https://doi.org/10.15388/vu.thesis.771
https://orcid.org/0000-0002-6913-6884

VILNIUS UNIVERSITY

Dalia Smalakytė

Insights into Type III CRISPR-Cas Immunity: Mechanisms of Csm6 Deactivation and a Comprehensive Characterization of the Lon-SAVED Effector

DOCTORAL DISSERTATION

Natural Sciences,
Biochemistry (N 004)

VILNIUS 2025

The dissertation was prepared between 2019 and 2025 at Vilnius University, Life Sciences Center, Institute of Biotechnology.
The research was supported by Research Council of Lithuania.

The dissertation is defended on an external basis.

Academic Consultant – Dr. Gintautas Tamulaitis (Vilnius University, Natural Sciences, Biochemistry – N 004)

Dissertation Defense Panel:
Chairman – Prof. Dr. Edita Sužiedėlienė (Vilnius University, Natural Sciences, Biochemistry – N 004).
Members:
Prof. Dr. Rolandas Meškys (Vilnius University, Natural Sciences, Biochemistry – N 004),
Dr. Patrick Pausch (Vilnius University, Natural Sciences, Biochemistry – N 004),
Dr. Constantinos Patinios (Vilnius University, Natural Sciences, Biochemistry – N 004),
Dr. Raymond Staals (Wageningen University, The Netherlands, Natural Sciences, Biochemistry – N 004).

The dissertation shall be defended at a public meeting of the dissertation defence panel at 14:00 on 13th of June 2025 in Room R401 of Vilnius University Life Sciences Center.
Address: Saulėtekio al. 7, LT-10257, Vilnius, Lithuania
Tel. +370 5 223 4419; e-mail: info@gmc.vu.lt

The text of this dissertation can be accessed at Vilnius University library, as well as on the website of Vilnius University:
www.vu.lt/lt/naujienos/ivykiu-kalendorius

https://doi.org/10.15388/vu.thesis.771
https://orcid.org/0000-0002-6913-6884

VILNIAUS UNIVERSITETAS

Dalia Smalakytė

III tipo CRISPR-Cas priešvirusinės apsaugos sistemos efektorių Csm6 deaktyvacijos ir Lon-SAVED veikimo mechanizmų tyrimas

DAKTARO DISERTACIJA

Gamtos mokslai,
Biochemija (N 004)

VILNIUS 2025

Disertacija rengta 2019-2025 metais Vilniaus universitete, Gyvybės Mokslų centre, Biotechnologijos institute.
Mokslinius tyrimus rėmė Lietuvos mokslo taryba.

Disertacija ginama eksternu.

Mokslinis konsultantas – dr. Gintautas Tamulaitis (Vilniaus universitetas, gamtos mokslai, biochemija – N 004)

Disertacijos gynimo taryba:
Pirmininkė – Prof. dr. Edita Sužiedėlienė (Vilniaus universitetas, gamtos mokslai, biochemija - N004).
Nariai:
Prof. dr. Rolandas Meškys (Vilniaus universitetas, gamtos mokslai, biochemija – N 004),
Dr. Patrick Pausch (Vilniaus universitetas, gamtos mokslai, biochemija – N 004),
Dr. Constantinos Patinios (Vilniaus universitetas, gamtos mokslai, biochemija – N 004),
Dr. Raymond Staals (Vageningeno universitetas, Nyderlandai, gamtos mokslai, biochemija – N 004).

Disertacija ginama viešame gynimo tarybos posėdyje 2025 m. birželio mėn. 13 d. 14 val. Vilniaus universiteto Gyvybės mokslų centro R401 auditorijoje.
Adresas: Saulėtekio al. 7, LT-10257, Vilnius, Lietuva
Tel. +370 5 223 4419; el. paštas info@gmc.vu.lt

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu:
https://www.vu.lt/naujienos/ivykiu-kalendorius

CONTENTS

ABBREVIATIONS

[S] < [E]	single turnover conditions
[S] > [E]	multiple turnover conditions
3′2′-cGAMP	3′-5′, 2′-5′ cyclic-GMP-AMP
3′3′-cGAMP	3′-5′, 3′-5′ cyclic-GMP-AMP
6H	six α-helices
A1/A2/A3/A4	$1^{st}/2^{nd}/3^{rd}/4^{th}$ adenosine
a.u.	arbitrary units
aa	amino acid
Acb	anti-CBAS
Apyc	anti-Pycsar
A_n>p	linear oligoadenylate with cyclic 2′-3′ phosphate
A_2>p/A_3>p/A_4>p/ A_5>p/A_6>p	linear di/tri/tetra/penta/hexaadenylate with cyclic 2′-3′ phosphate
A_2p/A_4p	linear di/tetraadenylate with 3′ phosphate
ATP/pppA	adenosine tri-phosphate
ATPase	adenosine 5′-triphosphatase
Avs	antiviral STAND
bGSDM	bacterial gasdermin
BLI	biolayer interferometry
BSA	bovine serum albumin
$cA_3/cA_4/cA_5/cA_6$	cyclic tri/tetra/penta/hexaadenylate
Cad1/CAAD	CRISPR-Cas-associated adenosine deaminase
CalpL	CRISPR associated Lon protease
Cami1	CRISPR-Cas-associated mRNA interferase 1
Can	CRISPR ancillary nuclease
cA_n	cyclic oligoadenylate
Cap	CD-NTase associated protein
CARF	CRISPR-associated Rossmann Fold
Cas	CRISPR associated
CBASS	cyclic oligonucleotide-based signaling system
CCa	*Candidatus* Cloacimonas acidaminovorans
CD-NTase/Cdn	cGAS/DncV-like nucleotidyltransferase
CFU	colony forming unit
cGAS-STING	cyclic GMP-AMP synthase–stimulator of interferon genes
CHAT	Caspase HetF Associated with Tprs
CHB	chromatography buffer
Craspase	CRISPR-guided caspase

CRISPR	clustered regularly interspaced short palindromic repeats
Crn	CRISPR ring nuclease
crRNA	CRISPR RNA
Cryo-EM	cryogenic electron microscopy
DNase	deoxyribonuclease
DTT	dithiothreitol
DUF	domain of unknown function
Ec	*Escherichia coli*
ECF	extracytoplasmic function
EDTA	ethylenediaminetetraacetic acid
Ei	*Enterococcus italicus*
ESI	electrospray ionization
GFP	green fluorescent protein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEPN	higher eukaryotes and prokaryotes nucleotide-binding
HPLC-MS	high-performance liquid chromatograph –mass spectrometry
HTH	helix-turn-helix
IPTG	isopropyl β-D-thiogalactopyranoside
LB	Luria Broth
Ll	*Lactococcus lactis*
Ln	*Limisphaera ngatamarikiensis*
MALS	multi-angle light scattering
NAD^+	nicotinamide adenine dinucleotide
NADase	nicotinamide adenine dinucleotide nucleosidase
NCBI	National Center for Biotechnology Information
Ni-NTA	nickel- nitriloacetic acid
NucC	nuclease, CD-NTase associated
OD_{600}	optical density at a wavelength of 600 nm
PAG	polyacrylamide gel
PAGE	polyacrylamide gel electrophoresis
PCaspase	prokaryotic caspase
Pct	particles
PDB	Protein Data Bank
PMSF	phenylmethylsulfonyl fluoride
pp_i	pyrophosphate
$pppA_n$	linear adenylate with triphosphate at the 5′-end

Pycsar	pyrimidine cyclase system for antiphage resistance
RNAP	RNA Polymerase
RNase	ribonuclease
SAVED	SMODS-associated and fused to various effector domains
SB	storage buffer
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SMODS	second messenger oligonucleotide or dinucleotide synthetase
Ss	Sulfurihydrogenibium spp. YO3AOP1
ssDNA	single stranded DNA
ssRNA	single stranded RNA
St	*Streptococcus thermophilus*
STAS	sulphate transporter and anti-sigma factor antagonist
Tad	Thoeris anti-defence
TBE	Tris/Borate/EDTA
Tc^R	tetracycline resistance gene
TEV	Tobacco Etch Virus
TIR	Toll/interleukin-1 receptor
TM	transmembrane
To	*Thermococcus onnurineus*
TPR	tetratricopeptide repeat
Tris	trisaminomethane
Tt	*Thermus thermophilus*
wHTH	winged helix turn helix
wt	wild type

INTRODUCTION

The ongoing arms race between bacteria and mobile genetic elements, such as bacteriophages (phages), has driven bacteria to evolve complex defense systems. These systems detect phage infection and coordinate effector responses to prevent viral replication. CRISPR-Cas systems are adaptive immune mechanisms that protect prokaryotes against invasive mobile genetic elements (Nussenzweig and Marraffini, 2020). While all CRISPR-Cas types use a crRNA-guided protein or protein complex to recognize and degrade foreign nucleic acids complementary to the guide RNA, type III CRISPR-Cas systems add an extra layer of defense: a cyclic oligoadenylate (cA_n, n=3–6)-based signaling pathway that activates auxiliary effector proteins (Stella and Marraffini, 2024).

When type III CRISPR-Cas interference complex, Csm in subtypes III-A and III-D or Cmr in subtype III-B, detects complementary foreign RNA, it activates the Palm domain of the Cas10 subunit, triggering the synthesis of cA_n from ATP (Kazlauskiene et al., 2017; Niewoehner et al., 2017). These cA_n molecules then bind to accessory proteins, activating their effector domains. In type III-A/B/D systems, the most common architecture of accessory proteins includes a CRISPR-Cas associated Rossmann fold (CARF) domain fused to various effector domains (Mascher, 2023; Österberg et al., 2011). CARF-containing effectors have been extensively characterized in recent years (Steens et al., 2022; Stella and Marraffini, 2024). However, some type III CRISPR-Cas systems feature proteins with a sensory SAVED domain (SMODS-associated and fused to various effector domains) instead of CARF (Makarova et al., 2020a; Steens et al., 2024). SAVED domains are closely linked to another prokaryotic defense system known as the cyclic oligonucleotide-based antiphage signaling system (CBASS) (Slavik and Kranzusch, 2023). Like type III CRISPR-Cas systems, CBASS relies on cyclic nucleotide signaling and effector protein activation, employing an abortive infection mechanism. However, experimental data on the role of the SAVED domain in CRISPR-Cas immunity remain scarce.

cA_n production stops once the target RNA is degraded by the Csm/Cmr complex (Kazlauskiene et al., 2017; Rouillon et al., 2018). However, the produced cA_n molecules can continue activating auxiliary effectors, potentially leading to cellular toxicity. To prevent this, additional regulation of the signaling pathway is crucial to avoid excessive activation of effectors and subsequent cellular damage. Recent studies have shown that type III CRISPR-Cas systems achieve this regulation by using specialized signaling molecule-hydrolyzing enzymes known as CRISPR ring nucleases (Crn)

(Athukoralage et al., 2018). Crn proteins contain a CARF fold, suggesting that CARF-effectors may also possess intrinsic ring nuclease activity similar to Crns.

Objects of the Thesis

The objects of this thesis are the type III-A CRISPR-Cas auxiliary effectors Csm6 and Csm6′ (StCsm6 and StCsm6′) from *Streptococcus thermophilus* DGCC8004, and the tripartite effector CalpL-CalpT-CalpS (CCaCalpL-CalpT-CalpS) from *Candidatus* Cloacimonas acidaminovorans str. Evry. The CalpL protein in this system is a Lon-SAVED fusion effector, the CalpT protein is an anti-σ factor and the CalpS protein is a σ factor.

Goals of the Thesis

The goals of the thesis were:

(i) to elucidate the mechanism of cA_6 signaling deactivation by StCsm6 and StCsm6′ in the *S. thermophilus* system,

(ii) to determine the detailed molecular mechanism of the CCaCalpL-CalpT-CalpS tripartite effector in the type III-A CRISPR-Cas defense.

To achieve these goals, the following **objectives** were established:

1. To evaluate the synthesis of cA_6 and other cA_n by type III-A CRISPR-Cas system in bacterial cells, with and without StCsm6 and StCsm6′.
2. To characterize the enzymatic activities of the CARF and HEPN domains of StCsm6 in degrading cA_6 and other cA_n *in vitro*.
3. To determine the molecular requirement for the activation of CCaCalpL effector and its interaction with the substrate CCaCalpT-CalpS.
4. To elucidate the mechanism of the release of the σ factor CCaCalpS from the CCaCalpT-CalpS complex.
5. To determine the deactivation mechanism of the CCaCalpL effector, by testing its ability to degrade cA_4.
6. To provide structural rationale for the activation and regulation of Lon-SAVED architecture effector CCaCalpL.

Scientific novelty and practical value

In 2017, type III CRISPR-Cas systems were discovered to produce cyclic oligoadenylate (cA_n) signaling molecules, which activate diverse auxiliary CARF- and SAVED-effector proteins to enhance CRISPR-Cas defense. While it was established that cA_n synthesis stops upon target RNA cleavage (Kazlauskiene et al., 2017), the regulation of cA_n-activated effectors remained unresolved.

CARF-domain effectors are now well characterized and include RNA nucleases (Csm6 and Csx1), DNA nucleases, deaminases, translation inhibitors, and transmembrane proteins (Stella and Marraffini, 2024). In contrast, the molecular mechanisms of SAVED domain containing-effectors remained largely unexplored, with recently gained insights being limited to the SAVED-CHAT effector (Steens et al., 2024).

This study aimed to elucidate the full regulatory mechanism of type III CRISPR-Cas systems by investigating the deactivation of CARF-effector Csm6 and the regulation of SAVED-effector CalpL. The key scientific contributions of this work include:

1. Regulation of cA_6 signaling: By monitoring cA_n production in *E. coli* cells expressing the *S. thermophilus* type III CRISPR-Cas system, we showed that StCsm6 and StCsm6' degrade their activator cA_6 and other cA_n species. Further *in vitro* studies revealed that the CARF domain of StCsm6 acts as a cA_6-specific ring nuclease, while its HEPN domain exhibits non-specific RNase activity, cleaving cA_6 and other cA_n at high substrate concentrations. This work highlights the function of the CARF domain as a built-in off-switch for both the effector and the CRISPR-Cas system - similar to standalone ring nucleases. It also reveals the interplay between the CARF and HEPN domains in the regulation of cA_n signaling.

2. Regulation of SAVED-effector CalpL: Studying the CalpL effector from *Ca.* C. acidaminovorans, we discovered that it degrades its activator, cA_4, through a unique three-step mechanism, generating two linear intermediates (A_4>p – linear tetraadenylate with cyclic 2′-3′ phosphate and A_4p – linear tetraadenylate with 3′ phosphate). This differs from known mechanisms of ring nucleases and is the first demonstration of enzymatic activity of the SAVED domain.

3. Mechanism of the tripartite CCaCalpL-CalpT-CalpS effector: Through biochemical, biophysical, and toxicity assays in *E. coli*, we demonstrated that the cA_4-activated Lon protease and SAVED domain fusion effector CCaCalpL specifically cleaves the CCaCalpT anti-σ factor in the CCaCalpT-CalpS anti-σ/σ factor pair. We further discovered that cA_4 binding induces CalpL filament formation, which is essential for CCaCalpT-CalpS substrate binding. Within the filament, CCaCalpT cleavage exposes a C-degron sequence, targeting it for degradation by cellular proteases and thereby releasing the σ-factor CCaCalpS, which binds RNA polymerase and induces growth arrest in *E. coli*. Structural studies showed that CCaCalpL forms filaments upon binding to cA_4 or its cleavage products (A_4>p and A_4p) explaining why filament formation is crucial for both the protease and ring nuclease activities of CCaCalpL effector.

These findings not only reveal the complexity of type III CRISPR-Cas systems and their auxiliary effectors but also underscore the stringent regulatory mechanisms of prokaryotic defense systems. Additionally, it contributes to the now growing notion that type III CRISPR-Cas systems based on cA_n signaling do not operate through an abortive infection mechanism. This knowledge has important implications for the strategic use of phages in phage therapy to combat antibiotic-resistant bacteria – a growing global threat to human health and the sustainability of food and agricultural systems.

The major findings presented for defense in this thesis:

1. StCsm6 and StCsm6' degrade cA_6 and other cyclic oligoadenylates (cA_n) produced by the StCsm complex in *E. coli*.
2. The CARF domain of StCsm6 acts as a cA_6-specific ring nuclease, while the HEPN domain exhibits non-specific RNase activity, degrading cA_6 and other cA_n.
3. cA_4 binding induces the formation of CCaCalpL filament, enabling CCaCalpT-CalpS binding and CCaCalpT cleavage by a neighboring CCaCalpL subunit.
4. CCaCalpL-mediated CCaCalpT cleavage exposes a C-degron, targeting the cleavage product for degradation by cellular proteases. Degradation of CCaCalpT leads to the release of CCaCalpS, which subsequently binds to bacterial RNA polymerase.
5. The SAVED domain of CCaCalpL functions as a cA_4-specific ring nuclease, degrading its activator through a sequential three-step mechanism, generating two intermediates, A_4>p and A_4p.
6. The formation of the CCaCalpL filament, triggered by cA_4 or its cleavage intermediates, is crucial for the protease and ring nuclease activities of the Lon-SAVED effector.

1. LITERATURE OVERVIEW

Prokaryotes are in constant contact with mobile genetic elements such as phages, plasmids, transposons, integrative conjugative elements (Gogarten and Townsend, 2005). Some mobile genetic elements are beneficial to their host by providing useful adaptive traits such as antibiotic resistance or virulence factors (Frost et al., 2005; Rankin et al., 2011), while other such as virulent phages are extremely harmful to the host because their replication ends in lysis of the host cell (Buckling and Rainey, 2002). Therefore, to combat destructive phage infections, prokaryotes have developed a range of defense mechanisms. In recent years more than 130 novel defense systems have been identified (Doron et al., 2018; Gao et al., 2020; Millman et al., 2022; Vassallo et al., 2022).

In general, antiphage defense systems sense phage infection and orchestrate an effector response to prevent viral proliferation. However, the modes of action and complexity of the system vary widely. For example, some defense systems, such as restriction-modification or most types of CRISPR-Cas systems, directly recognize and cleave invaders DNA, while others, such as the antiviral STAND (Avs), recognize the highly conserved phage proteins and oligomerize to activate diverse effector domains that serve as abortive infection modules that kill infected cells to prevent the spread of phages through a population (Gao et al., 2022; Mayo-Muñoz et al., 2023). Thus, as mentioned above, the phage detection and effector function can be physically coupled and performed by a single protein or inhibitor complex, but physically decoupled systems with more complicated regulatory networks that transmit the signal between sensor and effector parts of the system are also quite prominent. Defense systems, such as CBASS, type III CRISPR-Cas, Pycsar and Thoeris, utilize nucleotide-centric signal transduction pathways to activate effector proteins (Hobbs and Kranzusch, 2024). In other cases, such as bacterial gasdermin (bGSDM) and type III-E CRISPR-Cas proteolytic cleavage is used for signal transduction (Johnson et al., 2022; Yang and Patel, 2022).

In this literature overview I will focus on signal transduction pathways containing defense systems with a particular emphasis on nucleotide-based immune signaling of type III CRISPR-Cas. Furthermore, I will take a look into signal transduction by proteolytic cleavage in antiviral defense and in control of global cellular processes such as transcription.

1.1. Cyclic oligoadenylate signaling in type III CRISPR-Cas systems

CRISPR-Cas (clustered regularly interspaced short palindromic repeats; CRISPR associated) systems represent a vast highly diverse family of antiphage defense systems that utilize RNA guides to target invading nucleic acids (Nussenzweig and Marraffini, 2020). Currently CRISPR-Cas systems are categorized into 2 classes, 7 types (I-VI and recently proposed candidate type VII) and more than 30 subtypes (Altae-Tran et al., 2023; Makarova et al., 2020b; Yang et al., 2024). The CRISPR-Cas system consists of the CRISPR array and genes encoding Cas proteins. The CRISPR array itself consists of repeated DNA sequences – repeats – with unique sequences – spacers – interspersed between them. The repeat sequences are conserved within the species, while the spacer sequences are of extrachromosomal origin and represent previous encounters with phages or other mobile genetic elements (Mojica et al., 2005).

CRISPR-Cas systems work in three stages: adaptation, biogenesis, and interference (Figure 1.1). Adaptation happens during the first encounter with a phage, then short fragments of invaders nucleic acid are processed and incorporated into the CRISPR array by Cas1-Cas2 (and, in some cases, auxiliary Cas proteins) adaptation module (Lee and Sashital, 2022). In the second stage CRISPR array is transcribed and processed to short crRNAs consisting of single spacer flanked by repeat fragments that are incorporated into the Cas interference complexes and serve as programmable guides (Charpentier et al., 2015). Interference occurs during reinfection when crRNA guided Cas interference complex recognizes complementary invading nucleic acids. Depending on the type of the system the target nucleic acid can be either DNA (types I, II, IV, V) or RNA (types III, VI, candidate VII) (Altae-Tran et al., 2023; Makarova et al., 2020b; Nussenzweig and Marraffini, 2020) (Figure 1.1). Moreover, the molecular mechanism of the interference step is unique for each system type (Figure 1.1). While all CRISPR-Cas types use a crRNA-guided protein or protein complex to target and, in most cases, degrade foreign nucleic acids therefore falling under the direct immunity pathway, most subtypes of the type III CRISPR-Cas systems use an additional layer of defense – a cyclic oligoadenylate (cA_n, n=3-6)-based signaling pathway to activate accessory effector proteins (Stella and Marraffini, 2024). Therefore, type III CRISPR-Cas systems can be classified as signal transducing immune systems. Furthermore, type III CRISPR-Cas systems appear to be structurally and mechanistically the most complex and are proposed to be the most likely ancestor of other CRISPR-Cas types (Koonin and Makarova, 2022).

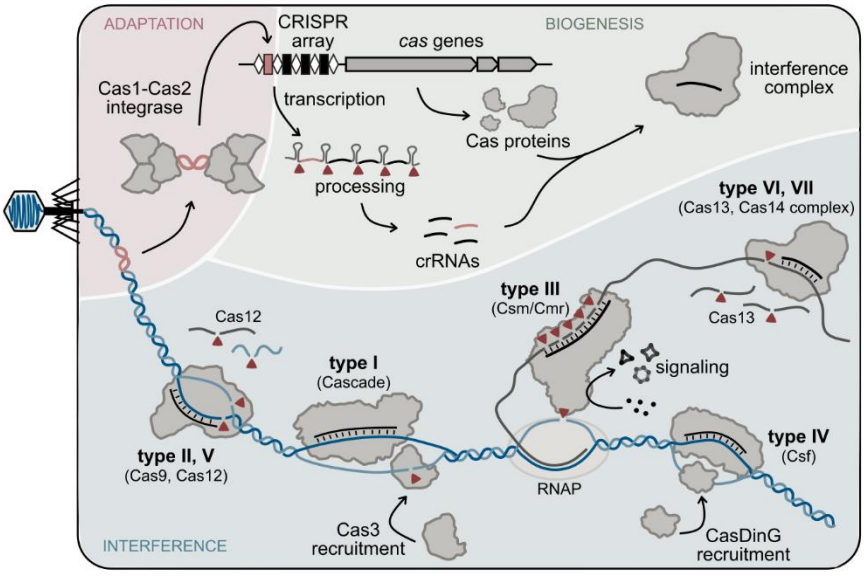


Figure 1.1. The general mechanism of CRISPR-Cas adaptive immunity. CRISPR-Cas systems respond to the invading nucleic acids in three stages: adaptation, biogenesis, and interference. During primary infection Cas1-Cas2 integrase complex captures the fragment of invading DNA and incorporates it into CRISPR array as a spacer. This step is known as adaptation. During biogenesis step Cas proteins are expressed and the CRISPR array is transcribed and processed into crRNAs that contain sequence of a single spacer each. Mature crRNAs are bound by Cas proteins forming crRNA-guided interference complexes. The molecular mechanisms of the interference step are unique for each type. The interference complex of type II (Cas9), type V (Cas12) and type I (Cascade) cleave invading DNA that is complementary to crRNA spacer region. Cas9 cleaves dsDNA using Nuc and RuvC domains, while Cas12 relies on single RuvC domain for cleavage of both DNA strands (Chylinski et al., 2014; Swarts and Jinek, 2019). Most Cas12 additionally possess non-specific single stranded DNA (ssDNA) and single stranded RNA (ssRNA) trans-cleavage activity (Fuchs et al., 2022; Urbaitis et al., 2022; Yan et al., 2019). Type I Cascade interference complexes recruit Cas3 nuclease for target DNA cleavage (Hochstrasser et al., 2014). Type VI (Cas13) and candidate type VII (Cas14-complex) interference complexes target invader's RNA. To specifically cleave bound target RNA Cas14 possess nuclease domains of metallo-β-lactamase family while Cas13 utilizes nuclease activity of HEPN domain (Altae-Tran et al., 2023; van Beljouw et al., 2023; Yang et al., 2024). Activated Cas13 possess nonspecific trans-cleavage activity on ssRNA (van Beljouw et al., 2023). Type III interference complex (Csm/Cmr) recognizes and cleaves invader's transcript and additionally degrades foreign ssDNA that is exposed in the transcription bubble as well as produce signaling molecules for activation of ancillary effectors (Stella and Marraffini, 2024; van Beljouw et al., 2023). Type IV interference complex, Csf, does not cleave the invading DNA, but recruits CasDinG helicase, which presumably interferes with the transcription of the foreign DNA (Čepaitė et al., 2024; Cui et al., 2023). However, CasDing-HNH nuclease domain fusion variants have been also identified (Altae-Tran et al., 2023).

Type III CRISPR-Cas systems belong to class 1 systems and are further subdivided into six (A-F) subtypes (Figure 1.2) (Koonin and Makarova, 2022; Makarova et al., 2020b). Recently, a new candidate subtype III-UAS has been proposed (Altae-Tran et al., 2023). Due to historic reasons, type III-A and III-B systems are the most thoroughly characterized and commonly referred to as typical type-III systems. While new studies on types III-E and III-D emerge (Hu et al., 2022; Kato et al., 2022b; Liu et al., 2022; Schwartz et al., 2024; van Beljouw et al., 2021), type III-F and III-C as well as a newly proposed III-UAS remain biochemically and structurally uncharacterized.

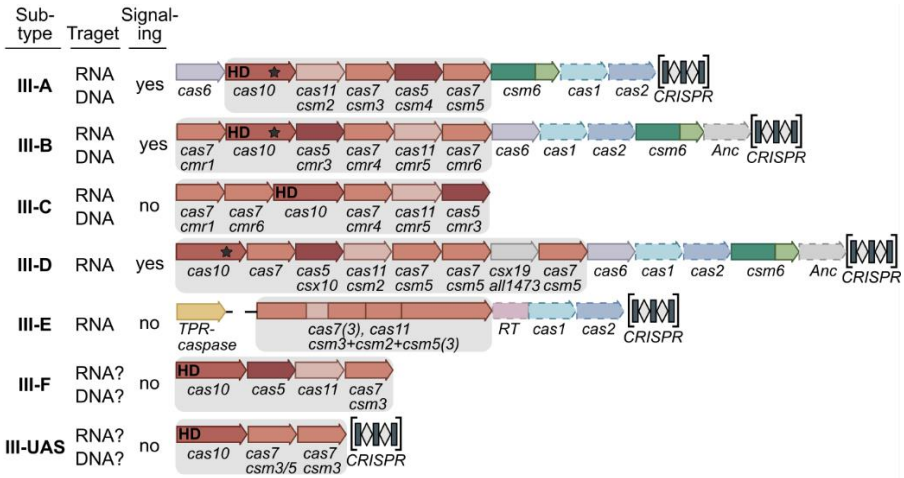


Figure 1.2. Current classification of type III CRISPR-Cas systems. Representative CRISPR–cas loci of each type III subtype is shown. Homologous genes are color-coded and identified by the systematic family name and a commonly used legacy name (the legacy name is given under the systematic name). Dashed lines indicate dispensable modules. Functionally uncharacterized genes are colored grey. The grey shading indicates the subunits forming the interference complex. The presence of the HD nuclease domain and cyclase GGDD motif in the Cas10 subunit is indicated by the "HD" or by the star, respectively. Anc – ancillary effector. Adapted from (Makarova et al., 2020b) with additional modifications based on (Altae-Tran et al., 2023; Koonin and Makarova, 2022)

1.1.1.The interference complex

The interference complexes of type III CRISPR-Cas systems, except for III-E, are composed of crRNA and multiple proteins and contain the signature protein Cas10 (Figure 1.2) (Altae-Tran et al., 2023; Koonin and Makarova, 2022; Makarova et al., 2020b). Cas10 (alternatively called Csm1 or Cmr2) is the large subunit of the complex and together with other proteins of Cas11, Cas7 and Cas5 families comprise the interference complex called Csm (types III-A, III-D and III-F) or Cmr (III-B and III-C) (Molina et al., 2020). The Cas7

family subunits Csm3/Cmr4 have a conserved RNA recognition motif core and possess RNase activity (Li et al., 2024; Molina et al., 2020). The Cas10 subunit, depending on the type and, in some instances, even on a particular case, may have the HD phosphodiesterase domain, which enables the Cas10 to cleave ssDNA, and/or two Palm domains forming a composite active site for production of signaling cA_n molecules (Molina et al., 2020; van Beljouw et al., 2023). The interference complexes of type III-A and III-B systems commonly possess all three activities: RNase, DNase and synthase, while the type III-D interference complex lacking the HD active site was demonstrated to cleave RNA and produce cA_n (Hoikkala et al., 2024; Koonin and Makarova, 2022; Makarova et al., 2020b; Molina et al., 2020). Type III-C, III-F and III-UAS are predicted to possess RNase and DNase activities since these types encode Csm3/Cmr4 homology subunits and Cas10 with HD domain (Altae-Tran et al., 2023; Hoikkala et al., 2024; Koonin and Makarova, 2022; Makarova et al., 2020b; Molina et al., 2020). Even though type III-E systems lack the signature Cas10, and the effector module is composed of a single Cas7-Cas11 fusion protein, these systems are still categorized as type III based on sequences similarity of Cas7 and Cas11 elements and will be reviewed in Chapter 1.5.1 (Makarova et al., 2020b).

The interference complexes Csm and Cmr of typical type III-A and type III-B CRISPR-Cas systems, respectively, have been the subjects for extensive structural studies (Benda et al., 2014; Goswami et al., 2024; Guo et al., 2019; Huo et al., 2018; Jia et al., 2019a, 2019c; Osawa et al., 2015; Sofos et al., 2020; Staals et al., 2014; You et al., 2019). The essential core of the interference complex is the mature crRNA. The crRNA is produced from the transcript of the CRISPR array by Cas6 protein cleavage at repeat sequences (Charpentier et al., 2015; Hatoum-Aslan Asma et al., 2014). The precursors of crRNAs are further truncated, presumably by cellular or CRISPR-Cas-associated nucleases, to the mature crRNAs, which consist of so-called 5′-tag of 5′-handle (8 nt), derived from the repeat, and the unique spacer sequence (Chi and White, 2024; Chou-Zheng and Hatoum-Aslan, 2022; Walker et al., 2017). The 5′-handle of crRNA is anchored by Csm4/Cmr3 and Cas10 subcomplex, while multiple copies of Csm3/Cmr4 bind the rest of crRNA forming the backbone of the interference complex (Guo et al., 2019; Huo et al., 2018; Jia et al., 2019c; Osawa et al., 2015; You et al., 2019). Small subunits Csm2/Cmr5 arrange alongside the Csm3/Cmr4 filament, while Csm5, in case of Cmr complex Cmr6 and Cmr1, caps the filament at the 3′-end of crRNA. The amount of the backbone subunits varies depending on the length of crRNA. For example, the type III-A interference complex of *Streptococcus thermophilus* (St) from different *S. thermophilus* strains, ND03

and DGCC8004, were found to have crRNA of different lengths, 35 nt and 40 nt, respectively, and in turn form the StCsm complexes with slightly different stoichiometries $1_{crRNA(35nt)}$:1_{Cas10}:2_{Csm2}:3_{Csm3}:1_{Csm4}:1_{Csm5} or $1_{crRNA(40nt)}$:1_{Cas10}:3_{Csm2}:5_{Csm3}:1_{Csm4}:1_{Csm5} (Guo et al., 2019; Mogila et al., 2019; Tamulaitis et al., 2014; You et al., 2019). The structure of the StCsm complex from *S. thermophilus* strain ND03 is shown in figure 1.3 A.

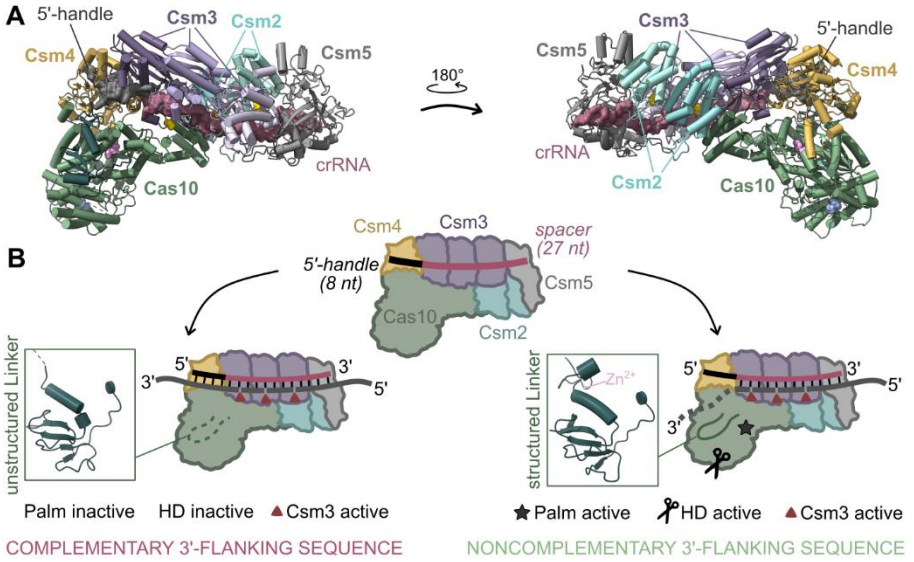


Figure 1.3. StCsm interference complex. (**A**) Structure of the StCsm complex from *S. thermophilus* ND03 strain (PDB ID: 6IFN) (You et al., 2019). Cas10 colored green, The Linker region - dark green, Csm4 – dark yellow, Csm5 – light grey, Csm2 – shades of cyan, Csm3 – shades of purple. Target RNA and crRNA are depicted as surface and spacer part of crRNA is colored burgundy, crRNA 5-handle -dark grey, target RNA - white. HD active site is depicted as blue spheres, GGDD motif – pink spheres, the catalytic D residue in Csm3 active site – yellow spheres. (**B**) Schematic representation of 35 nt crRNA containing StCsm complex and its activation upon binding the target RNA containing either complementary or noncomplementary 3′-flanking sequence. In both cases the closeup view of the Cas10 Linker region is depicted (PDB ID: 6IFL and 6IFY) (You et al., 2019). Subunits and crRNA parts are color-coded.

Since the StCsm complex has been extensively studied and possesses all three enzymatic activities (Guo et al., 2019; Kazlauskiene et al., 2017, 2016; Mogila et al., 2019; Tamulaitis et al., 2014; You et al., 2019), I will use it to illustrate the molecular mechanism of action of the interference complex of a typical type III CRISPR-Cas. Upon phage infection, the StCsm complex binds the phage transcript that is complementary to the spacer portion of the crRNA, with the Csm5 subunit playing a pivotal role in the initial phase of target RNA binding (Mogila et al., 2019). The target RNA binding triggers the structural

changes: the Csm2 subunits rotate away from the Csm3 subunits, Csm1 rotates toward the 5′-end of the crRNA, generating a wider binding channel to harbor the crRNA-target RNA duplex (You et al., 2019). Furthermore, the enzymatic activities of Cas10 are activated: the HD domain starts to degrade ssDNA presumably formed in the transcription bubble and the Palm domains start to synthesize cA_n from ATP (Figure 1.3 B) (Kazlauskiene et al., 2017, 2016). However, the Cas10 activation is additionally dependent on the lack of complementarity between 5′-handle of crRNA and 3′-flanking sequence of the target transcript (Kazlauskiene et al., 2017, 2016). The non-complementary 3′-flanking sequence swings away from the crRNA and is positioned in the cleft of Cas10 and causes structural rearrangement of the Cas10 Linker region, especially in the part containing the zinc finger motif, whereas the complementary 3′-flanking sequence does not interact with Cas10 directly, therefore the Linker remains disordered (Figure 1.3 B) (You et al., 2019). It is thought that this stabilization of the Linker may cause subtle conformational changes in the HD and Palm domains switching on the nuclease and cA_n synthesis activities. The non-complementarity rule between 5′-handle and 3′-flanking sequence serves as a mechanism for self versus non-self discrimination, protecting the cell from accidental activation of Cas10 when CRISPR array is transcribed in the opposite direction (Kazlauskiene et al., 2017, 2016; Molina et al., 2020; You et al., 2019).

The Cas10 activities are switched off in a timely manner when the bound transcript is cleaved by the Csm3 subunits (Kazlauskiene et al., 2017, 2016; Tamulaitis et al., 2014). Since each Csm3 subunit possesses RNase activity, the transcript is cleaved at multiple positions in 6 nt intervals and the Csm2 subunits are important for coordinating the cleavage at multiple sites further from Cas10 (Mogila et al., 2019; Tamulaitis et al., 2014). Interestingly, the target RNA cleavage by the Csm3 subunits occurs within seconds, whereas Cas10 activation lasts up to hours (Irmisch et al., 2024; Kazlauskiene et al., 2017, 2016). Recently, it has been shown that this prolonged stimulation is dependent on the release of the cleaved target ends (Irmisch et al., 2024). Since DNase and synthase activities are thought to be switched on by the same structural linker-loop motif in the Cas10 subunit, it is reasonable to speculate that the retention of the cleaved target ends also stimulates cA_n production (You et al., 2019).

1.1.2. Production of signaling molecules

For the synthesis of signaling cA_n molecules Cas10 utilizes a composite active site formed by a pair of Palm domains (Palm1 and Palm2) (Goswami

et al., 2024; Jia et al., 2019a; Kazlauskiene et al., 2017; Molina et al., 2020). *In vitro*, most Cas10 synthesize a mixture of cA_n species (Foster et al., 2020; Grüschow et al., 2021, 2019; Kazlauskiene et al., 2017; Nasef et al., 2019; Rouillon et al., 2018). Each cA_n species triggers unique downstream effectors, and bioinformatic studies based on the abundance of associated specific effectors revealed that the predominant cA_n in type III CRISPR-Cas signaling is cA_4, with cA_6 and cA_3 being significantly less abundant (Hoikkala et al., 2024). Structural studies of mainly cA_4 producing *Thermococcus onnurineus* Csm (ToCsm) complex and mainly cA_6 producing *Lactococcus lactis* Csm (LlCsm) provided mechanistic insights on cA_n of various lengths production by the Palm domains (Goswami et al., 2024; Jia et al., 2019a).

Each Palm domain has an adenosine binding pocket where the adenine base is specifically recognized (Goswami et al., 2024; Jia et al., 2019a). ATP situated at the cleft formed primarily by Palm1 ("acceptor" pocket) is called acceptor ATP, whereas ATP at the cleft formed primarily by Palm2 ("donor" pocket) is called donor ATP. The conserved GGDD motif protrudes from the Palm2 domain and the aspartate coordinates the divalent metal ion and likely activates the 3′-OH group of the acceptor ATP to initiate an in-line nucleophilic attack on the α-phosphate of the donor ATP, forming a 3′–5′ phosphodiester bond (Goswami et al., 2024). The ATP moiety of the linear di-adenylate with triphosphate at the 5′-end ($pppA_2$) intermediate synthesized in the first step, relocates to the Palm2 domain and occupies the "donor" pocket, while a new ATP binds to the Palm1 domain, and the in-line nucleophilic attack occurs again, resulting in the formation of $pppA_3$ intermediate (Goswami et al., 2024; Jia et al., 2019a). This process is repeated to generate linear $pppA_n$ of different lengths. Thus, unlike all conventional nucleotide polymerases that catalyze nucleotide addition to the 3′ end of a growing chain (Kimura et al., 2016), Cas10 catalyzes non-templated addition of nucleotides to the 5′-end of the growing oligoadenylate chain.

The $pppA_n$ is then cyclized by the intramolecular attack on the α-phosphate of the first ATP positioned in the "donor" pocket of the Palm2 domain by 3′-OH of terminal AMP positioned in the "acceptor" pocket of Palm1 domain, therefore both chain elongation and cyclization are performed by the same active site and mechanism (Goswami et al., 2024; Jia et al., 2019a). Figure 1.4 depicts the schematic representation of cA_n (n=2-4) synthesis by Palm domains.

