-CTTTCGGT, was obtained by the Quick-Change Mutagenesis (29).

The HEPN domain (172–428 amino acids of StCsm6) expression plasmid pHEPN was constructed in 2 stages. First, to produce an intermediate plasmid, PCR amplified *csm6* gene was cloned into pET-Duet-1 (Novagen) derivative containing an N-terminal His₉₉ tag and a TEV protease cleavage site coding sequence via Acc651 and AvrII sites. Additional sequence encoding for four Gly residues was introduced between the TEV site and *csm6* by Phusion Site-Directed Mutagenesis (Thermo Fisher Scientific). Finally, inverse PCR cloning (Thermo Fisher Scientific) was used to omit 1–171 codons from *csm6* in the intermediate plasmid.

pHEPN plasmid containing R371A, H376A mutations in HEPN domain was generated by Phusion Site-Directed Mutagenesis (Thermo Fisher Scientific) using pHEPN as the template.

To generate pCARF plasmid, PCR amplified fragment containing coding sequences of TEV protease cleavage site, four Gly coding sequence, CARF domain (1–169 codons of StCsm6) and PCR fragment obtained from pETDuet-1 (Novagen) derivative containing an N-terminal His₁₀₉ tag fused with maltose binding protein (MBP) were assembled using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). MBP fusion to CARF was used to increase solubility of the CARF domain.

To generate pdCARF and pdCARF-Csm6 plasmids containing D12A mutation in *csm6* gene, mutation carrying primers and pCARF and pCsm6 plasmid (10) as the template, respectively, were used to obtain PCR fragments. These PCR fragments were assembled using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs).

**Protein expression and purification**

WT StCsm complex was expressed and purified as described previously (27). Briefly, pCas/Csm plasmid was co-expressed in *E. coli* BL21 (DE3) with pCRISPR_S3, containing four identical spacers S3 interspaced by five repeats of the *S. thermophilus* DGCC8004 CRISPR2 system (27),
and pCsm2N-Tag, encoding the Csm2 gene with an N-terminal StrepII-Tag (27) (see Supplementary Table S1). Cells were harvested and disrupted; soluble StCsm complex was captured by Strep-affinity chromatography via tagged Csm2 protein and subjected to size exclusion chromatography.

Cas10 subunit of StCsm complex was expressed and purified as described previously (8). Briefly, pCas10 plasmid was expressed in E. coli DH10B (ara-) Cells were harvested and disrupted, Cas10 protein was purified using HisTrap and StrepTrap affinity columns (GE Healthcare).

Full-length WT and mutant StCsm6 and StCsm6’ proteins were expressed and purified as described previously (10). Briefly, E. coli DH10B (ara-) was transformed with relevant plasmids (pCsm6, pdCARF-Csm6, pdHEPN-Csm6 or pCsm6’). Cells were grown in LB medium supplemented with an appropriate antibiotic. Cells were harvested and disrupted by sonication, soluble proteins were purified by subsequent His-affinity and Strep-affinity chromatography.

StCsm6 WT and mutant HEPN and CARF domains were expressed in E. coli BL21 (DE3) transformed corresponding plasmid (pCARF, pHEPN, pdCARF or pHEPN). Bacteria were grown at 37°C in LB medium supplemented with ampicillin (50 μg/ml) to the mid-log phase. Expression was induced using 1 mM IPTG and cell suspension was further cultured overnight at 16°C. Cells were harvested and disrupted by sonication, soluble proteins were captured on the HisTrap column (GE Healthcare) using Chromatography buffer (20 mM HEPES (pH 8.0), 0.5 M KCl, 8 mM 2-mercaptoethanol) and 5–500 mM imidazole gradient. Further, to remove the His10-MBP-tag from CARF domain and His-tag from HEPN domain, eluted fractions were pooled and incubated with TEV protease (using a 1:50 (w/w) protease:target protein ratio) at 4°C overnight while dialyzing against the Chromatography buffer. Cleaved parts were removed by His-affinity chromatography. Flow-through proteins were dialyzed against 10 mM Tris–HCl (pH 8.0) buffer containing 300 mM KCl, 1 mM DTT, 0.1 mM EDTA and 50% (v/v) glycerol, and stored at −20°C.

Gel filtration analysis of StCsm6 domains
Analytical gel filtration was carried out at room temperature on an AKTA FPLC system (GE Healthcare) using a Superdex 75 10/300 GL column (GE Healthcare), pre-equilibrated with the Gel filtration buffer (20 mM HEPES (pH 8.0), 500 mM KCl, 8 mM 2-mercaptoethanol). CARF and HEPN samples at 0.7–0.8 mg/ml loading concentration were prepared in 100 μl of the Gel filtration buffer. Elution from the column was monitored by measuring absorbance at 280 nm. The apparent molecular weights of proteins were evaluated based on the elution volume using a series of standards (Gel filtration Calibration Kit from GE Healthcare).

Plasmid interference assay
Escherichia coli BL21 (DE3) was transformed first with two plasmids encoding the CRISPR–Cas system: (i) the relevant pCRISPR plasmid variant—targeting (pCRISPR-Tc) or non-targeting (pCRISPR-S3) and (ii) WT or mutant plasmid pCas/Csm. Such cells were then grown overnight in liquid LB media supplemented with 2% glucose, streptomycin (25 μg/ml) and chloramphenicol (30 μg/ml), and subsequently transformed with the target pTarget_Tc or control pControl plasmids (see Supplementary Table S1). 1/20 dilutions and non-diluted samples of the transformants were plated on LB media agar plates, supplemented with streptomycin (25 μg/ml), chloramphenicol (30 μg/ml) and ampicillin (50 μg/ml). The cells were grown overnight at 37°C and the resultant colonies were counted. Transformation results of pControl were used to normalize respective results obtained with pTarget_Tc. Arbitrary units of CFU were calculated as CFU(pTarget_Tc)/CFU(pControl).

Bacterial culture preparation for cA_n analysis by HPLC-MS
Escherichia coli BL21 (DE3) cells were first transformed with the two plasmids encoding the CRISPR–Cas system: (i) the relevant pCRISPR variant – targeting (pCRISPR-Tc) or non-targeting (pCRISPR-S3) and (ii) WT or mutant plasmid pCas/Csm. Such cells were subsequently transformed with pTarget_Tc, pTarget_Tchandle mutant or pControl_Tc plasmids and grown overnight at 37°C in liquid LB media supplemented with 1.5% glucose, streptomycin (25 μg/ml), ampicillin (50 μg/ml), and chloramphenicol (30 μg/ml). Fresh LB medium was inoculated with the overnight culture (1/20 (v/v)) and was further incubated at 37°C. When the culture reached the mid-log phase, expression of the CRISPR–Cas system was induced by addition of 1 mM IPTG. Cell suspension aliquots of 400 ml were taken at indicated time points and pelleted by centrifugation. Before HPLC–MS analysis, biomasses were tested for the presence of the plasmids (at 4 h after induction) by PCR. The presence of intact Cas10 GGDD motif in Palm domain and the Tc-targeting CRISPR region was verified by sequencing the corresponding PCR products.