Compared to cA_4-producing Cas10 variants, the Palm2 domain of cA_6-producing Cas10 contains an additional adenosine ring binding pocket (Goswami et al., 2024). The presence of the third adenosine binding pocket results in a different trajectory of the third adenosine in the $pppA_n$ chain,

allowing the synthesis of longer $pppA_n$ chains. The 3′ proximal Ap units that are not placed into the acceptor pocket or into additional adenosine binding pocket (in case of cA_6-producing Cas10) seem to form very few contacts with the protein and are labile (Goswami et al., 2024; Jia et al., 2019a). The mobility of 3′ proximal Ap units may result in the accidental placement of the 3′ terminal ATP into the Palm1 acceptor pocket, resulting in premature cyclization. This would explain how a mixture of various ring sizes cA_n, observed in many cases, is produced.

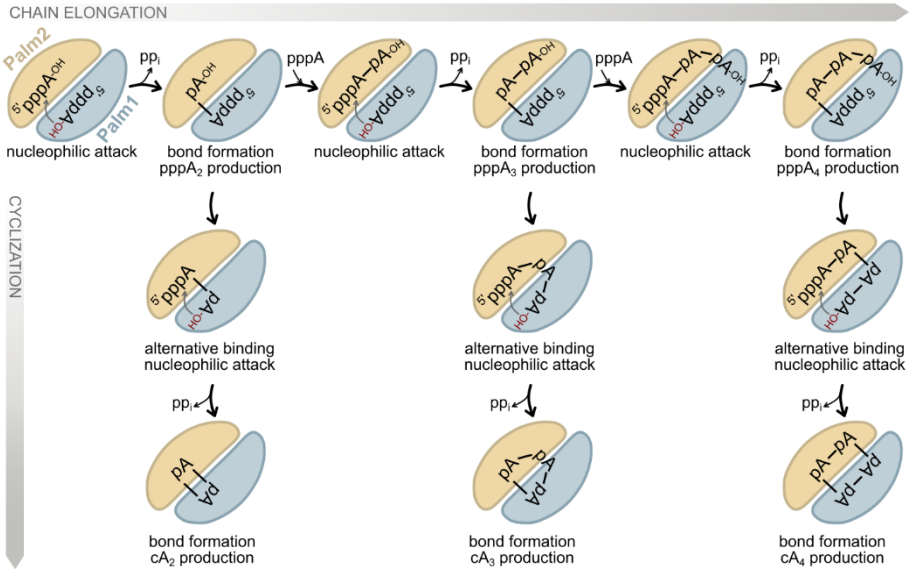


Figure 1.4. Schematic representation of sequential synthesis of cA_n by the Cas10 subunit of the Csm complex. The synthesis of cA_n can be divided into $pppA_n$ chain elongation and cyclization stages. First, one ATP (pppA) molecule is bound by the "donor" pocket in Palm2 domain while the other ATP is positioned in the "acceptor" cleft within Palm1. The 3′-OH of the ATP bound by Palm1 is activated by the aspartate from hallmark GGDD motif and initiates the nucleophilic attack on the α-phosphate of ATP accommodated in the Palm2 domain, forming a $pppA_2$ intermediate. The ATP moiety of the $pppA_2$ intermediate relocates to the "donor" pocket in Palm2 domain and a new ATP molecule occupies Palm1 "acceptor" site. Again, 3′-OH of the ATP bound by Palm1 performs nucleophilic attack on the α-phosphate of ATP moiety of $pppA_2$ intermediate residing in the Palm2 forming the second $pppA_3$ intermediate. By repeating these steps, the $pppA_n$ chain elongates. However, each $pppA_n$ intermediated can occupy an alternative binding conformation and position 3′-pA moiety into Palm1 "acceptor" pocket and 5′-pppA moiety into Palm2 "donor" pocket. The resulting intramolecular nucleophilic attack results in cyclization of $pppA_n$ producing cA_n. pp_i – inorganic phosphate. Based on (Goswami et al., 2024; Jia et al., 2019a).

Recently, a new type of signaling molecule produced by Cas10 has been identified, shifting the paradigm that type III CRISPR-Cas systems utilize

only cA_n molecules joined by 3′-5′ phosphodiester bonds. The Cas10 subunit of type III-B Cmr interference complex from *Bacteroides fragilis* synthesizes the SAM-AMP signaling molecule by conjugating ATP to S-adenosyl methionine (SAM) via a phosphodiester bond (Chi et al., 2023). The SAM-AMP synthesis is also performed by the two Palm domains, but the 'acceptor' pocket in the Palm1 contains two additional negatively charged residues, allowing the accommodation of SAM.

1.1.3. Activation of type III CRISPR-Cas accessory effectors

Type III CRISPR-Cas systems that produce signaling molecules encode auxiliary effector proteins within their CRISPR loci (Figure 1.2). These effector proteins typically consist of two key components: a sensory domain responsible for recognizing cyclic oligoadenylate (cA_n) signals and an effector domain that mediates downstream functional activities. The effector domains can be diverse, including nucleases (e.g., HEPN, PD-D/ExK, PIN domains), adenosine deaminases, proteases, adenylyl cyclase-like CYTH domains, potentially NAD^+-degrading Toll/interleukin-1 receptor protein (TIR) domains, nitrilases, or can contain transmembrane segments (Makarova et al., 2020a). Upon binding of cA_n to the sensory domain, the effector domain is activated to perform its dedicated function. The production of cA_n by the interference complex and the subsequent activation of auxiliary effectors has been demonstrated to play a dominant role in type III CRISPR-Cas immunity, whereas the DNase activity of Cas10 confers phage immunity only at low multiplicity of infection, and the RNase activity of the Csm3/Cmr4 subunits serves more as a regulatory mechanism rather than an interference step (K. A. Johnson et al., 2024).

For cA_n binding, auxiliary effectors most commonly use the CRISPR-associated Rossmann Fold (CARF) domain as a sensory domain. The CARF domain dimer forms a composite cA_n binding pocket, resulting in the dimeric nature of CARF domain-containing proteins (Athukoralage and White, 2021; Makarova et al., 2020a). In certain cases, auxiliary effectors use the SMODS-associated and fused to various effector domains (SAVED) domain instead of the CARF domain. The SAVED domain represents a highly divergent form of two CARF domains fused into a single polypeptide chain (Athukoralage and White, 2022; Lowey et al., 2020; Makarova et al., 2020a; Steens et al., 2022).

Based on phylogenetic and domain architecture analysis, CARF domains can be divided into 9 major (CARF1-9) and 13 minor (CARF m1-13) clades, while SAVED domains fall into 7 clades (SAVED1-7) (Makarova et al.,

2020a). All CARF groups, except for CARF6 and a few minor clades, demonstrate a strong association with type III-A/B/D CRISPR-Cas systems, while only SAVED4, SAVED5 and SAVED6 tend to associate with type III CRISPR-Cas systems. SAVED domain-containing effectors are more common in the other nucleotide signaling system, CBASS, which is described in detail in Chapter 1.2. However, some type III CRISPR-Cas axillary effectors do not have canonical CARF/SAVED sensory domains but instead possess unique folds for signal molecule binding (Chi et al., 2023; Grüschow et al., 2024; Hoikkala et al., 2024; Lau et al., 2020; Mayo-Muñoz et al., 2022). In this chapter, I will focus on the type III CRISPR-Cas ancillary effectors that have been structurally and biochemically characterized and are briefly summarized in Figure 1.5.

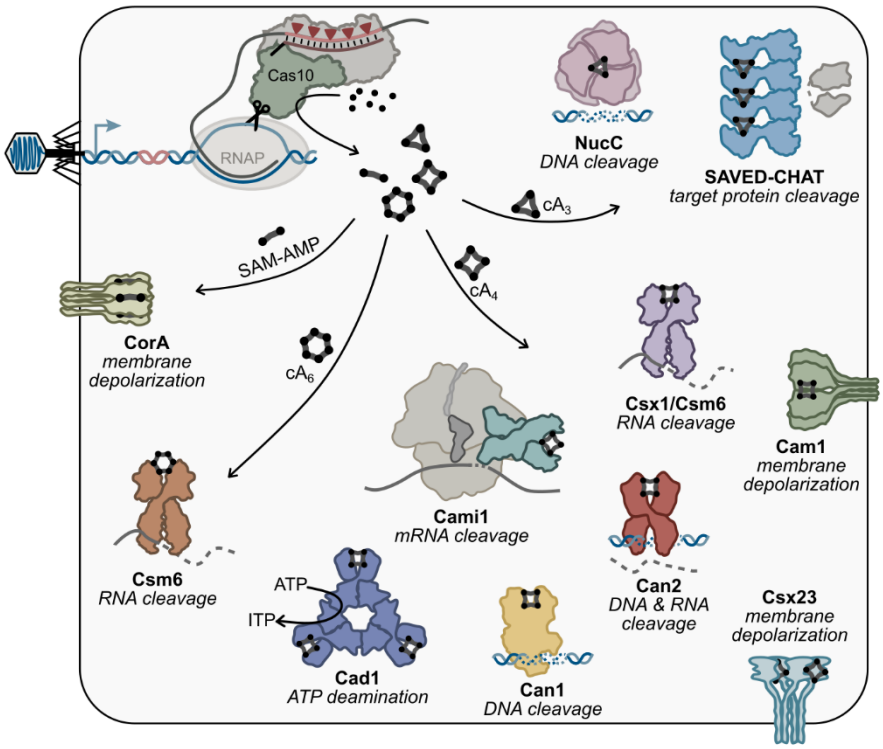


Figure 1.5. Schematic representation of characterized type III CRISPR-Cas ancillary effectors. Upon phage infection the Csm/Cmr interference complex binds target transcript and produces signaling molecules. Most characterized ancillary effectors recognize cA_4 (Csx1/Csm6, Can1, Can2, Cami1, Cam1, Csx23 and Cad1). Distinct family of Csm6 binds cA_6. NucC and SAVED-CHAT recognize cA_3 activator, while CorA binds a unique SAM-AMP second messenger. Activated effectors can degrade nucleic acids (Csm6/Csx1, Can1, Can2, Cami1, NucC), disrupt membrane integrity (CorA, Cam1, Csx23), modify nucleotides (Cad1) or target specific proteins (SAVED-CHAT).

Nucleolytic effectors

The most common architecture of the type III CRISPR-Cas-associated effector proteins is an N-terminal sensory CARF domain fused to a C-terminal higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domain, often they also contain the middle α-helical ('6H' or HTH) domain (Makarova et al., 2020a). Historically, these effectors are called Csm6 or Csx1 depending on the associated subtype of type III CRISPR-Cas system. Type III-A associated CARF-HEPN have been named Csm6 and type III-B and III-D associated were denoted Csx1 (Molina et al., 2020). The HEPN domain possesses non-specific RNase activity, which is switched on upon cA_n binding to the sensory CARF domain (Kazlauskiene et al., 2017; Niewoehner et al., 2017).

Since two CARF domains are required to form a composite cA_n binding pocket, Csm6/Csx1 proteins are dimers in solution, with an intriguing case of *Sulfolobus islandicus* Csx1, which forms a larger oligomer, a trimer of homodimers (Makarova et al., 2020a; Molina et al., 2019; Steens et al., 2022). Csm6/Csx1 proteins form a phylogenetically and structurally diverse group that is scattered among different CARF clades (CARF1, CARF2, CARF9) proposed by (Makarova et al., 2020a). Csx1 proteins recognize cA_4 as the activator (Athukoralage et al., 2020a; Foster et al., 2020; Grüschow et al., 2019; Han et al., 2018; Molina et al., 2019), while some Csm6 bind cA_4 and others cA_6 (Du et al., 2024; Grüschow et al., 2019; Jia et al., 2019b; Kazlauskiene et al., 2017; Niewoehner et al., 2017). Furthermore, the extent of conformational changes required to activate the RNase activity of the HEPN domain upon cA_n binding also varies among Csm6/Csx1 proteins. For example, cA_4 binding to the CARF domain of *Thermus thermophilus* Csm6 induces significant conformational changes, including the reorganization of the HEPN catalytic motif R-X_{4-6}-H, positioning it into a catalytically competent conformation (Du et al., 2024). In contrast, upon cA_4 binding *Thermococcus onnurineus* Csm6 exhibits minimal local conformational changes (Jia et al., 2019b). cA_6-binding Csm6 proteins confer a low level of non-specific RNase activity even in the absence of an activating molecule, therefore they are thought to exist in a dynamic conformational equilibrium between active and inactive forms, and upon cA_6 binding, Csm6 is stabilized in the activated conformation, unleashing efficient non-specific RNase activity of the HEPN domain (Foster et al., 2019; Kazlauskiene et al., 2017; Niewoehner et al., 2017; Steens et al., 2022). Activated HEPN domains of Csm6 cleaves ssRNA after purines (Foster et al., 2019; Kazlauskiene et al., 2017), *Pyrococcus furiosus* Csx1 prefers adenosines (Sheppard et al., 2016), while *Sulfolobus islandicus* Csx1 preferential cleavage site is in between two

cytosine (Molina et al., 2019). *In vivo*, the non-selective RNase activity of cA_n-activated Csm6/Csx1 proteins is thought to cause growth arrest in infected cells (Rostøl and Marraffini, 2019; Steens et al., 2022), however, a recent study contradicts this view, suggesting that Csm6 predominantly cleaves phage transcripts adjacent to the crRNA target site due to increased local concentrations of diffusible cA_n (K. A. Johnson et al., 2024).

A few other CARF-effectors capable of degrading nucleic acids upon cA_4 binding have also been characterized. These include CRISPR ancillary nuclease 1 (Can1) and Can2 (sometimes referred to as Card1 - cA_n-activated ssRNase and ssDNase 1) and CRISPR-Cas-associated mRNA interferase 1 (Cami1) (McMahon et al., 2020; Mogila et al., 2023; Rostøl et al., 2021; Zhu et al., 2021). Both Can1 and Can2 effectors consist of a CARF domain fused to a PD-D/ExK nuclease domain, but while Can2 proteins are dimers, Can1 contains two CARF domains in a single polypeptide chain and is therefore a monomer (McMahon et al., 2020; Rostøl et al., 2021; Zhu et al., 2021). Upon cA_4 binding, *Thermus thermophilus* Can1 triggers a conformational change that restores the composite DNA cleavage site and in turn Can1 starts nicking supercoiled dsDNA at random sites, presumably slowing viral replication (McMahon et al., 2020). Can2 effectors target both nucleic acids, DNA and RNA. Can2 from *Treponema succinifaciens* degrades ssRNA and ssDNA and thus causes growth arrest, presumably by introducing DNA lesions in both the host and phage genomes (Rostøl et al., 2021), while Can2 from *Sulfobacillus thermosulfidooxidans* and *Thioalkalivibrio sulfidiphilus* have been shown to degrade ssRNA and to progressively nick supercoiled dsDNA, providing the immunity against phages but not causing toxicity to the host cell (Zhu et al., 2021).

The Cami1 effector, composed of the CARF domain fused to the RelE toxin-like module via a winged helix-turn-helix (wHTH), specifically targets mRNA molecules that are being translated (Mogila et al., 2023). Upon cA_4 binding to the CARF domain, the middle wHTH linker domain undergoes a rotational motion that allows Cami1 to be captured by the ribosomal stalk protein, positioning the RelE toxin-like domain into the ribosomal A site for mRNA cleavage. The activation of Cami1 in *E. coli* cells interferes with translation and causes cell dormancy.

Interestingly, the non-CARF nuclease effector NucC (nuclease, CD-NTase associated), has also been found to be associated with type III CRISPR-Cas systems (Lau et al., 2020). Unlike the Csm6/Csx1, Can1-2, and Cami1 nucleases, which use two CARF domains to form a composite cAn binding site to recognize even-numbered cA_n (cA_4 and cA_4), NucC lacks a well-defined sensory domain and instead forms a trimer with a unique threefold

symmetric pocket to recognize the odd-numbered cA_3 activator (Lau et al., 2020). Activator binding induces the oligomerization of two trimers into an active hexamer, triggering significant conformational changes and activating potent DNase activity of the PD-D/ExK endonuclease-like domains. The degradation of dsDNA leads to cell death, preventing phage propagation within a bacterial culture (Lau et al., 2020; Mayo-Muñoz et al., 2022). NucC homologs are more commonly found associated with CBASS, which utilize a wide variety of nucleotide signals (see Chapter 1.2), suggesting the possibility that the type III CRISPR-Cas systems developed the ability to synthesize cA_3 and co-opted NucC (Stella and Marraffini, 2024).

Nucleotide modifying effector

Even though the most extensively studied group of type III CRISPR-Cas auxiliary effectors are CARF-nucleases, CARF domain can be fused to various other enzymatic effector domains (Makarova et al., 2020a). Recently the first non-nuclease enzymatic CARF effector, a CRISPR-Cas-associated adenosine deaminase (Cad1 or CAAD), has been characterized (Baca et al., 2024a; Li et al., 2025). Cad1 forms a trimer of homodimers, in a way resembling *Sulfolobus islandicus* Csx1 (Baca et al., 2024a; Li et al., 2025; Molina et al., 2019). Interestingly, Cad1 binds and can be activated by two species of cA_n. Both cA_4 and cA_6 activate the adenosine deaminase activity of the protein, initiating deamination of ATP, which leads to the growth arrest in cells by depleting the nucleotide pool (Baca et al., 2024a; Li et al., 2025). Surprisingly, the activation mechanism varies between two characterized homologues. In apo-form *Limisphaera ngatamarikiensis* Cad1 exists in dynamic conformational equilibrium of the inactive and active states, with inactive state being predominant. Upon cA_4/cA_6 binding Cad1 undergoes dramatic conformational changes and is locked in the active state (Li et al., 2025). However, the Cad1 homologue from *Bacteriodales* bacterium undergoes only minimal conformational changes upon activator binding (Baca et al., 2024a).

Membrane disrupting effectors

CARF domains are found fused to non-enzymatic effector domains such as transmembrane (TM) helixes (Hoikkala et al., 2024; Makarova et al., 2020a). The protein of such architecture (N-terminal TM domain fused with C-terminal CARF via a short linker) has been recently characterized and termed cA_n-activated membrane protein 1 (Cam1) (Baca et al., 2024b). It has been predicted that upon cA_4 binding a tetrameric complex comprising two separate CARF dimers and four N-terminal transmembrane α-helices organizes into a pore causing membrane depolarization. Similarly, other type

III CRISPR-Cas associated effectors Csx23 and CorA that contain TM helices have been shown to cause the disruption of membrane integrity (Chi et al., 2023; Grüschow et al., 2024). However, instead of the sensory CARF domain, Csx23 uses a unique cytoplasmic C-terminal domain for cA_4 recognition, while CorA utilizes soluble N-terminal domain to bind SAM-AMP signaling molecule.

SAVED-effectors

As mentioned above, instead of the sensory CARF domain, some type III CRISPR-Cas auxiliary effectors may contain the SAVED domain as a cA_n sensor (Makarova et al., 2020a). Although at least three SAVED clades (SAVED4, SAVED5 and SAVED6) demonstrate strong association with type III CRISPR-Cas systems, only a rare SAVED-CHAT (Caspase HetF Associated with Tprs) protease fusion effector belonging to the SAVED2 group has recently been characterized (Hoikkala et al., 2024; Makarova et al., 2020a; Steens et al., 2024). Since the CARF-like subdomains of the SAVED domain are not identical, SAVED-CHAT is able to form a deep binding pocket for the twofold symmetry-deficient cA_3 activator (Steens et al., 2024). Furthermore, upon activator binding SAVED-CHAT oligomerizes in head-to-tail fashion, forming filaments and burying cA_3 within the intrafilament interface between two SAVED domains. This oligomerization allows the CHAT-CHAT interdomain interactions that allosterically activate protease activity. This activation through oligomerization is common for SAVED domain containing proteins that are prominent in CBASS systems (see Chapter 1.2.1). Activated SAVED-CHAT specifically cleaves and activates another protease PCaspase (prokaryotic caspase), which in turn cleaves a multitude of proteins (more on caspase-like proteases see Chapter 1.5) (Steens et al., 2024). However, within the operon between the genes of *SAVED-CHAT* and *pCaspase*, additional protein genes, which could be potential downstream effectors or regulators, are encoded. Therefore, this SAVED-CHAT effector seems to function as a part of a multi-component effector system, however the full mechanism remains to be elucidated. Intriguingly, the existence of various type III CRISPR-Cas-associated multi-component effectors that can function as CRISPR-activated signaling cascades has also been suggested based on bioinformatic works (Altae-Tran et al., 2023).

As is evident from the examples discussed, the most extensively studied group of type III CRISPR-Cas auxiliary effectors are nucleases, with emerging studies on non-enzymatic and novel enzymatic effectors and potential multicomponent signaling cascades. Although auxiliary effectors are diverse, a few alternative activation mechanisms by cA_n binding can be

generalized: (i) cA_n binding can induce dramatic global conformational changes that alter the conformation of the effector domain, (ii) in some cases only minimal conformational changes are sufficient to activate the effector domain, in addition, (iii) activation by shifting the dynamic equilibrium between inactive and active states toward the active state is employed by some effectors, as well as (iv) activation by oligomerization. In general, this highlights the universality of cA_n signaling to activate highly diverse effectors and trigger diverse physiological responses.

Intriguingly, it is relatively common for multiple auxiliary effectors to co-occur in a single type III CRISPR-Cas locus (Hoikkala et al., 2024). It is hypothesized that co-occurrence of two or more effectors, each activated by the same cA_n species, would provide broader defense by simultaneously targeting different biomolecules. Even though most Cas10 *in vitro* produces a variety of cA_n species (Foster et al., 2020; Grüschow et al., 2019; Kazlauskiene et al., 2017; Nasef et al., 2019; Rouillon et al., 2018), co-occurrence of effectors recognizing different cA_n signals is very rare, with only one known example where a cA_3-binding NucC enzyme is encoded in the same type III CRISPR-Cas locus as the cA_4-recognizing Csx23 effector (Grüschow et al., 2024; Hoikkala et al., 2024).

Furthermore, bacteria can encode multiple CRISPR-Cas systems of the same or different types (Kelleher et al., 2024; Pinilla-Redondo et al., 2022; Xia et al., 2022). It is thought that multiple type III CRISP-Cas systems encoded in the same genome may share auxiliary effectors, since systems with the intact Palm domains but lacking known effectors or credible effector candidates tend to be found in genomes encoding multiple type III CRISPR-Cas systems (Hoikkala et al., 2024). However, this inter-locus signaling is not limited to the systems of the same type. Two thirds of *Sulfolobales* genomes harbor a type III system encoding cA_n-producing machinery and a type I system encoding cA_n-activatable CRISPR Apern 3 (Csa3) family transcription factors Csa3a and Csa3b (Xia et al., 2022). Csa3 are dimeric proteins with the CARF-wHTH architecture and are involved in the regulation of the spacer acquisition to the type I CRISPR array, as well as the regulation of transcription of the type III CRISPR-Cas genes (Charbonneau et al., 2021; Xia et al., 2022; Qing Ye et al., 2020). The regulatory functions of the Csa3 effectors are complex, but it highlights the crosstalk between type III and type I CRISPR-Cas systems. Furthermore, the crosstalk between type III CRISPR-Cas and CBASS seems theoretically possible, since the odd-numbered activator-recognizing effectors such as NucC, SAVED-CHAT are found associated with both CBASS and type III-CRISPR Cas systems (Hoikkala et

al., 2024; Lau et al., 2020; Steens et al., 2024), although a specific example of this crosstalk remains to be documented.

1.2. CBASS

Cyclic oligonucleotide-based anti-phage signaling system (CBASS) is one more example of antiviral defense systems that rely on nucleotide-based secondary messengers. The core of the CBASS operons is composed of two components: (i) cGAS/DncV-like nucleotidyltransferase (CD-NTase), which senses phage infection and produces the signaling molecules, and (ii) the CD-NTase associated proteins (Cap) effector, which recognizes the signal and promotes the death of the infected cell (Lowey et al., 2020; Whiteley et al., 2019). Based on the organization of the CBASS operons, CBASS are classified into four types (Millman et al., 2020) (Figure 1.6 A). The type I CBASS are compact two-gene systems that encode only the core elements, CD-NTase and effector protein. Other CBASS types encode additional accessory genes of the proposed regulatory proteins. The type II CBASS encode ancillary genes with ubiquitin associated domains, the type II CBASS harbor ancillary genes with HORMA (Hop1p, Rev7p and Mad2) and TRIP13 (thyroid hormone receptor 13) domains, while the ancillary proteins of type IV have nucleotide-modifying domains.

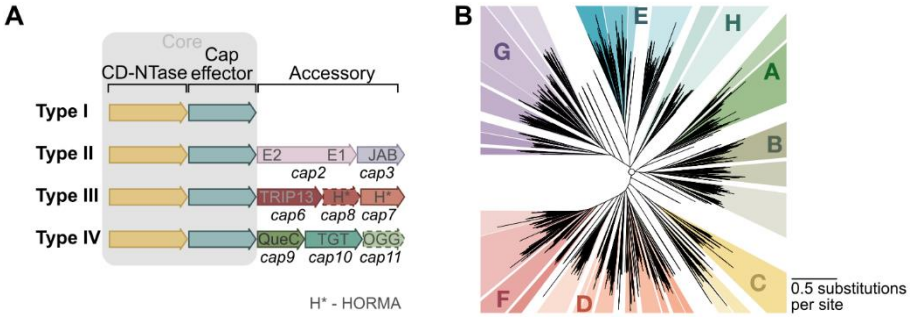


Figure 1.6. Classification of CBASS. (**A**) All CBASS encode core elements, CD-NTase and an effector protein. They are classified into four types. Type I CBASS contains the core elements only, while other types encode accessory proteins. (**B**) Clustering of CD-NTases into eight clades (A-H) based on sequence alignments. Similar colored and lettered clades group CD-NTases that are 10% identical. Similar color shades represent groups of CD-NTases that are 25% identical (Whiteley et al., 2019).

Intriguingly, CBASS systems are ancestors of the animal cyclic GMP-AMP synthase–stimulator of interferon genes (cGAS-STING) signaling pathway, which controls the cellular response to cytosolic dsDNA - a signal of damaged of infected cell (Hobbs and Kranzusch, 2024; Slavik and

Kranzusch, 2023). The cGAS enzyme recognizes and binds cytosolic dsDNA, and this binding activates the enzyme to synthesize the nucleotide second messenger 2′-5′, 3′-5′ cyclic-GMP-AMP (2′3′-cGAMP), which is recognized by the adaptor protein STING, in turn recruiting the kinase TBK1 (TANK-binding kinase 1) and the transcription factors IRF3 (interferon regulatory factor 3) and NF-kB (nuclear factor k-light-chain-enhancer of activated B cells) to activate type I interferon expression, thereby initiating an immune response (Ablasser and Chen, 2019). It is a clear parallel with bacterial CBASS, where bacterial cGAS homologues CD-NTase produces cyclic di- or trinucleotide secondary messengers to activate Cap effectors (Hobbs and Kranzusch, 2024; Slavik and Kranzusch, 2023).

At the sequence level bacterial CD-NTases are quite diverse and can be clustered into eight clades (A-H) (Figure 1.6 B) (Whiteley et al., 2019). To conveniently denote the clade to which CD-NTase belongs, the shorter abbreviation Cdn with clade identification A-H is commonly used, for example, CD-NTase of the clade E is known as CdnE.

Bacterial CD-NTases can utilize a broad variety of nucleotides to produce signaling molecules. They were found to produce 3'-5' linked (each 3′-5′ linkage is denoted as 3′): cyclic di-UMP (3′3′c-di-UMP), cyclic CMP-UMP (3′3′c-CMP-UMP), cyclic UMP-AMP (3′3′c-UMP-AMP), cyclic UMP-GMP (3′3′cUMP-GMP), cyclic di-AMP (3′3′c-di-AMP), cyclic GMP-AMP (3′3′cGAMP), cyclic di-GMP (3′3′c-di-GMP), cyclic AMP-AMP-GMP (3′3′3′cAAG), cyclic AMP-AMP-AMP (3′3′3′cA_3 or cA_3) (Cohen et al., 2019; Morehouse et al., 2022; Whiteley et al., 2019; Qiaozhen Ye et al., 2020). Furthermore, some CD-NTases join nucleotides by 2′-5′ linkage. For example, the CdnD from *Acinetobacter baumannii* produces 2′-5′, 3′-5′, 3′-5′ cyclic AMP-AMP-AMP (2′3′3′-cAAA) and CdnG from *Asticcacaulis sp*. produces 3′-5′, 2′-5′ cGAMP (3′2′cGAMP) (Fatma et al., 2021; Lowey et al., 2020). This is a striking difference from type III CRISPR-Cas systems where 3′-5′ linked cyclic oligoadenylates are predominant signaling molecules (see Chapter 1.1.2).

While in type III CRISPR-Cas systems it is well defined that phage infection is sensed by the binding of phage transcripts that are complementary to crRNA (see Chapter 1.1), the activation of CD-NTase in CBASS remains poorly understood. Based on the available data the activation mechanisms of CD-NTases diverge into: (i) activation by direct binding of a specific phage cue, or (ii) monitoring levels of key cellular metabolites to indirectly sense phage infection (Figure 1.7).

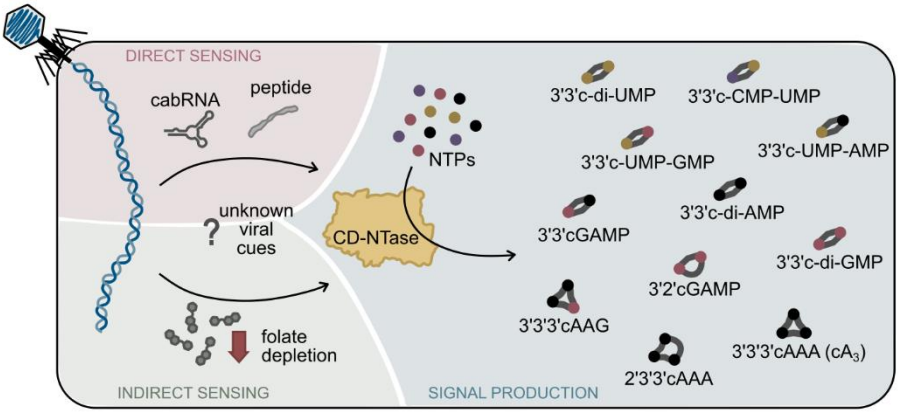


Figure 1.7. Activation of CD-NTases and the repertoire of signaling molecules they produce. CD-NTases can sense viral cues, such as structured cabRNA or viral peptides, by directly binding them or they can sense the infection indirectly by monitoring the cellar levels of metabolites, such as folate, in the cytoplasm. However, activators for the majority of CBASS remain unknown. CD-NTases of different clades can produce various signaling molecules with nucleotides linked by 3′-5′ (3′) or 2′-5′ (2′) phosphodiester bonds.

Examples of activation by direct binding are type I CBASS from *Staphylococcus schleiferi* and type III CBASS from *Escherichia coli* and *Pseudomonas aeruginosa*. The CdnE of *S. schleiferi* utilize a positively charged surface to bind a highly structured RNA, termed CBASS-activating bacteriophage RNA (cabRNA), produced by staphylococcal phages Φ80α-vir and ΦNM1γ6 (Banh et al., 2023). This binding activates CdnE to synthesize the 3′2′cGAMP secondary messenger. *E. coli* CdnC and *P. aeruginosa* CdnD of type III CBASS are activated by the binding of the foreign peptide motif (Qiaozhen Ye et al., 2020). However, the binding is not performed by these CD-NTases directly, accessory HORMA domain-containing Cap7 protein recognizes and binds the foreign peptide. The activating peptide-Cap7 complex then binds to the CD-NTase and activates it to produce cA_3 signaling molecules. The unintended activation of CD-NTase by accidental binding of Cap7 is regulated by Cap6 (a Trip13-like ATPase) protein, which disassembles unintended Cap7-CD-NTase complexes.

Another activation strategy is employed by *Vibrio cholerae* DncV, which is thought to monitor the levels of a cellular metabolite, folate (Brenzinger et al., 2024; Severin et al., 2023; Zhu et al., 2014). It is hypothesized that in uninfected cells, folate-like molecules bind DncV and lock it in an inactive state, whereas during phage infection, rapid viral genome replication depletes folate, leading to dissociation of the DncV-folate complex and the activation of DncV. However, an additional level of regulation is observed for *Vibrio*

cholerae DncV, as it belongs to the type II CBASS, which encodes the accessory regulatory proteins Cap2 and Cap3. In a heterologous host a ubiquitin transferase-like regulatory protein, Cap2, conjugates DncV to an unidentified target molecule, which primes DncV for 3′3′-cGAMP synthesis by an as yet unknown mechanism (Ledvina et al., 2023). Since priming is independent of phage infection, a phage cue must be detected to fully activate DncV and Cap3 may reverse it by sequence-specific proteolytic cleavage. In other type II CBASS from *Bacillus* species Cap2 conjugates CD-NTase to the host protein PspA (phage shock protein A) locking CD-NTase in inactive state, until Cap3 liberates it during phage infection, suggesting that Cap3 has a more complex role than a negative regulation (Krüger et al., 2024).

1.2.1. Activation of Cap effectors

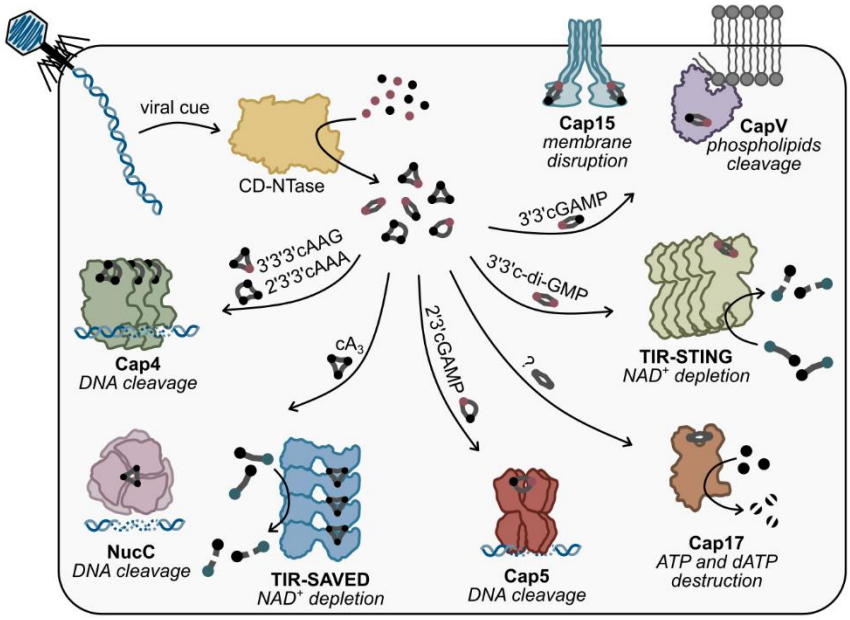


Figure 1.8. Schematic representation of characterized CBASS ancillary effectors. CD-NTase senses phage infection and produces signaling molecules. Signaling molecule-activated effectors target DNA, metabolites or cell membrane integrity causing cell death before phage can finish its replication cycle.

In CBASS systems the signal produced by CD-NTase is recognized by the CD-NTase associated protein (Cap) effectors. Many Cap effectors have a well-defined two domain, sensor-effector, architecture (Millman et al., 2020). The sensor domain is responsible for cyclic-nucleotide messenger binding, while the enzymatic or toxin-like effector domain executes cell death upon activation. CD-NTases can produce various signaling molecules, therefore in

no surprise the sensory domains of Cap effectors are diverse. Cap effectors utilize evolutionary unrelated SAVED, STING or a novel β-barrel fold as sensory domains (Duncan-Lowey et al., 2021; Lowey et al., 2020; Morehouse et al., 2020). The most common sensory domain is SAVED - Cap proteins with a SAVED domain are found in about one third of the CBASS operons (Burroughs et al., 2015). However, some Cap effectors do not contain known nucleotide-binding domains for second messenger recognition (Duncan-Lowey et al., 2021), suggesting the existence of yet uncharacterized recognition modules.

At the effector domain level, the most abundant effector domains are non-enzymatic TM domains (encoded in >40% of CBASS operons), with nucleases (encoded in >20% of CBASS operons) and phospholipases (encoded in ~20% of CBASS operons) domains also being relatively abundant (Duncan-Lowey et al., 2021). In this chapter, I will focus on structurally and biochemically characterized Cap effectors, which are briefly summarized in Figure 1.8.

SAVED domain containing effectors

The SAVED sensory domain is most commonly found fused to HNH nuclease-like domains, PD-(D/E)XK endonuclease-like domains and TM domains, with the effector domain prevalence of 50%, 26% and 20%, respectively. In rare cases (~1-2%), SAVED is found fused to protease (Caspase-like or metallopeptidase-like), TIR or calcineurin domain (Burroughs et al., 2015; Lowey et al., 2020).

To date three CBASS effectors containing SAVED sensory domain have been characterized. Two nuclease-SAVED fusion proteins, Cap4 ((PD-D/ExK)-SAVED) and Cap5 (HNH-SAVED), and NAD^+ degrading TIR-SAVED effector (Fatma et al., 2021; Hogrel et al., 2022; Lowey et al., 2020; Rechkoblit et al., 2024). All these effectors bind two-fold symmetry lacking activators. Cap4 homologues from *Enterobacter cloacae* and *Acinetobacter baumannii* recognize the odd-numbered activators 3′3′3′cAAG and 2′3′3′cAAA respectively (Lowey et al., 2020). TIR-SAVED responds to cA_3 (all 3′-5′ linkages) and Cap5 effectors recognize cyclic dinucleotide 3′2′cGAMP (Fatma et al., 2021; Hogrel et al., 2022; Rechkoblit et al., 2024). As the SAVED domain consists of two non-identical CARF-like subdomains fused in a single polypeptide chain, the formed activator binding pocket allows recognition of two-fold symmetry lacking molecules (Figure 1.9) (Lowey et al., 2020; Rechkoblit et al., 2024).

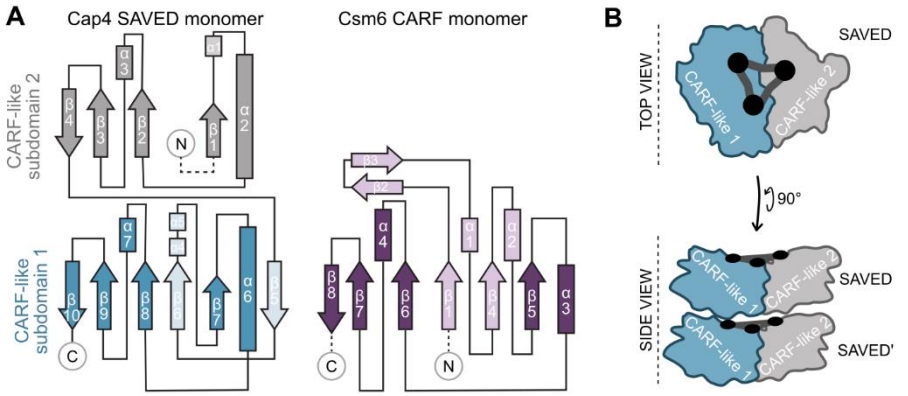


Figure 1.9. Binding of twofold symmetry-lacking activators by the SAVED domain. (**A**) Comparison of CBASS Cap4 SAVED domain and type III CRISPR-Cas Csm6 CARF domain. Topology diagrams of the *A. baumannii* Cap4 SAVED domain and the *T. onnurineus* Csm6 CARF domain are depicted. Each CARF-like subdomain of SAVED (blue and gray) shares a common core CARF domain topology (darker shades). Reproduced from (Lowey et al., 2020). (**B**) Schematic representation of the top and side views of the odd-numbered activator, such as cA_3, binding to the SAVED domain. Non-identical CARF-like subdomains allow the formation of a binding pocket lacking twofold symmetry (top view). Upon activator binding, the SAVED domains stack in a head-to-tail fashion, burying the activator in the domain interface.

Furthermore, the binding of the activator triggers oligomerization of the effector proteins (Figure 1.9 B). In the activator-bound structure, *Microbacterium ketosireducens* TIR-SAVED forms ordered filaments with cA_3 activator sandwiched between the SAVED domains that are organized in a head-to-tail fashion (Hogrel et al., 2022). Upon binding of signaling molecules, Cap5 forms an activated tetramer (dimer-of-dimers), in which pairs of SAVED domains are reorganized to stack together in a head-to-tail fashion with activator locked between the "top" and "bottom" SAVED domains (Fatma et al., 2021; Rechkoblit et al., 2024). Oligomerization is the key aspect to unleashing the enzymatic activity of the effector domain. In the TIR-SAVED effector, filament formation reconstitutes a composite active site of the TIR, allowing it to cleave NAD^+ molecules to ADP ribose and nicotinamide depleting cell's resources of this crucial metabolite (Hogrel et al., 2022). Similarly, activation by formation of higher oligomers is observed for the Cap4 effectors (Lowey et al., 2020). In the case of Cap5, the head-to-tail stacking of the SAVED domains triggers a dimer to tetramer transition, bringing together two HNH domains and converting them to a catalytically competent state (Rechkoblit et al., 2024). Activated nuclease-SAVED fusion proteins indiscriminately cleave phage and host dsDNA, killing the cell before phage replication can be finished (Fatma et al., 2021; Lowey et al., 2020; Rechkoblit et al., 2024).

STING domain containing effectors

The signal-sensing STING adaptor proteins are one of the key proteins in the eukaryotic cGAs-STING signaling pathway (Ablasser and Chen, 2019). The sensory STING domain containing Cap effectors are also found in bacterial CBASS, although they are rare: found in < 5% of CBASS operons. In these instances, STING domain is found fused with TIR or TM domain (Morehouse et al., 2020). Only TIR-STING fusion effector has been characterized (Ko et al., 2022; Morehouse et al., 2022, 2020). TIR-STING employs a homodimeric V-shaped ligand binding pocket, structurally similar to the eukaryotic counterparts, to recognize a two-fold symmetric 3′3′cGAMP activator. Activator binding drives STING domain dimers to oligomerize into long filaments establishing interactions between TIR domains that trigger NADase activity (Morehouse et al., 2022). Similar to TIR-SAVED effector, upon activation TIR-STING effector depletes NAD^+ molecules in turn killing the infected cell.

Effectors with novel or unestablished nucleotide binding domains

Many CBASS operons encode effectors without a previously characterized nucleotide-binding domain (Duncan-Lowey et al., 2021). For example, a cA_3 activated PD-D/ExK-like nuclease NucC, which is found associated with both CBASS and type III CRISPR-Cas systems, lacks a well-defined sensory domain, but utilizes a unique fold to bind the activator and induce the nuclease activity through oligomerization (Lau et al., 2020) (see also Chapter 1.1.3). Recently it has been discovered that β-barrel fold is capable of cyclic nucleotide binding, and it acts as a sensory domain in the TM-(β-barrel) fusion protein Cap15 (Duncan-Lowey et al., 2021). The Cap15 C-terminal β-barrel domain contains a solvent-exposed pocket for activator binding at the top of the β-barrel and two potential protein-protein interaction interfaces. Activator binding triggers the formation of higher order Cap15 oligomers of undetermined length. *In vivo* activation of Cap15 results in the collapse of the inner membrane of the cell in turn killing the cell. Based on the nature of Cap15 TM segments it is more likely that Cap15 causes membrane disruption through formation of raft-like assemblies, than through formation of cell membrane piercing pores.

Some type II CBASS contain a CapV (cGAMP-activated phospholipase in *Vibrio*) effector, which possesses a phospholipase activity upon activation by 3′3′cGAMP (Cohen et al., 2019; Severin et al., 2018). Upon activations CapV has been shown to degrade membrane phospholipids, triggering the cell death. However, at the sequence level no defined nucleotide binding domain was

identified and studies of CapV effectors lack structural data, therefore the exact activation mechanism remains unknown.

The activation of the CBASS effector Cap17, which contains a predicted cyclic nucleotide-binding domain annotated as Pfam PF18178, also remains to be elucidated, as a recent study on this effector lacks structural data (Rousset et al., 2023). Cap17 contains a C-terminal domain of the purine nucleoside phosphorylase family which upon activation acts as an ATP nucleosidase, degrading ATP and dATP to block phage replication (Rousset et al., 2023).

As can be seen from the examples discussed, most CBASS effectors target essential cellular elements such as DNA, the cell membrane, or an important metabolite, NAD^+ or ATP and dATP, to kill the cell and thus stop phage replication (Figure 1.8). CBASS effectors appear to be modular systems that allow different combinations of sensor and effector domains. Despite this variability, the activation mechanism is uniform – signaling molecule binding induced oligomerization.

1.3. Other defense systems with nucleotide-based signaling pathway

1.3.1. Pycsar

Around 2.4% of sequenced bacteria encode the pyrimidine cyclase system for antiphage resistance (Pycsar) antiviral defense system which is composed of two genes, one encoding the pyrimidine cyclase PycC and the other encoding either TM domain-containing immune effector PycTM or TIR domain-containing immune effector PycTIR (Tal et al., 2021). Upon phage infection PycC is activated to synthesize signaling molecules cyclic nucleotide monophosphates. Based on sequence similarities PycC forms five (A-E) major clades and PycC from clades A-D synthesizes 3′,5′-cyclic uridine monophosphate (cUMP), while PycC belonging to clade E synthesizes 3′,5′-cyclic cytidine monophosphate (cCMP) secondary messenger (Figure 1.10).

The immune effector PycTM is commonly found encoded next to the cCMP-producing PycC and when activated PycTM causes abnormal membrane protrusions that lead to the cell death (Figure 1.10) (Tal et al., 2021). The PycTIR effectors associate with the cUMP producing PycC and upon activation cause cell death by depleting NAD^+ (Figure 1.10). Similar to TIR domain containing CBASS effectors, TIR-SAVED and TIR-STING, activated PycTIR also forms filamentous structures, highlighting the common mechanism to activate TIR-domain effectors (Hogrel et al., 2022; Morehouse et al., 2022; Tal et al., 2021).

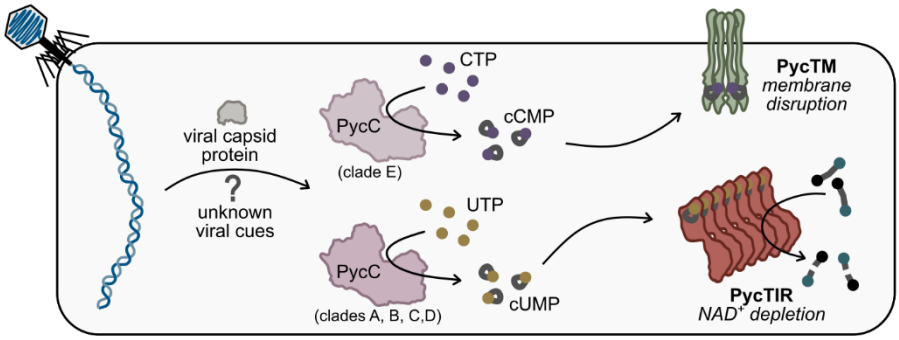


Figure 1.10. Schematic representation of Pycsar immunity. The pyrimidine cyclase PycC senses viral infection and starts to synthesize cyclic nucleotide monophosphates. PycC belonging to the clade E synthesizes cCMP from CTP molecules. Then cMP activates PycTM effector, which disrupts the membrane integrity. PycC of the clades A-D produces cUMP signal, which activates PycTIR effector. Upon cUMP binding PycTIR forms activated filaments and degrades cellular NAD^+ recourse.

The cyclase, PycC, is thought to respond to a molecular cue, possibly a viral capsid protein, that is produced at a late stage of viral replication (Tal et al., 2021). It has been suggested that abortive infection-inducing host defense pathways, such as Pycsar, are reserved as a last resort response and are activated only in the final stages of phage replication, i.e., during capsid assembly stage (Hobbs and Kranzusch, 2024).

1.3.2. Thoeris

Similar to other nucleotide signaling based antiviral systems, the two core elements of the Thoeris defense system are (i) an infection sensing and signal-producing protein and (ii) an effector protein (Hobbs and Kranzusch, 2024). Thoeris systems are classified into four types (I - IV) (Doron et al., 2018; Ofir et al., 2021; Rousset et al., 2025; Sabonis et al., 2024; van den Berg et al., 2024) (Figure 1.11 A). All types encode at least one TIR domain-containing protein, ThsB, which, in a striking contrast to the TIR domain-containing proteins of the CBASS and Pycsar systems, is not an NAD^+-degrading effector, but rather a signal-producing protein (Hogrel et al., 2022; Leavitt et al., 2022; Manik et al., 2022; Morehouse et al., 2022; Ofir et al., 2021; Tal et al., 2021; Tamulaitiene et al., 2024). In type III Thoeris, the core ThsB protein additionally contains the nuclease-like domain NucS (van den Berg et al., 2024).

During phage infection, ThsB recognizes a viral cue and utilizes its TIR domain to synthesize signaling molecules from NAD^+ substrate (Figure 1.11

B). ThsB of type I Thoeris systems produces a cyclic ADP-ribose isomer containing a 1″-3′ ribose-ribose linkage (1″–3′ gcADPR) (Leavitt et al., 2022; Manik et al., 2022; Ofir et al., 2021; Tamulaitiene et al., 2024), type IV Thoeris relies on ADP-ribose isomer containing 1″-N7 ribose-adenine bond (N7-cADPR) signaling (Rousset et al., 2025), while ThsB of type II Thoeris conjugates NAD^+ and histidine forming His-ADPR signaling molecule (Sabonis et al., 2024).

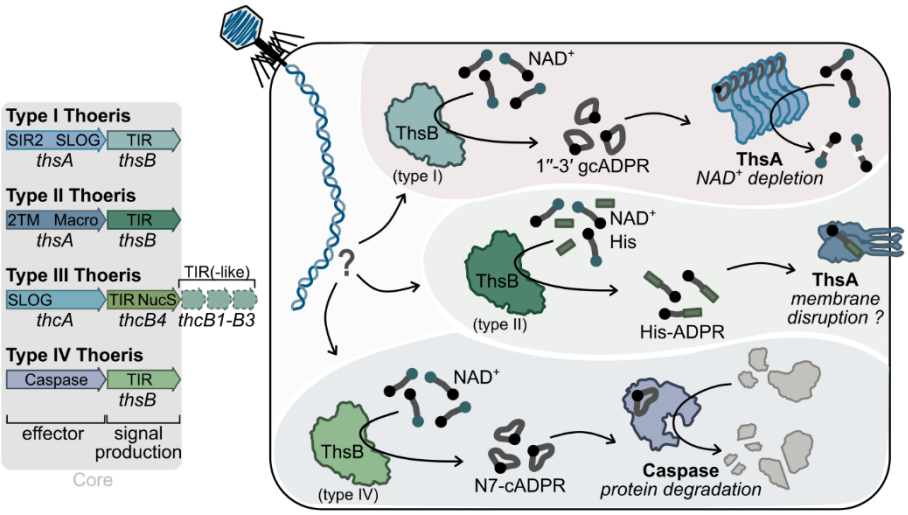


Figure 1.11. Overview of the mechanisms of Thoeris immunity. Schematic representation of type I-IV Thoeris operons (left) and the mechanisms of characterized type I, II and IV Thoeris immune signaling (right). Upon phage infection TIR domain-containing protein ThsB recognizes yet unidentified viral cues and synthesizes signaling molecules. The type I ThsB produces 1”-3′gcADPR signal, type II ThsB synthesize His-ADPR molecules, while type IV produces N7-cADPR. The recognition of the cognate signaling molecule triggers NADase activity of ThsA in type I systems and protease activity of the Caspase-like effector of type IV Thoeris. In type II Thoeris His-ADPR-activated ThsA kills the cell, possibly by membrane disruption.

Produced signaling molecules are further recognized by the effector proteins to trigger their activity (Figure 1.11 B). The nucleotide-binding Smf/DprA-LOG (SLOG) domain of type I ThsA recognizes and binds 1″–3′ gcADPR, triggering signaling molecule-induced oligomerization and conformational changes that reposition the active site residues of the sirtuin (SIR2) domain, which rapidly degrades NAD^+ (Leavitt et al., 2022; Ofir et al., 2021; Tamulaitiene et al., 2024). Type II ThsA utilizes nucleotide binding Macro domain for His-ADPR signal recognition (Sabonis et al., 2024). Since type II ThsA contains TM domains, it is likely that upon His-ADPR binding ThsA causes cell death through inner membrane disruption. While type III

Thoeris remains to be characterized, the type IV Thoeris functions by activating a promiscuous caspase-like protease that, when activated, cleaves several essential proteins in *E. coli*, such as the ribosomal protein RplE and the elongation factor Tu (EF-Tu) (Rousset et al., 2025).

Thus, the Thoeris defense system produces diverse signaling molecules and utilizes a variety of effector modules.

1.4. Disruption of nucleotide-based signaling

On average, bacteria encode six antiphage defense systems per genome (Tesson et al., 2022). These systems may be activated by different phages or may function as multiple lines of defense against the same phage at different stages of infection (Tesson and Bernheim, 2023). CBASS, Thoeris, and Pycsar are thought to act as a second line of defense during late stages of infection by activating effector proteins that target essential cell components and in turn killing the infected cell to provide population-level protection (Hobbs and Kranzusch, 2024). However, it is in the interest of the host to limit the activation of effector proteins to avoid excessive toxicity and to recover after infection is cleared (Stella and Marraffini, 2024). In type III CRISPR-Cas systems, this is achieved by regulating the production and the lifetime of signaling molecules through target RNA cleavage and degradation of signaling molecules, respectively (Athukoralage and White, 2022, 2021; Stella and Marraffini, 2024). However, for other nucleotide signaling based defense systems, no specific signal degrading enzymes encoded by the host bacteria are known.

From a phage perspective, interfering with immune signaling is an excellent strategy for evading antiviral defenses. Not surprisingly, phage-encoded anti-defense proteins targeting signaling molecules have been discovered to overcome the immunity of all characterized bacterial nucleotide-based signaling systems (Murtazalieva et al., 2024). In this chapter, I will discuss the regulation of signaling in type III CRISPR-Cas and will review anti-defense proteins that specifically target signaling molecules of the type III CRISPR-Cas, Pycsar, CBASS, and Thoeris systems.

1.4.1. Switching off the type III CRISPR-Cas signaling

In type III CRISPR-Cas systems cA_n signaling is first controlled at the synthesis level - cA_n production is halted once the target RNA is cleaved by the interference Csm/Cmr complex (Figure 1.12) (Grüschow et al., 2019; Kazlauskiene et al., 2017; Nasef et al., 2019; Rouillon et al., 2018). Signaling

molecules can be further cleared in a timely manner by dedicated CRISPR ring-nucleases (Crn), or in some cases by the effector itself using its sensory CARF domain, which possesses a ring-nuclease activity (Athukoralage and White, 2022, 2021; Patel et al., 2022). Since cA_4 molecules seem to be the most common second messengers in type III CRISPR-Cas signaling (Hoikkala et al., 2024), to no surprise cA_4-cleaving Crn and self-limiting CARF effectors have been extensively characterized. Three Crn families (Crn1-3) of CARF and CARF-unrelated folds and self-limiting CARF-effectors belonging to different CARF clades have been structurally and biochemically characterized (Figure 1.12) (Athukoralage et al., 2020b, 2020c, 2019, 2018; Du et al., 2024; Jia et al., 2019b; Li et al., 2025; Makarova et al., 2020a; Mogila et al., 2023; Samolygo et al., 2020).

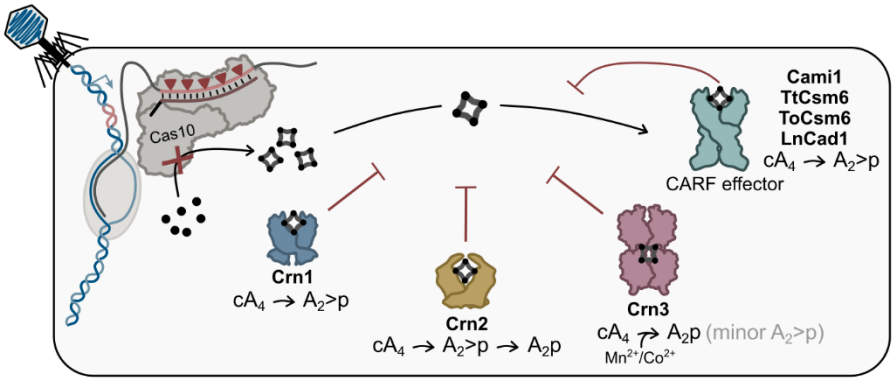


Figure 1.12. Overview of cA_4 signaling termination in type III CRISPR-Cas systems. Production of cA_4 by the Cas10 Palm domains is switched off following the target transcript cleavage by the interference complex (red triangles). Specialized cA_4-specific ring nucleases Crn1, Crn2, and Crn3 terminates signaling by cA_4 cleavage. Furthermore, CARF domain of some characterized CARF effectors (Cami1, TtCsm6, ToCsm1, LnCad1) possess ring nuclease activity. Ring nuclease reaction scheme is indicated for each enzyme family. A_2>p - linear di-adenylates with 2′,3′-cyclic phosphate; A_2p - linear di-adenylates with 3′- phosphate.

CRISPR ring-nuclease 1 (Crn1)

Minimal dimeric CARF domain proteins of the CARF7 or CARF_m13 clades that specifically catalyze the cA_4 cleavage into linear di-adenylates with 2′,3′-cyclic phosphate (A_2>p) in a metal-independent process have been named CRISPR ring-nuclease 1 (Crn1) (Athukoralage et al., 2018; Du et al., 2023; Molina et al., 2022, 2021). Crn1 proteins do not contain a fused effector domain and instead are fused only to a short wHTH domain or contain a short C-terminal (Ct) insertion (Du et al., 2023; Makarova et al., 2020a; Molina et al., 2022, 2021). Extensive structural studies of the *Sulfolobus islandicus* Crn1 proteins Sis0811 and Sis0455 of CARF-wHTH and CARF-Ct architectures, respectively, revealed that both Crn1 undergo drastic conformational changes

upon cA_4 binding (Molina et al., 2022, 2021). Substrate binding to the Sis0811 triggers a cork-screw motion of wHTH, which in turn alters the conformation of the CARF domains and positions key catalytic residues to proceed with cA_4 cleavage (Figure 1.13 A) (Molina et al., 2021). Upon cA_4 binding, two monomers of Sis0455 undergo a jaw-like motion that traps the substrate in the catalytic pocket with Ct inserts forming the closure (Figure 1.13 B) (Molina et al., 2022).

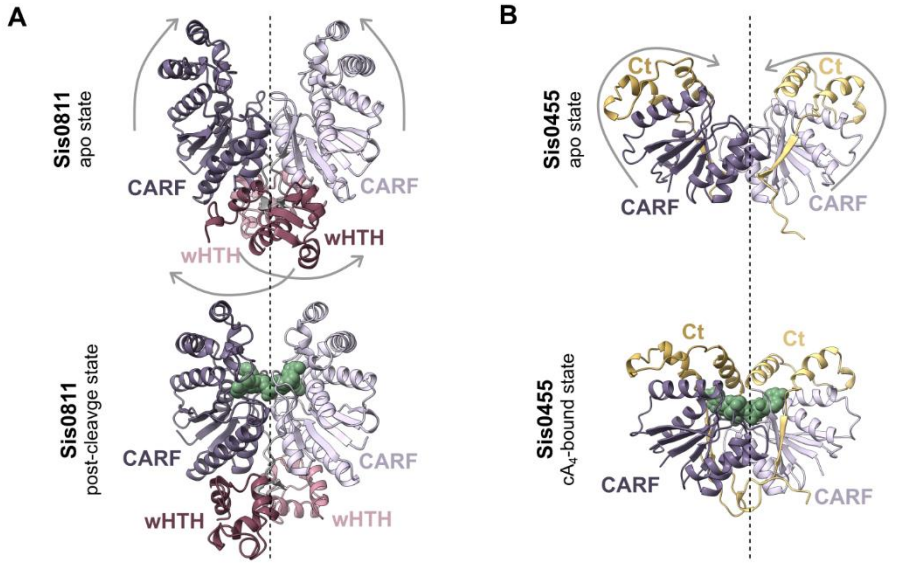


Figure 1.13. Architectures of Crn1 (**A**) Structure of Sis0811 - a Crn1 of CARF-wHTH architecture, in apo (PDB ID: 7PQ2) and post-cleavage (PDB ID: 7PQ3) states (Molina et al., 2021). (**B**) Structure of Sis0455 - a Crn1 of CARF-Ct insertion architecture, in apo (PDB ID: 7Z56) and cA_4-bound (PDB ID: 7Z55) states (Molina et al., 2022). CARF domains are colored shades of purple, wHTH – shades of burgundy, cA_4 is shown as green spheres. Arrows indicate movement upon cA_4 binding.

Each CARF domain of Crn1 has two conserved catalytic motifs (motif-I (xGT**S**) and motif-II (TxGx**K**)) that are involved in ring nuclease activity (Figure 1.14) (Makarova et al., 2020a). The motif-I residues are involved in the coordination of the scissile phosphate and sterically positioning the ribose 2′-OH group of the ribose for the inline nucleophilic attack and later stabilizing the 2′-3′cyclic phosphate, while the lysine from motif-II is responsible for the initial substrate binding and for stabilization the pentacovalent phosphorus formed in the transition state (Du et al., 2023; Molina et al., 2022, 2021).

Although both CARF domains of dimeric Crn1 are identical and contribute the identical residues to form a composite active site for cA_4 cleavage, the

cleavage reactions occur in non-concerted manner - some linear tetra-adenylate with cyclic-2′3′ phosphate (A_4>p) intermediate is observed in the reactions (Athukoralage et al., 2018; Du et al., 2023; Molina et al., 2022). Moreover, Crn1 enzymes of different architectures are found encoded by the same host and appear to have different affinities for cA_4 and different cA_4 degradation rates (Molina et al., 2022), which may reflect the regulatory needs of a given type III CRISPR system (Athukoralage and White, 2021).

Self-limiting CARF-effectors

Some CARF domains that are fused to effector domains have been found to possess metal independent cA_4-specific ring nuclease activity similar to Crn1 proteins (Athukoralage et al., 2019; Du et al., 2024; Jia et al., 2019b; Li et al., 2025; Mogila et al., 2023). In this case degradation of the activating molecule would provide a "timer" mechanism allowing a timely deactivation of the effector. CARF-effectors belonging to different CARF clades have been demonstrated to possess ring nuclease activity: Csm6 proteins of CARF1 and CARF9 groups (Athukoralage et al., 2019; Du et al., 2024; Jia et al., 2019b; Makarova et al., 2020a), Cad1 of CARF5 clade (Li et al., 2025; Makarova et al., 2020a) and Cami1 of CARF7 group (Makarova et al., 2020a; Mogila et al., 2023).

Since Cami1 (CARF-wHTH-RelE) belongs to the same CARF7 group as Crn1 of the CARF-wHTH architecture, it is not surprising that Cami1 utilizes the same cA_4-specific cleavage mechanism (Athukoralage et al., 2018; Mogila et al., 2023; Molina et al., 2021). The Cad1 proteins belong to the CARF5 clade and possess similar ring nuclease active site motifs as the CARF7 and CARF_m13 groups, however, only a homolog of Cad1 from *Limisphaera ngatamarikiensis* has been shown to be capable of cA_4 cleavage (Figure 1.14) (Baca et al., 2024a; Li et al., 2025; Makarova et al., 2020a). In the case of Csm6, at least two signature variants of catalytic motifs can be associated with ring nuclease activity (Figure 1.14) (Makarova et al., 2020a). *Thermus thermophilus* Csm6, which belong to the CARF1 clade, contain catalytic motifs, that are similar to CARF7, CARF5 and CARF_m13 groups (Athukoralage et al., 2019; Du et al., 2024; Makarova et al., 2020a). Contrary *Thermococcus onnurineus* Csm6, which belong to the CARF9 clade, contains significantly different motif-I and motif-II and requires the conserved W14, which coordinates and locks the scissile phosphate in a conformation compatible with in-line attack by ribose 2′-OH, for cA_4 cleavage (Figure 1.14) (Jia et al., 2019b; Makarova et al., 2020a). Interestingly, *T. onnurineus* Csm6 crystalizes with cA_4 bound by both domains CARF and HEPN, suggesting that nonspecific HEPN RNase activity could also be involved in cA_4 cleavage

(Jia et al., 2019b). In accordance, other nonspecific RNase, membrane-associated DHH-DHHA1 family nuclease (MAD), has been demonstrated to be capable of cA_4 degradation (Zhao et al., 2020).

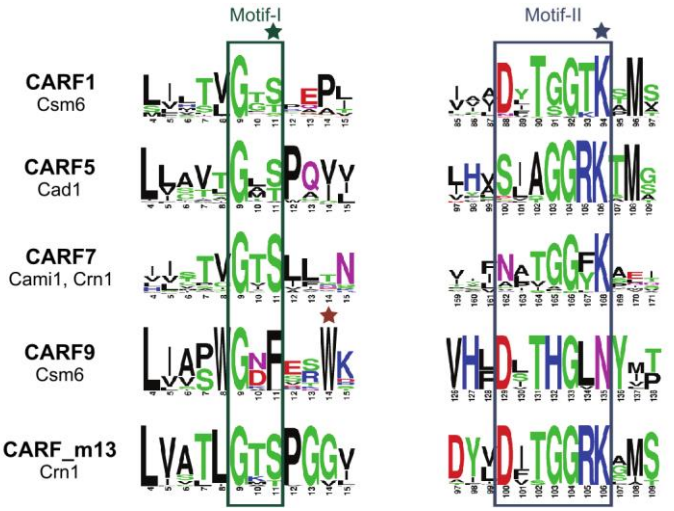


Figure 1.14. Ring nuclease motifs of CARF domains. Conserved ring nuclease motifs of CARF1, CARF5, CARF7, CARF9 and CARF_m13 group proteins, reproduced from (Makarova et al., 2020a). Stars indicate catalytic residues.

Interestingly, to this date only effectors targeting transient molecules such as RNA and ATP have been observed to possess ring nuclease activity (Athukoralage et al., 2019; Du et al., 2024; Jia et al., 2019b; Li et al., 2025; Mogila et al., 2023). CARF domains of cA_4-activatable effectors targeting DNA (Can1 and Can2), or membrane (Cam1) do not possess this regulatory mechanism (Baca et al., 2024b; McMahon et al., 2020; Rostøl et al., 2021; Zhu et al., 2021). Moreover, the importance of ring nuclease activity for cell survival has been illustrated in Cami1systems, where mutations impairing ring nuclease activity caused elevated cell toxicity (Mogila et al., 2023), highlighting the need to control activation of CARF-effectors for cells to recover.

Crn2

Unlike Crn1 and self-limiting CARF-effectors, Crn2 uses the CARF-unrelated DUF1874 fold for cA_4 binding and cleavage (Athukoralage et al., 2020b; Samolygo et al., 2020). Although Crn2 homologs were first identified as phage-encoded anti-CRISPR protein neutralizing type III CRISPR-Cas immunity (AcrIII-1), standalone Crn2 is also found in bacterial genomes associated with type III CRISPR-Cas systems, moreover, Crn2 can also be fused to ring nuclease-deficient Csx1 effectors (Athukoralage et al., 2020b;

Samolygo et al., 2020; Zhang et al., 2024). The only characterized homologue of a standalone Crn2 is the viral AcrIII-1 (Athukoralage et al., 2020b) (Chapter 1.4.2), in bacteria only the Csx1-Crn2 fusion protein have been characterized structurally and biochemically (Samolygo et al., 2020; Zhang et al., 2024). Csx1-Crn2 from *Marinitoga piezophile* consists of the C-terminal Crn2 domain and the N-terminal Csx1-like domain, which is further subdivided into CARF, HTH and HEPN domains (Samolygo et al., 2020; Zhang et al., 2024). Csx1-Crn2 adopts a tetrameric (dimer-of-dimer) structure: the two monomers associate to form a dimer through the Csx1 domains, and two such dimers further dimerize through the Crn2 domains (Zhang et al., 2024). Each Crn2 dimer adopts a basket-like shape and the binding of cA_4 triggers significant local conformational changes: the mobile loops completely enclose the cA_4 molecule within the catalytic pocket, allowing a conserved glutamate E536 from the mobile loop to coordinate the imidazole side chain of the catalytic H495 residue, which may act as a general acid, facilitating in-line attack of the 2′-OH (Zhang et al., 2024).

Crn3 (Csx3)

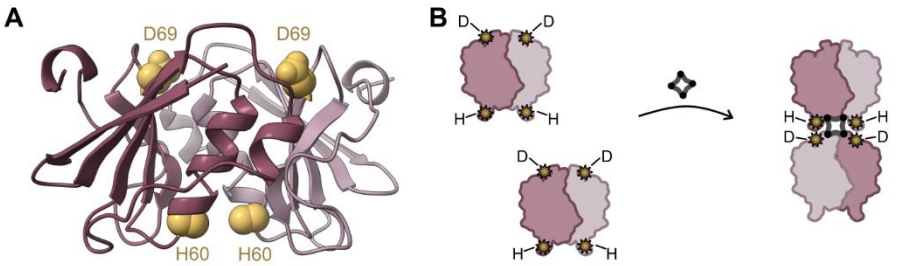


Figure 1.15. Formation of a composite active site by Crn3. (**A**) The main active site residues (yellow spheres) are located on opposite sides of the Crn3 dimer (PDB ID: 6YUD). Each protein in the dimers is colored different shades of burgundy. (**B**) Schematic representation of the reconstitution of a composite active site at the interface of two dimers. Active site residues are depicted as yellow stars and indicated.

Crn3 (also known as Csx3) is the only divalent metal ions-dependent cA_4-specific ring nuclease characterized to date (Athukoralage et al., 2020c; Brown et al., 2020). Crn3 is a sulphate transporter and anti-sigma factor antagonist (STAS) protein that forms dimers in solution that are somewhat similar to CARF (Kaur et al., 2020; Makarova et al., 2020a). The critical catalytic residues (H60 or D69) are located on opposite sides of the Crn3 dimer (Figure 1.15 A), therefore upon cA_4 binding, a composite ring nuclease active site must be formed by stacking Crn3 dimers head-to-tail (Figure 1.15 B) (Athukoralage et al., 2020c). This stacking potentially leads to filament formation. It is thought that conserved H60 are involved in Mn^{2+} coordination,

and a metal-activated hydroxyl ion initiates nucleophilic attack on the phosphodiester bond, therefore the primary product of cA_4 cleavage is A_2p rather than A_2>p (Athukoralage et al., 2020c).

Other enzymes in degradation of signaling molecules

The potential regulation of SAM-AMP signaling has been also described (Chi et al., 2023). Type III CRISPR-Cas systems encoding CorA effector, which has been demonstrated to recognize SAM-AMP signaling and form channel-like pores in cell membrane (Chapter 1.1.3), also encode the DHH family phosphodiesterase - nuclease NrN or a DEDD family nuclease (Chi et al., 2023). The *B. fragilis* NrN protein has been shown to specifically degrade SAM-AMP (Chi et al., 2023). It is possible that specialized NrN and DEDD family phosphodiesterases are an 'off switch', analogous to the ring nucleases that degrade cA_4 molecules. However, the obligatory associated effector CorA requires the presence of NrN to function *in vivo* (Chi et al., 2023), suggesting that specialized NrN and DEDD family phosphodiesterases are required to prevent desensitization of the CorA channel at high activator concentrations (Chi et al., 2023; Velisetty and Chakrapani, 2012). Therefore, signal-degrading enzymes may have a more subtle regulatory function than simply turning off the signal.

1.4.2. Anti-defense proteins

As bacteria encode many different antiviral defense mechanisms, phages have also evolved to evade bacterial defense systems. To do this, phages encode different anti-defense proteins that target specific defense systems (Murtazalieva et al., 2024). In nucleotide-based signaling systems, signaling molecules appear to be a particularly vulnerable part. Phages have evolved proteins that can degrade or sequester second messengers, thereby undermining the immunity of defense systems that rely on nucleotide-based signaling (Athukoralage et al., 2020b; Cao et al., 2024; Hobbs et al., 2022; Huiting et al., 2023; Jenson et al., 2023; Leavitt et al., 2022; Yirmiya et al., 2024) (Figure 1.16).

Signal-cleaving anti-defense proteins

Reducing the effective concentration of signaling molecules generally disrupts the signal. For example, the viral Crn2 homolog, an anti-CRISPR protein neutralizing type III CRISPR-Cas immunity (AcrIII-1), is an efficient cA_4-specific ring nuclease that cleaves cA_4 ~100 times faster than canonical Crn1, allowing the phage to subvert type III CRISPR-Cas immunity (Athukoralage et al., 2020b). Accordingly, the phage-encoded protein of the

compact 2H-phosphoesterase family, known as anti-CBASS 1 (Acb1) protein, degrades at least one adenosine-containing signaling molecules and subverts CBASS immunity (Hobbs et al., 2022). Acb1 captures the substrate in a U-shaped ligand binding pocket and locks it in by a flexible C-terminal lid. This distorts the molecule and positions the scissile phosphate over an active site HxT/HxT tetrad for acid-base catalysis. The cleavage reaction results in a linear product with 5'-OH and 3'-phosphate (Hobbs et al., 2022).

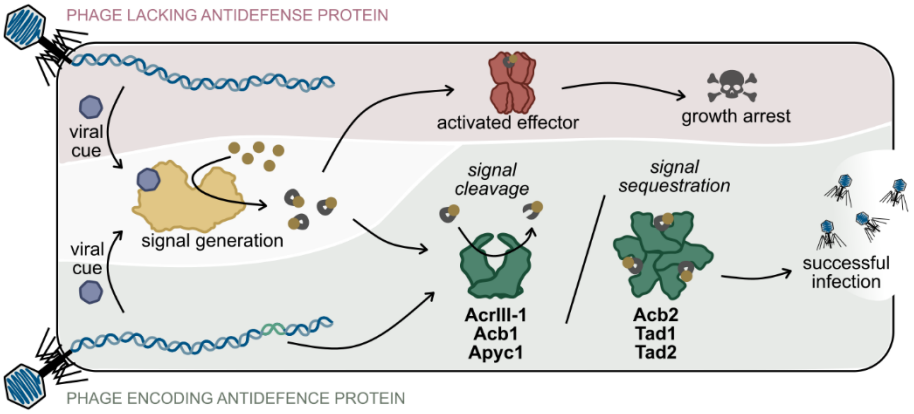


Figure 1.16. Overview of phage anti-defense strategies to evade nucleotide-centric immune systems. When a cell is infected by a phage lacking anti-defense protein, the signaling molecules generated by nucleotide-centric defense system activates associated effectors triggering cell dormancy or abortive infection. When a cell is infected by a phage that encodes protein that can degrade (AcrIII-1, Acb1 and Apyc1) or sequester (Acb2, Tad1, Tad2) signaling molecules, associated effectors are not triggered, cell do not undergo premature cell death and phage successfully finishes its replication cycle.

To evade Pycsar immunity phages utilize a homomeric β-lactamase (MBL) fold phosphodiesterase, an anti-Pycsar 1 (Apyc1) proteins with a highly conserved HxHxDH catalytic motif (Hobbs et al., 2022). The Zn^{2+}-dependent Apyc1 protein cleaves cUMP and cCMP to UMP and CMP with 5′-phosphate. Intriguingly, Apyc1 homologs are found encoded in bacteria, raising the possibility that host Apyc1 enzymes may play a role in regulating Pycsar defense (Hobbs et al., 2022).

Signal-sequestering anti-defense proteins

A different strategy to reduce the effective concentration of signaling molecules is utilized by a widespread Acb2 proteins and Thoeris anti-defense proteins Tad1 and Tad2. These proteins functions as molecular “sponges” sequestering immune signals and preventing the activation of corresponding effector proteins (Cao et al., 2024; Huiting et al., 2023; Jenson et al., 2023; Leavitt et al., 2022; Yirmiya et al., 2024). Acb2 forms hexametric complexes,

that can bind up to three cyclic di nucleotide signaling molecules (Cao et al., 2024; Huiting et al., 2023; Jenson et al., 2023). Interestingly, many Acb2 homologs have an additional binding surface that can accommodate cyclic trinucleotides. Therefore this "sponging" mechanism can antagonize the cA_3 based type III CRISPR-Cas signaling (Cao et al., 2024).

Tad1 are homodimers forming two ligand-binding pockets at the top and bottom ends of the Tad1 complex and has been demonstrated to bind multiple isoforms of gcADPR, 1″–2′ gcADPR and 1″–3′ gcADPR (Leavitt et al., 2022), while Tad2 forms a homotetrameric assembly with two ligand-binding pockets to enclose 1″–3′ gcADPR (Yirmiya et al., 2024). Furthermore, some Tad2 homologs can sequester the signaling molecules, His-ADPR, of type II Thoeris systems (Sabonis et al., 2024).

1.5. Signal transduction by proteases

1.5.1. Proteases in antiviral defense systems

Nucleotide-based signaling is not the only means of transmitting the signal from the infection sensor to the effector of the antiviral defense system. Some defense systems, such as bacterial gasdermins (bGSDMs) and type III-E CRISPR-Cas systems, utilize proteolytic cleavage to activate downstream effectors (Johnson et al., 2022; Yang and Patel, 2022). In addition, other antiviral defense systems, such as Borvo, PD-λ-2 systems and some homologs of Lamasu and antiviral STAND (Avs), are known to encode proteases or protease-like domains (Gao et al., 2022; Millman et al., 2022; Vassallo et al., 2022). However, the defense mechanisms of these systems remain to be elucidated.

To date, only bacterial antiviral defense systems containing caspase-like proteases have been more thoroughly characterized (Johnson et al., 2022; Rousset et al., 2025; Steens et al., 2024; Yang and Patel, 2022). Prokaryotic caspase-like proteins (sometimes referred to as orthocaspases or metacaspases) are classified in the MEROPS peptidase database as belonging to CD clan C14 along with metazoan caspases (Rawlings et al., 2018). In animals, caspases are known to play a central role in innate immunity by forming large multiprotein complexes called inflammasomes, sensing pathogen-associated molecular patterns or danger-associated molecular patterns, and activating proteolytic activity to cleave their substrates such as the pro-inflammatory cytokines pro-IL-1β and pro-IL-18 or gasdermin D to induce pyroptotic cell death (Van Opdenbosch and Lamkanfi, 2019). Furthermore, caspases are responsible for programmed cell death, known as

apoptosis. These caspases are functionally subdivided into initiator and executioner caspases (Julien and Wells, 2017; Zhuang et al., 2022). During apoptosis the initiatory caspases undergo proximity-induced autoactivation and then can activate executioner caspases by cleaving the corresponding zymogens. Activated executioner caspases can cleave several hundred substrates, leading to the characteristic morphological changes of apoptosis (Julien and Wells, 2017).

Parallels to proteolytic activities of caspases in animal innate immunity and programmed cell death can be found in characterized bacterial antiviral defense systems. Similarly to inflammatory caspases, prokaryotic caspase-like proteins transduce the signal to the effectors in bGSDM and type III-E CRISPR-Cas systems (Johnson et al., 2022; Yang and Patel, 2022). Caspase-like proteases that act as executioners can be found in signaling molecule-activated proteolytic effectors of type IV Thoeris and type III-B CRISPR-Cas (Rousset et al., 2025; Steens et al., 2024) (see also Chapters 1.1.3 and 1.3.2).

Bacterial gasdermin (bGSDM) systems

Prokaryotes encode an antiviral defense system known as bacterial gasdermins (bGSDM), which is an analog of pyroptotic cell death inducing eukaryotic gasdermins and works by inducing cell death through the formation of membrane disrupting pores (Johnson et al., 2022). The bGSDM systems consist of two genes, one encoding the bGSDM protein and another encoding protease domain containing protein. Upon phage infection, the protease component is activated, and it specifically cleaves bGSDM, removing the inhibitory C-terminal fragment, allowing bGSDM to undergo conformational changes and oligomerize in the membrane to form pores (A. G. Johnson et al., 2024; Johnson et al., 2022). The pore-pierced membrane loses its integrity and the cell dies before phage is able to finish its replication cycle.

Even though the molecular cues that activate bGSDM systems are yet to be identified, the protease domain-containing component is implicated in sensing the infection since they are commonly found fused to NACHT domains or repeat domains such as tetratricopeptide repeat (TPR), leucine-rich repeat (LRR) or WD40 repeat that are known to be involved in pathogen recognition and inflammasome function in human innate immunity (Johnson et al., 2022; Shi et al., 2020). Protease domains of various types can be associated with bGSDM, but in most cases the proteases domains are caspase-like peptidases (Johnson et al., 2022).

Type III-E CRISPR-Cas

The caspase-like protease TPR-CHAT (tetratricopeptide repeat fused to Caspase HetF Associated with TPRs), which structurally resembles

eukaryotic separase (Cui et al., 2022), is responsible for signal transduction in the type III-E CRISPR-Cas systems. The type III-E CRISPR-Cas systems lack the type III signature Cas10 protein and thus differ significantly from the canonical type III CRISPR-Cas systems reviewed in Chapter 1.1. The genomic organization of type III-E systems includes a large chimeric gene encoding the Cas7-11 polypeptide, a *csx29* gene encoding a caspase-like protease TPR-CHAT, and three accessory genes that encode an alternative extracytoplasmic function (ECF)-like sigma factor CASP-σ (sometimes referred to as RpoE or σ^E) and two additional proteins Csx31 and Csx30 (Figure 1.17 A) (Özcan et al., 2021; Yang and Patel, 2022).