To demonstrate that different transcript can be used for cA_n synthesis in cells, E. coli BL21 (DE3) was transformed first with (i) targeting pCRISPR_GFP or non-targeting pCRISPR_S3 plasmid and (ii) WT or mutant plasmid pCas/Csm. Next, such cells were transformed with the target plasmid pTarget_GFP or pControl_GFP plasmid and grown overnight at 37°C in LB medium supplemented with streptomycin (25 μg/ml), ampicillin (50 μg/ml), chloramphenicol (30 μg/ml) and 1.5% glucose. Fresh LB medium was inoculated with the overnight culture (1/20 (v/v)), and bacteria were grown at 37°C. When they reached the mid-log phase, 1 mM IPTG was added and cells were grown for 4 h before being pelleted by centrifugation.

For phage-induced cA_n assays, pCRISPR_MS2 or pCRISPR_S3 and pCas/Csm, dCsm3_dHDCas10ΔCsm6ΔCsm6 plasmids were co-expressed in E. coli NovaBlue (DE3), grown at 37°C in LB medium, supplemented with tetracycline (10 μg/ml), streptomycin (25 μg/ml), chloramphenicol (30 μg/ml). Fresh LB medium was inoculated with the overnight culture (1/20 (v/v)), and bacteria were grown at 37°C. When they reached the mid-log phase, 1 mM IPTG was added,
and after additional 30 min the cells were infected by phage suspension. The cells were collected by centrifugation at 16 h post induction.

ca₆ detection in bacterial lysates

cA₆-containing fraction was extracted from 100 mg (±5%) bacterial pellet by addition of 2 ml 80% ice-cold methanol prior to sonication, thorough vortexing and centrifugation (16 000 g at 4°C) for 10 min. The supernatants were transferred to new tubes and lyophilized to dryness. The samples were subsequently resuspended in 30 μl solvent A (10 mM ammonium acetate in water, pH 7.0) just before analysis.

2–5 μl of each sample was injected using a Vanquish Horizon UPLC (Thermo Fisher Scientific) or an Agilent 1290 Infinity HPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with an Agilent Zorbax Eclipse Plus C18 column (2.1 × 150 mm, 1.8 μm) with a 50 mm guard-column, both kept at 40°C. The chromatographic gradient was run at a flow rate of 300 μl/min with the following solvent composition of solvent A and solvent B (acetonitrile): 97% A from 0–2 min, 97–60% A from 2–8 min, 60–10% A from 8–12 min, 10% from 12–15, before equilibration for 3 min with the initial conditions. The flow from the HPLC was coupled to a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific) or Agilent Q-TOF 6530b system (Agilent Technologies) operated in negative ion mode. Data was extracted using Agilent Masshunter Profinder 10.0 or Thermo Xcalibur v. 3.1 with a mass precision of 10 ppm. The cA₆s were annotated based on accurate mass, retention time from synthesized standards and fragment ions. The molar amount of cA₆s was estimated based on comparison to standards of known quantity. The molar concentration of cA₆s was calculated assuming that each E. coli cell is 1 fl in volume and its wet weight is 1 pg (30).

In vitro cA₆ synthesis assay

The cA₆ synthesis reactions were initiated by adding 10 mM CoCl₂ or 1 mM Mg(CH₃COO)₂ into a mix of 0.2 μM StCsm complex, 0.2 or 10 μM of target RNA S3/1 (27) and 500 μM ATP in the reaction buffer Y (33 mM Tris-acetate (pH 7.6 at 37°C), 66 mM K-acetate, 0.1 mg/ml BSA) or in the reaction buffer T (20 mM HEPES (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA) and carried out at 37°C for 1 or 16 h. The reactions were stopped by adding 15 mM EDTA.

Reaction products were subjected to HPLC-MS analysis.

For time-course experiments, 0.2 μM StCsm complex was mixed with 10 μM target RNA S3/1 (27), 1 mM Mg(CH₃COO)₂, 0.5 mM ATP and 10 nM [α-³²P]-ATP in the Reaction buffer H. 7 μl aliquots were taken at indicated time points and reactions were quenched by addition of 7 μl 1× RNA loading dye (Thermo Fisher Scientific) supplemented with additional (6.25 mM) EDTA. Radioactively labeled products were analyzed by denaturing PAGE (24% gels, 19:1 acrylamide:bis-acrylamide ratio) and visualized by autoradiography.

α³²P-labeled cA₆ for subsequent direct visualization of its hydrolysis products was prepared by mixing 0.2 μM StCsm complex with 10 μM target RNA S3/1 (27), 1 mM Mg(CH₃COO)₂, 0.5 mM ATP and 1 μM [α-³²P]ATP in the reaction buffer H and incubating at 37°C for ~16 h. For the control cA₆ hydrolysis reactions with StCsm complex and Cas10 subunit (see Supplementary Figure S4) the synthesized α³²P-labeled cA₆ was used without further purification. For all other reactions, synthesis products were separated by denaturing PAGE (24% gels, 19:1 acrylamide:bis-acrylamide ratio) and cA₆ was purified from the gel by phenol extraction and ethanol precipitation.

Synthesis and purification of cA₆

200 nM StCsm was mixed with 200 nM target RNA S3/1 (27) and 17 μM of pppA₆ in the Reaction buffer Y and incubated for 20 min at 37°C before initiating the reaction by adding 10 mM CoCl₂. The reactions were performed for 1.5 h at 37°C. The reaction products were fractionated by HPLC (Waters Breeze) on the Discovery HS C18 Column (15 cm × 10 mm, 5 μm) (Sigma-Aldrich Supelco) pre-equilibrated with solvent A (100 mM TEAA (pH 7.0)) at room temperature and 1 ml/min flow rate with a linear gradient of B (60% CH₃CN in buffer A) in A (0–100% of B over 100 ml). Fractions containing cA₆ were pooled and the samples were concentrated on a vacuum concentrator (Eppendorf).