The effector module of Type III-E CRISPR-Cas is composed of the crRNA, Cas7-11 and Csx29 (TPR-CHAT) proteins and is called CRISPR-guided caspase (Craspase) (van Beljouw et al., 2021). Similarly to interference complexes of canonical type III CRISPR-Cas systems, the Cas7-11 part of Craspase recognizes and cleaves target RNA based on the complementarity to crRNA (Cui et al., 2022; Kato et al., 2022a, 2022b; Liu et al., 2022; Özcan et al., 2021; van Beljouw et al., 2021). Furthermore, the TPR-CHAT is activation is achieved upon binding of the target RNA, which contains non-complimentary 3′-flanking sequence, therefore the complementarity between the 5′-handle of the crRNA and the 3′-flanking sequence of the target RNA serves as a mechanism to discriminate RNA of self *versus* non-self origin (Cui et al., 2022; Kato et al., 2022a; Liu et al., 2022). Binding of the cognate target RNA induces the rearrangement of the active site of TPR-CHAT proteases into an active proteolytic state (Cui et al., 2022; Kato et al., 2022a; Liu et al., 2022). The activated protease domain specifically cleaves Csx30, which was proposed to function as an anti-σ factor to its cognate ECF-like σ factor CASP-σ (Cui et al., 2022; Kato et al., 2022a; Liu et al., 2022; Strecker et al., 2022). Typically, ECF σ factors are bound to their cognate anti-σ factors until an environmental signal triggers their release, allowing ECF σ factors to bind to RNA polymerase and activate the transcription of a set of genes (Mascher, 2023). Similarly, after Csx30 cleavage, CASP-σ is thought to be released and functions as a σ factor initiating the expression of genes, such as of CRISPR adaptation genes (Strecker et al., 2022). Although a simple release of σ factor from its cognate anti-σ factor following cleavage of the anti-σ factor seems trivial, it has been observed that all three accessory proteins, cleaved Csx30, Csx31 and CASP-σ, are essential for protection against phages (Liu et al., 2022). Therefore, the exact mechanism of how CASP-σ is released from the complex after Csx30 cleavage and the role of the Csx30 cleavage fragments, Csx30-C and Csx30-N, as well as the role of Csx31, remain to be determined.

Furthermore, similar to typical type III CRISPR-Cas systems, the activation of TPR-CHAT is temporal - the cleavage of target RNA by Cas7-like domains deactivates the protease activity (Cui et al., 2022; Kato et al., 2022a). The proposed mechanism of action of type III-E CRISPR-Cas systems is schematically depicted in Figure 1.17 B.

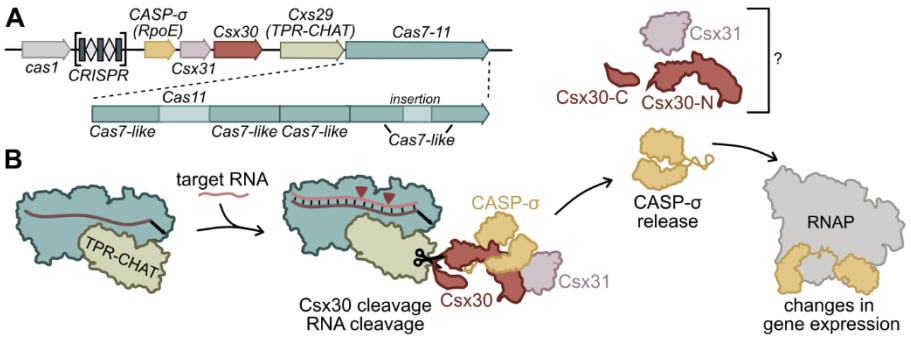


Figure 1.17. Type III-E CRISPR-Cas system. (**A**) Representative CRISPR–cas locus of type III-E CRISPR-Cas system based on the type III-E CRISPR-Cas from *D. ishimotonii* (Özcan et al., 2021). The type III-E CRISPR-Cas system is composed of adaptation genes, CRISPR array, gene encoding large polypeptide Cas7-11, *csx29* gene encoding TPR-CHAT protease and three accessory genes *csx30*, *csx31* and *CASP-σ*. The large Cas7-11 polypeptide is composed of 5 domains: 4 Cas7-like domains, one of which contains insertion, and Cas11-like domain. (**B**) Mechanism of action of type III-E CRISPR-Cas system. Cas7-11 together with Csx29 (TPR-CHAT) and crRNA forms the interference complex known as Craspase. Upon target RNA binding Csx29 is activated and cleaves it substrate Csx30. This cleavage releases CASP-σ, which forms holoenzyme with RNA polymerase and initiates expression of specific genes. The role of the accessory protein Csx31 and the cleavage products of Csx30 remain unknown. Craspase also possesses RNase activity and similarly to typical type III CRISPR-Cas systems cleave target RNA in turn deactivating protease activity of Csx29.

1.5.2. ATP-dependent proteases in transcription modulation

Bacteria must respond rapidly not only to viral infections, but also to all kinds of environmental stresses, such as changes in temperature or nutrient levels. This could be readily achieved by degrading regulatory and key cellular proteins by ATP-dependent (also known as AAA+) proteases (Gur et al., 2011). After target recognition, these proteases aggressively shred proteins with little sequence specificity (Gur et al., 2011; Philipps-Wiemann, 2018). In the context of this thesis, I will briefly review examples of the ATP-dependent bacterial proteases and some well-studied examples of how the ATP-dependent proteases modulate a global cellular process – transcription.

ATP-dependent proteases and substrate recognition

Gram-negative bacteria encode five ATP-dependent proteases: ClpAP, ClpXP, FtsH, HslUV and Lon, while Gram-positive bacteria encode three additional proteases: ClpCP, ClpEP and proteasome (Gur et al., 2011). Since the most extensively characterized proteases are of *E. coli*, here I will focus on ATP-dependent proteases of Gram-negative bacteria. All ATP-dependent proteases share a common mechanism where an ATPase component unfolds substrates and translocates them into a proteolytic chamber of the protease component (Gur et al., 2011; Mahmoud and Chien, 2018) (Figure 1.18 A). ClpAP, ClpXP, HslUV are multi-component proteases as the ATPase and protease components are encoded by separate polypeptide chains. Meanwhile Lon and FtsH are single-component enzymes – protease and ATPase domains are part of a single polypeptide chain. Exceptionally, unlike the other proteases, FtsH is associated with the inner membrane through transmembrane helical segments (Figure 1.18 B).

Since proteolysis is an irreversible process, proteases must be strictly selective of their substrates. ATP-dependent proteases recognize their substrates by short substrate sequences called degradation tags or degrons (Baker and Sauer, 2006) (Figure 1.18B). The most extensively studied example of specific substrate recognition through degron is the SsrA degron recognition by ClpXP. It is a part of protein quality control mechanism, where *E. coli* stalled ribosome rescue system, SsrA–SmpB, adds SsrA degron of 11 amino acids (AANDENYALAA) in a process called *trans*-translation to the C-terminus of proteins stuck in stalled ribosomes (Keiler et al., 1996). C-terminal alanines of the SsrA degron are critical for its recognition by ClpX through three sets of loops that are located in the ClpX pore (Fei et al., 2020; Ghanbarpour et al., 2023). After degron recognition the substrate is directed to the proteolytic chamber of ClpP and the substrate is further translocated with very low sequence specificity (Barkow et al., 2009; Ghanbarpour et al., 2023). Most SsrA-tagged proteins are degraded by ClpXP, however other ATP dependent proteases ClpAP, Lon and FtsH can also contribute to degradation (Langklotz et al., 2012). There is variation in the SsrA-tag sequences among bacterial species, but these tags commonly end in LAA (Ahlawat and Morrison, 2009; Flynn et al., 2001). Even though ClpXP is capable of binding the SsrA-tagged proteins on its own, the adaptor protein SspB delivers and stimulates degradation of the SsrA-tagged substrates by ClpXP (Flynn et al., 2001; Wah et al., 2002). Furthermore, SspB also interacts with some other ClpXP substrates, for example the N-terminal fragment of RseA (N-RseA) (Levchenko et al., 2005). Surprisingly, SsrA degron and the C-terminal end of N-RseA do not share sequence similarity upstream from

terminal (L/V)AA motif and are bound by opposite SspB orientations (Mahmoud and Chien, 2018). Thus, adaptor proteins can accommodate multiple substrates.

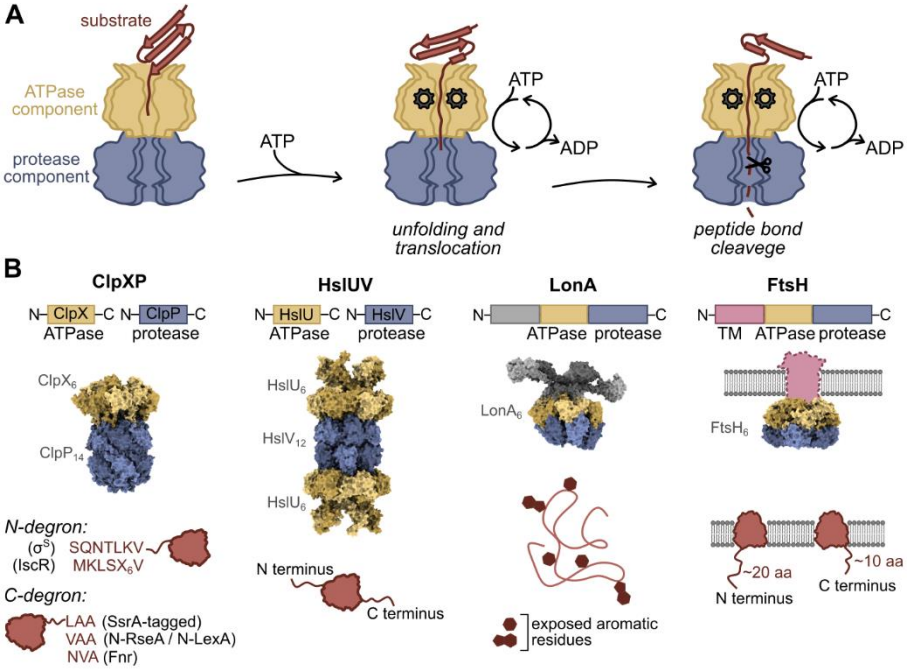


Figure 1.18. ATP-dependent proteases. (**A**) General mechanism of action of ATP-dependent proteases. Substrate is recognized by the ATPase component. ATP hydrolysis powers the unfolding and translocation of the substrate into the proteolytic chamber of the protease component, where peptide bonds are hydrolyzed. (**B**) Examples of ATP dependent proteases of Gram-negative bacteria. Organization of ATPase and protease components (top) and structures of ATP-dependent proteases (middle): *E. coli* ClpXP (PDB ID: 9C88), *E. coli* HslUV (PDB ID: 1G4B), *M. taiwanensis* LonA (PDB ID: 8K3Y) and cytosolic part of membrane associated *E. coli* FtsH (PDB ID: 7WI4). ClpXP consists of a tetradecameric proteolytic core (ClpP) capped with an ATPase hexamer (ClpX). Same applies to ClpAP (not shown). HslUV is composed of the dodecameric HslV protease core capped by the hexameric HslU ATPase components. FtsH and Lon proteases form hexamers. FtsH is associated with the inner membrane. Common degrons recognized by these proteases (bottom).

Other ATP-dependent proteases also require degrons for substrate binding, but there are some differences in the nature of degron sequences. For example, HslUV protease is activated at increased temperatures and requires degrons on both termini of the substrate protein (Baytshtok et al., 2021; Sundar et al., 2010). While the membrane-bound FtsH protease degrades membrane proteins with exposed unstructured N or C-terminus of ~20 or 10 amino acids, respectively (Bittner et al., 2017; Langklotz et al., 2012). FtsH with the help of chaperones like DnaK/J and GroES/EL as adaptor proteins can also degrade

cytoplasmic proteins carrying internal degrons (Langklotz et al., 2012). *E. coli* Lon protease, belonging to the LonA subfamily, recognizes sequences (either internal or terminal) rich in aromatic residues that become accessible in unfolded proteins but are hidden in most native states (Gur and Sauer, 2009). Therefore, Lon protease broadly recognizes misfolded proteins and is very important for protein quality control, since Lon can degrade about half of all *E. coli* proteins when they are misfolded (Gur et al., 2011) Intriguingly, Lon can be allosterically activated. The increase in misfolded proteins results in Lon activation, that degrades not only unfolded proteins but also fully active native regulatory proteins, such as replication initiator DnaA in *Caulobacter* (Jonas et al., 2013; Mahmoud and Chien, 2018). Furthermore, upon amino acid starvation, *E. coli* Lon protease has been shown to bind polyphosphate, which stimulates its activity in degrading specific substrates, such as ribosomal proteins (Kuroda et al., 2001; Mahmoud and Chien, 2018). Moreover, Lon can also be allosterically regulated by binding to DNA, therefore it is thought to react to DNA damage (Mahmoud and Chien, 2018).

Protease in transcription regulation

As mentioned, ATP-dependent proteases are involved in signal transduction pathways that regulate transcriptional outcomes. In general, protein synthesis is an energetically costly process, hence protein degradation can be seen as a wasteful way to control protein levels. However, compared to downregulation of protein synthesis at the transcriptional level, proteolysis provides a more rapid response required under stress conditions (Gur et al., 2011). Proteases may achieve transcription regulation by degrading or releasing alternative σ factors or degrading transcriptional regulators.

Some alternative σ factors are rapidly degraded under some conditions and stabilized under others to perform transcription initiation from the corresponding promoters. For example, the alternative σ factor σ^{32}, which controls the expression of a large number of proteins at elevated temperatures, is synthesized at normal temperatures, but its half-life is only ~1 min as it is rapidly degraded by FtsH protease with assistance of chaperones (Herman et al., 1995; Langklotz et al., 2012; Tatsuta et al., 2000). At elevated temperatures, the availability of chaperones becomes scarce as they are occupied by rescuing denatured proteins, leading to the accumulation of σ^{32}. Similarly, the levels of an alternative σ factor, σ^{S} (also known as RpoS), are kept low during exponential growth by ClpXP-mediated proteolysis, and the stabilization of σ^{S} during the stationary phase is achieved by inhibiting the adaptor protein RssB, which is essential for substrate recognition by ClpXP (Battesti et al., 2013; Micevski et al., 2015; Zhou et al., 2001).

Another example of transcription modulation achieved by proteolytic cleavage is the release of ECF σ factor, σ^E (Figure 1.19). This alternative σ factor is activated through the cascade of proteolytic cleavages and regulates the expression of the periplasmic protein quality control system (Ades, 2008; Dartigalongue et al., 2001; Mascher, 2023; Österberg et al., 2011). At normal conditions σ^E is bound by the cytoplasmic domain of the transmembrane anti-σ factor RseA and upon envelope stress, periplasmic protease DegS becomes activated and cleaves the periplasmic domain of RseA (Walsh et al., 2003). This in turn activates the membrane-embedded protease RseP to further cleave the transmembrane domain of RseA and releases the N-terminal cytoplasmic domain (still bound to σ^E) of RseA (N-RseA) from the membrane (Heinrich and Wiegert, 2009). After this cleavage the C-terminal degron sequence ending with VAA is formed on newly formed C-terminus of N-RseA (Flynn et al., 2004). Exposed degron targets N-RseA for degradation by ClpXP releasing σ^E from the complex.

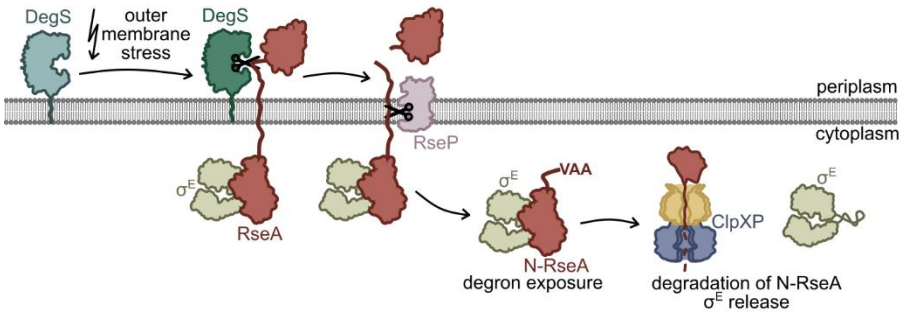


Figure 1.19. Release of the ECF σ factor, σ^E, by a cascade of proteolytic cleavages. Upon outer membrane stress, the membrane-associated periplasmic protease DegS is activated. Activated DegS cleaves the transmembrane anti-σ factor RseA at the periplasmic side. This cleavage is further sensed by the membrane-embedded protease RseP, which cleaves RseA releasing the N-terminal fragment of RseA (N-RseA) bound to σE from the membrane. After RseP cleavage the degron sequence is exposed on a newly formed N-RseA C-terminus. This degron is recognized by ClpXP and N-RseA is degraded in turn releasing σ^E.

Interestingly, the membrane associated anti-σ factor and ECF σ factor pair, DdvA-DdvS, has been observed to regulate an antiviral CBASS-CRISPR-Cas defense island in *Myxococcus xanthus*. (Bernal-Bernal et al., 2024, 2018). The DdvA-DdvS follows a pathway that is somewhat similar to that of RseA-σE: the periplasmic domain of an anti-σ factor, DdvA, adopts a TPR-CHAT protease–like architecture and, upon yet unknown viral infection cue, undergoes autoproteolysis, which in turn induces additional RseP-mediates intramembrane proteolysis that subsequently destabilizes the interaction of the

DdvA cytoplasmic domain with DdvS, releasing the σ factor (Bernal-Bernal et al., 2024).

Proteases also degrade regulatory proteins and thus affect transcription. For example, for the cell to progress through the cell cycle, a positive transcriptional regulator of the capsule synthesis, RcsA, is degraded by Lon protease (Gur et al., 2011). Moreover, degradation of the master regulator CtrA by adaptor proteins-assisted ClpXP is essential in *Caulobacter crescentus* for swarmer cells transition to replication-capable stalked cell type (Jenal and Fuchs, 1998; Mahmoud and Chien, 2018; Skerker and Laub, 2004). Proteolysis is also essential for the regulation of the SOS response – a bacteria response to DNA damage. Under normal conditions transcription repressor LexA inhibits the expression of the SOS response genes that are important for error prone repair processes (Butala et al., 2008). Upon DNA damage the self-cleavage of LexA is induced, which results in the formation of C-degron ending with VAA at the N-fragment of LexA targeting it for degradation by ClpXP (Pruteanu and Baker, 2009). Proteolysis is also important for the recovery from the SOS response. For example, following recovery from DNA damage, DNA repair proteins, which were expressed during SOS, such as UvrA and UmuD, as well as the cell division repressor SulA, are degraded by proteases, allowing the cell to begin dividing (Pruteanu and Baker, 2009).

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Chemicals

Chemicals used in this study were purchased from Merck, Roth, Fluka and Thermo Fisher Scientific of the highest purity grade available. Radiolabeled nucleotides α-^{32}P-ATP and γ-^{32}P-ATP were purchased from Revvity.

2.1.2. Commercial kits, proteins and dyes

Phusion High-Fidelity and DreamTaq DNA polymerases, "FastDigest" restriction endonucleases (Acc65I, AvrII and others), T4 polynucleotide kinase (PNK); T4 DNA ligase, FastAP thermosensitive alkaline phosphatase, RiboLock RNase inhibitor, 2× RNA Loading Dye, PageRuler Unstained Protein Ladder (protein sizes in kDa: 200, 150, 120, 100, 85, 70, 60, 50, 40, 30, 25, 20, 15, 10) were purchased from Thermo Fisher Scientific. Bovine serum albumin (BSA) used to supplement the reactions was obtained from Thermo Fisher Scientific, while BSA used for biolayer interferometry (BLI) experiments was obtained from Merck. NEBuilder HiFi DNA Assembly Master Mix and *E. coli* RNA Polymerase Core Enzyme were purchased from New England Biolabs.

"GeneJET PCR Purification Kit", "GeneJET Plasmid Miniprep Kit", "T7 High Yield Transcription Kit" and "GeneJET RNA Cleanup and Concentration Micro Kit" were purchased from Thermo Fisher Scientific.

Gel filtration Calibration Kit was obtained from GE Healthcare.

Nucleic acid dye "SYBR Gold" obtained from Invitrogen.

InstantBlue Coomassie Protein Stain was purchased from Abcam.

All these products were used according to the manufacturer's instructions.

2.1.3. Bacterial strains and media

E. coli strain DH5α [F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYAargF) U169, hsdR17(rK- mK+), λ–] was used for cloning.

E. coli strain BL21(DE3) [F– ompT gal dcm lon hsdSB(rB- mB-) λ(DE3)] was used for interference assay of *S. thermophilus* type III-A CRISPR-Cas system, expression of *S. thermophilus* type III-A CRISPR-Cas system for cA_n detection in cells, expression of StCsm6 isolated wt and mutant

CARF, HEPN, dCARF and dHEPN domains, expression of *Ca.* C. acidaminovorans Csm surveillance complex.

E. coli strain DH10B [F- endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-] was used for Csm6 expression.

E. coli strain ER2267 [(F′ proA+B+ lacIq Δ(lacZ)M15 zzf::mini-Tn10 (KanR)/ Δ(argF-lacZ)U169 glnV44 e14-(McrA-) rfbD1? recA1 relA1? endA1 spoT1? thi-1 Δ(mcrC-mrr)114::IS10] (New England Biolabs) was used for cultivation of MS2 phage.

E. coli NovaBlue (DE3) [endA1 hsdR17(rK12 – mK12 +) supE44 thi-1 recA1 gyrA96 relA1 lac (DE3) F'[proA+ B+ lacIq ZΔM15::Tn10] (TetR)] was used for expression of *S. thermophilus* type III-A CRISPR-Cas system for cA_n detection in cells infected with MS2 phage.

E. coli strain ER2566 [F- λ- fhuA2 [lon] ompT lacZ::T7 gene 1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr- 73::miniTn10-TetS)2 R(zgb-10::Tn10)(TetS) endA1 [dcm]] was used for expression of *Ca.* C. acidaminovorans proteins CCaCalpL, CCaCalpT and CCaCalpS, for plasmid interference assay of CCaCalpT-CalpS system and the investigation of $CCaCalpT_{24N}$ degron.

If not stated otherwise, *E. coli* bacteria were cultivated in LB medium (1 % peptone, 0.5 % yeast extract, 0.5 % NaCl in water, pH 7.0 (25 °C)) or on plated LB agar, which also contained 1.5 % agar-agar.

Time-course measurements of GFP degradation in *E. coli* cells were performed in M9 minimal medium (47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM NH_4Cl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$).

Trace elements for M9 minimal medium (31 µM $FeCl_3$, 130 µM EDTA, 6.2 µM $ZnCl_2$, 0.8 µM $CuSO_4$, 0.4 µM $CoCl_2$, 1.6 µM H_3BO_3, 0.1 µM $MnCl_2$).

2.1.4. Proteins and nucleic acids

Plasmids used in this study are listed in Appendices 1 and 2:

Appendix 1 – plasmids used in the study of cA_6 signaling regulation.

Appendix 2 – plasmids used in the study of tripartite effector system.

Information on used RNA molecules is listed in the Appendix 3 – RNAs used in this study.

2.1.5. Buffers and other solutions

Unless stated otherwise, pH is indicated at 25°C.

Chromatography buffers:

CHB 1: 20 mM Tris-HCl (pH 8.5), 0.5 M NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA

CHB 2: 20 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 8 mM 2-mercaptoethanol

CHB 3: 20 mM HEPES-KOH (pH 8.0), 0.5 M KCl, 8 mM 2-mercaptoethanol

CHB 4: 20 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 8 mM 2-mercaptoethanol

CHB 5: 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 8 mM 2-mercaptoethanol

SEC-MALS: 20 Tris-HCl (pH 8.0), 100 NaCl, 1 mM DTT

Protein storage buffers:

SB 1: 10 mM Tris-HCl (pH 8.5), 300 mM NaCl, 1 mM DTT, 0.1 mM EDTA, and 50% (v/v) glycerol

SB 2: 10 mM Tris–HCl (pH 8.0), 300 mM KCl, 1 mM DTT, 0.1 mM EDTA and 50% (v/v) glycerol

SB 3: 10 mM Tris-HCl (pH 7.5), 300 mM NaCl, 1 mM DTT, 0.1 mM EDTA and 50% (v/v) glycerol

SB 4: 10 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM DTT, 0.1 mM EDTA and 50% (v/v) glycerol

SB 5: 25 mM Tris-HCl (pH 7.5), 250 mM NaCl, 1 mM DTT and 50% (v/v) glycerol

HPLC-MS solvents:

A1: 10 mM ammonium acetate in water, pH 7.0

B1: acetonitrile

A2: 5 mM ammonium acetate in water, pH 7.0

B2 (5 mM ammonium acetate in methanol, pH 7.0)

A3 (1% formic acid in water)

B3 (1% formic acid in acetonitrile)

Reaction buffers for activity assays:

Reaction buffer Y: 33 mM Tris-acetate (pH 7.6 at 37°C), 66 mM KCH_3COO

Reaction buffer H: 20 mM HEPES-KOH (pH 7.5), 50 mM KCl

Reaction buffer T: 40 mM Tris, 20 mM acetic acid (pH 8.1), 1 mM EDTA

Electrophoresis buffers, loading dyes and gels:

2× RNA Gel Loading Dye (Thermo Fisher Scientific): 95 % formamide 0.025 % SDS 0.025 % bromophenol blue 0.025 % xylene cyanol FF 0.025 % ethidium bromide 0.5 mM EDTA

0.5× TBE buffer: 44.5 mM Tris-HCl, 44.5 mM H_3BO_3 (pH 8.3), 1 mM EDTA

Denaturing 15% PAG: 8 M urea, 15 % acrylamide/N,N'-methylenebisacrylamide (w/w; 29:1) in 0.5× TBE buffer

Denaturing 24% PAG: 6 M urea, 24 % acrylamide/N,N'-methylenebisacrylamide (w/w; 19:1) in 0.5× TBE buffer

Denaturing 30% (w/v; 19:1) PAG: 6 M urea, 30 % acrylamide/N,N'-methylenebisacrylamide (w/w; 19:1) in 0.5× TBE buffer

4× SDS loading dye: 200 mM Tris, 400 mM DTT, 8% (w/v) SDS, 40% (v/v) glycerol and 6 mM bromophenol blue

Protein electrophoresis buffer: 25 mM Tris-HCl (pH 8.3), 190 mM glycine, 0.1 % SDS.

4 % concentrating PAG: 4 % acrylamide/N,N'-methylenebisacrylamide (w/w; 37.5:1), 125 mM Tris-HCl (pH 6.8), 0,1 % SDS;

12 % fractionating PAG: 12 % acrylamide/N,N'-methylenebisacryl-amide (w/w; 37.5:1), 375 mM Tris-HCl (pH 8.8), 0,1 % SDS

15 % fractionating PAG: 15 % acrylamide/N,N'-methylenebisacryl-amide (w/w; 37.5:1), 375 mM Tris-HCl (pH 8.8), 0,1 % SDS

2.1.6. Other items

Magnetic beads "Dynabeads™ His-Tag Isolation and Pulldown" were purchased from Invitrogen. Copper 300 mesh R1.2/1.3 holey carbon grids were obtained from Quantifoil. Octet Ni-NTA Biosensors were purchased from Sartorius.

2.2. Methods

2.2.1. Homology search

The HHPred tool (Zimmermann et al., 2018) was used to search for homology of *Ca.* C. acidaminovorans proteins of interest (NCBI Reference Sequences: WP_044278915.1 (CalpS), WP_015424588.1 (CalpT), WP_015424587.1 (CalpL), WP_015424586.1 (Cad1)) in structural and domain databases: PDB, Pfam-A, COG_KOG, NCBI (CD). The search was performed with default settings.

The DaliSearch tool (Holm et al., 2023) was used to search for remote structural homologs of the CCaCalpL SAVED domain and full-length CCaCalpL. The SAVED domain (219-503 aa) or SAVED C-subdomain (357-504 aa) from the solved cA_4-bound CCaCalpL filament structure and a single CCaCalpL subunit from the A_4p-bound CCaCalpL filament structure was used as a query.

2.2.2. Structure prediction

The CCaCalpS protein and the CCaCalpT-CalpS heterodimer were modeled with AlphaFold2 (Evans et al., 2022; Jumper et al., 2021) under the ColabFold (Mirdita et al., 2022) framework using default parameters. The SsCalpT-CalpS heterodimer and the complex of CalpL filament, AAAA RNA and CCaCalpT-CalpS heterodimer were modeled with AlphaFold3 (Abramson et al., 2024).

2.2.3. Sequences alignments

Alignment of Csm6 sequences

A set of 10 Csm6 protein sequences (NCBI Reference Sequences: WP_014621552.1, WP_014621551.1, WP_007208953.1, EZS13226.1, WP_002502662.1, AGA14268.1, WP_044359384.1, WP_133445588.1, WP_089831386.1, WP_123144061.1) were aligned using MUSCLE (Edgar, 2004) and alignment was visualized using ESPript 3.0 (Robert and Gouet, 2014).

Alignment of CalpL, CalpT, CalpS sequences

Pairwise sequence alignment of CalpL (WP_012459369.1), CalpT (WP_012459368.1) and CalpS (WP_012459367.1) proteins of *Sulfurihydrogenibium* spp. YO3AOP1 (Rouillon et al., 2023) and *Ca.* C. acidaminovorans (this study) was performed using EBI EMBOSS Needle (Madeira et al., 2024) with default settings.

Alignment of SAVED4 sequences

A set of 73 recently classified SAVED4 clade (Makarova et al., 2020a) sequences were used for sequence analysis. Only sequences longer than 450 amino acids and containing SAVED4 fused to Lon protease domain were used for the analysis. A non-redundant set of 16 Lon-SAVED4 was generated by MMseqs2 (Steinegger and Söding, 2017) and these sequences were aligned using MUSCLE (Edgar, 2004). Alignment was visualized using ESPript 3.0 (Robert and Gouet, 2014).

2.2.4. Cloning and mutagenesis

General considerations: primers were purchased from Metabion or IDT; PCRs were performed using Phusion HF DNA polymerase; PCR products were purified using “GeneJET PCR Purification Kit”; *E. coli* DH5α strain was used for cloning and plasmid amplification; colony PCRs were performed

using DreamTaq polymerase; plasmids were purified using "GeneJET Plasmid Miniprep Kit".

Production of plasmids for the study of cA_6 signaling regulation.

Plasmids and derivatives of plasmids used in previous studies (Kazlauskiene et al., 2017, 2016; Mogila et al., 2019; Tamulaitis et al., 2014) (Cas/Csm, pCas/Csm_dCsm3, pCas/Csm_dCsm3-dHDCas10, pCas/Csm_dCsm3-dHDCas10_ΔCsm6′ΔCsm6, pCas/Csm_dCsm3-dHD Cas10_dPalm-Cas10_ΔCsm6′ΔCsm6, pCRISPR_Tc, pCRISPR_S3, pCRISPR_MS2, pTarget_Tc, pTarget_Tc^{mut}, pTarget_ctrl, $pCsm2^{N}$-tag, pCsm6, pCsm6′, pdHEPN-Csm6) were supplied by Dr. G. Tamulaitis, Dr. M. Kazlauskiene and Dr. I. Mogila.

The pHEPN expression plasmid encoding the HEPN domain of StCsm6 (172–428 codons) was constructed by cloning PCR amplified *csm6* gene into pET-Duet-1 derivative containing an N-terminal His_6-tag and a TEV protease cleavage site coding sequence via Acc65I and AvrII sites. An additional sequence encoding four glycine residues was introduced between the TEV site and *csm6* by site-directed mutagenesis resulting in the intermediate plasmid. Another round of site-directed mutagenesis was used to omit 1–171 codons from *csm6* in the intermediate plasmid.

pdHEPN plasmid

containing R371A, H376A mutations in HEPN domain was generated by site-directed mutagenesis using pHEPN as the template.

The pCARF plasmid was generated by the Gibson assembly method. PCR amplified fragment containing coding sequences of the TEV protease cleavage site, four glycines coding sequence, CARF domain (1–169 codons of StCsm6) and PCR fragment obtained from pETDuet-1 derivative containing an N-terminal His_{10}-tag fused with maltose binding protein (MBP) were assembled.

To generate pdCARF and pdCARF-StCsm6 plasmids containing D12A mutation in *csm6* gene, mutation carrying primers and pCARF and pCsm6 plasmid as the template, respectively, were used to obtain PCR fragments. These PCR fragments were assembled by Gibson assembly method.

Sequence integrity of all constructed plasmids was confirmed by Sanger sequencing.

Production of plasmids for the study of tripartite effector system

The coding sequences of the *Candidatus* Cloacimonas acidaminovorans strain Evry proteins of interest and the repeat sequence, were obtained from the NCBI reference genome sequence NC_020449.1. All gene sequences of the proteins have been optimized for expression in *E. coli*.

Synthetic gene fragments encoding *Ca.* C. acidaminovorans *cas10* (fused to TEV-cleavable His_6-StrepII-His_6 N-tag), *csm2*, *csm3*, *csm4* and *csm5* genes were purchased from Twist Bioscience and cloned into the pCDFDuet-1 plasmid via Gibson assembly to obtain pCCaCsm plasmid. The pCRISPR plasmid encoding the minimal CRISPR region (repeat-spacer-repeat) and *cas6* gene was constructed by cloning a synthetic gene fragment into pETDuet-1 backbone using the Gibson assembly method. The gene fragment encoding the minimal CRISPR region was purchased from GenScript Biotech, and the gene fragment encoding *cas6* was purchased from Twist Bioscience.

Production of pCCaCalpL^{N}-tag and pCCaCalpL^{C}-tag was ordered from Twist Bioscience. These plasmids were produced by synthesizing *calpL* gene fused with TEV-cleavable His_6-StrepII-His_6 tag at the N- or C-terminal and cloning it into pBAD/HisA vector (Invitrogen). To generate plasmids encoding mutant versions (S154A; K330A; R358A; K361A; S359A; H396A; H476A and R358E-K361E) of *calpL*, site-directed mutagenesis was used with primers carrying the desired mutations.

Synthetic gene fragments encoding *calpS* and *calpT* genes were purchased from Twist Bioscience and were used as PCR templates to generate desired fragments for further cloning. The Gibson assembly strategy was used to clone *calpS* into pETDuet-1 vector derivative fusing the protein with TEV-cleavable His_6-StrepII-His_6 tag at the C-terminus (pETCCaCalpS-^{C}Tag). Similarly, the pET-CCaCalpT-^{C}Tag, a pETDuet-1 vector encoding *calpT* fused to a His_6-StrepII-His_6 tag at the C-terminus, was generated. The Gibson assembly was also used to generate pCDF-CCaCalpT, a pCDFDuet-1 vector encoding *calpT* without the tag. pSAVED-^{C}Tag encoding a C-tagged *calpL* gene spanning 209-506 codons (with an additional I209A mutation) was generated from a plasmid pCalpL-^{C}Tag using site-directed mutagenesis to remove the corresponding parts of the *calpL* gene.

For the *E. coli* survivability assay pCCaCalpS, pCCaCalpT, pCCaCalpT$_{24N}$, pCCaCalpT$_{11C}$ and pCCaCalpT$_{split}$ vectors were generated. The pCCaCalpS vector was produced by cloning the *calpS* gene into the pCDFDuet-1 vector using the Gibson assembly method. The pCCaCalpT vector encodes the entire *calpT* gene spanning 1-299 codons and was generated by site-directed mutagenesis of the pET-CCaCalpT-^{C}Tag vector, excluding the segment encoding the affinity tag. The pCalpT$_{24N}$ vector encodes the first 204 codons of the *calpT* gene, while the pCalpT$_{11C}$ vector encodes codons 205-299 (with an additional methionine for initiation) of the *calpT* gene. These vectors were generated from the pCCaCalpT plasmid using site-directed mutagenesis to remove the corresponding parts of the *calpT* gene. The pCCaCalpT$_{split}$ plasmid encodes codons 1-204 (CalpT$_{24N}$) and 205-299

($CalpT_{11C}$) of the *calpT* gene, separated by a 45 nt spacer encoding a ribosome binding site and an additional methionine as an initiation codon for the expression of the $CCaCalpT_{11C}$. To produce the $pCCaCalpT_{split}$ plasmid, a whole plasmid PCR fragment was generated from the pCCaCalpT template and assembled with a 77 nt bridging oligo using the Gibson assembly method.

The $pCalpT_{24N}$(mut) plasmid encodes the first 204 codons of *calpT* with an A204D mutation at the end of the potential degron sequence and was constructed by site-directed mutagenesis using $pCCaCalpT_{24N}$ as a template. pGFP was constructed by cloning the GFPmut1 (Cormack et al., 1996) gene into the pETDuet-1 vector via the Gibson assembly. Subsequently, pGFP was used as a template in a site-directed mutagenesis strategy to generate plasmids encoding GFP fused to the degron sequences: (i) wt-$CCaCalpT_{24N}$ C-degron PETLQLAA, (ii) mutant $CCaCalpT_{24N}$ C-degron PETLQLAD, or (iii) partial ssrA C-degron ENYALAA, generating pGFP-(wt)$CalpT_{24N}$-degron, pGFP(mut)$CalpT_{24N}$-degron and pGFP-ssrA-degron plasmids, respectively.

Sequence integrity of all constructed plasmids was confirmed by either Sanger sequencing or whole plasmid sequencing (SeqVision, Lithuania).

2.2.5. Preparation of bacterial lysates for cA_n detection

E. coli BL21 (DE3) cells were first transformed with the two plasmids encoding the *S. thermophilus* 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_Tc^{mut} or pTarget_ctrl 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 StCsm complex 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 synthetase motif in the Palm domain and the Tc-targeting CRISPR region was verified by sequencing the corresponding PCR products.

For phage-induced cA_n assays, the MS2 (ATCC® 15597B1) phage sample was obtained from Dr. Andris Dislers (Latvian Biomedical Research and Study Center) as a kind gift. Prior to the experiment phage was propagated in *E. coli* ER2267 strain according to the manufacturer's protocol of MS2 propagation on soft agar covered double layer agar plates. Plasmids

pCRISPR_MS2 or pCRISPR_S3 and pCas/Csm_dCsm3_dHDCas10_ ΔCsm6′ΔCsm6 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 an additional 30 minutes the cells were infected with MS2 phage suspension. The cells were collected by centrifugation at 16 h post induction.

2.2.6. Protein expression and purification

Proteins were expressed in indicated *E. coli* strains (Table 2.1). *E. coli* cells were transformed or co-transformed with relevant plasmids (Table 2.1, Appendices 1 and 2) and were grown in liquid LB medium supplemented with appropriate antibiotics at 37°C with 200 rpm shaking. At the mid-log phase (OD_{600} 0.6 to 0.8) expression was induced using the relevant inductor and cells were further cultured at indicated temperature for indicated time (Table 2.1). Cells were harvested by centrifugation and pellets were resuspended in appropriate chromatography buffer (Table 2.1) supplemented with 0.1 mM phenylmethylsulfonyl fluoride (PMSF), except for CCaCalpL for which protease inhibitor was not used, and disrupted by sonication. Soluble fractions were loaded on relevant columns (Table 2.1). All protein purifications were performed using ÄKAT avant chromatography system (Cytiva). All proteins were dialyzed against the relevant storage buffer and stored at -20°C (Table 2.1).

2.2.7. Determination of the oligomeric state

SEC analysis of isolated CARF and HEPN domains

Analytical SEC was carried out at room temperature on an ÄKTA FPLC system (GE Healthcare) using a Superdex 75 10/300 GL column (GE Healthcare), preequilibrated with the chromatography buffer 3 (CHB 3). CARF and HEPN 100 µl loading samples were prepared by diluting proteins with the CHB 3 buffer to the final loading concentrations of 0.7–0.8 mg/ml. 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).

SEC-MALS analysis of CCaCalpL-CalpT-CalpS components

SEC-MALS analysis was performed on the HPLC system (Waters Breeze) combined with a miniDAWN MALS detector (Wyatt) and an Optilab

refractive index detector (Wyatt). CCaCalpL, CCaCalpT-CalpS, CCaCalpT or CCaCalpS protein samples were diluted with the SEC-MALS buffer to a final concentration of 0.3 mg/ml. 200 mL of diluted samples were loaded to a Superdex 200 Increase 10/300 GL column (Cytiva) equilibrated with SEC-MALS buffer. Data acquisition and evaluation were performed using ASTRA 7 software (Wyatt).

2.2.8. Preparation of RNAs

The S3/1, NS and CCaTarget RNAs were obtained by *in vitro* transcription using TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific). First, DNA fragments containing a T7 promoter in front of the desired RNA sequence were produced either by PCR using pUC18_S3/1 (Tamulaitis et al., 2014) or by annealing appropriate complementary oligodeoxynucleotides. Then, PCR fragments were purified using "GeneJET PCR Purification Kit" (Thermo Fisher Scientific) and subjected to *in vitro* transcription reaction, while assembled DNA oligoduplexes were used without additional purification. *In vitro* transcription was performed according to the manufacturer's instructions. Synthesized RNAs were purified using "GeneJET RNA Cleanup and Concentration Micro Kit" according to manufacturer's instructions.

NS RNA substrate used for StCsm6 RNase activity assay was dephosphorylated using FastAP thermosensitive alkaline phosphatase (Thermo Fisher Scientific) and ^{32}P 5′-labeled with γ-^{32}P-ATP (Perkin Elmer/Revvity) using PNK (Thermo Scientific) according to manufacturer's instructions.

2.2.9. cA_n synthesis *in vitro*

cA_n synthesis using StCsm complex

The cA_n synthesis reactions were initiated by adding 10 mM $CoCl_2$ or 1 mM $Mg(CH_3COO)_2$ into a mix of 0.2 µM StCsm complex, 0.2 or 10 µM of target RNA S3/1 and 500 µM ATP in the reaction buffer Y or H both supplemented with 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.

cA_n synthesis using CCaCsm complex

The mixture of 0.2 µM of CCaCsm complex, 10 µM of target RNA CCaTarget or NS RNA and 60 mM of ATP in the reaction buffer H, supplemented with 0.1 mg/ml BSA and 1 mM of $Mg(CH_3COO)_2$, was incubat-

Table 2.1. Conditions of protein expression and purification

Protein	Expression				Purification				Storage buffer
	E. coli strain	Plasmid	Inductor	Grown post-induction	Step	Column / enzyme	Chrom. buffer	Elution	
StCsm complex (wt and mutant)	BL21 (DE3)	pCas/Csm / pCas/Csm-dCsm3 pCRISPR_S3 and pCsm2-^{N}Tag	1 mM IPTG	4 h at 37°C	1	StrepTrapHP	CHB 1	25 mM d-desthiobiotin	SB 1
					2	Superdex 200 10/300 GL	CHB 1	CHB 1	
StCsm6′ and StCsm6 (wt and mutant)	DH10B	pCsm6 / pdCARF-Csm6 / pdHEPN-Csm6 / pCsm6′	0.2% L-(+)-arabinose	4 h at 37°C	1	HisTrapHP	CHB 2	5–500 mM imidazole gradient	SB 2
					2	StrepTrapHP	CHB 2	25 mM d-desthiobiotin	
CARF domain / HEPN domain (wt and mutant)	BL21 (DE3)	pCARF / pdCARF / pHEPN / pdHEPN	1 mM IPTG	~16 h at 16°C	1	HisTrapHP	CHB 3	5–500 mM imidazole gradient	SB 2
					2	TEV (1:50 (w/w) TEV:target)			
					3	HisTrapHP	CHB 3	Target protein in flow-through	
CCaCsm complex	BL21 (DE3)	pCCaCsm and pCCaCRISPR	0.5 mM IPTG	~16 h at 16°C	1	StrepTrapXT	CHB 4	50 mM D(+)-biotin	SB 3
					2	TEV (1:25 (w/w) TEV:target)			
					3	HiTrap Heparin	CHB 5	50–1000 mM NaCl gradient	
CCaCalpL N-taged	DH10B	pCCaCalpL-^{N}Tag	0.2% L-(+)-arabinose	4 h at 37°C	1	HisTrapHP	CHB 2	5–500 mM imidazole gradient	SB 4
					2	StrepTrapXT	CHB 2	50 mM D(+)-biotin	
CCaCalpL (wt ant mutants)	ER2566	pCCaCalpL-^{C}Tag / p(S154A / K330A / R358A / K361A / S395A / H396A / H476A / R358E-K361E)CCaCalpL-^{C}Tag	0.2% L-(+)-arabinose	4 h at 37°C	1	HisTrapHP	CHB 2	5–500 mM imidazole gradient	SB 4
					2	StrepTrapXT	CHB 2	50 mM D(+)-biotin	
					3	TEV (1:25 (w/w) TEV:target)			
					4	HisTrapHP	CHB 2	Target protein in flow-through	
CCaCalpT-CalpS heterodimer / CCaCalpT	ER2566	pETCCaCalpS-^{C}Tag and pCDF-CCaCalpT / pET-CalpT-^{C}Tag	0.5 mM IPTG	4 h at 37°C	1	HisTrapHP	CHB 2	5–500 mM imidazole gradient	SB 5
					2	StrepTrapXT	CHB 2	50 mM D(+)-biotin	
					3	Superdex 200 10/300 GL	CHB 2	CHB 5	
CCaCalpS	ER2566	pET-CalpS-^{C}Tag	0.5 mM IPTG	4 h at 37°C	1	HisTrapHP	CHB 2	5–500 mM imidazole gradient	SB 5
					2	StrepTrapXT	CHB 2	50 mM D(+)-biotin	

ed at 37°C for 75 min. The reaction was stopped by freezing the sample. The reaction products were subjected to HPLC-MS analysis.

Production of radiolabeled cA_6

The α^{32}P-labeled cA_6 for subsequent direct visualization of its hydrolysis was prepared by mixing 0.2 µM StCsm complex with 10 µM target RNA S3/1, 1 mM $Mg(CH_3COO)_2$, 0.5 mM ATP and 1 µM α-^{32}P-ATP in the reaction buffer H and incubating at 37°C for ~16 h. Synthesis products were separated by denaturing PAGE (24% gel) and cA_6 was purified from the gel by phenol extraction and ethanol precipitation.

Production of radiolabeled cA_4

The α^{32}P-labeled cA_4 for subsequent direct visualization of its hydrolysis was prepared by mixing 0.2 µM CCaCsm complex, 10 µM CCaTarget RNA, 1 mM $Mg(CH_3COO)_2$, 0.2 mM ATP and 1 µM α^{32}P-ATP in the reaction buffer H and incubating at 37°C for 4 hours. Synthesis products were separated by denaturing PAGE (24% gel) and cA_4 was purified from the gel by phenol extraction and ethanol precipitation.

2.2.10. Ring nuclease assay

For TtCsm6

TtCsm6 (Kazlauskiene et al., 2017) at the indicated TtCsm6 concentrations was incubated with 20 µM of synthetic cA_4 (Biolog) or cA_6 (Biolog) in the reaction buffer Y supplemented with 0.1 mg/ml BSA (Thermo Fisher Scientific), and 0.5 U/ml RiboLock RNase inhibitor (Thermo Fisher Scientific) for 120 min. Reactions were stopped by freezing and subjected to HPLC-MS analysis.

For StCsm6′, StCsm6 or isolated CARF and HEPN domains

Unless stated otherwise, cA_n hydrolysis by StCsm6, StCsm6′, CARF domain or HEPN domain was conducted in the reaction buffer T supplemented with 0.5 U/µl RiboLock RNase Inhibitor (Thermo Fisher Scientific) and contained 20 µM of synthetic cA_6 (Biolog), synthetic cA_4 (Biolog), linear A_6>p (ChemGenes) or ~30 µM of StCsm-produced cA_n 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 cA_6, 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 α^{32}P-cA_6. Multiple turnover reactions additionally

contained 10 µM of unlabeled cA_6 (Biolog). Reactions were started by adding an indicated amount (0-10 µM) of protein and were carried out at 37°C. aliquots were taken at indicated time points and reactions were quenched by mixing with equal amount of 2× RNA Gel Loading Dye (Thermo Fisher Scientific). Reaction products were analyzed by denaturing PAGE and visualized by autoradiography.

The analysis of cA_6 cleavage reaction products by denaturing PAGE was performed using 24% denaturing gels of 16.5 cm × 22.0 cm format. First a pre-electrophoresis was performed for 30 min at 600V, then samples were loaded on the gel and electrophoresis was carried out at 600V for 4 hours. Both stages were performed at room temperature. Gels were subsequently visualized by autoradiography.

For single turnover reactions, k_{obs} were determined by fitting a single exponential to the substrate depletion data. For k_{cat} calculation, the reaction rates were determined from the linear parts of the reaction progress curves by linear regression.

For CCaCalpL

The cA_n hydrolysis assay was performed in the reaction buffer Y supplemented with 1 mM DTT, 0.1 mg/ml BSA (Thermo Fisher Scientific), and 0.5 U/ml RiboLock RNase inhibitor (Thermo Fisher Scientific) containing 10 µM of synthetic cA_3 (Biolog), synthetic cA_4 (Biolog), synthetic cA_6 (Biolog), linear A_4>p (ChemGenes), or linear A_4p (Metabion). Reactions were initiated by adding 1 µM of protein (wt or mutant versions of CCaCalpL) and incubated at 37°C for the indicated time. Reactions were stopped by freezing the sample. For identification of the cleavage products formed in a time course experiment of cA_4 cleavage, aliquots were taken at the indicated time points and the reactions were quenched by heating at 85°C for 10 min. The reaction products were subjected to HPLC-MS analysis.

To directly visualize the hydrolysis of cA_4, reactions were performed in the reaction buffer Y supplemented with 1 mM DTT, 0.1 mg/ml BSA (Thermo Fisher Scientific), and 0.5 U/ml RiboLock RNase inhibitor (Thermo Fisher Scientific) and containing 50 nM CCaCsm-produced radiolabeled α^{32}P-cA_4. Multiple turnover reactions additionally contained 10 µM of unlabeled cA_4 (Biolog). Single and multiple turnover reactions were started with the addition of 1 µM protein (wt or mutant CCaCalpL or isolated SAVED domain) and performed at 37°C. At the indicated time points, 5 ml aliquots were taken and the reactions were quenched by mixing with an equal amount of 2× RNA Loading Dye (Thermo Fisher Scientific) at 85°C. The reaction products were subjected to analysis by denaturing PAGE.

Analysis by denaturing PAGE was performed using 30% denaturing gel of 21 cm × 55 cm. First, a pre-electrophoresis was performed for 30 min at 1500V, then samples were loaded on the gel and electrophoresis was carried out at 2500V for 14 hours. Both stages were performed at 25°C. Gels were subsequently visualized by autoradiography.

Non-linear regression analysis of multiple-turnover cA_4 cleavage reaction data was performed using Kyplot 2.0 software (Yoshioka, 2002). The reaction course was approximated by simultaneous fitting equations (1-4) that describe sequential conversion of cA_4 to A_4>p, A_4p and A_2>p + A_2p over time with the rate constants k_1, k_2 and k_3, respectively, to experimental data:

$$[cA_4] = [cA_4]_0 \times e^{(-k_1 \times t)} \quad (1)$$

$$[A_{4>}p] = [cA_4]_0 \times \left[\frac{k_1}{(k_1 - k_2)}\right] \times \left[\left(e^{(-k_2 \times t)} - e^{(-k_1 \times t)}\right)\right] \quad (2)$$

$$[A_4p] = [cA_4]_0 \times k_1 \times \frac{k_2}{[(k_1 - k_2) \times (k_1 - k_3) \times (k_2 - k_3)]} \times \left[(k_2 - k_3) \times e^{(-k_1 \times t)} - (k_1 - k_3) \times e^{(-k_2 \times t)} + (k_1 - k_2) \times e^{(-k_3 \times t)}\right] \quad (3)$$

$$[A_{2>}p] + [A_2p] = [cA_4]_0 - [cA_4] - [A_{4>}p] - [A_4p] \quad (4)$$

The observed turnover rates k_{obs1}, k_{obs2} and k_{obs3} for the three reaction stages were then calculated by multiplying k_1, k_2 and k_3 by the total substrate concentration $[cA_4]_0$ (10 µM) and dividing by the total enzyme concentration (1 µM) used in the reaction. The determined cA_4, A_4>p and A_4p turnover rates are presented as the mean ± 1 standard deviation of k_{obs1}, k_{obs2} and k_{obs3} values determined from three independent experiments.

2.2.11. HPLC-MS analysis

Detection of cA_n in bacterial lysates

cA_n-containing fraction was extracted from 100 mg (±5%) bacterial pellet (Chapter 2.2.5.) 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 A1 just before analysis. Each sample (2–5 µl) was injected using a Vanquish Horizon UHPLC (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 A1 and solvent B1 (acetonitrile): 97% A1 from 0–2 min, 97–60% A1 from 2–8 min, 60–10% A1 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 cA_n were annotated based on accurate mass, retention time from synthesized standards and fragment ions. The molar amount of cA_n was estimated based on comparison to standards of known quantity. The molar concentration of cA_n was calculated assuming that each *E. coli* cell is 1 fl in volume and its wet weight is 1 pg (Sajed et al., 2016).

Analysis of cA_n synthesis and 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 DiscoveryÒHS C18 column (7.5 cm 3 2.1 mm, 3 mm). Elution was performed with a linear gradient of solvents A2 and B2 at a flow rate of 0.3 ml/min at 30°C as follows: 0-3 min, 0% B2; 3-23 min, 15% B2; 23-25 min, 40% B2, 25-26 min 80% B2, 26-29 min 100% B2. Ionization capillary voltage was set to 5000 V, fragmentor to - 150 V. Molecules were annotated based on accurate mass, retention time of synthesized standards and fragment ions.

Analysis of CCaCalpT cleavage products

Samples were analyzed on an integrated HPLC/ESI-MS system (Agilent 1290 Infinity) equipped with a Poroshell 300SB-C8 column (2.1x75 mm, 5 mm) by elution with a linear gradient of solvents A3 and B3 at a flow rate of 0.4 ml/min at 30°C as follows: 0-1 min, 2% B3; 1-6 min, 2-98% B3; 6-7 min, 98% B3; 7-9 min, 98-2% B3; 9-10 min, 2% B3. High resolution mass spectra of protein products were acquired on an Agilent Q-TOF 6520 mass analyzer (100-3200 m/z range, positive ionization mode). Results were analyzed using Agilent MassHunter Qualitative Analysis B.05.00 software.

2.2.12. RNA hydrolysis assays

RNA hydrolysis assay

RNA hydrolysis reactions by StCsm6′, StCsm6 and HEPN domain 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 NS RNA and 0–1 μM cA_6 (Biolog) or 0–100 nM cA_5 (synthesized and purified by Dr. M. Kazlauskiene and Dr. A. Šilanskas). 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% gel) and visualized by autoradiography.

RNA hydrolysis competition assay

Reactions were performed at 25°C and contained 10 nM of 5′-^{32}P-radiolabeled NS RNA and 0-500 μM of cA_6 in the reaction buffer Y with 0.5 U/μl RiboLock RNase Inhibitor (Thermo Fisher Scientific) and 0.1 mg/mL BSA (Thermo Fisher Scientific). Reactions were initiated by addition of StCsm6 to the final concentration of 10 nM. Aliquots of the samples were collected at indicated time intervals and quenched by mixing with 2× RNA Gel Loading Dye (Thermo Fisher Scientific). The reaction products were separated by denaturing PAGE (15% gel) and visualized by autoradiography.

RNA hydrolysis constants k_{obs} were determined by fitting a single exponential to the substrate depletion data.

2.2.13. Protease assay

Protease reactions were performed by pre-incubating 5 μM of CCaCalpL (when CCaCalpL mutant was combined in the oligomerization assay, 2.5 uM of each mutant was used) and 5 μM of CCaCalpT-CalpS dimer (or CCaCalpT monomer) in the reaction buffer Y supplemented with 2 mM DTT at 37°C for 5 min. The reaction was started by the addition of 25 μM (unless otherwise stated) of cA_3, cA_4, A_4>p or A_4p. Linear A_2>p and A_2p mix was prepared by incubating cA_4 with wt-CCaCalpL for 4 hours and stopping the reaction by heating at 85°C for 10 minutes and removing proteins by centrifugation. The supernatant containing A_2>p and A_2p was used in the protease reaction. Samples were incubated at 37°C for 120 min (or other time point if specified otherwise) and the reaction was stopped by mixing with 4× SDS loading dye. Reaction products were separated by SDS-PAGE (15% gel) and visualized by Coomassie staining. Gels were imaged by GelDoc Go System (Bio-Rad) or using the densitometry scanning mode on an Amersham Typhoon Biomolecular Imager (Cytiva).

2.2.14. Cryo-EM

Sample preparation

The CCaCalpL complexes with oligoadenylates were prepared by mixing 10.5 mM (0.66 mg/ml) of the N-tagged wt-CCaCalpL with 30 µM of cA_4, A_4>p or A_4p in the reaction buffer Y supplemented with 1 mM DTT and 2.5% glycerol just before applying on the freshly glow-discharged copper 300 mesh R1.2/1.3 holey carbon grids (Quantifoil), in a Vitrobot Mark IV (FEI) at 4°C with a waiting time of 0 s and a blotting time of 5 s under 95% humidity conditions. Grids were plunge-frozen in liquid ethane cooled at liquid nitrogen temperature. A_4>p and A_4p samples were prepared on holey carbon grids overlaid with graphene oxide. The graphene oxide coating was performed as described in (Bokori-Brown et al., 2016). Briefly, the glow-discharged (60 s, 50 mA) 1.2/1.3 Cu 300 mesh grids (Quantifoil) were coated with graphene oxide by applying 3 mL of aggregate-free supernatant of 10-fold diluted graphene oxide dispersion (Merck; 2 mg/mL) on the carbon-coated side of the grid and incubating for 1 min. To remove the excess of graphene oxide, the grids were blotted with filter paper, washed three times with water, and dried on filter paper.

Data collection and image processing

The cryo-EM data for the wt-CCaCalpL with bound oligoadenylates were collected using a Glacios microscope (Thermo Fisher Scientific), running at 200 kV and equipped with a Falcon 3EC Direct Electron Detector in the electron counting mode (Vilnius University). Images were recorded with EPU (v.3.2) at a nominal magnification of 392,000, corresponding to a calibrated pixel size of 1.10 Å per pixel, using an exposure of 0.80 e/Å^2 s^{-1}, in 30 frames and a final dose of 29.7 e/Å^2, over a defocus range of -1.0 to -2.0 µm. Patch motion correction, CTF estimation, micrograph curation, blob picking, and particle extraction were performed in real-time in CryoSPARC Live (v.4.2.1) (Punjani et al., 2020, 2017). Further data processing was performed using standard CryoSPARC (v.4.2.1) (Punjani et al., 2020, 2017).

2,604,122 particles of the wt-CCaCalpL-A_4p complex were extracted (box size 240 pixels) from 3,127 accepted micrographs. After 2D classification, the selected particles (898,101) were subjected to heterogeneous refinement using four volumes obtained from an ab-initio reconstruction. Class 1 possessing higher FSC resolution (4.22 Å, 488,970 pct) was further subjected to non-uniform refinement followed by 3D classification into 3 classes. After 3D classification, particles from one class (175,828) were used for the final reconstruction using non-uniform refinement, local refinement, reference-

based motion correction, followed by the final rounds of non-uniform and local refinement.

1,426,639 particles of the wt-CCaCalpL-cA_4 complex were extracted (box size 240 pixels) from 2,471 accepted micrographs. After 2D classification, the selected particles (687,353) were subjected to heterogeneous refinement using three volumes obtained from an ab-initio job. Class 2 possessing higher FSC resolution (4.22 Å, 339,216 pct) was further subjected to a second round of heterogeneous refinement into 3 classes. Particles from class 2 possessing higher FSC resolution (4.22 Å, 132,224 pct) were subjected to another round of 2D classification. Selected filament particles (124,470 pct) were used for the final reconstruction using non-uniform and local refinement.

3,472,781 particles of the wt-CCaCalpL-A_4>p complex were extracted (box size 240 pixels) from 3,127 accepted micrographs. After 2D classification, the selected particles (1,358,761 pct) were subjected to heterogeneous refinement using four volumes obtained from an ab-initio reconstruction. Class 1 possessing higher FSC resolution (4.22 Å, 559,559 pct) was further subjected to 3D classification to 3 classes. After 3D classification, particles from class 0 (215,062) were used for the final reconstruction using non-uniform and local refinement.

The global resolution and sphericity values for all reconstructions were estimated using 3DFSC v.3.0 software (Tan et al., 2017) according to the Fourier shell correlation of 0.143 criterion (Rosenthal and Henderson, 2003). The local resolution was estimated in CryoSPARC (v.4.3.0) (Punjani et al., 2020, 2017).

Model building and validation

The initial protein model for wt-CCaCalpL-A_4p complex was generated using AlphaFold (Jumper et al., 2021) under the ColabFold (Mirdita et al., 2022) framework using default parameters and MMseqs2 to search for homologues into the ColabFold database, and manually modified using Coot (v.0.9.8.1) (Emsley et al., 2010) against the map sharpened using phenix.auto_sharpen (v.1.20.1-4487) (Liebschner et al., 2019) A_4p was built manually. Model refinement was performed using phenix.real_space_refine (v.1.20.1-4903) (Liebschner et al., 2019) The final model of the wt-CCaCalpL-A_4p complex covers protein residues 4-13, 18-97, 104-166, 171-197, 201-505 of chain A; residues 4-30, 36-62, 66-96, 104-189, 201-422, 426-489, 492-506 of chain B; residues 203-351, 363-504 chain C; and two A_4p molecules. This model was further used as the initial model for wt-CCaCalpL-cA_4 and wt-CCaCalpL-A_4>p complexes, which were rebuilt and refined using similar procedures. The final model of the wt-CCaCalpL-cA_4 complex covers

protein residues 202-352, 357-504 of chain A; residues 202-229, 233-273, 279-337, 344-353, 356-422, 426-486, 496-506 of chain B a and one cA_4 molecule. The final model of the wt-CCaCalpL-A_4>p complex covers protein residues 90-95, 107-–112, 118–164, 174–188, 202-353, 357-505 of chain A; residues 154-167, 207-228, 231-337, 344-352, 358-419, 426-486, 494-504 of chain B and residues 220-300, 307-326, 333-350, 384-419, 422-473, 480-502 of chain C and two A_4>p molecules. The statistics of the three-dimensional reconstruction and model refinement are summarized in Appendices 4-7. The molecular graphics figures were prepared with ChimeraX (v. 1.5) (Meng et al., 2023).

2.2.15. Protein interaction assays

Biolayer Interferometry

Biolayer interferometry (BLI) experiments were performed using the Octet K2 system (Sartorius). Experiments were performed at 37°C in 96-well plate format using 200 mL reagent volumes and a stage rotation rate set of 1000 rpm.

CCaCalpL interaction with CCaCalpT-CalpS

The His_6-StrepII-His_6-tagged CCaCalpT-CalpS heterodimer was used as the ligand and the H396A or R358/K361E variants of the CCaCalpL protein as analyte. The Ni-NTA biosensors (Sartorius) were hydrated in the reaction buffer Y supplemented with 1 mM DTT, 0.01% Triton X-100, 10 mM imidazole and 0.1% BSA (Merck) for 10 minutes before use. After the initial baseline step of 120 seconds, the CCaCalpT-CalpS heterodimer was immobilized on biosensors at a concentration of 200 nM for 120 seconds. A secondary baseline of 120 seconds was then performed to stabilize the signal. The 60-second association step was recorded by transferring the ligand-bound biosensors to wells containing the analyte CCaCalpL at concentrations of 250 nM, 500 nM, 1000 nM and 2000 nM. The association step was followed by a dissociation phase of 180 seconds. The reactions were supplemented with 4 mM of cA_4 in the appropriate measurement series. A biosensor without CCaCalpT-CalpS ligand was used as a reference. Before each measurement in the series with increasing analyte concentration, the biosensors were regenerated by three cycles of incubation for 5 seconds in 10 mM glycine (pH 1.7), followed by 5 seconds in the reaction buffer. The biosensors were then recharged by incubation in 10 mM $NiCl_2$ for 60 seconds. The mean amplitude of the CCaCalpL binding response, normalized to the amplitude of the CCaCalpT-CalpS ligand loading response, is reported from three replicates for each CCaCalpL concentration.

CCaCalpS interaction with EcRNAP

The His_6-StrepII-His_6-tagged CCaCalpS, CCaCalpT-CalpS heterodimer, or CCaCalpT-CalpS heterodimer after cleavage with cA_4-activated CCaCalpL, was used as the ligand and *E. coli* RNA polymerase (EcRNAP) core enzyme (New England Biolabs) as the analyte. The binding assay was performed in the *E. coli* RNA Polymerase reaction buffer (New England Biolabs) supplemented with 0.01% Triton X-100, 10 mM imidazole and 0.1% BSA (Merck). Measurement steps were performed as described above for the CCaCalpL and CCaCalpT-CalpS binding experiment, except that the association step was extended to 90 seconds, and only a single concentration of RNAP (100 mU/mL) was used. The averaged BLI sensorgrams of three technical replicates for each immobilized protein is reported.

Interaction assays using Dynabeads

CCaCalpT-CalpS interaction within the heterodimer after cleavage by CCaCalpL

The protease reaction was performed by pre-incubating 5 µM of CCaCalpL and 5 µM of CCaCalpT-CalpS dimer in the reaction buffer Y supplemented with 2 mM DTT at 37°C for 5 minutes and starting the reaction by adding 25 µM of cA_4. Samples were incubated at 37°C for 120 minutes in a final volume of 45 µL. The subsequent purification using magnetic beads was performed as described above for the CCaCalpS interaction with *E. coli* RNAP experiment, except reaction buffer Y with additives was used instead of reaction buffer H.

2.2.16. *E. coli* survivability assays

Spot assay

E. coli ER2566 cells were co-transformed with pCalpS or pCDFDuet-1 (Novogen) plasmid in combination with (i) pETDuet-1 (Novogen), (ii) pCalpT, (iii) $pCalpT_{24N}$, (iv) $pCalpT_{11C}$, or (v) $pCalpT_{split}$, plated on LB media agar plates, supplemented with streptomycin (25 µg/ml), carbenicillin (50 µg/ml) and 1% glucose and grown overnight at 37°C. A single colony of each variant was selected and cultivated overnight at 37°C in liquid LB supplemented with streptomycin (25 µg/ml), carbenicillin (50 µg/ml) and 1% glucose. The overnight cultures were serially diluted in fresh LB media and then plated onto LB media agar plates supplemented with streptomycin (25 µg/ml), carbenicillin (50 µg/ml) and 0.15 mM IPTG or 1% glucose. Bacteria were incubated at 37°C for 16 hours. The experiment was performed in triplicate.

Time course experiment

Similar to the spot assay, *E. coli* ER2566 cells were co-transformed with the appropriate plasmids and the overnight cultures were obtained. The overnight cultures were diluted 1:200 in the fresh LB containing streptomycin (25 µg/ml) and carbenicillin (50 µg/ml) and incubated in a 96-well plate for 1 hour at 37°C with shaking. Then 0.1 mM IPTG was added to induce the protein expression. After induction bacterial growth was measured by recording OD_{600} every 10 minutes using the CLARIOstar Plus Microplate Reader (BMG Labtech). The mean of three biological replicates is presented for each experiment.

2.2.17. GFP degradation in *E. coli* assays

GFP degradation in E. coli assays

GFP degradation assays, the plate spot assay and the time-course measurements, were performed according to (Klimecka et al., 2021) and were carried out as described below.

Plate spot assay of GFP degradation

E. coli ER2566 cells were transformed with pGFP, pGFP-(wt)CalpT$_{24N}$-degron, pGFP-(mut)CalpT$_{24N}$-degron, or pGFP-ssrA and plated on LB medium agar plates, supplemented with carbenicillin (50 µg/ml) and grown overnight at 37°C. A single colony of each variant was selected and grown overnight at 37°C in liquid LB supplemented with carbenicillin (50 µg/ml). The overnight cultures were serially diluted in fresh LB media and then plated onto LB media agar plates supplemented with carbenicillin (50 µg/ml) and 0.075 mM IPTG. Bacteria were grown at 37°C for 14 hours. Images of GFP-fluorescing cells on the agar plates were captured using the Amersham Typhoon Biomolecular Imager (Cytiva). Three biological replicates were performed.

Time-course measurements of GFP degradation

The overnight cultures (prepared as described for plate spot assay) were diluted 1:50 with fresh LB supplemented with ampicillin (50 µg/ml) and were grown to the mid-log phase. GFP-degron expression was induced with 0.005 mM IPTG, and the cell suspension was further cultured overnight at 18°C. Bacterial cultures were diluted 1:10 in M9 minimal medium supplemented with 5% (w/v) glucose, 10 µg/ml thiamine, 10 µg/ml biotin, trace elements, and 100 µg/ml spectinomycin in a 96-well plate. Spectinomycin was used to stop protein translation. The OD_{600} and fluorescence (excitation at $\lambda = 470$ nm; emission at $\lambda = 512$ nM) were measured every 10 minutes for 12 hours at 30°C

using a CLARIOstar Plus Microplate Reader (BMG Labtech). Fluorescence results were normalized first by optical density and then by setting the values at zero time points to 1. Results represent the mean values of three biological replicates.

3. RESULTS AND DISCUSSION

3.1. Regulation of cA_6 signaling

In 2017, my colleagues in a research group led by Dr. Gintautas Tamulaitis discovered a cyclic oligoadenylate (cA_n) signaling pathway involved in bacterial antiviral defense (Kazlauskiene et al., 2017). This discovery was made using a model type III-A CRISPR-Cas system from *Streptococcus thermophilus* (St) strain DGCC8004. This system comprises adaptation genes (*cas1* and *cas2*), a CRISPR array, the *cas6* gene encoding the crRNA-processing protein, interference complex genes (*cas10*, *csm2*, *csm3*, *csm4*, and *csm5*), and two auxiliary CARF-HEPN architecture nucleases (*csm6* and *csm6'*) (Figure 3.1 A) (Tamulaitis et al., 2014).

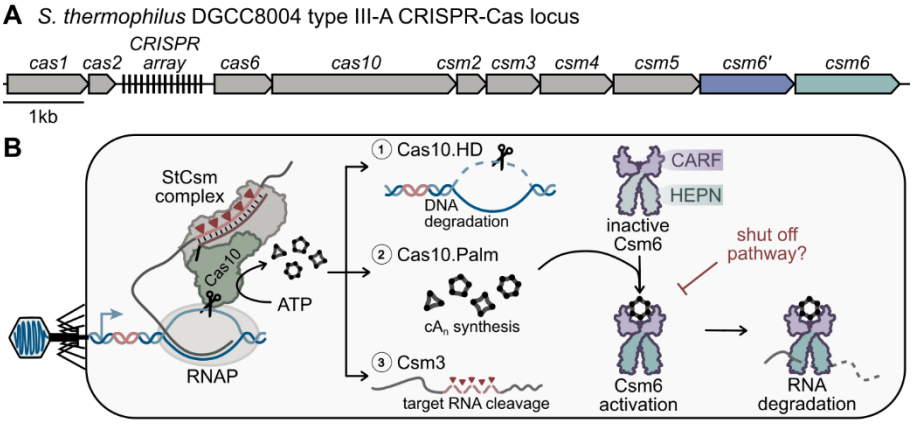


Figure 3.1. *S. thermophilus* type III-A CRISPR-Cas system. (**A**) The type III-A CRISPR-Cas locus in the genome of *S. thermophilus* DGCC8004 strain (GenBank ID: KM222358.1). Genes of associated effectors *csm6'* and *csm6* are colored blue and teal. (**B**) The mechanism of action of *S. thermophilus* type III-A CRISPR-Cas system. The StCsm interference complex binds a phage transcript that is complementary to the spacer sequence of the crRNA. This binding triggers the enzymatic activities of the large subunit Cas10: HD nuclease domain starts to degrade ssDNA formed in the transcription bubble (1), and Palm synthase domain starts to synthesize cA_n from ATP (2). CARF domain of the associated effectors StCsm6 and StCsm6′ binds cA_6 and this initiates nonspecific degradation of RNA by activated HEPN RNase domain. The enzymatic activities of Cas10 are switched off after StCsm complex-bound target RNA is cleaved by Csm3 subunits (3), however it is not known how the pool of synthesized cA_6 and other cA_n is cleared and cA_6 signaling is terminated.

Upon viral RNA binding, the interference complex is activated, leading to the production of cyclic oligoadenylates (cA_n, n = 3–6) by the Cas10 subunit (Figure 3.1 B). These cA_n act as secondary messengers to activate auxiliary ribonucleases. Specifically, StCsm6 and StCsm6′ function as non-specific

RNases that are stimulated by cA_6 (Kazlauskiene et al., 2017). Interestingly, while StCsm6 was activated by cA_6, a homologous ribonuclease from *Thermus thermophilus* (TtCsm6) responded to cA_4 instead (Kazlauskiene et al., 2017). In addition, it was demonstrated that the cA_n synthesis activity of StCsm is switched-off upon cleavage of the target RNA by the Csm3 subunit (Kazlauskiene et al., 2017).

In vitro the StCsm complex synthesizes a mixture of cA_n from ATP, with cA_3 being the major product. However, only the minor synthesis product cA_6 was found to activate the StCsm6 RNase (Kazlauskiene et al., 2017). These results raised several important questions: (i) which cA_n are produced in the bacterial cell? (ii) does the type III CRISPR-Cas act as an abortive defense mechanism leading to dormancy or cell death, or are cA_n degraded once their synthesis is stopped? (iii) if cA_n are degraded, what mechanisms regulate their removal?

A breakthrough came in 2018 when M. White's group discovered that type III CRISPR-Cas systems encode specialized CARF-domain proteins called CRISPR ring nucleases (Crn), which specifically hydrolyze cA_4 (Athukoralage et al., 2018). This led us to hypothesize that the CARF-HEPN ribonucleases StCsm6 and TtCsm6 might also degrade cA_6 or cA_4 *via* their CARF domains.

To test this, we analyzed whether TtCsm6 could hydrolyze its activator, cA_4 (Figure 3.2 A). Initial TtCsm6 experiments were performed by Dr. Irmantas Mogila. Using high-performance liquid chromatography-mass spectrometry (HPLC-MS), which was performed by Audronė Rukšėnaitė, we found that TtCsm6 converted cA_4 into linear di-adenylate with a cyclic-2′,3′ phosphate (A_2>p) after 2 hours. This reaction matched the cleavage pattern of the Crn1 ring nuclease (Athukoralage et al., 2018), confirming that cA_4-activated TtCsm6 can also function as a slow-acting ring nuclease. Our findings were later corroborated by independent studies showing that both TtCsm6 and *Thermococcus onnurineus* Csm6 (ToCsm6) specifically degrade cA_4 (Athukoralage et al., 2019; Jia et al., 2019b).

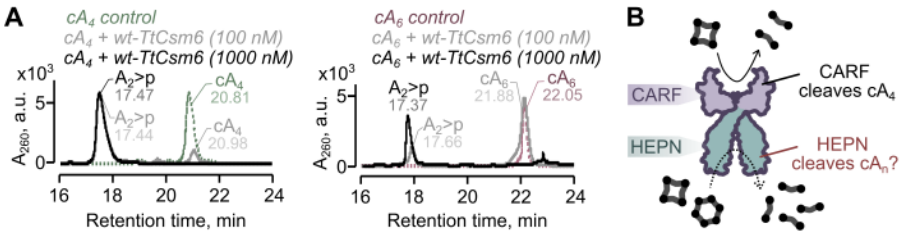


Figure 3.2. Ring nuclease activity of cA4-specific TtCsm6 (**A**) TtCsm6 cA_4 and cA_6 cleavage reactions analyzed by HPLC-MS. Reactions contained 20 µM cA_4 or cA_6 and 0, 100 or 1000 nM wt-TtCsm6 and were incubated for 120 min at 37°C. (**B**)

Schematic representation of Csm6 putative ring nuclease activities. The involvement of HEPN domain in cA_n cleavage remains elusive.

Unexpectedly, we also observed that at high protein concentrations, TtCsm6 can degrade cA_6 in addition to cA_4. Since cA_6 is too large to fit into the CARF-domain binding pocket, we hypothesized that Csm6 ribonucleases could use their HEPN domain to degrade all cA_n produced by the type III CRISPR-Cas Csm complex (Figure 3.2 B). In support of this, the crystal structure of ToCsm6 revealed that cA_4 binds to both the CARF and HEPN domains (Jia et al., 2019b). To further investigate the cA_6 signaling pathway, its deactivation mechanism, and the role of the CARF and HEPN domains of StCsm6/StCsm6' in cA_n degradation, we performed a comprehensive analysis of cA_n synthesis and degradation in the cA_6-specific *S. thermophilus* type III-A CRISPR-Cas system.

3.1.1. Production and degradation of cA_n in bacteria cells

First, we sought to resolve the discrepancy between the cA_n species produced by the StCsm interference complex and the species that activate the associated effectors StCsm6 and StCsm6'. Given the possibility that the *in vitro* synthesis conditions may not be optimal, and no direct evidence of cA_n production in cells existed, we sought to address the cA_n synthesis in a more natural environment. To this end, we conducted a study on cA_n production in a heterologous *E. coli* host, henceforth referred to as a cA_n production *in vivo*.

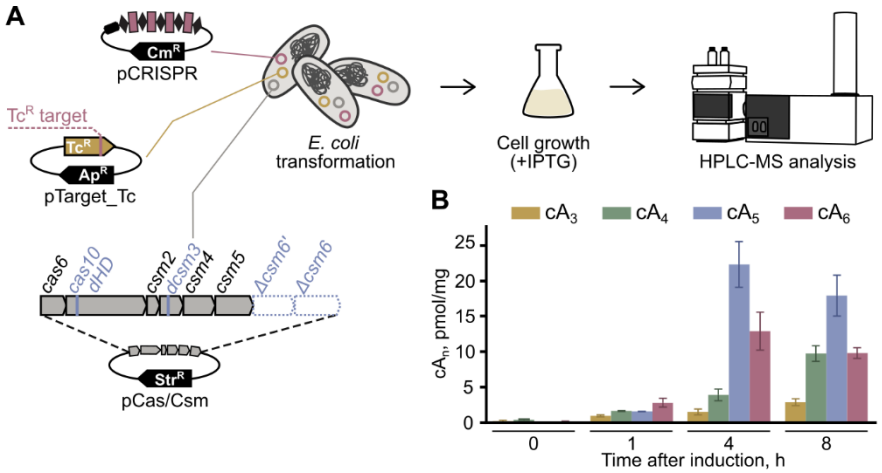


Figure 3.3. cA_n synthesis by StCsm complex *in vivo* analyzed by HPLC-MS. (A) Experimental setup. *E. coli* cells were cotransformed with pCRISPR_Tc, mutant variant of pCas/Csm (dHD-Cas10, dCsm3, ΔCsm6′ΔCsm6) and the target plasmid pTarget_Tc. Cells were grown in liquid LB supplemented with appropriate antibiotics. The expression of the StCsm interference complex was induced using

IPTG and samples of the cells were taken at defined time points. Cells were lysed and extracted metabolites were submitted for HPLC-MS analysis. (**B**) HPLC-MS analysis of cA_n (n=3-6) production. The amount of cA_n in pmol per mg of wet cell culture is plotted for at least two replicates. The error bars indicate standard deviation between the at least two replicates.

To examine the population of cA_n produced *in vivo*, we expressed *S. thermophilus* type III-A CRISPR–Cas system in *E. coli* BL21 (DE3) laboratory strain and used the HPLC-MS technique to analyze the cA_n metabolites (Figure 3.3 A). For expression of *S. thermophilus* type III-A CRISPR-Cas system, *E. coli* cells were co-transformed with plasmids pCas/Csm, pCRISPR_Tc and pTarget_Tc. pCas/Csm encoded genes of the proteins required for crRNA maturation and the StCsm interference complex assembly. To avoid the possible toxicity and to increase the detectible cA_n signal, we used pCas/Csm variant, where all enzymatic activities of the StCsm complex, except that of the Cas10 Palm synthase, have been inactivated by mutations (D33A in Csm3 subunit - dCsm3, D16A in HD domain of Cas10 - dHD-Cas10) and the genes of associated effectors *csm6′* and *csm6* were removed (ΔCsm6′ΔCsm6). The pCRISPR_Tc carried a CRISPR array containing 4 copies of spacer targeting the transcript of tetracycline resistance gene (Tc^R) gene, which was produced constitutively from plasmid pTarget_Tc. The expression of pCas/Csm and pCRISPR_Tc were induced by adding IPTG. Samples of the *E. coli* cells were collected before the addition of the inductor (time point 0 h) and after 1, 4 and 8 hours. The metabolites of the cells were extracted and submitted to the HPLC-MS analysis. Metabolite extraction and HPLC-MS were performed by Dr. Jesper F. Havelund and Dr. Nils J. Færgeman (University of Southern Denmark). The profile of cA_n metabolites revealed that *in vivo* StCsm interference complex produces various cA_n, from n=3 to n=6 (Figure 3.3 B), and cA_2 was produced only in trace amounts. However, in a striking difference to the production profile *in vitro* where cA_5 and cA_6 were minor products (Kazlauskiene et al., 2017), *in vivo* synthesis equilibrium was shifted towards cA_5 and cA_6 species (Figure 3.3 B). The amount of cA_5 and cA_6 was increasing during the first 4 h post-induction and reached 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. Notably, at 8 h post-induction the amount of cA_5 and cA_6 had slightly decreased while cA_3 and cA_4 kept increasing (Figure 3.3 B). It is possible that constant synthesis of cA_n results in decrease in total cellular ATP concentration and this may shift the cA_n synthesis reaction equilibrium toward shorter reaction products. Indeed, a recent *in vitro* study confirmed that ATP concentration determines the length of cA_n (Jungfer et al., 2025).