In vitro cA₆ hydrolysis assay

Unless stated otherwise, cA₆ hydrolysis assay was conducted in the Reaction buffer T (40 mM Tris, 20 mM acetic acid (pH8.1 at 25°C), 1 mM EDTA, 0.1 mg/ml BSA) supplemented with 0.5 U/μl RiboLock RNase Inhibitor (Thermo Fisher Scientific) containing 20 μM of synthetic cA₆ (Biolog), synthetic cA₆ (Biolog), linear A₆ (Genes) or ~30 μM of StCsm-produced cA₆s mixture. Reactions were started by adding an indicated amount (0–10 μM) of protein and incubated at 37°C. Reactions were stopped by freezing and subjected to HPLC-MS analysis.

To directly visualize the hydrolysis of cA₆, reactions were conducted in the Reaction buffer T supplemented with 0.5 U/μl RiboLock RNase Inhibitor (Thermo Fisher Scientific) and contained 50 nM of StCsm-produced radioactively labeled α³²P-cA₆. Multiple turnover reactions contained additionally 10 μM of unlabeled cA₆ (Biolog). Reactions were started by adding an indicated amount (0–10 μM) of protein and were carried out at 37°C. 5 μl aliquots were taken at indicated time points and reactions were quenched by mixing with equal amount of 2× RNA Loading Dye (Thermo Fisher Scientific). Reaction products were analysed by denaturing PAGE (24% gels, 19:1 acrylamide:bis-acrylamide ratio) and visualized by autoradiography. For single turnover reactions, ko₉ were determined by fitting a single exponential to the substrate depletion data. For kcat calculation, the reaction rates were determined from the linear parts of the reaction progress curves by linear regression.

For the control cA₆ hydrolysis reactions (presented in Supplementary Figure S4), 10 μM of α³²P-labelled cA₆ was taken from the synthesis mixture and incubated with 10 μM of unlabeled cA₆ (Biolog), 10 nM of protein (StCsm or Cas10) in the Reaction buffer Y in the absence or presence of target RNA S3/1 (27) and 10 mM of Mg²⁺, Mn²⁺,
Co2+ or Ni2+ at 37°C for 8.5 h. Reaction products were separated by denaturing PAGE (24% gels, 19:1 acrylamide:bis-acrylamide ratio) and visualized by autoradiography.

**In vitro RNA hydrolysis assay**

RNA hydrolysis reactions were conducted in the reaction buffer T, supplemented with 0.5 U/µl Ribolock RNase Inhibitor (Thermo Fisher Scientific) and additional 1 mM EDTA. Reactions contained 10 nM of 5′-32P-radiolabeled RNA NS (27) and 0–1 µM cA6 (Biolog) or 0–100 nM cA5 synthesized from the pppA5 precursor. Reactions were started by adding 0–10 µM StCsm6 or its isolated domains and were carried out at 37°C. The reaction products were separated by denaturing PAGE (15% gels, 29:1 acrylamide:bis-acrylamide ratio) and visualized by autoradiography.

**Electrophoretic mobility shift assay**

0.2 nM of 5′-32P-radiolabeled RNA NS (27) was incubated with different amounts (0–3 µM) of CARF domain or (0–1 µM) of dHEPN domain in the Reaction buffer T for 10 min at room temperature. Reaction mixtures were analyzed by native PAGE and visualized by autoradiography.

**RNA hydrolysis competition assay**

RNA hydrolysis reactions were performed at 25°C and contained 10 nM of 5′-32P-radiolabeled RNA NS (27) and 0–500 µM of cA6 in reaction buffer Y. Reactions were initiated by addition of StCsm6 to the final concentration of 10 nM. 5 µl aliquots of the samples were collected at indicated time intervals and quenched by mixing with 2× RNA loading dye (Thermo Fisher Scientific). The reaction products were separated by denaturing PAGE (15%, 29:1 acrylamide:bis-acrylamide ratio) and visualized by autoradiography. RNA hydrolysis constants kobs were determined by fitting a single exponential to the substrate depletion data.

**HPLC-MS analysis of Csm6 cA6 hydrolysis products**

To analyze the cA6 hydrolysis products, electrospray ionization mass spectrometry (ESI-MS) was performed in negative mode using an integrated HPLC/ESI-MS system (1290 Infinity, Agilent Technologies/Q-TOF 6520, Agilent Technologies) equipped with a Supelco Discovery® HS C18 column (7.5 cm × 2.1 mm, 3 µm). Elution was performed with a linear gradient of solvents A (5 mM ammonium acetate in water, pH 7.0) and B (5 mM ammonium acetate in methanol, pH 7.0) at a flow rate of 0.3 ml/min at 30°C as follows: 0–2 min, 0% B; 2–22 min, 20% B; 22–25 min, 50% B, 25–29 min 100% B. Ionization capillary voltage was set to 5000 V, fragmentor— to 150 V. The cA6 hydrolysis products were annotated based on accurate mass, retention time from synthesized standards and fragment ions.

**Treatment of StCsm6 cA6 hydrolysis products with P1 nuclease**

11 µl of the StCsm6 cA6 hydrolysis reaction or 20 µM of control compound (A6>p or A3) was incubated with 10 mU P1 nuclease (Sigma) in the P1 Reaction buffer (10 mM Tris-acetate (pH 7.1 at 37°C), 1 mM Zn-acetate) at 37°C for 1 h. Reaction was stopped by adding 5mM EDTA to the reaction mixture and heating at 75°C for 10 min. 10 µl of this reaction mix was diluted with water and subsequently subjected to HPLC-MS analysis.

**5′-Labeling of StCsm6 cA6 hydrolysis products**

11 µl of StCsm6 cA6 hydrolysis reaction or 20 µM of control compound (A2 or A3) was incubated with 1 U T4 polynucleotide kinase (PNK) and 500 µM ATP in the reaction buffer PNK A (50 mM Tris–HCl (pH 7.6 at 25°C), 10 mM MgCl2, 5 mM DTT, 1 mM spermidine) at 37°C for 30 min. Reaction was stopped by adding EDTA to the final concentration of 5 mM and heating at 75°C for 10 min. Such reaction mixture was subsequently subjected to HPLC-MS analysis.

**Prediction of the StCsm6 CARF active site**

Blastp (31) was used to find StCsm6-CARF domain homologs with 20–28% sequence identity (maximum target sequences 20,000). Multiple alignment was performed with MultAlin (32) and the figure was prepared with ESPript 3.0 (33).

topp (34) was used to superpose StCsm6-CARF domain model (10) and ToCsm6-cA4 structure (25). Figure of StCsm6 CARF-cA4 model was prepared using PyMOL (35).

**Quantification and statistical analysis**

Data analysis was performed using Kryplot 2.0 software (36). Determined cA6 hydrolysis rate constants are presented as the optimal value ±1 standard deviation.

**RESULTS**

**Plasmid interference by S. thermophilus type III-A CRISPR–Cas system in the heterologous E. coli host**

We repurposed the S. thermophilus Type III-A CRISPR–Cas system expressed in heterologous E. coli strain (27) to investigate CRISPR immunity in vivo using plasmid interference assay (Figure 1A). In brief, E. coli host contained two recombinant plasmids: pCas/Csm plasmid expressing all Cas/Csm proteins, except for easl and eas2, and pCRISPR_Tc plasmid carrying a CRISPR array containing 4 copies of spacer targeting the transcript of tetracycline resistance gene (TcR). We transformed these cells with pTarget_Tc plasmid and plated transformants with respective antibiotic combinations (Figure 1B). As a control, we used E. coli cells containing the pCas/Csm plasmid and a pCRISPR_S3 plasmid that carried 4 copies of non-targeting S3 spacer (Supplementary Table S1). In the presence of WT StCsm complex, transformation of pTarget_Tc plasmid into E. coli expressing Tc-targeting crRNA resulted in ~100-fold fewer transformants, as compared to the non-targeting S3 crRNA expressing cells (Figure 1C). This shows that StCsm-mediated immunity is manifested in the heterologous E. coli host and it depends on target transcription as
it was suggested by in vitro experiments (8). Notably, even though both cas/csm genes and CRISPR array were under an IPTG-inducible promoter, the plasmid interference manifested in the absence of IPTG, indicating that traces of the StCsm complex produced due to the promoter leakage are sufficient to provide immunity against exogenic DNA.

To establish the role of different enzymatic activities in plasmid immunity, the interference assay was conducted with pCas/Csm variants (Supplementary Table S1) that carried inactivating mutations: D33A in the Csm3 RNase subunit (dCsm3), D575A and D576A in the Cas10 Palm domain (dHD-Cas10), D16A in Cas10 HD-nuclease domain, or both StCsm6′ and StCsm6 deletion variants (ΔCsm6′/ΔCsm6). In the case of dCsm3, the transformation efficiency was ~10-fold lower in comparison to the WT StCsm (Figure 1C) suggesting that the Csm3-dependent RNase activity is not required for immunity and might even down-regulate it by constantly cutting the target RNA, required for Cas10 activation. Meanwhile disruption of any other enzymatic activity levelled target plasmid transformation efficiency to that of non-targeting plasmid (Figure 1C) suggesting that both DNA degradation and cA₅₆ signaling pathways are required for S. thermophilus type III-A CRISPR–Cas mediated immunity. Intriguingly, simultaneous inactivation of both Csm3 RNase and Cas10 DNase secured efficient plasmid interference, implying that in the absence of Csm3 RNase activity, the cA₅₆-dependent pathway alone is able to provide plasmid immunity, as Csm3 inactivation results in a constant and a high production of cA₅₆ (10,12).

Monitoring cA₅₆ production in vivo

We have previously shown that StCsm complex in vitro synthesizes cA₅₆s (n = 2–6) from ATP. cA₃ was identified as the major reaction product; however, only the minor cA₆ product activated StCsm6 RNase (10). To get an insight into StCsm6 RNase activation mechanisms in the cell, we decided to examine cA₅₆ metabolites produced in vivo in the heterologous E. coli host expressing the S. thermophilus type III-A CRISPR–Cas system. To monitor amount of the cyclic adenylates produced in E. coli we employed high-performance liquid chromatography–mass spectrometry (HPLC-MS) technique. In these experiments we used the same E. coli host as in the interference assay (Figure 1A), except that the inducer IPTG was added to the cell culture. To prevent RNA and DNA degradation by the CRISPR–Cas system and ensure cell viability, we used a pCas/Csm variant where all proteins’ activities except that of the Cas10 Palmsynthase have been inactivated by mutations (dCsm3, dHD-Cas10, ΔCsm6′/ΔCsm6) (Supplementary Table S1). Such setup enabled monitoring of cA₅₆ level in the host constitutively expressing the targeted TcR gene in the Target_Tc plasmid. Targeted HPLC-MS analysis reported various cA₅₆s, from n = 2 up to 6 but the equilibrium was significantly shifted towards cA₅ and cA₆ species (Figure 2A and Supplementary Figure S1A). The cA₅ and cA₆ amount was significantly increased during the first 4 h after IPTG induction to 22.3 ± 4.6 and 12.9 ± 3.8 pmol/mg of wet cell culture, which translates to 22.3 ± 4.6 and 12.9 ± 3.8 μM concentration in the E. coli cell. The amount of cA₅ and cA₆ detected in the heterologous E. coli host is comparable to the amount of the c-di-AMP metabolite (13.8 ± 1.1 pmol/mg) present in the Streptococcus suis host (37). Interestingly, after 8 hours post-induction, the amount of cA₅ and cA₆ had slightly decreased while cA₃ and cA₄ kept increasing (Figure 2A). It cannot be excluded that in...
Figure 2. cAₙ synthesis by StCsm complex in vivo analyzed by HPLC-MS. (A) E. coli harboring plasmids (as indicated in Figure 1A and Supplementary Table S1) encoding the TcR transcript and expressing StCsm complex mutant (dCsm3, dHD-Cas10 and ΔCsm6 ΔCsm6) targeting TcR were grown in the presence of inducer IPTG. The samples were taken at defined time points after induction and cAₙ amount was analyzed by HPLC-MS. (B) Analysis of cAₙ produced in E. coli cells, which express StCsm complex mutant (dCsm3, dHD-Cas10 and ΔCsm6 ΔCsm6) targeting MS2 transcripts, after 16 h of phage MS2 infection. (C) HPLC-MS analysis of cAₙ amount changes over time in vivo in the absence (left) or presence (right) of StCsm6 and StCsm6′ encoding genes. The inset depicts the same data on a different scale.

this case the decrease in total cellular ATP concentration shifts the ATP polymerization:cyclisation reaction equilibrium to shorter reaction products.

We have previously shown that for in vitro cAₙ synthesis the following conditions must be met: (i) spacer sequence of crRNA must be complementary to the protospacer sequence of target RNA, (ii) the 3′-flanking sequence of the target RNA must not base-pair with the 5′-handle of cr-RNA and (iii) Cas10 in the effector complex must contain an intact GGDD-active site in the Palm domain (10). HPLC-MS analysis of cA₆ metabolites in E. coli (Supplementary Figure S2A) confirmed the cAₙ synthesis requirements established in vitro for S. thermophilus and other type III systems (10,11,13–15).