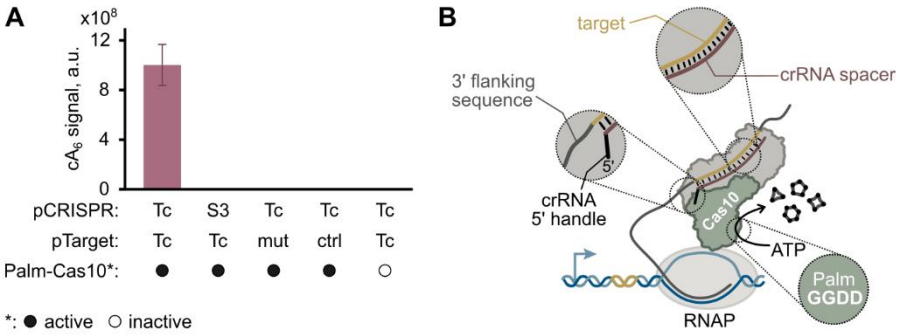


Figure 3.4. Confirmation of cA_n synthesis requirements *in vivo*. (**A**) HPLC-MS analysis of cA_6 synthesis by *E. coli* cells harboring indicated pTarget, and pCRISPR variants and expressing the StCsm complex (dCsm3, dHD-Cas10, ΔCsm6′ΔCsm6) with intact Palm-Cas10 (filled circle) or deficient dPalm-Cas10 domain (empty circle). Plasmid pTraget_Tc^{mut} encodes the Tc^R with mutated 3′-flanking sequence of the Tc target so that it is complementary to the 5′-handle of crRNA of the StCsm complex. Plasmid pTarget_ctrl does not contain the Tc^R gene. The cA_6 signal is plotted for two replicates. The error bars indicate the standard deviation between the two replicates. (**B**) Schematic representation of the requirements for cA_n synthesis: (i) spacer sequence of crRNA must be complementary to the target RNA, (ii) the 3′-flanking sequence of the target RNA must not base-pair with the 5′-handle of crRNA and (iii) Cas10 in the effector complex must contain an intact GGDD-active site in the Palm domain.

Next, we employed cA_n production *in vivo* assay to confirm the cA_n synthesis requirements that were previously established *in vitro* for *S. thermophilus* and other type III systems (Han et al., 2018; Kazlauskiene et al., 2017; Nasef et al., 2019; Niewoehner et al., 2017; Rouillon et al., 2018). We used our *in vivo* assay with pCas/Csm variants: (i) dCsm3, dHD-Cas10, ΔCsm6′ΔCsm6 or (ii) dCsm3, dHD-Cas10, dPalm-Cas10, ΔCsm6′ΔCsm6; and pTarget variants: (i) Tc, Tc^{mut} or ctrl. Plasmid pTarget_Tc^{mut} encodes the Tc^R with mutated 3′-flanking sequence of the Tc target so that it is complementary to the 5′-handle of crRNA of the StCsm interference complex. Plasmid pTarget_ctrl does not contain the Tc^R gene. By monitoring the production of cA_6 *in vivo* we corroborate the findings that (i) spacer sequence of crRNA must be complementary to the target RNA, (ii) the 3′-flanking sequence of the target RNA must not base-pair with the 5′-handle of crRNA and (iii) Cas10 in the effector complex must contain an intact GGDD-active site in the Palm domain (Figure 3.4).

Additionally, we tested whether the cA_6 signaling pathway is induced in response to phage infection. We targeted StCsm interference complex against RNA coliphage MS2 by changing the spacers in the pCRISPR plasmid. We induced the expression of type *S. thermophilus* III-A CRISPR-Cas system and infected those cells with the phage. Cells were collected 16 hours post-

infection and the presence of the main activator cA_6 was analyzed using HPLC-MS (Figure 3.5 A). As expected, accumulation of cA_6 was observed in the cells carrying the MS2 targeting spacer (Figure 3.5 B), demonstrating that the exogenic nucleic acid which triggers the production of cA_n by the type III-A CRISPR–Cas system may also derive from a phage.

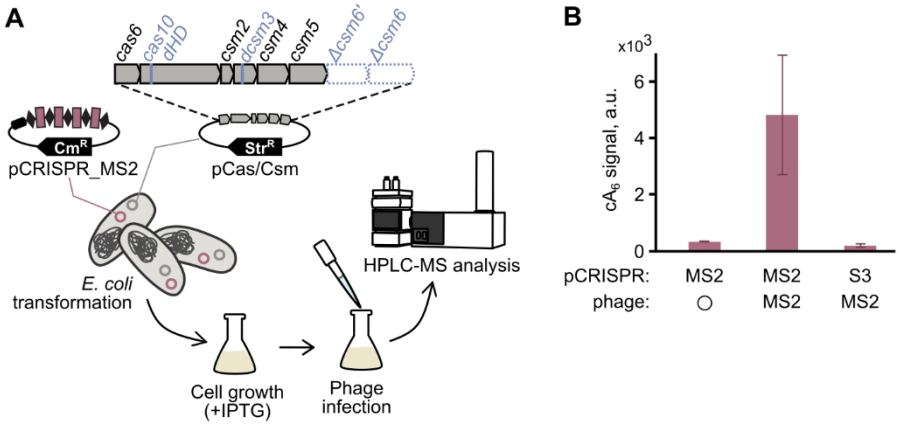


Figure 3.5. cA_6 synthesis in response to phage infection. (**A**) Experimental setup. *E. coli* cells were co-transformed with MS2 phage targeting pCRISPR_MS2 and mutant variant of pCas/Csm (dHD-Cas10, dCsm3, ΔCsm6′ΔCsm6). The expression of the StCsm complex was induced using IPTG and phage suspension was added. Samples of the infected cells were taken at 16 h time point. Cells were lysed and extracted metabolites were submitted for HPLC-MS analysis. (**B**) HPLC-MS analysis of cA_6 production by StCsm complex in response to MS2 phage infection. The cA_6 signal is plotted for two replicates. The error bars indicate the standard deviation between the two replicates.

Next, we addressed the main question: could StCsm6 or/and StCsm6′ be involved in degradation of the cA_6 activator similar to the cA_4 degrading TtCsm6 and ToCsm6 (Athukoralage et al., 2019; Jia et al., 2019b). To get first insights, we monitored the production of cA_n *in vivo* in the cells expressing *S. thermophilus* type III system together with StCsm6′ and StCsm6 and compared the results with the production in cells lacking the associated effectors as used in previous experiments (Figure 3.3 B). *E. coli* cells expressing StCsm6′ and StCsm6 exhibited drastically reduced levels of cA_6 (Table 3.1), suggesting that these proteins might possess ring nuclease activity. Moreover, the levels of other cA_n were also diminished (Table 3.1), implying that StCsm6 and StCsm6′ proteins may be responsible for the degradation of all cA_n species in the cells.

Taken together, the analysis of cA_n metabolites in *E. coli* cells demonstrates that in response to exogenic nucleic acids, that derive from a phage infection or from a constitutively transcribed gene in the plasmid, the

S. thermophilus type III-A CRISPR-Cas system produces signaling molecules in the heterologous host. Production *in vivo* follows the requirements established *in vitro*, but the distribution of cA_n products differs. Furthermore, the expression of StCsm6 and StCsm6′ drastically reduces levels of cA_6 and other cA_n.

Table 3.1. HPLC-MS analysis of cA_n amount changes over time *in vivo* in the absence (Δ) or presence of StCsm6 and StCsm6′ encoding genes. Data represents mean values of at least two replicates ± standard deviation; n.d. – no signal detected in the sample.

cA_n	time, h	cA_n amount, pmol/mg	
		ΔStCsm6′ΔStCsm6	StCsm6′, StCsm6
cA_3	0	0.24 ± *0.15*	0.000 ± *0.001*
	1	0.94 ± *0.20*	0.005 ± *0.007*
	4	1.51 ± *0.56*	0.002 ± *0.003*
	8	2.87 ± *0.68*	0.017 ± *0.007*
cA_4	0	0.39 ± *0.10*	n.d.
	1	1.62 ± *0.13*	0.003 ± *0.005*
	4	3.88 ± *1.18*	n.d.
	8	9.78 ± *1.57*	n.d.
cA_5	0	0.04 ± *0.03*	n.d.
	1	1.58 ± *0.004*	n.d.
	4	22.33 ± *4.58*	n.d.
	8	17.89 ± *4.05*	n.d.
cA_6	0	0.12 ± *0.07*	0.126 ± *0.178*
	1	2.80 ± *0.88*	0.005 ± *0.007*
	4	12.91 ± *3.80*	n.d.
	8	9.79 ± *1.07*	n.d.

3.1.2. Cleavage of cA_6 by StCsm6 and StCsm6′ *in vitro*

To thoroughly characterize cA_6 cleavage by StCsm6 and StCsm6′ proteins *in vitro* we were in need of easily detectable substrate such as radiolabeled cA_6 or/and other cA_n variants. Since HPLC-MS analysis of cA_n produced in *E. coli* cells revealed a difference in the cA_n profile when compared to the *in vitro* experiments (Figure 3.3 B) (Kazlauskiene et al., 2017), this suggests that discrepancies in the size distribution could arise due to different reaction conditions. Therefore, we reevaluated cA_6 synthesis *in vitro*. Synthesis reactions were conducted using different molar ratios of the StCsm complex and target RNA, different reaction buffer compositions, and either Co^{2+} or Mg^{2+} ions as cofactors. The HPLC-MS analysis of reactions revealed a shift in the cA_n size distribution profile towards a larger ring size and ultimately to cA_6 in the presence of target RNA excess and Mg^{2+}. Prolonged incubation of the StCsm complex with ATP resulted in a slight increase in the yield of cA_3

and cA_4 (Figure 3.6), which is in line with the *in vivo* data (Figure 3.3 B). This and all subsequent HPLC-MS analyses were performed by Audronė Rukšėnaitė.

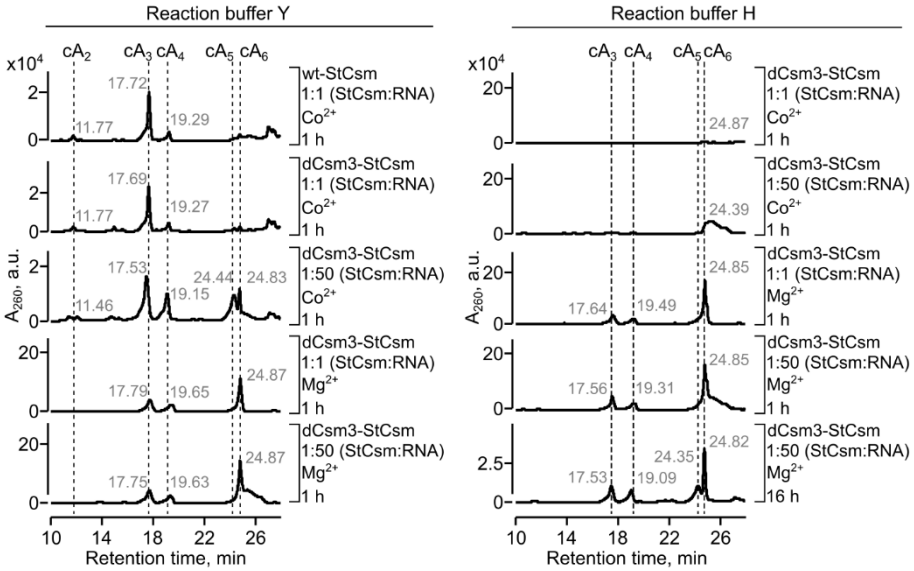


Figure 3.6. Optimization of the cA_6 synthesis conditions. The HPLC-MS analysis of cA_n produced by StCsm interference complex *in vitro*. Reactions were performed using wt or dCsm3-StCsm complex at StCsm:target RNA molar ratio 1:1 or 1:50 in indicated reaction buffer supplemented with indicated metal ion cofactor. Reactions were performed for either 1 or 16 hours.

Using optimized synthesis conditions, we produced radiolabeled cA_6 and incubated it with purified StCsm6 and StCsm6' proteins. In the absence of divalent metal ions both proteins degraded radiolabeled cA_6, generating faster migrating products in denaturing PAGE (Figure 3.7 A). By employing the HPLC-MS analysis we identified that in both cases the cA_6 cleavage reaction yielded a mixture of three products, namely (i) 2′,3′-cyclic AMP (A>p), (ii) terminal 2′,3′-cyclic phosphate-containing linear di-adenylate (A_2>p) and (iii) terminal 2′,3′-cyclic phosphate-containing linear tri-adenylate (A_3>p) (Figure 3.7 B). Formation of linear cyclic-2′,3′ phosphate products are not surprising since cyclic-2′,3′ phosphate products are usually generated by other metal independent RNases (Carte et al., 2014; Yang, 2011) as well as by CARF domain containing cA_4 ring nucleases (Athukoralage et al., 2019, 2018; Jia et al., 2019b).

StCsm6 and StCsm6′ are structural homologues and show ~34% amino acid identity (Kazlauskiene et al., 2017; Tamulaitis et al., 2014). However, StCsm6′ appears to be a shorter version comprising 386 aa, while StCsm6 is composed of 428 aa. Given that both proteins exhibited similar activity on the

cA_6 substrate, subsequent *in vitro* investigation was performed only on a longer homolog, StCsm6, for the sake of simplicity.

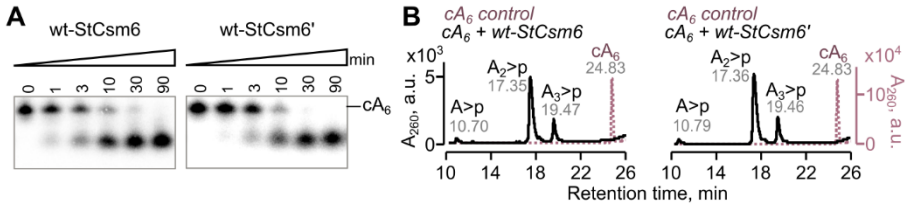


Figure 3.7. Cleavage of cA_6 by StCsm6 and StCsm6′ (**A**) cA_6 cleavage by StCsm6 and StCsm6′ analyzed by denaturing PAGE. Reactions contained 50 nM of labeled cA_6 and 10 nM of protein. (**B**) cA_6 cleavage by StCsm6 and StCsm6′ analyzed by HPLC-MS. cA_6 control depicted in light red. The 90 min reactions contained 20 µM of cA_6 and 0 or 10 nM of protein.

3.1.3. Activity of isolated CARF and HEPN domains of StCsm6

StCsm6 is comprised of two functional domains: CARF and HEPN connected by α-helical 6H domain. Since Crn enzymes, which are composed of CARF domains, and the CARF domain of some cA_4-dependent effectors cleave the cA_4 activator (Athukoralage et al., 2019, 2018; Jia et al., 2019b), we hypothesized that the CARF domain of StCsm6 has a similar function in cA_6 cleavage. On the other hand, the HEPN domain is an RNase, exhibiting a preference for the degradation of single-stranded RNA at AA or GA dinucleotides (Kazlauskiene et al., 2017). Consequently, it is plausible that HEPN domain may also be capable of degrading cA_6. Furthermore, a drastic reduction of other cA_n species was observed in *E. coli* cells in the presence of StCsm6 and StCsm6′ proteins (Table 3.1). This may be attributed to the RNase activity of HEPN domain.

First, to establish which domain is responsible for the cleavage of the activator cA_6, we engineered and purified individual CARF (1–169 aa) and HEPN (172–428 aa) domains of StCsm6 (Figure 3.8 A). Similarly to a full length StCsm6 protein which is a dimer in solution (Kazlauskiene et al., 2017), isolated CARF and HEPN domains remained dimers (Figure 3.8 B). Domain separation and the evaluation of the oligomeric state were performed by Augustė Rimaitė. Furthermore, the isolated HEPN domain retained residual RNase activity that is comparable with wt-StCsm6 in the absence of cA_6 activator (Figure 3.8 C). As expected, no nonspecific RNase activity was detected for isolated CARF domain (Figure 3.8 C).

Further, we submitted the isolated domains to the cA_6 degradation assay. Both CARF and HEPN domains were able to degrade radiolabeled cA_6 (Figure 3.9 A and B). Under single turnover conditions the CARF domain was

more efficient at degrading cA_6 than the HEPN domain (Figure 3.9 A), this may be attributed to the only residual activity of HEPN as observed in the RNA degradation assay (Figure 3.8 C).

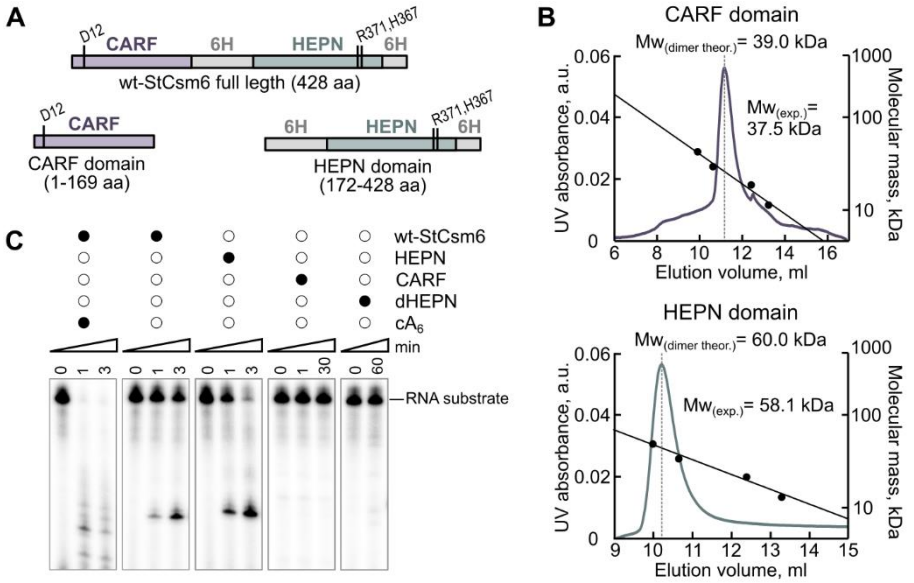


Figure 3.8. Biochemical characterization of isolated CARF and HEPN domains of StCsm6. (**A**) Schematic illustration of wt-StCsm6 protein and its CARF (1–169 aa) and HEPN (172–428 aa) domains. Residues important for activity are indicated. (**B**) Size-exclusion chromatography analysis of the CARF and HEPN domains. The experimental molecular weights were calculated based on the size markers: BSA (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and ribonuclease A (13.7 kDa). (**C**) Comparison of RNase activity of isolated domains with wt-StCsm6 and cA_6 activated wt-StCsm6. The reactions contained 10 nM α^{32}P-labelled RNA substrate, and 10 nM protein and 10 nM cA_6 if indicated. Reaction products were analyzed by denaturing PAGE. Protein concentration of 10 μM was used in the case of dHEPN (an isolated HEPN domain with R371A-H367A mutations).

We predicted the putative active site of CARF domain by aligning the sequences of the Csm6 homologues and by superposing StCsm6 CARF domain model (Kazlauskiene et al., 2017) with known cA_4 specific ring nuclease activity possessing *T. onnurineus* Csm6-cA_4 structure (PDB ID: 6O6V) (Jia et al., 2019b) (Appendices 8 and 9). We identified that D12 residue of StCsm6 is located near the cyclic oligoadenylate, and this residue is absolutely conserved in the Csm6 homologues (Figure 3.9 C). D12A mutation (dCARF variant) completely abolished the cA_6 hydrolysis activity of the CARF domain (Figure 3.9 A), indicating that this residue is crucial for cA_6 cleavage. Mutations of the HEPN active site residues R371A and H376A (dHEPN variant) required for RNase activity (Kazlauskiene et al., 2017) dramatically impaired hydrolysis of both cA_6 (Figure 3.9 A) and linear RNA

(Figure 3.8 C), indicating that the same active site in the HEPN domain is responsible for cA_6 and linear RNA cleavage.

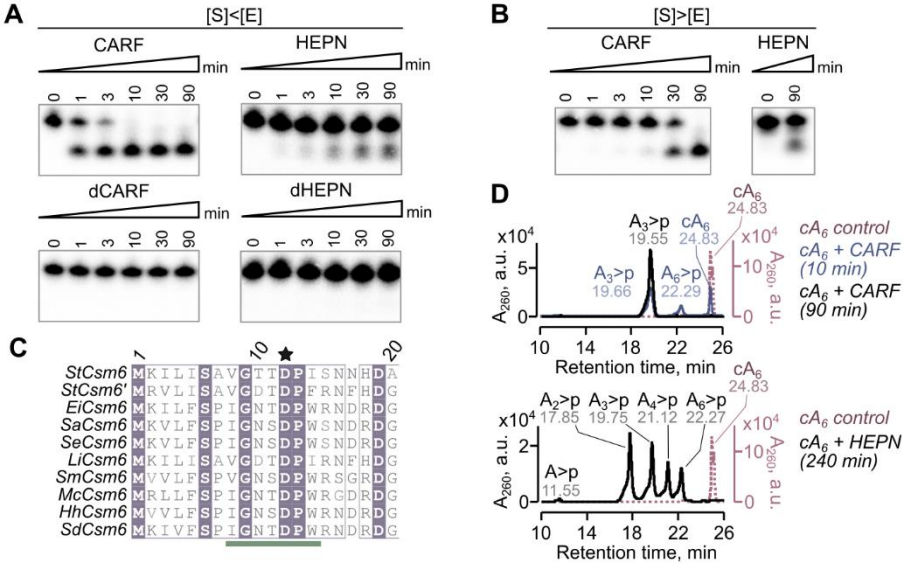


Figure 3.9. The cleavage of cA_6 by isolated CARF and HEPN domains. (**A**) Hydrolysis of cA_6 by isolated wt CARF and HEPN domains and their mutant dHEPN (R371A, H376A mutations) and dCARF (D12A mutation) variants performed under single turnover conditions ([S]<[E]). Reactions contained 50 nM of labeled cA_6 and 1 μM of CARF or dCARF and 10 μM of HEPN or dHEPN. Reaction products were analyzed by denaturing PAGE. (**B**) Hydrolysis of cA_6 by isolated CARF and HEPN domains under multiple turnover conditions ([S]>[E]). Reactions were performed using 1 μM of the CARF and HEPN domains, 50 nM labeled cA_6 and 10 μM of non-labeled cA_6. Hydrolysis products were analyzed by denaturing PAGE. (**C**) Multiple sequence alignment of the CARF domain fragment (1-20 aa) of Csm6 homologs. Amino acids of the loop located close to the predicted cA_6 binding site (residues 8-14) are underlined in green. A star indicates an absolutely conserved D residue. For the full alignment refer to the Appendix 8 (**D**) Analysis of the cA_6 cleavage products by HPLC-MS. Hydrolysis reactions were performed by mixing 20 μM cA_6 and 1 μM CARF or 10 μM HEPN domain and incubating for indicated time.

Since precise cleavage products cannot be identified in a denaturing 24% PAG, we analyzed reaction products generated by the CARF and HEPN domain-mediated cleavage of cA_6 by HPLC-MS (Figure 3.9 D). The analysis revealed that the CARF domain converts cA_6 to A_3>p final product. At the shorter reaction time (10 min) some linear A_6>p intermediate is also present in the reaction mixture but is further cleaved by CARF domain into the final reaction product, A_3>p. Presumably, each monomer of homodimeric CARF domain cleaves cA_6 producing A_3>p, similarly to the CARF ring nucleases and self-limiting CARF effectors that split cA_4 into two A_2>p (Athukoralage et al., 2019, 2018; Jia et al., 2019b). Since some intermediate product A_6>p is

observed in the reaction mixture, the reaction may occur non-concertedly. This is also in line with the characterized cA_4 cleaving enzymes (Athukoralage et al., 2019, 2018; Jia et al., 2019b).

Unlike the CARF domain which produces a single final cleavage product A_3>p, the HEPN domain degraded cA_6 into a mixture of A_6>p, A_4>p, A_3>p, A_2>p and A>p oligoadenylates (Figure 3.9 D). This product mix may be attributed to the RNase activity of the HEPN domain, which preferentially cleaves at AA and AG dinucleotides (Kazlauskiene et al., 2017), explaining the formation of A_n>p of various lengths. Together, this shows that both StCsm6 domains degrade cA_6 but generate different cleavage products.

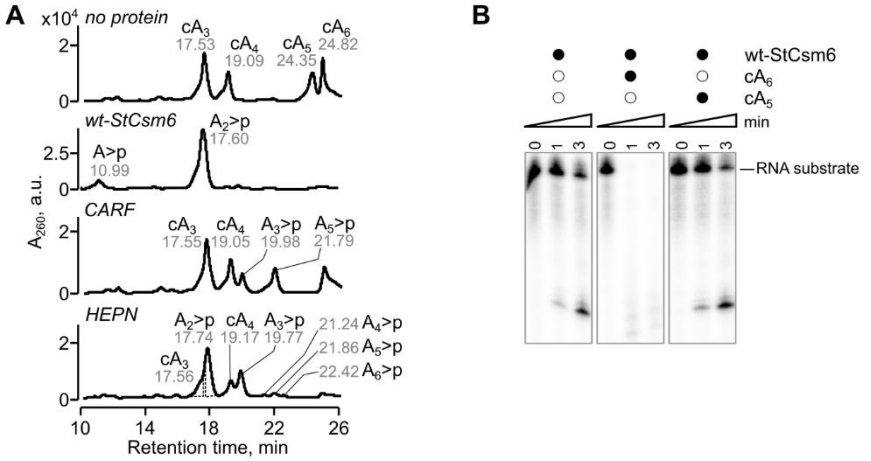


Figure 3.10. Hydrolysis of cA_n mix by the isolated CARF and HEPN domains. (**A**) HPLC-MS analysis of cA_n hydrolysis reactions. Reactions contained approximately 30 μM cA_n, 10 nM wt-StCsm6 or 1 μM CARF domain or 10 μM HEPN domain and were incubated for 16 h or 90 min in case of wt-StCsm. (**B**) RNA hydrolysis by wt-StCsm6 in the absence or presence of cA_6 or cA_5. Reactions contained 10 nM protein, 10 nM radiolabeled RNA and 100 nM cA_6 or cA_5. Samples were analyzed by denaturing PAGE.

Furthermore, in *E. coli* cells expressing StCsm6 and StCsm6′, not only cA_6 levels were dramatically reduced, but also the levels of all cA_n (Table 3.1), suggesting that StCsm6 is capable of degrading different cA_n produced by the StCsm interference complex. Indeed, *in vitro* wt-StCsm6 degraded cA_n mixture into A_2>p and some A>p products (Figure 3.10 A). Therefore, we further tested the activity of isolated CARF and HEPN domains on the mixture of cA_n. Since the HEPN domain lacked allosteric activation, the reaction occurred poorly, but the HEPN domain was still able to linearize/cleave all cA_n species (Figure 3.10 A). In line with cA_6 cleavage experiments (Figure 3.9 D), the isolated CARF domain converted cA_6 to A_3>p in the mixture of cA_n (Figure 3.10 A). Nevertheless, it failed to cleave cA_3 and cA_4 (Figure 3.10

A). Surprisingly, the CARF domain was able to linearize cA_5 to A_5>p (Figure 3.10 A). It is possible that cA_5 can adopt a conformation that allows the binding to the CARF domain. However, the ring of cA_5 is not symmetrical and only one phosphodiester bond that is positioned correctly in the ring nuclease active site can be hydrolyzed linearizing cA_5. However, cA_5 did not activate RNase activity of wt-StCsm6 (Figure 3.10 B), implying that the linearization of cA_5 by the CARF domain is not an intended reaction for regulation. Therefore, we can conclude that the CARF domain possesses a cA_6-specific ring nuclease activity, while HEPN domain degrades cA_6 and other cA_n as a result of its RNase activity.

3.1.4. The role of CARF and HEPN in full-length StCsm6

Next, we examined a possible allosteric connection between the CARF and HEPN domains in cA_6 degradation by full-length StCsm6 protein. To dissect possible allosteric interactions between the domains, we first compared cA_6 cleavage by wt-StCsm6 and mutant variants dCARF-StCsm6 and dHEPN-StCsm6, which carried a D12A mutation in CARF or R371A and H376A mutations in the HEPN active site, respectively (Figure 3.11 A and B).

Under single turnover ([S]<[E]) conditions, dHEPN-Csm6 cleaved cA_6 at a rate (k_{obs}= 0.27 ± 0.10 min^{-1}) that was comparable to that of wt-StCsm6 (k_{obs}= 0.19 ± 0.04 min^{-1}) and about 2-fold slower than that of the isolated CARF domain (k_{obs}= 0.47 ± 0.18 min^{-1}) (Figures 3.11 A and 3.9 A), suggesting that at low cA_6 concentrations, the CARF domain functions as a ring nuclease, and the contribution of the HEPN domain to cA_6 degradation is negligible. Indeed, under the same conditions, the dCARF-Csm6 variant barely cleaved cA_6, similarly to the isolated HEPN domain (Figures 3.11 A; 3.9 A). Low cA_6 degradation of dCARF-Csm6 variant may be caused by a preferential binding of cA_6 to the CARF domain and poor binding of cA_6 to the HEPN domain. Since we were unable to determine cA_6 binding affinity by the HEPN domain directly, we evaluated cA_6 binding using an RNA cleavage competition assay. We mixed full-length wt-StCsm6 and substrate RNA at a 1:1 ratio, and measured RNA cleavage rate (k_{obs}) in the presence of increasing concentrations of cA_6 (Figure 3.11 B). At low cA_6 concentrations, wt-StCsm6 hydrolyzed RNA very slowly because cA_6 concentrations were too low to be bound by the CARF domain and therefore the HEPN domain was not allosterically stimulated. However, the RNA hydrolysis rate increased at 10 nM of cA_6, implying cA_6 binding to the CARF domain and allosteric activation of the RNase activity of the HEPN domain. Further increase of cA_6 concentration up to 10 μM did not appear to affect the RNA hydrolysis rate

and inhibition of RNA hydrolysis was observed only at cA_6 concentrations above 100 µM (Figure 3.11 B), suggesting that HEPN has lower binding affinity for cA_6 than linear ssRNA.

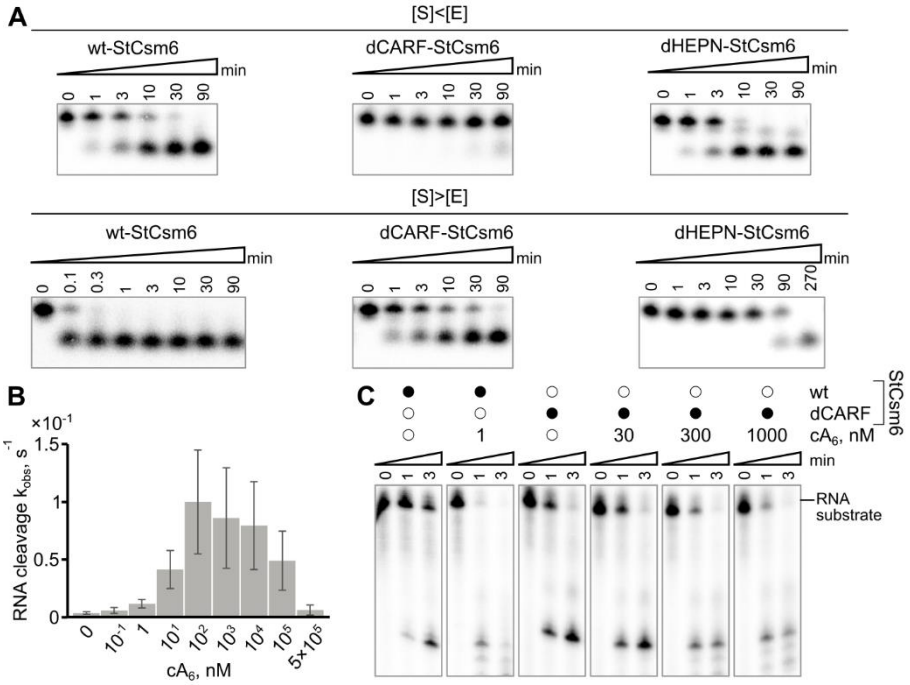


Figure 3.11. Hydrolysis of cA_6 by the full-length StCsm6 variants. (**A**) cA_6 hydrolysis under single turnover ([S]<[E]) and multiple turnover ([S]>[E]) conditions analyzed by denaturing PAGE. Single turnover reactions contained 50 nM labeled cA_6 and 1 µM of wt-StCsm6, dHEPN-StCsm6 or dCARF-Csm6. Multiple turnover reactions were supplemented with an additional 10 µM non-labeled cA_6. (**B**) RNA hydrolysis competition assay. RNA hydrolysis rate of wt-StCsm6 (10 nM labeled RNA, 10 nM StCsm6) was monitored in the presence of increasing concentration (0–500 µM) of cA_6. The mean values of the reaction rate constants (k_{obs}) are depicted. The error bars indicate standard deviation of three experiments. (**C**) Dependence of RNA hydrolysis by dCARF-StCsm6 on cA_6 concentration. Reactions contained 10 nM protein, 10 nM radiolabeled RNA and 0-1000 nM cA_6. Reactions containing wt-StCsm6 served as a control. Samples were analyzed by denaturing PAGE.

Under multiple turnover ([S]>[E]) conditions wt-StCsm6 rapidly degrades cA_6 with a turnover rate of >20 min^{-1} (Figure 3.11 A). It is likely that at high concentrations cA_6 binds both to the CARF and HEPN domains, thereby allosterically activating effective degradation of cA_6 by the HEPN RNase activity. In contrast, isolated CARF domain and dHEPN-StCsm6 variant cleaved cA_6 slower (k_{cat} <0.1 min^{-1}), suggesting that activated HEPN is a much more efficient ring nuclease than CARF. The dCARF-StCsm6 mutant degraded cA_6 faster than isolated HEPN domain (Figures 3.9 B and 3.11 A),

indicating that cA_6 binding but not cleavage at the CARF domain is required for allosteric activation of the HEPN nuclease. The slower cA_6 cleavage by the dCARF-StCsm6 mutant as compared to wt-StCsm6 may result from compromised cA_6 binding affinity, caused by the D12A mutation in the CARF domain. In agreement, an increased cA_6 concentration (up to 300 nM) was required for activation of RNA degradation by the dCARF-StCsm6 mutant (Figure 3.11 C).

Taken together, our results demonstrate that the CARF domain is relatively slow but cA_6-specific ring nuclease, whereas the HEPN domain, upon allosteric activation, exhibits potent RNase activity and degrades ssRNA and cA_6 present in high concentrations.

3.1.5. Regulation of type III-A CRISPR–Cas immunity by StCsm6

Evaluation of cA_n production by StCsm complex in *E. coli* cells and analysis of StCsm6 ring nuclease activities *in vitro* allowed us to update the model of *S. thermopilus* type III-A CRISPR-Cas immunity by elucidating the regulation of cA_6 signaling pathway (Figure 3.12). At an early infection stage, in response to the target RNA binding the StCsm interference complex starts to produce cA_n of various ring sizes. The cA_6 acts as a secondary messenger that allosterically activates RNase activity of StCsm6 through binding at the sensory CARF domain. This initiates RNA degradation. The sensory CARF domain also functions as a ring nuclease that slowly converts cA_6 to A_3>p thereby switching off the activity of HEPN nuclease. This would provide a "timer" mechanism for the HEPN domain activity control.

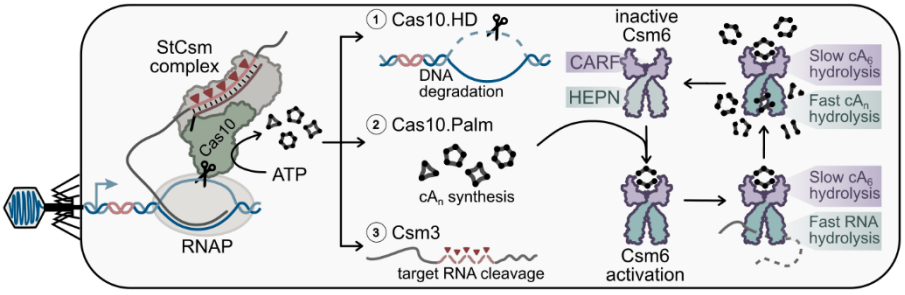


Figure 3.12. An updated interference mechanism of *S. thermophilus* type III-A CRISPR-Cas system. Recognition of the invasive transcript by the StCsm interference complex triggers enzymatic activities of the large subunit Cas10: HD nuclease domain starts to degrade ssDNA formed in the transcription bubble (1), and Palm synthase domain starts to synthesize cA_n from ATP (2). CARF domain of the associated effector StCsm6 binds cA_6 and this initiates nonspecific degradation of RNA by the activated HEPN RNase domain. The StCsm6 activation is regulated by CARF domain which acts as cA_6-specific ring nuclease that slowly converts cA_6 into

linear A_3>p. Moreover, cleavage of the target RNA by Csm3 subunit of StCsm complex (3) switching off cA_n synthesis and accumulated cA_n are degraded by cA_6-activated HEPN domain thereby restoring the pre-infection cell state.

Although the activated HEPN domain shows a clear preference for RNA over cA_n, it can switch to cA_n degradation at low RNA and high cyclic oligoadenylate concentrations, presumably during the later stages of phage infection when local RNA concentration decreases and cA_n concentration increases. After cleavage of the foreign transcript by the StCsm interference complex and the dissociation of cleavage products, the Cas10 subunit is switched off and cA_n synthesis is terminated, while accumulated cA_n are degraded by cA_6-activated HEPN domain. Thus, the signaling molecules are eliminated and the signaling pathway is shut down. Such a self-regulation mechanism would prevent the cell death and ensure the recovery of the host in the post-infection stage. Indeed *S. thermophilus* type III-A system, activation of auxiliary effectors, StCsm6 and StCsm6′, is sufficient and essential for maintaining immunity against phage infection without killing the cell (K. A. Johnson et al., 2024).

3.1.6. Structural insights into cA_6 cleavage

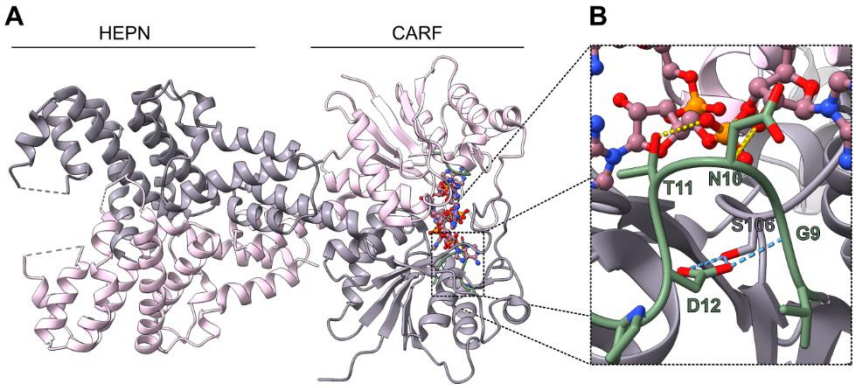


Figure 3.13. Stabilization of the "catalytic loop". (A) Structure of cA_6 bound StCsm6′ PDB ID: 8PE3 (McQuarrie et al., 2023). StCsm6′ is a dimer, monomers comprising a dimer are colored pink and purple. (**B**) Close up view of the catalytic loop (green) of CARF domain of StCsm6′. The side chain OH group of T11 forms hydrogen bond with OH group of scissile phosphate, N10 main chain amide interacts with 2′-OH of the ribose (yellow dotted lines). The side chain of D12 may form stabilizing hydrogen bonds with main chain amide of G9 or with side chain of conserved S106 (cyan dotted lines).

In parallel with our study, on which Chapter 3.1 is based, the article reporting the crystal structure and the cA_6 cleavage by *Enterococcus italicus*

Csm6 (EiCsm6) was published (Garcia-Doval et al., 2020). In addition, a comprehensive structural characterization of StCsm6' was reported following our study (McQuarrie et al., 2023). In general, the results of both studies on Csm6 homologues are in line with our findings. It was confirmed that both CARF and HEPN domains are involved in cA_6 degradation, with the major contribution, at low cA_6 concentrations, being the CARF ring nuclease activity (Garcia-Doval et al., 2020; McQuarrie et al., 2023).