To probe whether cA₆ is produced in response to phage infection, we performed HPLC-MS analysis of cAₙ metabolites in the phage-infected cells. Heterologous E. coli host carrying pCas/Csm and pCRISPR_MS2 plasmids (Supplementary Table S1) was infected with RNA coliphage MS2 (27). Targeted HPLC-MS analysis revealed accumulation of cA₅ and cA₆ in response to MS2 phage infection (Figure 2B and Supplementary Figure S1B). Taken together, analysis of cAₙ metabolites in E. coli cells demonstrates that the type III-A CRISPR–Cas system in the heterologous host produces signaling molecules in response to exogenic nucleic acids that derive either from plasmid or phage.

Optimization of cA₆ synthesis in vitro

HPLC-MS analysis of cAₙs produced in E. coli cells revealed an altered cAₙ profile in comparison to in vitro experiments (10), suggesting that discrepancies in the size distribution could arise due to different reaction conditions. Therefore, we re-evaluated cA₆ synthesis in vitro using a different StCsm: target RNA ratios, buffers and Co²⁺ or Mg²⁺ ions as cofactors (Supplementary Figure S3A). We found a significant shift in the cAₙ size distribution profile towards a larger ring size and ultimately to cA₆ in the presence of target RNA excess and Mg²⁺. Interestingly, prolonged incubation of StCsm with ATP resulted in a slightly increased cA₃ and cA₄ yield in vitro (Supplementary Figure S3A, lane x), in line with in vivo HPLC-MS data (Figure 2A, 8 h point), even though Csm3 D33A mutation abolishes target RNA cleavage.

The cAₙ-dependent RNA-degradation pathway has to be switched-off once the exogenic nucleic acid is eliminated, raising the question of how cyclic oligoadenylate concentration is regulated in the cell. HD domain phosphodiesterases, that are widespread in bacteria, are often involved in cA₂ degradation in the cell (38). While the Cas10 subunit of StCsm contains an HD domain, which is involved in ss-DNA degradation (8), we did not detect any cAₙ degradation by either isolated Cas10 or the StCsm complex (Supplementary Figure S4A). It has been recently shown that some CARF domain-containing proteins act as ring nucleases that degrade cA₄ molecules (22–25). We reasoned that the CARF domain of StCsm6 could also degrade the cA₆ activator. Analysis of cAₙ amount in the cell provided indirect support for this hypothesis: E. coli cells expressing WT StCsm6′ and StCsm6 exhibited drastically reduced levels of cA₆ and other cAₙ (Figure 2C and Supplementary Figure S1C) suggesting that StCsm6 proteins are responsible for cAₙ degradation in the cells.
To investigate whether Csm6 proteins degrade cA₆s, we isolated individual StCsm6 and StCsm6' proteins and studied their activities in vitro. In the absence of divalent metal ions both proteins degraded radiolabeled cA₆, generating faster migrating products in denaturing PAGE (Figure 3B and Supplementary Figure S4B). HPLC-MS analysis revealed a mixture of three cleavage products, namely cAMP (2′,3′-cyclic AMP (A>p) or 3′,5′-cyclic AMP), and linear or cyclic di- and tri-adenylates (Figure 3B and Supplementary Figure S4B). To differentiate between cyclic (cA₆) and terminal 2′,3′-cyclic phosphate-containing linear oligoadenylates (Aₖ>p), we further treated these products with P1 nuclease and polynucleotide kinase (PNK). Subsequent HPLC-MS analysis confirmed that StCsm6 reaction products are converted to adenosine and AMP by P1 nuclease, suggesting an asymmetric structure of the substrate. Moreover, these products could be 5′-labeled by PNK to produce di- and tri-adenylate phosphates, confirming the presence of a free terminal 5′-OH group (Supplementary Figure S5). Together, these experiments are consistent with a linear structure of the products, suggesting that StCsm6 in the absence of metal ions degrades cA₆ into a mixture of A>p, A₃>p and A₅>p. This is not surprising, since cyclic 2′,3′-phosphate products are usually generated by other metal-independent RNases like ribonuclease A, Cas6 and tRNA splicing endonuclease (39,40).

cA₆ degradation by isolated CARF and HEPN domains

According to HPLC-MS analysis, E. coli cells in the presence of StCsm6 and StCsm6' RNases exhibit drastically reduced levels of cA₆ and other cAₖs (Figure 2C and Supplementary Figure S1C), suggesting their involvement in cA₆ degradation. StCsm6 is comprised of CARF and HEPN domains. It has been previously shown that the CARF domain of other, cA₄-dependent, Csm6 proteins is able to cleave the cA₄ activator (24,25), implying a similar function for the StCsm6 CARF domain. On the other hand, the HEPN domain of StCsm6 protein is a RNAse, which digests RNA even in the absence of the cA₆ activator albeit at high concentrations (10); therefore, it is possible that it could also degrade cA₆. To establish which domain is responsible for cA₆ cleavage, we engineered and purified individual CARF (1–169 aa) and HEPN (172–428 aa) domains of StCsm6 (Figure 3A and Supplementary Figure S6A) that both remained dimers in solution (Supplementary Figure S6B) like WT StCsm6 (10). Isolated HEPN domain degraded linear RNA at the same rate as WT StCsm6 in the absence of activator (Supplementary Figure S7).

Next, we analyzed cA₆ degradation activity of individual CARF and HEPN domains. Surprisingly, both CARF and HEPN domains were able to degrade radiolabeled cA₆; however, under single turnover conditions the CARF domain was much more efficient at degrading cA₆ than the HEPN domain (Figures 3C and D). Mutations of HEPN active site residues R371A and H376A (10) required for RNAse activity dramatically impaired hydrolysis of both cA₆ (Figure 3D) and linear RNA (Supplementary Figure S7C), indicating that the same active site is responsible for cA₆ and linear RNA cleavage.

To predict the putative active site of CARF in StCsm6, we superposed StCsm6 CARF domain model (10) with T. omnurum Csm6-cA₄ structure (PDB ID: 6O6V) (25) and identified that conserved D12, S105 and T107 residues of StCsm6 are located near the cyclic oligoadenylate (Supplementary Figure S8). While S105 and T107 are important for cA₆ binding (10), D12A mutation (dCARF variant) completely abolished the cA₆ hydrolysis activity of the CARF domain (Figure 3C), indicating that this residue is important for cA₆ cleavage.

Next, we analyzed reaction products generated by CARF and HEPN domain-mediated cleavage of cA₆. HPLC-MS analysis revealed that the CARF domain converts cA₆ to A₃>p; however, linear A₅>p is also present in the reaction mixture (Figure 3C) and can be further cleaved by CARF into the final reaction product, A₅>p (Supplementary Figure S7D). Presumably, each monomer of homodimeric CARF domain cuts cA₆ producing A₅>p, similarly to the ring nucleases that split cA₄ into two A₂>p (22). The reaction occurs non-concertedly since A₆>p product is identified in the reaction mixture. Unlike CARF domain, HEPN domain degraded cA₆ into a mixture of A₅>p, A₇>p, A₉>p and A₇>p or 3′-phosphoadenylates (Figure 3D). Together, this shows that both StCsm6 domains function as metal-independent ring nucleases that degrade cA₆ but generate different cleavage products.

cA₆ degradation by the full-length StCsm6

We found that isolated CARF and HEPN domains of StCsm6 both are capable of degrading cA₆; moreover, binding of cA₆ to the CARF domain allosterically stimulates RNAse activity of the HEPN domain (10). Therefore, we next examined possible allosteric connection between the CARF and HEPN domains in cA₆ degradation by the WT StCsm6. To dissect possible allosteric interactions between the domains, we compared cA₆ cleavage by WT StCsm6 and mutants that were deficient in cA₆ cleavage in either CARF (dCARF-Csm6) or HEPN (dHEPN-Csm6) domains (Figure 3E).

Under enzyme excess (E>>S) conditions, dHEPN-Csm6 cleaved cA₆ at the rate (kₙ₁(Csm6-Csm6) = 0.27 ± 0.10 min⁻¹) that was comparable to the WT StCsm6 (kₙ₁(WT) = 0.19 ± 0.04 min⁻¹) and about 2-fold slower than the isolated CARF domain (kₙ₁(CRF) = 0.47 ± 0.18 min⁻¹) (Figures 3B, C and E). This suggests that at low cA₆ concentrations CARF domain of StCsm6 functions as a ring nuclease, and the contribution of HEPN domain is negligible. Indeed, dCARF-Csm6 variant cleaved cA₆ very slowly, similarly to the isolated HEPN domain (Figures 3D and E).

Low cA₆ degradation activity of the HEPN domain (Figure 3D) may be due to poor binding of cA₆. We were unable to determine cA₆ binding affinity of the HEPN domain directly; therefore, we indirectly evaluated cA₆ binding using a cleavage competition assay. We mixed full-length WT StCsm6 and RNA at a 1:1 ratio, and measured the cleavage rate of RNA in the presence of increasing concentrations of cA₆ (Figure 4A). At low cA₆ concentrations StCsm6 hydrolyzed RNA very slowly because HEPN was not allosterically activated by cA₆ binding in the CARF domain. RNA hydrolysis rate increased significantly at 10 nM of cA₆, im-
Figure 3. cA₆ hydrolysis by CARF and HEPN domains of StCsm6 in vitro. (A) Schematic illustration of WT StCsm6 protein and its CARF (1–169 aa) and HEPN (172–428 aa) domains. Residues important for activity are indicated. In vitro hydrolysis of cA₆ by (B) WT StCsm6, (C) isolated CARF domain, (D) isolated HEPN domain and (E) full-length StCsm6 with R371A, H376A (dHEPN-Csm6) and D12A (dCARF-Csm6) mutations, as analyzed by PAGE and HPLC-MS. All cA₆ hydrolysis reactions analyzed by PAGE contained 50 nM of ³²P-labeled cA₆ and 1 μM of WT, CARF, dHEPN-Csm6 or dCARF-Csm6, or 10 μM of HEPN. MS reactions contained 20 μM of non-labeled cA₆ and 10 nM of WT, 1 μM of CARF or 10 μM of HEPN.

Applying cA₆ binding in the CARF domain and allosteric activation of the RNase activity of HEPN domain. Further increase of cA₆ concentration up to 100 nM did not change RNA hydrolysis rate. Inhibition of RNA hydrolysis was observed at cA₆ concentrations above 100 μM (Figure 4A), suggesting that HEPN has lower binding affinity towards cA₆ than CARF domain.

Therefore, we next analyzed StCsm6 cA₆ cleavage under multiple turnover (S > E) conditions. Under these conditions WT StCsm6 rapidly degraded cA₆ with a turnover rate of > 20 min⁻¹ (Figure 4B). It is likely that at high concentrations cA₆ binds both to CARF and HEPN, thereby allosterically activating effective degradation of cA₆ by the HEPN nuclease. In contrast, isolated CARF domain and HEPN-deficient dHEPN-Csm6 cleaved cA₆ significantly slower ($k_{cat} < 0.1$ min⁻¹). Taken together, these results suggest that an activated HEPN is a much more efficient ring nuclease than CARF. The dCARF-Csm6 mutant with impaired cA₆ cleavage at the CARF domain degraded cA₆ much faster than isolated HEPN domain (Figure 4B), indicating that cA₆ binding but not cleavage is required for activation of the HEPN ring nuclease. The slower cA₆ cleavage
Figure 4. cA6 hydrolysis by the HEPN domain of StCsm6 in vitro. (A) RNA hydrolysis competition assay. RNA hydrolysis rate of WT StCsm6 (10 nM 32P-labeled RNA, 10 nM StCsm6) was monitored in the presence of increasing concentration (0–500 μM) of cA6. (B) cA6 hydrolysis under multiple turnover conditions. 1 μM of full-length WT StCsm6 or its CARF and HEPN mutant variants were mixed with 50 nM α32P-labeled cA6 and 10 μM of non-labeled cA6 and its hydrolysis products were analyzed by PAGE. (C) cA6 hydrolysis by WT StCsm6. A mixture of cA6s (≈30 μM) produced by the StCsm complex was incubated with 10 nM WT StCsm6 for 90 min and analyzed by HPLC-MS.

by the dCARF-Csm6 mutant as compared to WT may result from compromised cA6 binding affinity, caused by the D12A mutation in the CARF domain. In agreement with impaired cA6 binding, an increased cA6 concentration (up to 300 nM) was required for activation of RNA degradation by the dCARF-Csm6 mutant (Supplementary Figure S7E).

cA6 degradation by StCsm6

HPLC-MS analysis revealed that in E. coli host expressing StCsm6 and StCsm6’ RNases, the level of all cA6s is dramatically decreased (Figure 2C and Supplementary Figure S1C), suggesting that StCsm6 is capable of degrading different cA6s including cA6 produced by the Csm effector complex. We found that after prolonged incubation StCsm6 converted a mixture of cA6s to A3>p and A>p (Figure 4C). Isolated CARF domain failed to cleave cA3 and cA6 in the mixture of cA6 but was able to convert cA6 to A3>p and linearize cA5 (Supplementary Figure S9A). However, cA5-bound CARF did not activate RNase activity of the HEPN domain of WT StCsm6 (Supplementary Figure S9B). Isolated HEPN domain was not as effective in converting cA6s to A3>p as the full-length StCsm6 (Figure 4C and Supplementary Figure S9A).