In addition, the structures of activator-bound StCsm6′ and EiCsm6 suggest the mechanism of cA_6 cleavage by the CARF domain. The protein itself does not promote cA_6 cleavage by deprotonation of the nucleophile or protonation of the leaving group and the cA_6 cleavage is thought to be mediated by steric factors that force the ligand to adopt a conformation compatible with in-line nucleophilic attack of the ribose 2′-OH on the scissile phosphate (Garcia-Doval et al., 2020; McQuarrie et al., 2023).

Interestingly, a conserved residue D12, which we have mutated and observed inactivation of the ring nuclease activity of the CARF domain of the StCsm6, does not form direct contact with cA_6 in homologous proteins (Garcia-Doval et al., 2020; McQuarrie et al., 2023). However, D12 is a part of the conserved G-X-[T/S]-DP motif (Figure 3.9 C) present in a loop following the first beta strand of the CARF domain (8-14 aa), hereafter referred to as the catalytic loop. In the StCsm6′-cA_6 structure the catalytic loop residues N10 and T11 form hydrogen bonds with cA_6 utilizing main chain or ride chain atoms, respectively (McQuarrie et al., 2023). These contacts probably force the cA_6 into appropriate conformation for hydrolysis, thus, the conformation of the catalytic loop might be the driving force for the catalysis. D12 potentially could be a key player in keeping the catalytic loop in place as D12 may form hydrogen bonds with the main chain amide of G9 or the side chain of conserved S106 (S112 in EiCsm6; S105 in StCsm6) (Figure 3.13), these bonds could be essential in stabilizing the catalytic loop. This would explain our observation that upon mutating D12 to alanine the CARF domain loses its ring nuclease activity (Figure 3.9 and 3.11). Furthermore, S105A mutation of StCsm6 (would correspond to S106 in StCsm6′ and S112 in EiCsm6) has been reported to diminish the cA_6 binding (Kazlauskiene et al., 2017), further supporting the possible stabilization of the catalytic loop by hydrogen bond between aspartate and serine.

3.2. The mechanism of a tripartite effector

In recent years, numerous new type III CRISPR-Cas CARF-effectors have been characterized (reviewed in Chapter 1.1.3), with even more predicted

through bioinformatics (Altae-Tran et al., 2023; Hoikkala et al., 2024; Makarova et al., 2020a). Among them, CARF-nucleases remain the most extensively studied, while the molecular mechanisms of SAVED-domain effectors have remained largely unexplored.

As a second focus of my doctoral studies, we investigated the protease Lon-SAVED effector together with two additional proteins encoded by the same operon in the type III-A CRISPR-Cas system of *Candidatus* Cloacimonas acidaminovorans (CCa).

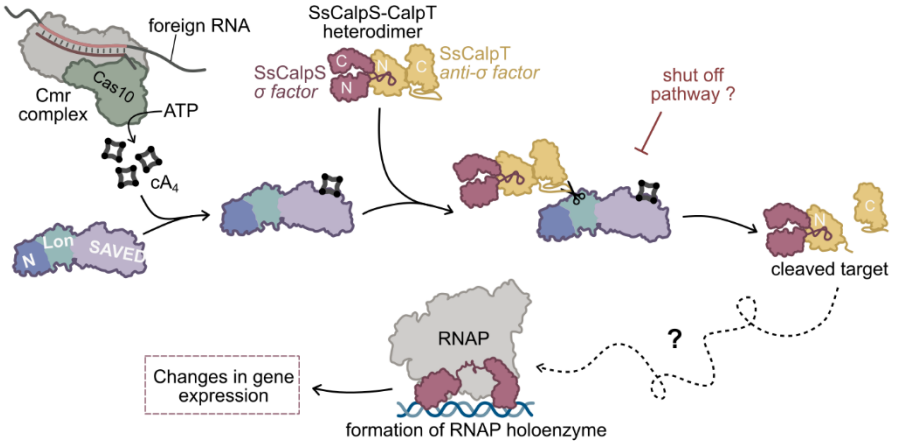


Figure 3.14. Proposed mechanism of the *Sulfurihydrogenibium* spp. type III-B CRISPR-Cas associated tripartite effector CalpL-CalpT-CalpS. Detection of the invader RNA triggers cA_4 synthesis by the type III-B CRISPR-Cas complex. Binding of cA_4 to the SAVED domain of SsCalpL triggers the protease activity of the Lon domain. Activated SsCalpL cleaves the anti-σ factor CalpT within the SsCalpT-CalpS heterodimer. This cleavage somehow leads to the release of the σ factor SsCalpS to bind DNA-directed RNA polymerase (RNAP).

During our study Rouillon et al. characterized a homologous effector from *Sulfurihydrogenibium spp*. YO3AOP1 (Ss) (Rouillon et al., 2023). They identified a Lon-SAVED protein, named CalpL (CRISPR-Cas associated Lon protease), as part of a tripartite effector system, together with the anti-σ factor CalpT and the ECF-like σ factor CalpS. They showed that upon activation by cA_4, SsCalpL protease cleaves SsCalpT to release the σ factor SsCalpS, which then binds to RNA polymerase (RNAP) and alters gene expression (Figure 3.14) (Rouillon et al., 2023). However, the mechanism of CalpT cleavage by CalpL, the sequence of events leading to the release of CalpS once CalpT is cleaved, and the regulation of this tripartite system remained unresolved.

To address these gaps, we conducted an in-depth characterization of CalpL-CalpT-CalpS system from *Candidatus* Cloacimonas acidaminovorans (CCa), with a particular focus on its regulatory mechanism.

3.2.1. CCaCalpL-CalpT-CalpS effector system

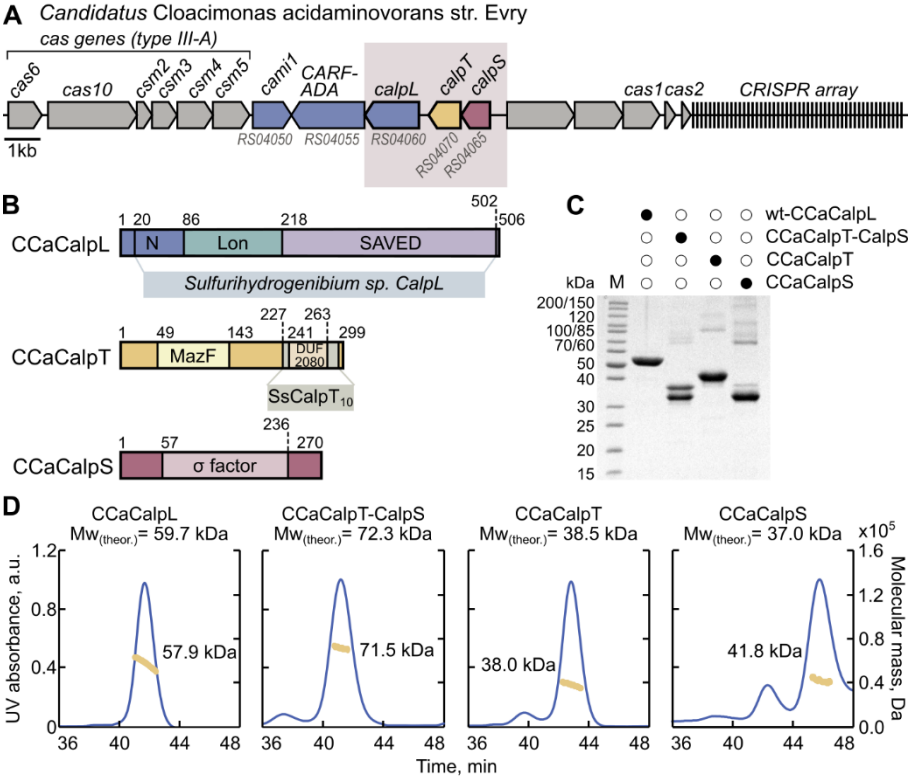


Figure 3.15. Tripartite effector system encoded by *Ca*. Cloacimonas acidaminovorans. (**A**) Schematic representation of the *Ca*. Cloacimonas acidaminovorans type III-A CRISPR-Cas locus (NCBI ID: NC_020449.1). Effector protein genes are colored blue. The pink background highlights genes of the *calpL-calpT-calpS* effector system. (**B**) Schematic representation of CCaCalpL, CCaCalpT, and CCaCalpS proteins with regions of homology identified using HHpred. (**C**) SDS-PAGE of purified wt-CCaCalpL, CCaCalpT-CalpS (CalpS C-tagged), CCaCalpT (C-tagged), CCaCalpS (C-tagged) proteins. M, protein molecular weight marker. (**D**) SEC-MALS analysis of purified proteins: CCaCalpL (C-tag removed by TEV), CCaCalpT-CalpS (C-tag on CCaCalpS), CCaCalpT (C-tag), and CCaCalpS (C-tag).

The type III-A CRISPR-Cas system of *Ca*. C. acidaminovorans possesses an extended arsenal of auxiliary effectors. It encodes a characterized cA_4-activated auxiliary effector Cami1 (Mogila et al., 2023), a homolog of Cad1 (Baca et al., 2024a; Li et al., 2025), and a homolog of the tripartite effector that we were interested in (Rouillon et al., 2023) (Figure 3.15 A). The homology search confirmed that *CLOAM_RS04060* encodes a Lon-SAVED fusion protein (Figure 3.15 B), which shares 26.5% identity and 46.6% similarity in amino acid sequence with SsCalpL, thus named CCaCalpL. Correspondingly, *CLOAM_RS04065* and *CLOAM_RS04070* were confirmed

to encode CCaCalpT (CCaCalpT, 42.7% similarity to SsCalpT) and CCaCalpS (CCaCalpS, 33.9% similarity to SsCalpS). We obtain the components of the CCaCalpL-CalpT-CalpS effector system by expressing the CCaCalp proteins and purifying them from *E. coli* (Figure 3.15 C). Subsequent SEC with multi-angle light scattering (SEC-MALS) analysis revealed that individually purified CCaCalpL, CCaCalpT and CCaCalpS are monomers in solution, but co-expressed together CCaCalpT and CCaCalpS form a stable heterodimer (Figure 3.15 D).

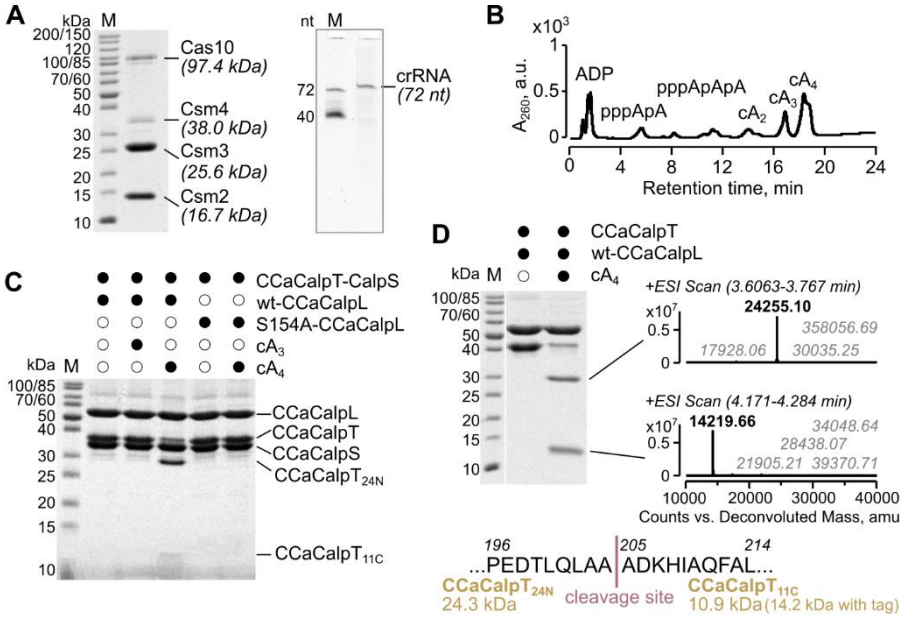


Figure 3.16. Identification of CCaCalpL activator and CCaCalpT cleavage position. (**A**) Composition of purified CCaCsm interference complex. Protein composition of the complex was analyzed by SDS-PAGE and Coomassie staining (left). M, protein molecular weight marker. Nucleic acid co-purifying with the complex was analyzed by denaturing PAGE and staining with SybrGold (right). M, DNA oligonucleotides of a known length, used as length markers. (**B**) The HPLC-MS analysis of cA_n produced by the CCaCsm complex in the presence of target RNA. (**C**) CCaCalpL protease activity analyzed by SDS-PAGE and Coomassie staining. 120 min reactions contained 5 μM CCaCalpL (wt or mutant), 5 μM CCaCalpT-CalpS and 0 or 25 μM cA_4 or cA_3. (**D**) Identification of the CCaCalpL cleavage site within the CCaCalpT. Reaction products were analyzed by SDS-PAGE and Coomassie staining and by HPLC-MS. Depiction of CCaCalpT cleavage site in polypeptide chain is shown at the bottom.

As we sought to identify the activator recognized by the SAVED domain of CCaCalpL, we first examined cA_n production by the native CCaCsm complex *in vitro*. We expressed the CCaCsm interference complex in *E. coli* and purified it by affinity chromatography (Figure 3.16 A). Although the

purified complex lacked the Csm5 subunit, which is important for the target RNA binding (Mogila et al., 2019), and contained an unmature 72 nt crRNA, in the presence of a high target RNA concentration (10 µM), the complex synthesized cA_3 and cA_4, with the major product being cA_4 (Figure 3.16 B). In accordance, cA_4 has been previously found to activate CCaCami1 - one of *Ca*. C. acidaminovorans effectors (Mogila et al., 2023). However, since different accessory effectors can be activated by different cA_n in the same bacteria (Grüschow et al., 2024, 2021), we tested both cA_3 and cA_4 as potential activators for CCaCalpL. We subjected CCaCalpT-CalpS to the CCaCalpL proteolysis reaction and SDS-PAGE analysis revealed that cA_4, but not cA_3 activated the CCaCalpL protease to cleave the CCaCalpT protein in the CCaCalpT-CalpS heterodimer (Figure 3.16 C). Furthermore, mutation of the conserved S154 to alanine in the predicted Lon protease active site completely abolished the CCaCalpL protease activity (Figure 3.16 C).

To identify the precise cleavage position within the CCaCalpT protein, we subjected the C-tagged version of a CCaCalpT to cleavage by CCaCalpL and analyzed the cleavage products using HPLC-MS (Figure 3.16 D). CCaCalpL cleaved CCaCalpT in an alanine-rich motif ^{202}LAAA between the amino acids A204 and A205, yielding a 24.3 kDa N-terminal fragment ($CCaCalpT_{24N}$), and a 10.9 kDa C-terminal fragment ($CCaCalpT_{11C}$).

We next tested whether CCaCalpS is able to interact with the DNA-directed RNA polymerase (RNAP) as a σ factor should (Figure 3.15 B). Although we do not observe copurification of *E. coli* RNAP (EcRANP) when purifying the C-terminally tagged CCaCalpS protein (Figure 3.15 C), the superimposition of the AlphaFold2 model of CCaCalpS onto σ^E in the *E. coli* RNAP/σ^E transcription initiation complex (PDB ID: 6JBQ) (Fang et al., 2019), showed that the N- and C-terminal domains of CCaCalpS aligne with the promoter-binding domains σ_2 and σ_4 of σ^E, respectively (Figure 3.17 A). Therefore, we further investigated the CCaCalpS-EcRNAP interaction *in vitro* using the biolayer interferometry (BLI) technique (Figure 3.17 B and C). We immobilized C-tagged CCaCalpS on the Ni^{2+}-NTA biosensor and exposed it to the EcRNAP core enzyme, which resulted in a wavelength shift, confirming the interaction between CCaCalpS and EcRNAP (Figure 3.17 C). No EcRNAP binding occurred when the full-length CCaCalpT-CalpS heterodimer was immobilized on the biosensor (Figure 3.17 C), suggesting that a full-length CCaCalpT protein acts as an anti-σ factor and blocks the interaction between CCaCalpS and RNAP. Interestingly, immobilization of a pre-cleaved CCaCalpS-CalpT heterodimer did not result in EcRNAP binding (Figure 3.17 C). This may indicate that cleavage alone is not sufficient to disrupt the interaction of the proteins in the CCaCalpT-CalpS complex. Taken

together, these results confirm that CCaCalpL-CalpT-CalpS is a tripartite cA_4-activated effector system, in which CCaCalpL is a cA_4-dependent protease that cleaves CCaCalpT in the CCaCalpT-CalpS heterodimer. However, the subsequent release of CCaCalpS needed to be further investigated.

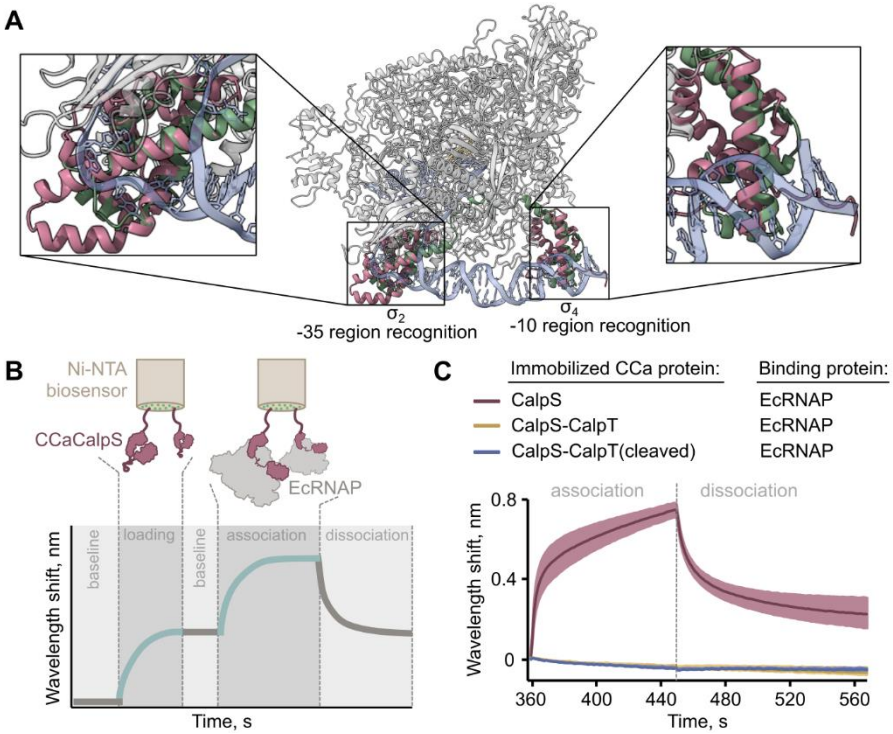


Figure 3.17. A σ factor like properties of CCaCalpS. (**A**) AlphaFold2 model of CCaCalpS superimposed on *E. coli* RNAP/σ^E transcription initiation complex (PDB ID: 6JBQ) (Fang et al., 2019). EcRNAP is colored light gray, σ^E – green, DNA – blue, RNA – yellow, CCaCalpS – red. The linker region between the N-part and C-part of CCaCalpS has been cut to allow the superposition onto the σ^2 and σ^4 domains of σ^E. (**B**) BLI experiment scheme of CCaCalpS binding to the EcRNAP. CCaCalpS, CCaCalpT-CalpS, or cleaved CCaCalpT-CalpS were immobilized on Ni-NTA biosensor (loading) to monitor RNAP binding (association). (**C**) Association and dissociation curves of the CCaCalpS binding to the EcRNAP BLI experiment. Data points are mean values from three experiments. The error bars indicate standard deviation of three measurements.

3.2.2. The mechanism of CCaCalpS release

The AlphaFold2 model of the CCaCalpT-CalpS heterodimer suggests that CCaCalpT interacts with CCaCalpS exclusively through the N-terminal portion of the protein, and that the cleavage site within CCaCalpT is distant from the CCaCalpT-CalpS interface (Figure 3.18 A). This may explain why no binding of EcRNAP was observed for the pre-cleaved CCaCalpT-CalpS

heterodimer (Figure 3.17 C). Therefore, we further tested whether CCaCalpS is released from the heterodimer after the CCaCalpT cleavage using magnetic beads to pull down the proteins that remain associated with CCaCalpS. Remarkably, CCaCalpS remained bound to both $CCaCalpT_{24N}$ and $CCaCalpT_{11C}$ (Figure 3.18 B), indicating that the CCaCalpT likely requires further processing in the cell to release CCaCalpS, which can then bind to RNAP.

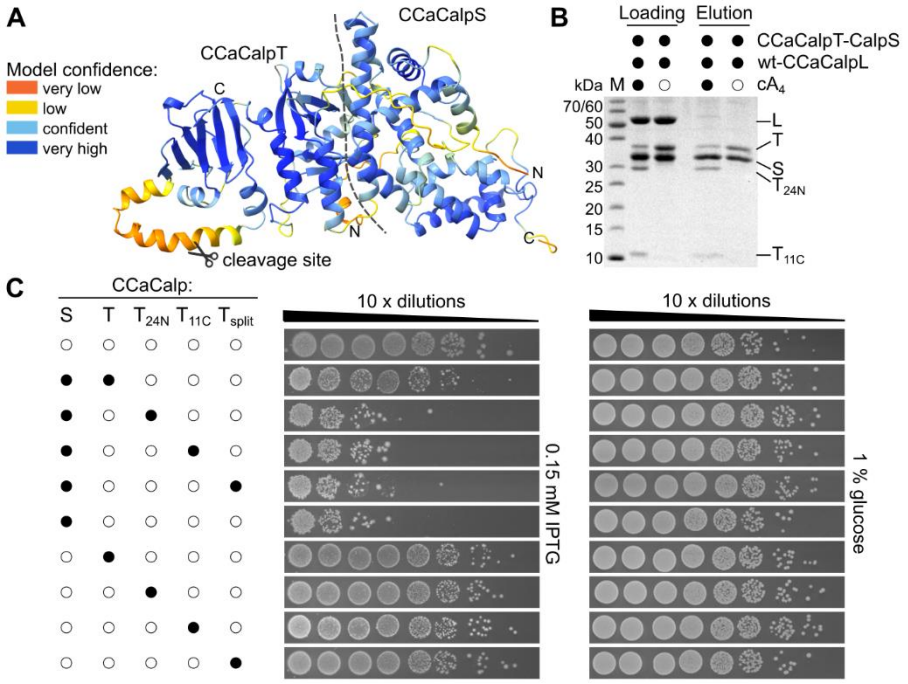


Figure 3.18. Studies on the release of CCaCalpS post-CCaCalpT clevage *in vitro* and in *E. coli*. (**A**) AlphaFold2 model of CCaCalpT-CalpS heterodimer colored by model confidence score: very low (pLDDT <50); low (70> pLDDT >50); confident (90> pLDDT >70); very high (pLDDT >90). Dotted line indicates interface. (**B**) Testing the CCaCalpT-CalpS heterodimer dissociation following CCaCalpL cleavage by pull-down assay using magnetic beads. Eluted CCaCalpS with tightly bound proteins was analyzed by SDS-PAGE. M, protein molecular weight marker. (**C**) Serial dilutions of *E. coli* expressing combinations of CCaCalpS with full-length CCaCalpT, $CCaCalpT_{24N}$, $CCaCalpT_{11C}$ or split CCaCalpT plated on LB media with 0.15 mM IPTG or 1% glucose.

Since we observed that expression of CCaCalpS caused a reduction in cell growth during cultivation of *E. coli* for the CCaCalpS purification, we reasoned that CCaCalpS, as a σ factor, initiates the transcription of some toxic genes or interferes with transcription by sequestering RNAP. Either way, the toxicity of CCaCalpS could be a detectable signal, allowing us to explore the effects of CCaCalpT cleavage in *E. coli* cells. We constructed an IPTG-

inducible vector system encoding *CCacalpS* and a copy of either (i) the full-length *CCacalpT*, (ii) the *CCacalpT*$_{24N}$, (iii) the *CCacalpT*$_{11C}$, or (iv) a split version of *CCaCalpT* that would be formed after its cleavage with CCaCalpL. We co-transformed *E. coli* with pairs of vectors, performed serial dilutions, and plated the cells on LB media supplemented with either IPTG or glucose. *E. coli* expressing CCaCalpS alone formed fewer colonies, while co-expression of full-length CCaCalpT reduced the toxicity of CCaCalpS (Figure 3.18 C). No toxicity was observed when CCaCalpT or its cleavage variants were expressed alone (Figure 3.18 C), confirming that CCaCalpS is a toxic component of the CCaCalpT-CalpS pair and further supporting the hypothesis that CCaCalpT functions as an anti-σ factor, preventing the interaction between the σ factor CCaCalpS and RNAP.

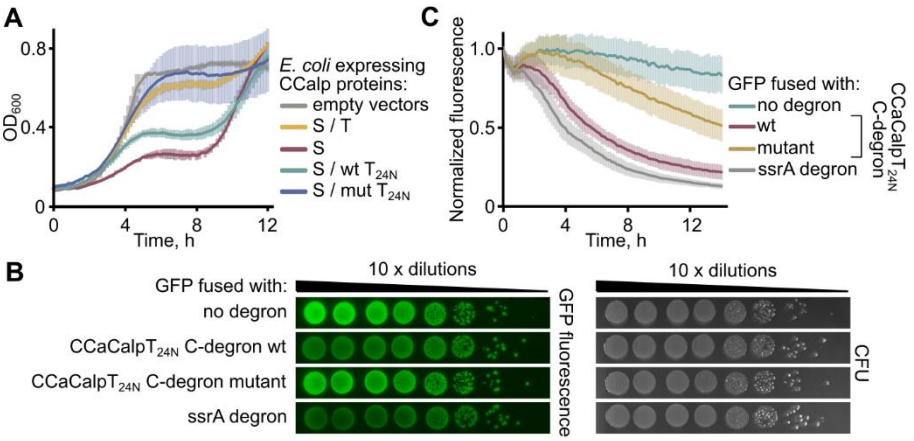


Figure 3.19. Investigation of CCaCalpT$_{24N}$ C-degron. (**A**) Growth curves of *E. coli* expressing CCaCalpS and wt CCaCalpT$_{24N}$ or mutant CCaCalpT$_{24N}$. All data points are presented as the mean values of three experiments, and the standard deviation is indicated by error bars. (**B**) GFP fluorescence interference assay in *E. coli. E. coli* cells, transformed with plasmids encoding GFP fused to different degrons (PETLQLAA – wt CCaCalpT$_{24N}$ C-degron, PETLQLAD - mutant CCaCalpT$_{24N}$ C-degron and ENYALAA - partial fragment of ssrA C-degron), were serially diluted and plated on LB medium with IPTG. CCaCalpT$_{24N}$ C-degron reduced GFP fluorescence (left) without affecting colony forming units (CFU) (right). (**C**) Quantitative degradation assay of GFP fused with degrons as in B in liquid *E. coli* cultures over time. The fluorescence signal of GFP was normalized to optical density. All data points are presented as the mean values of three experiments, and the standard deviation is indicated by error bars.

In a striking contrast to the *in vitro* data (Figure 3.17 C), the presence of CCaCalpT cleavage products had no effect on CCaCalpS toxicity in cells (Figure 3.18 C), suggesting that CCaCalpS is released from the CCaCalpT cleavage products *in vivo*. We hypothesized that the CCaCalpT$_{24N}$ may carry a degron at its newly formed C-terminus that targets it for degradation by

cellular proteases, thereby releasing CCaCalpS, a strategy observed in other ECF-like σ factor-anti-σ factors, such as RseA-σ^E (Ades, 2008; Flynn et al., 2004) (see Chapter 1.5.2). Based on the AlphaFold2 model (Figure 3.18 A), we predicted that the terminal 9 amino acids ^{196}PEDTLQLAA^{204} of CCaCalpT$_{24N}$ might be poorly structured and comprise the C-degron, as unstructured peptides ending with alanines at the C-terminus are commonly found in degrons that target proteins for degradation by the ClpXP unfoldase-protease complex (Fei et al., 2020; Flynn et al., 2004) (see Chapter 1.5.2). To test the CCaCalpT$_{24N}$ C-degron hypothesis, we mutated the C-terminal A204 residue in the putative CCaCalpT$_{24N}$ C-degron to aspartate, since this type of mutation in the terminal AA motif of a well-characterized C-terminal SsrA-degron significantly reduces the efficiency of target protein degradation by ClpXP (Flynn et al., 2001), and examined its effect on the CCaCalpS toxicity in *E. coli* by registering growth curves in liquid LB after IPTG induction (Figure 3.19 A). Cells expressing CCaCalpS alone or co-expressing CCaCalpS with the wt-CCaCalpT$_{24N}$ exhibited poor growth. However, co-expression of CCaCalpS with the CCaCalpT$_{24N}$ with mutated C-degron resulted in normal *E. coli* growth, comparable to cells coexpressing CCaCalpS and the full-length CCaCalpT.

To further confirm the presence of the degron, we employed the GFP reporter system, using expression vectors encoding *gfp* fused at the C-terminus to: (i) the wt C-degron of CCaCalpT$_{24N}$ (PEDTLQLAA), (ii) the mutant C-degron of CCaCalpT$_{24N}$ (PEDTLQLAD), and (iii) a partial SsrA C-degron (ENYALAA) as a positive control (Fei et al., 2020). First, *E. coli* cells were transformed with these vectors and serial dilutions of cells were plated on LB containing IPTG. While cell survival was not affected, a decrease in GFP fluorescence was observed in cells expressing GFP fused with wt-C-degron, but not to mutant C-degron (Figure 3.19 B). In addition, we confirmed the results of our spot assay by quantitatively analyzing the fluorescence levels of GFP-C-degron in liquid culture over time (Figure 3.19 C). These results demonstrate that the CCaCalpL cleavage product CCaCalpT$_{24N}$ is further degraded by cellular proteases and only after the degradation of CCaCalpT$_{24N}$, CCaCalpS is released from the heterodimer, causing growth arrest in *E. coli*.

3.2.2. Regulatory mechanism – CCaCalpL is a ring nuclease

Next, we focused on the potential regulation of the CCaCalpL-CalpT-CalpS effector. As our study on StCsm6 and other recent studies on auxiliary effectors have revealed that CARF domains often possess intrinsic ring nuclease activity (Athukoralage et al., 2019; Du et al., 2024; Garcia-Doval et

al., 2020; Jia et al., 2019b; Li et al., 2025; McQuarrie et al., 2023; Mogila et al., 2023), and more importantly, the CARF domain of the neighboring effector CCaCami1 possesses a cA_4-specific ring nuclease activity (Mogila et al., 2023), we hypothesized that type III CRISPR-Cas-associated effectors in *Ca*. C. acidaminovorans are not intended for abortive infection, and hence the SAVED domain of CCaCalpL may also degrade its activator cA_4, in a manner similar to CARF ring nucleases. We incubated a radioactively labeled cA_4, produced by the CCaCsm complex, with a CCaCalpL and analyzed the reaction products by denaturing PAGE. We found that CCaCalpL does indeed cleave cA_4, and no differences in cA_4 cleavage were observed in the presence of Mg^{2+}, Mn^{2+}, or EDTA (Figure 3.20 A). We further confirmed that this ring nuclease activity is specific to cA_4 by incubating CCaCalpL with cA_3, cA_4, and cA_6 and analyzing reaction products by HPLC-MS. As expected, CCaCalpL degraded only cA_4 and three cA_4 cleavage products were detected: A_4p (linear tetraadenylate with a 3′-phosphate), A_2>p (linear diadenylate with a 2′,3′-cyclic phosphate), and A_2p (linear di adenylate with a 3′-phosphate) (Figure 3.20 B). This demonstrates that CCaCalpL is a metal-independent ring nuclease specific to cA_4.

To further investigate the ring nuclease activity of CCaCalpL, we performed a time course experiment under different conditions. We found that under multiple turnover conditions ([S] > [E], 10 μM cA_4 and 1 μM CCaCalpL), cA_4 is first converted into the linear reaction intermediate A_4>p (linear tetraadenylate with a 2′,3′-cyclic phosphate), followed by the conversion of A_4>p into the second linear intermediate A_4p, which is eventually cleaved into equimolar amounts of linear dinucleotides A_2>p and A_2p (Figure 3.20 C right gel) (the identity of the products was confirmed by HPLC-MS analysis). The accumulation of large amounts (up to 9 μM or up to 90% of all reaction products, i.e., 9-fold higher than the CCaCalpL concentration) of both A_4>p and A_4p intermediates implies that (i) the reaction product dissociates from the CCaCalpL enzyme after each cleavage step, and (ii) the CCaCalpL reaction rates decrease by more than 10-fold for each subsequent reaction step. Indeed, fitting a set of equations describing a three-step sequential reaction yielded enzyme turnover rate estimates (k_{obs}) for all three stages: 34 min^{-1} for cA_4 conversion to A_4>p, 2.2 min^{-1} for A_4>p conversion to A_4p, and 0.10 min^{-1} for A_4p conversion to A_2>p + A_2p (Figure 3.20 C bottom graph). Dr. Giedrius Sasnauskas wrote the equations and fitted the data to them. Under single turnover conditions ([E] > [S], 1 μM CCaCalpL and 50 nM cA_4) the reaction followed the same mechanism, where cA_4 was relatively rapidly (on a scale of seconds) converted to A_4>p and A_4p, whereas A_4p was relatively slowly (on a scale of tens of minutes) converted into A_2>p

and A_2p (Figure 3.20 C left gel). Furthermore, we found that both A_4>p and A_4p stimulate the protease activity of CCaCalpL to a similar extent as cA_4, whereas CCaCalpL is not stimulated by the final reaction products A_2>p and A_2p (Figure 3.20 D). Taken together, this confirms that CCaCalpL possesses a regulatory mechanism - it cleaves the activator cA_4 and remains activated up until cA_4 is cleaved into final products.

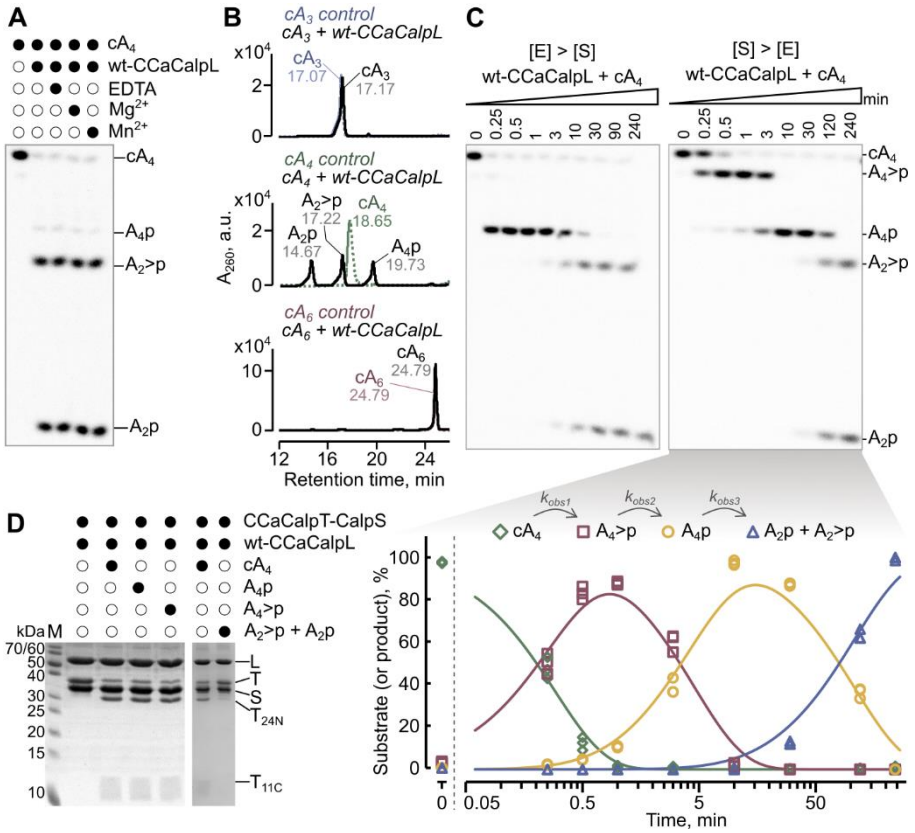


Figure 3.20. Ring nuclease activity of CCaCalpL. (**A**) Cleavage of radioactively labeled cA_4 by wt-CCaCalpL in the presence or absence of divalent metal ions. 90 min reactions contained 1 μM CCaCalpL and 50 nM labeled cA_4 in the reaction buffer supplemented with either 1 mM EDTA, 10 mM $Mg(CH_3COO)_2$ or 1 mM $MnCl_2$. (**B**) HPLC-MS analysis of cA_3, cA_4 and cA_6cleavage by wt-CCaCalpL. The reactions contained 1 μM wt-CCaCalpL and 10 μM cA_n. Colored dotted lines represent cA_n controls (cA_3 - blue, cA_4 - green, cA_6 - red). (**C**) Cleavage of labeled cA_4 by wt-CCaCalpL under single turnover ([E]>[S]) (left gel) and multiple turnover ([S]>[E]) (right gel) conditions analyzed by denaturing PAGE. Reactions contained 1 μM wt-CCaCalpL and 50 nM labeled cA_4. Multiple turnover reactions were additionally supplemented with 10 μM cA_4. Time course of cA_4, intermediate products A_4>p and A_4p and final products A_2>p and A_2p under multiple turnover conditions from three experiments are plotted (bottom graph). Averaged exponential fits (solid lines) to

substrate cleavage data were used to determine reaction rates: k_{obs1}= 34 ± 2 min^{-1}, k_{obs2}= 2.2 ± 0.1 min^{-1} and k_{obs3}= 0.092 ± 0.003 min^{-1}. (**D**) Wt-CCaCalpL protease reactions using cA_4, A_4>p, A_4p or A_2>p and A_2p as activators. 120 min reactions contained 5 μM CCaCalpL, 5 μM CCaCalpT-CalpS and 0 or 25 μM of adenylates, analyzed by SDS-PAGE and Coomassie staining. M, protein molecular weight marker.

3.2.3. The active site of the CCaCalpL ring nuclease

To gain more insight into the mechanism of cA_4 cleavage, we characterized CCaCalpL structurally by performing cryogenic electron microscopy (cryo-EM) imaging of wt-CCaCalpL samples in the presence of the activator cA_4 or its cleavage intermediates A_4>p or A_4p (Appendices 4-7). Cryo-EM was performed by Dr. Giedrius Sasnauskas and Dr. Giedrė Tamulaitienė. Cryo-EM analysis revealed that cA_4, A_4>p and A_4p-bound CCaCalpL proteins stack together and form ordered short filamentous structures (Figure 3.21 A), resembling several other SAVED domain-containing proteins, which form filaments upon activator molecule binding (Hogrel et al., 2022; Lowey et al., 2020; Steens et al., 2024).

The best quality cryo-EM map (2.75 Å resolution) was obtained for the A_4p-bound wt-CCaCalpL, which allowed us to model two adjacent full-length CCaCalpL subunits and a SAVED domain of the third subunit, with A_4p molecules bound between the adjacent SAVED domains (Figure 3.21 B right). The 2.97 Å map of the cA_4-bound CCaCalpL and 2.81 Å map of the A_4>p-bound CCaCalpL allowed us to model only the SAVED domains from two (cA_4 structure) or three (A_4>p structure) adjacent CCaCalpL subunits (Figure 3.21 B left and middle). Model building was performed by Dr. Giedrius Sasnauskas and Dr. Giedrė Tamulaitienė. There is no observable difference between the structures, and in the filament, the CCaCalpL monomers oligomerize in a head-to-tail fashion, with the adjacent SAVED domains sandwiching the activator between the opposite surfaces in a manner similar to other SAVED effectors (Figure 3.21 B) (Fatma et al., 2021; Hogrel et al., 2022; Rechkoblit et al., 2024; Steens et al., 2024). The first (primary) CCaCalpL subunit binds the activator in a deep pocket, while the adjacent (secondary) subunit CCaCalpL′ utilizes a raised patch with two phenylalanine "fingers" on the opposite domain side (Figure 3.21 C). The cA_4 ring adopts a distorted rectangular shape with different side lengths (Figure 3.21 D and E). All four adenine bases are inserted into tight protein pockets where they are stabilized by numerous contacts with amino acids from the primary CCaCalpL subunit, and R358′ from the adjacent CCaCalpL′ subunit (Figure 3.21 D). The

protein also forms extensive contacts to the phosphodiester backbone (Figure 3.21 E). Taken together, the structures explain the specificity of CCaCalpL to tetra-adenylates and show how sandwiching of cA_4 or its cleavage products A_4>p/A_4p between the SAVED domains enables CCaCalpL oligomerization.