Together, these results show that StCsm6 has two ring nucleolytic activities: the CARF domain, which is specific to cA6 and upon cleavage produces A3>p, and the HEPN domain, which upon allosteric activation indiscriminately degrades cA6 (n = 3–6) oligoadenylates down to A>p.

DISCUSSION

Type III CRISPR–Cas systems are complex, versatile and efficient barriers protecting host cells against foreign nucleic acids (41). We have previously shown that in the S. thermophilus type III-A CRISPR–Cas system, target RNA binding by the StCsm effector complex activates the Cas10 protein, whose HD domain starts ssDNA degradation, and the Palm domain initiates synthesis of cyclic cA6s (n = 2–6) (8,10). The cA6 produced in response to viral infection acts
as a signaling molecule that binds to the CARF domain of the stand-alone StCsm6 and StCsm6’ proteins and allosterically activates the non-specific ribonucleolytic activity of their HEPN domains (10). Here, we aimed to establish the control and regulation mechanisms of the cA₆-dependent StCsm6 RNase both in vitro and in vivo.

**cA₅ and cA₆ are major reaction products produced in the cell**

While the target plasmid degradation by the Cas10 DNase has been monitored in cells over time (3,20), the Cas10-mediated cA₆ synthesis in cells has not yet been analyzed. Using targeted HPLC-MS analysis we show that in response to the target sequence transcription from a plasmid or a phage, *S. thermophilus* type III-A CRISPR–Cas system in heterologous *E. coli* host produces cA₅, cA₆ of different ring size (Figure 2). The cyclic oligoadenylates with larger ring sizes (n = 5–6) were most abundant, including cA₆, that binds to the sensor CARF domain and activates StCsm6 and StCsm6’ RNases. The cA₆ size distribution detected in the cell matched that observed in vitro under high target RNA concentrations (Supplementary Figure S3). In addition to cA₅ and cA₆, significant amount of cA₄ and cA₃ was also detected (Figure 2A). Bacterial cGAS/DncV-like nucleotidyltransferases synthesize cyclic tri-nucleotide compounds (42) including cA₃ that activate restriction enzyme-like endonuclease effector NucE in the bacteriophage immunity pathway (43,44). It cannot be excluded that cA₃ produced by the type III systems could contribute to alternative pathways in bacterial signaling, such as NucE activation.

**Self-regulation of the cA₆ signaling pathway**

Plasmid interference experiments revealed (Figure 1C) that, in agreement with previous type III studies (12,20,45,46), both ssDNA degradation and cA₆ signaling pathway are required for *S. thermophilus* type III-A CRISPR–Cas mediated immunity. Notably, *E. coli* cells do not survive when they carry heterologous type III CRISPR–Cas system with catalytically deficient Csm3 and DNase variants (Figure 1C), indicating that cA₆-activated StCsm6 and StCsm6’ RNases alone enable plasmid immunity. Since cA₆ signaling pathway is critical for CRISPR–Cas type III immunity, it must be regulated to avoid host damage. First, cA₆ synthesis should be switched-off once the foreign transcript is eliminated, and next, cA₆ messengers produced in response to infection should be cleared from the cell. Csm3 RNase controls cA₆ level at the synthesis stage through degradation of target RNA that is required for initiation of cA₆ synthesis (10,14). Csm3 inactivation results in higher cA₆ production yield in vitro (10,12). In the cell, it should promote upsurge in cA₆ production and subsequent increase of StCsm6 RNase activity, which may cause dormancy (47,48) or cell death. Uncontrollably enhanced StCsm6 RNase activity presumably ensures plasmid immunity in the case of DNase-deficient variant of StCsm (Figure 1C).

HPLC-MS analysis of cell metabolites revealed that in the presence of StCsm6 and StCsm6’, cA₅s, including cA₆, are depleted from *E. coli* cells (Figure 2C and Supplementary Figure S1C). CARF ring nucleases were recently shown to be responsible for degradation of specific cA₄ messengers (22–25). However, in the *S. thermophilus* type III-A system cA₅s are degraded by StCsm6 independently of the ring size, suggesting a different cA₆ degradation pathway. Indeed, we show here that both CARF and HEPN domains of StCsm6 proteins function as ring nucleases but exhibit different binding specificities and produce distinct reaction products (Figures 3 and 4).

**StCsm6 CARF is a cA₆-specific ring nuclease**

Isolated StCsm6 CARF ring nuclease shows no activity either on linear RNA or cA₅ (Supplementary Figures S7B and S9A) but tightly binds cA₆ and converts it to A₃-p (Figures 3C and 4A). Similarly to other CARF ring nucleases (22,24) cleavage of cA₆ by StCsm6 CARF domain is slow (Figures 3 and 4). Identification of the linear A₆-p intermediate during the reaction (Figure 3D) suggests that both active sites in CARF dimer act independently. Similarly, it has been reported that *T. onnurineus* Csm6 CARF domain cleaves cA₄ first to produce A₄-p and subsequently to A₂-p (25). CARF is also able to bind and linearize cA₃ (Supplementary Figure S9A). However, cA₅ is not an optimal substrate for the CARF ring nuclease and is only cleaved by one of the active sites of the CARF domain dimer.

Aspartate D12 is proposed as a catalytic residue in the StCsm6 CARF domain since D12A mutation abolishes cA₆ cleavage activity (Figure 3C). Catalytic residues identified in ring nucleases are very diverse: D12 in Csm6 of *S. islandicus* and S9A in Csm6 of *T. thermophilus*, T10/T11 in Csm6 of *T. thermophilus*, W14 in Csm6 of *T. onnurineus* and Y14 in Csx1 of *S. islandicus* (23–25). Variation of catalytic residues in the active sites of CARF domains argues that there are multiple ways for positioning and activation of 2′-OH group of the ribose for nucleophilic attack that results in the phosphodiester bond cleavage during cyclic oligoadenylate (cA₄ and cA₅) degradation. Recent structural and biochemical study of *Enterococcus italicus* Csm6, published during preparation of this manuscript, also revealed that CARF domain is capable of cleaving cA₆ (49).

**cA₆-binding at CARF domain promotes HEPN-mediated degradation of RNA and cA₅**

In contrast to the CARF domain that is specific to cA₆, the HEPN domain shows clear preference for RNA degradation (Figure 4A and Supplementary Figure S7A). However, at high cA₆ concentrations, HEPN domain also becomes a ring nuclease that effectively degrades cA₆ down to A₂-p and A₃-p (Figures 3E and 4C). Unlike the CARF ring nucleases that exhibit strict specificity and cleave either cA₄ or cA₅ oligoadenylates, the HEPN domain degrades all cA₅s, including cA₆, produced by the Csm effector complex (Figure 4C). It first opens the ring of cA₆ and then chops linearized A₆-p like any other linear RNA (StCsm6 HEPN domain cuts linear RNA preferentially at AA or GA nucleotides (10)). Notably, an A₃-p molecule is bound to the HEPN domain in the structure of *T. onnurineus* Csm6 (25). The allosterically activated HEPN domain in StCsm6 is a much faster ring nuclease than the CARF domain (Figure 4B). The rate of cA₆ hydrolysis by the activated HEPN domain is similar to the cA₄ hydrolysis rate by the phage
Figure 5. Model of self-regulation of the cA₆ signaling pathway of type III-A CRISPR–Cas system. Recognition of the invasive transcript by the Csm complex triggers three catalytic activities: the Csm3-mediated cleavage of the invasive transcript, degradation of DNA template by the HD domain of Cas10, and synthesis of cA₂₋₆ by the Palm domain of Cas10. The cA₆ acts as a signaling molecule that binds to the CARF domain of the Csm6 ribonuclease and activates non-specific RNA degradation by the HEPN domain. CARF domain acts as ring nuclease that slowly converts cA₆ into linear A₈₋₆ and finally into A₃₋₆. Cleavage of the target RNA by Csm complex switches off cA₆ synthesis by the Cas10 subunit. cA₆ and A₃₋₆ remain bound to the CARF domain and activate the HEPN domain for degradation of cA₆ accumulated in the cell, thereby switching off signaling and restoring the pre-infection cell state.

ring nuclease, which is the fastest ring nuclease known to date and functions as viral anti-CRISPR (AcrIII-1) (26,50). During the preparation of the present manuscript, Foster et al. reported that Pyrococcus furiosus Csx1 HEPN is also capable of degrading cyclic oligoadenylates (51). It would be interesting to see if the allosterically activated HEPN domain degrades cA₆ in the cA₆-dependent CARF-HEPN nucleases (25,26,50) and whether majority of Csm6/Csx1 proteins are dual ring nucleases as StCsm6.

Csm6 as the regulator of type III-A CRISPR–Cas immunity

Our data and previous reports show that CARF-containing Csm6/Csx1 proteins play major role in the type III CRISPR–Cas immunity (46,52). Here we propose the following model of the type III-A Csm6-dependent CRISPR–Cas immunity (Figure 5). At an early infection stage, in response to the target RNA binding the Csm effector complex starts to produce cyclic oligoadenylates of various ring size. Cyclic hexa-adenylate cA₆ acts as a secondary messenger that allosterically activates Csm6 HEPN RNase through binding at the CARF domain and initiates RNA degradation. CARF domain also functions as a ring nuclease that slowly converts cA₆ to A₆₋₆ and finally to A₃₋₆ (Figure 3C). The slow turnover rate suggests that A₆₋₆ remains bound to the CARF domain and may continue to activate the HEPN domain (11) until it is converted to A₃₋₆ thereby switching off HEPN activity. This would provide a timer mechanism for the HEPN domain activity control.

The HEPN domain shows clear preference for RNA over cA₆₋₆ (Figure 4A and Supplementary Figure S7C) but can switch to cA₆ degradation at high cyclic oligoadenylate concentration presumably at the late stage of phage infection. Notably, high concentrations of cA₅ and cA₆ (22.3 ± 4.6 and 12.9 ± 3.8 μM) are detected in E. coli harbouring the S. thermophilus type III-A CRISPR–Cas system and its target (Figure 2A).

E. coli is able to survive in the presence of cA₆-activated Csm6 RNase, which suggests that Csm6 does not degrade all cellular transcripts as its activity is self-regulated through the concerted action of CARF and HEPN ring nucleases (Figures 1C, 2C and 4). In the early infection stage, the activated Csm6 RNase could be co-localized with the Csm complex due to increased local concentration of cA₆, which is synthesized in response to target RNA binding. After cleavage of the foreign transcript at the late stage of infection, the cA₆ synthesis is terminated (10,14) and accumulated cA₆₋₆ are degraded by cA₆-activated HEPN domain, thus completely eliminating the signaling molecules and switching off RNases (Figure 5). Such self-regulation mechanism would prevent the cell death and ensure the recovery of the host in the post-infection stage.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Česlovas Venclovas for discussions and Giedrius Sasnauskas for critical reading of the manuscript and comments.

FUNDING

Research Council of Lithuania [MIP-40/2013, APP-3/2016 to G.T.]. Funding for open access charge: Research Council of Lithuania.

Conflict of interest statement. M. K., G.T. and V. S. are inventors on patent applications related to Csm complex. V. S. is a founder of CasZyme and a chairman of the board.
cGAS-like enzymes synthesize diverse nucleotide signals. *Nature*, 567, 194–199.
43. 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.
44. 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. (2019) HORMA domain proteins and a Trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity. *Mol. Cell*, 77, 709–722.
45. Foster, K., Kalter, J., Woodside, W., Terns, R. M. and Terns, M. P. (2019) The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR–Cas systems. *RNA Biology*, 16, 449–460.
46. Hatoum-Aslan, A., Maniv, I., Samai, P. and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR–Cas system. *J. Bacteriol.*, 196, 310–317.
47. Meeske, A. J., Nakandakari-Higa, S. and Marraffini, L. A. (2019) Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature*, 570, 241–245.
48. Abudayeh, O. O., Gootenberg, J. S., Esletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A. et al. (2017) RNA targeting with CRISPR–Cas13. *Nature*, 550, 280–284.
49. Garcia-Doval, C., Schwede, F., Berk, C., Rostol, J. T., Niewoechner, O., Tejero, O., Hall, J., Marraffini, L. A. and Jinek, M. (2020) Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. *Nat. Commun.*, 11, 1596.
50. Samolygo, A., Athukoralage, J. S., Graham, S. and White, M. F. (2020) Fuse to defuse: a self-limiting ribonuclease-ring nuclease fusion for type III CRISPR defence. *Nucleic Acids Res.*, 48, 6149–6156.
51. Foster, K., Gruschow, S., Bailey, S., White, M. F. and Terns, M. P. (2020) Regulation of the RNA and DNA nuclease activities required for Pyrococcus furiosus Type III-B CRISPR–Cas immunity. *Nucleic Acids Res.*, 48, 4418–4434.
52. Deng, L., Garrett, R. A., Shah, S. A., Peng, X. and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol. Microbiol.*, 87, 1088–1099.