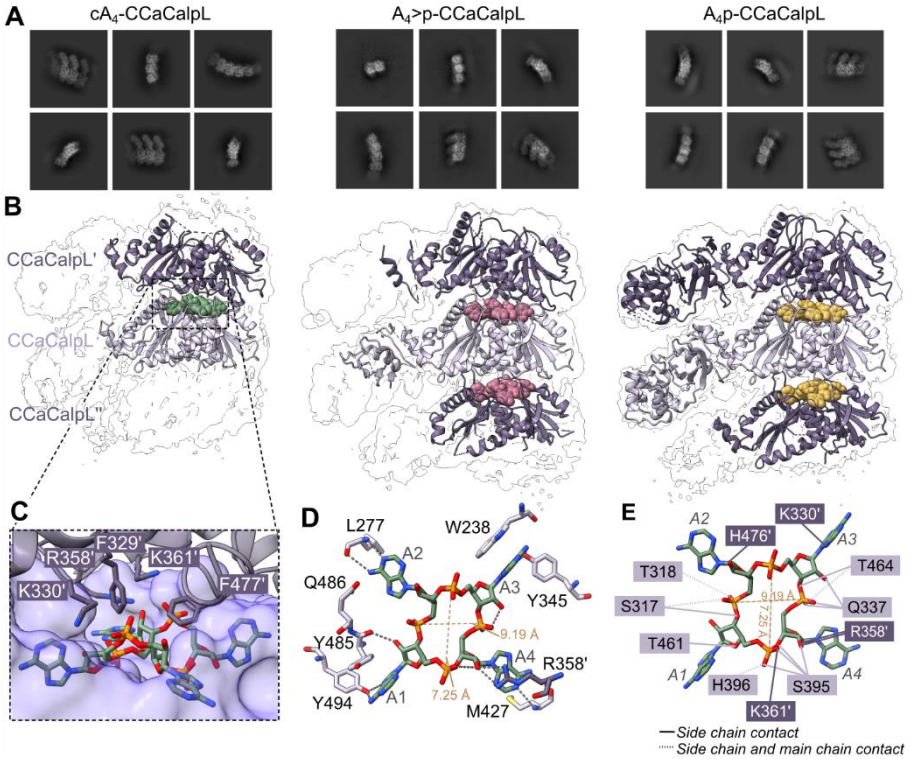


Figure 3.21. Structures of activator-bound CCaCalpL. (**A**) Representative 2D classes of wt-CCaCalpL filament bound to cA_4, A_4>p and A_4p. (**B**) Atomic model of CCaCalpL bound to cA_4 (green), A_4>p (red) and A4p (yellow). Adjacent proteins are colored dark and light purple. Atomic models are depicted on electron density maps. (**C**) Interactions between the neighboring SAVED domains and the sandwiched cA_4. Primary CCaCalpL subunit (light purple) binds cA_4 in a deep pocket, while the adjacent CCaCalpL′ subunit (dark purple) forms fewer contacts with cA_4. (**D**) Amino acid residues contacting the bases of cA_4: light purple for CCaCalpL, dark purple for the adjacent CCaCalpL′. (**E**) CCaCalpL contacts with the phosphodiester backbone of cA_4. Solid lines depict side chain contacts, dotted lines - side chain and main chain contacts. The residues from the CCaCalpL subunit are depicted in light purple squares, the residues of the adjacent CCaCalpL′ subunit are shown in dark purple squares.

Based on the structures of cA_4/A_4>p/A_4p-bound wt-CCaCalpL filaments, we have identified residues S395 and H396 of the primary CCaCalpL subunit and R358′ and K361′ of the secondary CCaCalpL′ as likely candidates for the catalytic center of the ring nuclease (Figure 3.22 A). These residues are in the

vicinity of the scissile phosphate, a phosphate between the 4th and 1st adenosines (A4 and A1), and based on their positions, they could play the following roles. Since R358′ makes hydrogen bonds with A4 2′-OH (Figure 3.22 A and Appendix 11), it presumably positions 2′-OH for a nucleophilic attack on the adjacent phosphodiester bond, resulting in the formation of a 2′,3′-cyclic phosphate at A4 and a 5′-OH group at A1. S395, H396 and K361′ form hydrogen bonds with the scissile phosphate (Figure 3.22 A and Appendix 11), thus they all may contribute to stabilization of the pentavalent transition state. The H396 residue is also well positioned for protonation of the 5′-O leaving group and is observed associated with 5′-OH in the CCaCalpL structures with A_4>p and A_4p products (Figure 3.22 A and Appendix 11).

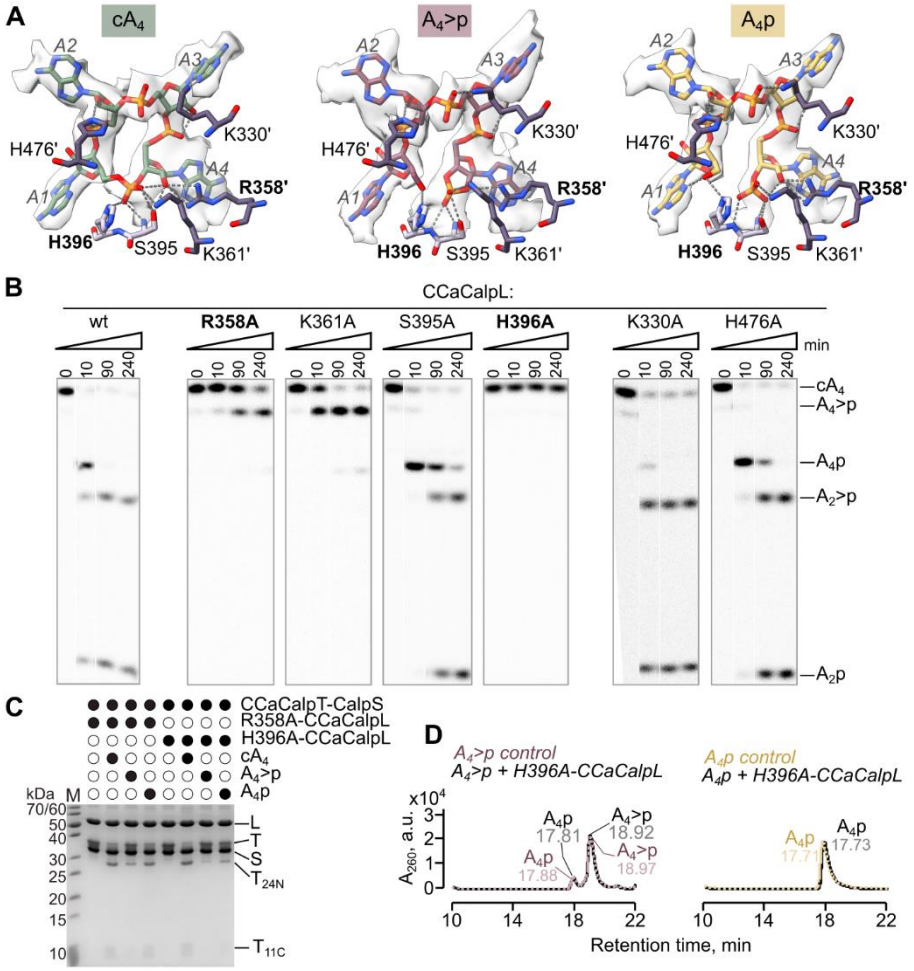


Figure 3.22. Determining the active site of SAVED ring nuclease. (**A**) Potential catalytic residues interacting with cA_4, A_4>p and A_4p. See also Appendix 11. (**B**) Cleavage of labeled cA_4 by CCaCalpL mutants of amino acids located near A1-A4 scissile phosphate: R358A, K361A, S395A, H396A, of amino acids located near A2-

A3 scissile phosphate: K330A and H476A under single turnover conditions analyzed by denaturing PAGE. For comparison the reaction with wt-CCaCalpL is also shown. The reactions contained 1 μM CCaCalpL and 50 nM labeled cA_4. For ease of comparison, the wt CCaCalpL gel from Figure 3.20 C is reused. (**C**) Protease reaction performed with R358A and H396A-CCaCalpL using cA_4, A_4>p or A_4p as activators. 120 min reactions contained 5 μM CCaCalpL, 5 μM CCaCalpT-CalpS and 0 or 25 μM cA_4, A_4>p or A_4p, analyzed by SDS-PAGE and Coomassie staining. M, protein molecular weight marker. (**D**) Cleavage reactions of intermediates A_4>p and A_4p by H396A-CCaCalpL analyzed by HPLC-MS.

Next, we constructed and purified the mutant CCaCalpL variants and evaluated the involvement of all the above residues in catalysis. R358A and H396A mutations had the strongest effect on cA_4 cleavage, with H396A mutant being completely inactive (Figures 3.22 B). Furthermore, protease activity of R358A-CCaCalpL was still stimulated by cA_4 and A_4>p, and to a lesser extent by A_4p, whereas protease activity of H396A-CCaCalpL was stimulated by cA_4, but was negligible in the presence of either A_4>p and A_4p (Figure 3.22 C), implying that both mutants had diminished ability to bind the linear tetra-adenylates. This implies that H396 and R358′ coordinates the open 5′-OH and 2′,3′-cyclic phosphate/3′-phosphate termini of the A_4>p and A_4p intermediates. Thus, the active site with H396 and R358′ catalytic residues cleave cA_4 and converts A_4>p to A_4p.

However, the mechanism of cleavage of the A2-A3 scissile phosphate and conversion of A_4p to the final products A_2>p and A_2p seems to be more complicated. Two different theories can be proposed: (i) the existence of a secondary active site or (ii) occasional A_4p cleavage by the same active site.

The theory of two active sites

In the solved structures there are a clear preferred binding orientation of A_4>p and A_4p intermediates relative to the H396 and R358′ catalytic residues (Figure 3.22 A). If we assume that the CCaCalpL has the second catalytic center close to the A2-A3 scissile phosphate the most likely catalytic residues for this would be H476′ and K330′ from the adjacent SAVED domain. H476′ and K330′ are both located in immediate proximity to the A2-A3 phosphate opposite of H396 (Figure 3.22 A). However, neither K330 nor H476 mutations abolished the formation of the dinucleotide products (Figures 3.22 B), ruling out their direct involvement in the catalysis. On the other hand, some CARF ring nucleases, such as EiCsm6, StCsm6, and StCsm6′, do not directly support cleavage by deprotonation of the nucleophile or protonation of the leaving group, and the catalysis is based solely on the confirmation that the signaling molecule is forced to adopt (McQuarrie et al., 2023). A similar mechanism may be utilized by the second active site in CCaCalpL SAVED domain.

Relatively slow conversion of A_4p to A_2>p + A_2p (Figure 3.20 C) may be attributed to a lower catalytic efficiency of this second catalytic center. This catalytic center, if present, should also cleave cA_4 in the presence of the H396A mutation that inactivates the first catalytic center. Since no cA_4 linearization was observed in the presence of H396A-CCaCalpL (Figure 3.22 B), we must assume that the second catalytic center becomes active only when bound to linear A_4p intermediate, but not the cyclic cA_4. However, no cleavage of intermediate A_4p was detected when incubating with H396A-variant (Figure 3.22 D), but this may be attributed to the poor binding of A_4p by this mutant as it is evident from protease activation assay (Figure 3.22 C). The ability to discriminate bound cA_4 from A_4p, could be attributed to subtle conformational differences between CCaCalpL bound to cA_4/A_4p that are not detectable in our structures due to limited resolution.

The theory of a single active site

We also considered an alternative mechanism utilizing only a single catalytic site with H396 and R358′ catalytic residues for cleavage of both scissile phosphates A1-A4 and A2-A3. In this case it is necessary to presume that the linear intermediate A_4p can bind to the SAVED domain in two orientations: (i) with the open 3′-phosphate and 5′-OH termini facing H396 (the preferred orientation, captured in the A_4p-bound CCaCalpL structure (Figure 3.22A), and (ii) the opposite orientation, which would place the A2-A3 phosphate in the proximity of H396 and R358′, and the 3′-phosphate/5′-OH termini in the proximity of K330′ and H476′. Orientation (ii) must be significantly less populated than (i), accounting for the low cleavage rate of the A_4p intermediate at the A2-A3 phosphate (Figure 3.20 C), and lack of density corresponding to an intact A2-A3 phosphodiester bond in the proximity of H396 in the A_4p-bound CCaCalpL structure (Figure 3.22 A). Nevertheless, the dynamic nature of the CCaCalpL filaments, which allows dissociation of cA_4 cleavage intermediates from the enzyme during the ring nuclease reaction (as discussed above, Figure 3.20 C), may allow multiple rounds of A_4p dissociation and re-binding required for occasional A_4p binding in the (ii) orientation leading to its cleavage. This mechanism would explain why H396A-CCaCalpL variant does not linearize cA_4 nor cleavage the intermediate A_4p (Figure 3.22 B and D).

In summary, we demonstrate here that CCaCalpL H396 and R358′ are the key catalytic residues responsible for the cA_4 ring opening reaction. This active site is formed by the residues from adjacent SAVED subunits within a filament and due to the dynamic nature of the CCaCalpL filaments, the same catalytic center may also be responsible for A_4p cleavage to the final products

A_2>p and A_2p, albeit we cannot completely rule out the presence of a second catalytic center.

3.2.4. The significance of CCaCalpL filament formation

Further, we attempted to evaluate the significance of CCaCalpL filament formation for ring nuclease activity, CCaCalpT-CalpS binding and CCaCalpT cleavage.

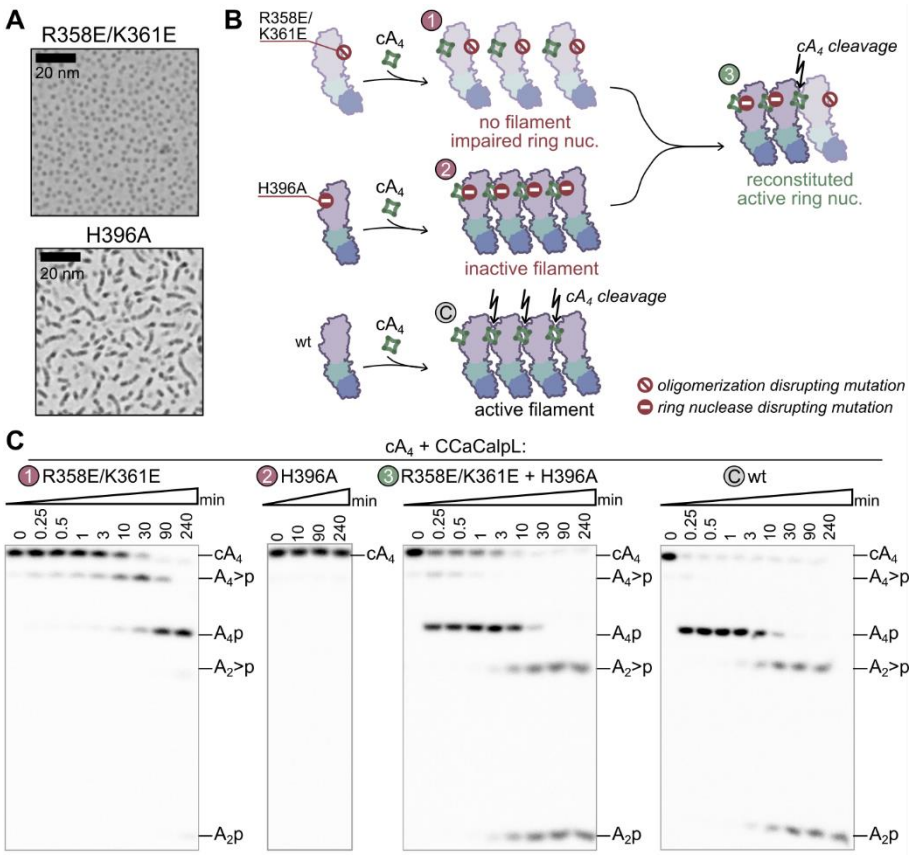


Figure 3.23. The importance of CCaCalpL oligomerization for ring nuclease activity. (**A**) Denoised cryo-EM micrographs of R358E/K361E and H396A CCaCalpL variants in the presence of cA_4. (**B**) Schematic representation of the reconstitution of the CCaCalpL interface for cA_4 cleavage. Mixing of ring nuclease-deficient (2) and filament-deficient (1) CCaCalpL variants should result in formation of dimers or R358E/K361E variant terminated filaments with reconstituted cA_4 ring nuclease active site. (**B**) The ring nuclease activity of the filament-deficient R358E/K361E-CCaCalpL (left) and the reconstituted CaCalpL dimer/terminated filament (R358E/K361E + H396A) (third form the left) analyzed by denaturing PAGE. Reactions were performed under a single turnover conditions and contained 1 μM CCaCalpL and 50 nM cA_4. For ease of comparison, the H396A and wt CCaCalpL gels from Figures 3.22 B and 3.20 C, respectively, are reused.

As the ring nuclease active site residues, H396 and R358′, are derived from adjacent SAVED domains within the CCaCalpL filament (Figure 3.22), we tested whether a restoration of the SAVED-SAVED interface is sufficient for ring nuclease activity. We constructed a filament-deficient CCaCalpL variant by introducing two R358E/K361E charge reversal mutations that disrupt the interface between the SAVED domains (Figure 3.23 A). The R358E/K361E interface mutations reduced the ring nuclease activity of CCaCalpL (>100-fold compared to wt-CCaCalpL), supporting the hypothesis that filament formation is required for the ring nuclease activity (Figure 3.23 C). Next, we combined the ring nuclease-deficient mutant H396A with the filament-deficient R358E/K361E-CCaCalpL and were able to restore cA_4 cleavage activity to wt-like levels by restoring the intersubunit interface of the SAVED domain (Figures 3.23 C). Although, a small amount of uncleaved substrate was observed over the first minutes of the reaction, likely due to cA_4 sequestration by the inactive H396A homo-oligomers, this demonstrates that CCaCalpL oligomerization is essential for efficient cA_4 cleavage.

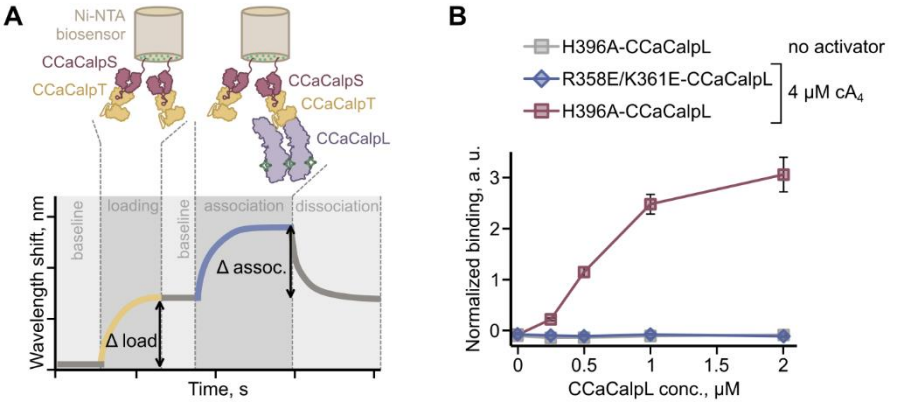


Figure 3.24. The importance of CCaCalpL oligomerization for CCaCalpT-CalpS binding. (**A**) Schematic representation of the setup of the BLI experiment. CCaCalpT-CalpS heterodimer was immobilized onto a Ni-NTA biosensor (loading). The loaded sensor was used to monitor CCaCalpL binding at various concentrations (association). (**B**) The normalized binding response equals (Δassociation)/(Δloading). All data points are presented as mean values from three experiments. The error bars indicate the standard deviation of three experiments.

Next, we tested the importance of CCaCalpL filament formation for protease target binding. We used BLI to evaluate the binding of CCaCalpT-CalpS to a filament-deficient R358E/K361E-CCaCalpL in the presence of cA_4 and to filament-forming H396A-CCaCalpL in the presence or absence of cA_4 (the ring nuclease mutant was used instead of wt-CCaCalpL to avoid rapid cA_4 cleavage during the assay) (Figures 3.23 A and 3.24). No CCaCalpL

binding to the sensor surface was detected with immobilized CCaCalpT-CalpS in the absence of cA_4, but the addition of cA_4 stimulated the binding of H396A-CCaCalpL, which was manifested as an increase in the BLI signal (Figures 3.24 B). In contrast, no binding was detected with the R358E/K361E mutant in the presence of cA_4, providing evidence that CCaCalpL filament formation is essential for its interaction with the target CCaCalpT-CalpS complex.

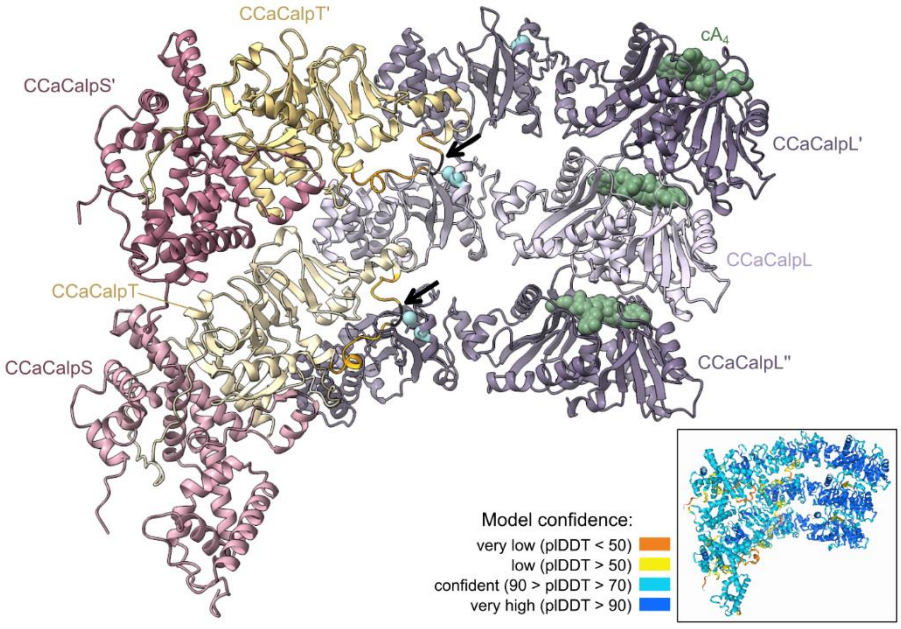


Figure 3.25. CCaCalpL interaction with CCaCalpT within a filament. AlphaFold3 model of CCaCalpL trimeric filament in complex with 3 molecules of AAAA RNA, 2 molecules of CCaCalpT and 2 molecules of CCaCalpS. CCaCalpS is colored shades of red, CCaCalpT – shades of yellow, CCaCalpL – shades of purple. Lon active site is depicted as cyan spheres. AAAA RNA is depicted as green spheres. The 194-210 aa loop of CCaCalpT is colored orange with a cleavage position colored black and indicated by arrows. The same AlphaFold3 model colored by the confidence (plDDT) values is shown in a box.

Finally, we sought to elucidate the mechanism of CCaCalpT cleavage by CCaCalpL. We attempted to obtain a cryo-EM structure of the cA_4/A_4p-activated CCaCalpL filament bound to CCaCalpT-CalpS or CCaCalpT, but we were unsuccessful. Therefore, we employed AlphaFold3 and obtained a model of three CCaCalpL subunits bound to three molecules of AAAA RNA, mimicking the activator, and two copies of CCaCalpT-CalpS heterodimer (Figure 3.25). Within the predicted filament the CCaCalpT′ contacts two adjacent CCaCalpL subunits: the ‘primary’ CalpL′ subunit with the high-confidence interface, 960 $Å^2$, and the ‘secondary’ CalpL subunit with the

lower-confidence interface, 1040 Å^2. The contact with the 'secondary' CalpL is primarily formed by the CCaCalpT′ loop (194-210 aa) inserted into the Lon protease domain. The cleavage site of the CCaCalpT is located within the 194-210 aa loop and resides in the proximity of Lon active site of CCaCalpL (Figure 3.25).

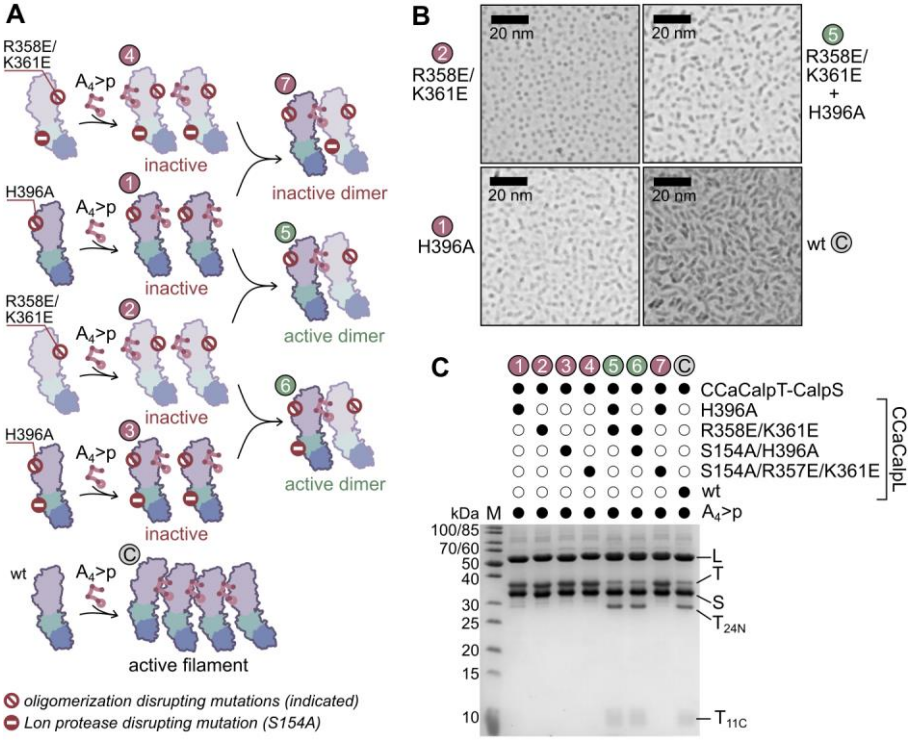


Figure 3.26. CCaCalpL oligomerization and CCaCalpT clevage mechanism. (**A**) Schematic representation of the assembly of the A_4>p-activated CCaCalpL dimers. (**B**) Denoised cryo-EM micrographs of CCaCalpL variantas: wt, R358E/K361E, H396A, and R358E/K361E mixed together with H396A in the presence of A_4>p. (**C**) Protease reactions of assembled dimers analyzed by SDS-PAGE. Reactions contained 5 μM CCaCalpL, 5 μM CCaCalpT-CalpS and 25 μM A_4>p.

To further back up our observations, we tested whether the CCaCalpL dimer, i.e. the minimal fragment of a filament, is sufficient for the Lon protease activity. For this we attempted to reconstitute an active CCaCalpL heterodimer from two inactive variants carrying mutations at the opposite surfaces of the SAVED domain (Figure 3.26 A): (i) the H396A mutant, which is unable to bind A_4>p and form filaments in the presence of A_4>p (Figure 3.22 C and 3.26 B), and (ii) the filament deficient R358E/K361E mutant (Figure 3.26 B). Although each individual mutant lacked protease activity,

their equimolar mix yielded an active CCaCalpL variant with wt-like CCaCalpT-CalpS cleavage activity (Figure 3.26 C, lanes 1, 2 and 5).

Next, to demonstrate that CCaCalpT binding and cleavage orientations, predicted by our AlphaFold3 model (Figure 3.25), are relevant, we mutated the Lon protease active site in H396A and R358E/K361E mutant proteins by introducing the S154A mutation and tested the CCaCalpL protease activity using different mutant combinations. Mutant combination H396A and R358E/K361E/S154A, which would be compatible with the *in-cis* cleavage (i.e. the same subunit predominately binds and cleaves the substrate) of CCaCalpT in the formed dimer, was inactive in the presence of A_4>p (Figure 3.26 C, lane 7). By contrast, mixing H396A/S154A and R358E/K361E mutants compatible with the *in-trans* cleavage (i.e. one subunit predominately binds the substrate while the other subunit performs the cleavage) yielded a dimeric CCaCalpL variant with wt-type-like protease activity (Figure 3.26 C, lane 6), confirming the proposed model of CCaCalpT cleavage within the activated CCaCalpL filament. Collectively, our findings demonstrate that CCaCalpL oligomerization is crucial for its interaction with the CCaCalpT-CalpS target, and both its protease and ring nuclease activities.

3.2.5. Mechanism of action of CalpL-CalpT-CalpS effector

Compared to the CalpL-CalpT-CalpS tripartite effector mechanism described at the beginning of Chapter 3.2 (Figure 3.14), our detailed characterization of CCaCalpL-CalpT-CalpS added the missing pieces to it. Figure 3.27 illustrates our proposed updated mechanism of the immunity provided by the tripartite effector. Initially, Csm-produced cA_4 binds to the SAVED domain of the CalpL monomer, triggering the formation of the CalpL filament and recruitment of the CalpT-CalpS. Within the filament, one CalpL subunit predominantly binds CalpT, while the adjacent subunit performs the proteolysis. CalpT cleavage products remain bound to CalpS, but the exposed C-degron of $CalpT_{24N}$ directs it for degradation by cellular proteases, releasing CalpS to bind RNAP and alter transcription. In addition, CalpL possesses a regulatory mechanism: the SAVED domain of CalpL acts as a ring nuclease, sequentially cleaving cA_4 in a three-step reaction. This cleavage serves as a "timer" to limit the duration of effector activation.

Importantly, in this study we demonstrated that cleavage of the anti-σ factor CalpT alone is not sufficient to release the σ factor CalpS from the CalpT-CalpS. The cleavage exposes the C-terminal degron on $CalpT_{24N}$, leading to its degradation, which then releases CalpS, a mechanism similar to other anti-σ/σ factor pairs (Mascher, 2023; Österberg et al., 2011). However,

the question of which genes are controlled by CalpS σ factors remains unanswered. Interestingly, σ factors are present in several other types of III CRIPSR-Cas associated effector systems. As discussed in Chapter 1.5.1, the type III-E interference complex utilizes direct association with the TRP-CHAT protease to specifically cleave Csx30 protein in the Csx30-CASP-σ pair, releasing CASP-σ, a σ factor, for further control of gene expression responsible for additional defense functions such as CRISPR spacer acquisition and transcription of other defense genes (Strecker et al., 2022). Moreover, the type III-B CRISPR-Cas associated multicomponent effector, which contains the cA_3-inducible SAVED-CHAT protease, also encodes a σ factor, although the function of this σ factor and the detailed mechanism of the entire multicomponent effector remains elusive (Steens et al., 2024) (see also Chapter 1.1.3).

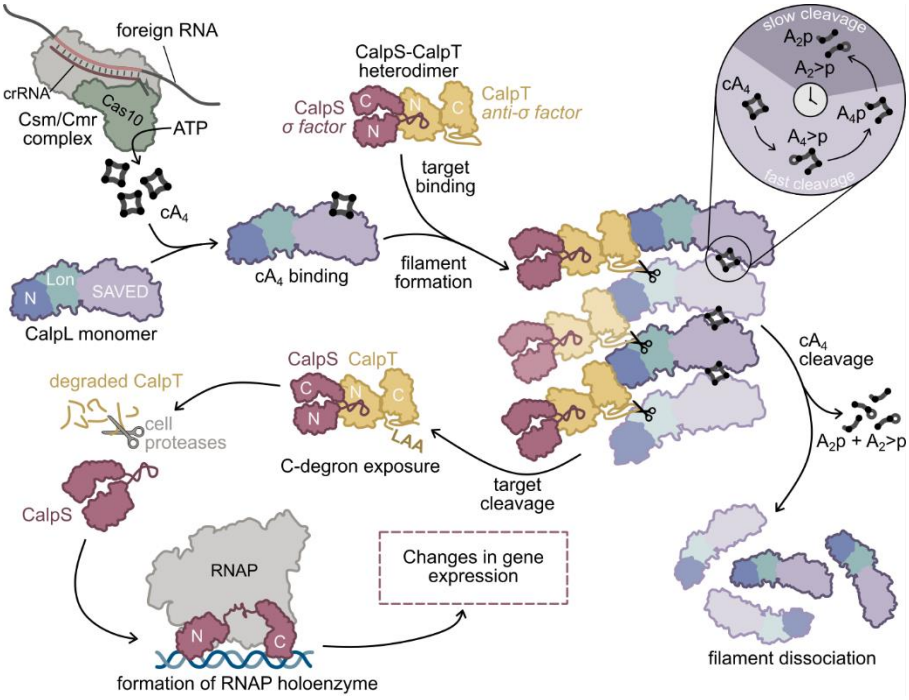


Figure 3.27. Mechanism of type III CRISPR-Cas antiviral defense mediated by the cA_4-activated tripartite CalpL-CalpT-CalpS effector system. Foreign RNA triggers cA_4 synthesis by the type III CRISPR-Cas complex. CalpL binds cA_4 by its SAVED domain, inducing filament formation. This allows the CalpL Lon domain to cleave CalpT in the CalpT-CalpS heterodimer. Cleaved CalpT is degraded by cellular proteases, releasing CalpS to bind RNA polymerase and act as a σ factor for transcription of specific genes. Activity of CalpL Lon protease is regulated through cA_4 cleavage by the CalpL SAVED domain.

Furthermore, the head-to-tail association of SAVED domains via a signaling molecule observed in the CCaCalpL filament appears to be a conserved feature among prokaryotic SAVED effectors. As discussed in Chapter 1.2.1, SAVED filament formation can assemble the composite active site of the effector, as seen with CBASS-associated TIR-SAVED (Hogrel et al., 2022) or allosterically activate effector domains, like in CRISPR type III-associated SAVED-CHAT (Steens et al., 2024) or enable cleavage of double-stranded DNA by Cap4 and Cap5 (Fatma et al., 2021; Lowey et al., 2020; Rechkoblit et al., 2024). In accordance, CalpL filament formation enables CalpT-CalpS substrate cleavage in the protease active site of the adjacent CalpL subunit within the filament.

Taken together, our study on CCaCalpL, together with complementary evidence from the parallel studies on the homologous SsCalpL-CalpT-CalpS, provides the most detailed mechanism of action of a type III CRISPR-Cas-associated multicomponent effector known to date.

3.3. Final remarks: An expanded repertoire of ring nucleases

During the course of the studies described in this dissertation, numerous CARF effectors were discovered to possess intrinsic cA_4-specific ring nuclease activity (reviewed in Chapter 1.4.1). In addition, two other protein folds, DUF1874 and STAS, have been identified as the core of Crn2 and Crn3 enzymes, respectively (Athukoralage et al., 2020b, 2020c; Samolygo et al., 2020) (see also Chapters 1.4.1. and 1.4.2). CARF and DUF1874 fold ring nucleases are homodimers, while Crn3 of STAS fold forms a composite active site at the interface of two dimers and they all use symmetrically arranged active sites to split cA_4 into two halves (Figure 3.28 A).

Similarly, the CARF domain of cA_6-specific Csm6 splits cA_6 into two halves - A_3>p molecules (Figure 3.28 A). Our research on StCsm6 and structural studies by other groups on StCsm6′ and EiCsm6 (Garcia-Doval et al., 2020; McQuarrie et al., 2023) have consistently shown that the CARF domain of cA_6-specific Csm6 promotes cleavage by forcing cA_6 to adopt the catalytically compatible conformation within the binding site. As not only the side chain but also the main chain atoms of the ammino acids located in the so-called catalytic loop forms interactions with cA_6, it is difficult to pinpoint the exact catalytic residues, thus no catalytic residues are indicated for cA_6-cleaving Csm6 in Figure 3.28 A. However, threonine from the catalytic loop may be an important candidate, as it uses its side-chain atoms to interact with the scissile phosphate (Garcia-Doval et al., 2020; McQuarrie et al., 2023). In addition, we have directly demonstrated that the nonspecific RNase activity

of the HEPN domain of StCsm6 contributes to the degradation of cA_6 and other cA_n at high signaling molecule concentrations, further expanding the repertoire of enzymes capable of cleaving cA_n.

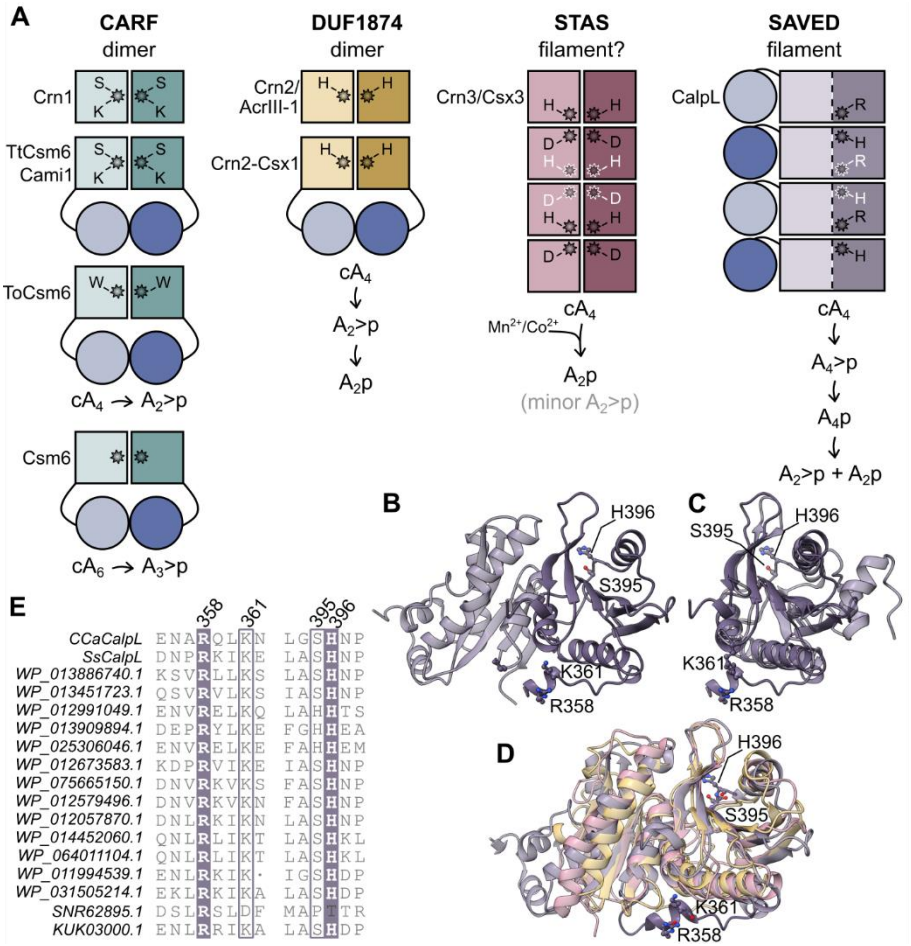


Figure 3.28. An expanded assortment of ring nucleases. (**A**) Diversity of ring nucleases. Ring nuclease domains of CARF, DUF1874, STAS and SAVED families are depicted as squares, circles mark various effector domains. Active sites are depicted as stars and catalytic residues (if present) are indicated. Dotted line indicates subdomains. Protein oligomeric states for each type of ring nucleases are specified. (**B**) SAVED domain of CCaCalpL. N-subdomain (202-352 aa) and C-subdomain (357-504 aa) are shown in dark purple and light purple, respectively. Active site residues are shown. (**C**) Superposition of the N-subdomain and C-subdomain of the CCaCalpL SAVED domain (colors as in A). (**D**) Superposition of the SAVED domains of CCaCalpL (purple) SAVED of CBASS-linked effectors: Cap5 (pink, PDB ID: 7RWK) (Fatma et al., 2021) and Cap4 (beige, PDB ID: 7YIB) (Lowey et al., 2020). The α-helix containing the residues of the CCaCalpL ring nuclease active site residues are absent from the Cap5 and Cap4 proteins is highlighted. Active site residues of SAVED CalpL are shown as in A. (**E**) Strict conservation of the ring

nuclease active site residues within Lon-SAVED4 proteins. For full alignment refer to Appendix 10.

Furthermore, studies of the regulatory mechanism of the tripartite CCaCalpL-CalpT-CalpS effector described in this thesis have added the SAVED domain of CCaCalpL to the list of cA_4-specific ring nucleases, making it the first known ring nuclease to use a three-step sequential cA_4 cleavage mechanism (Figure 3.28 A). CCaCalpL, together with other Lon-SAVED and TM-SAVED effectors, belongs to the SAVED4 family, whose members are often found in the vicinity of type III CRISPR-Cas systems (Makarova et al., 2020a). In contrast to the SAVED domains found in the effectors of CBASS systems, which interact specifically with asymmetric (odd-numbered) activators (Fatma et al., 2021; Hogrel et al., 2022; Lowey et al., 2020) (see also Chapter 1.2.1), CCaCalpL is specific for a symmetric (even-numbered) activator cA_4. Although, the superimposition of two subdomains of the SAVED domain of CCaCalpL generally confirmed the symmetric nature of the CARF pseudo-dimer (Figure 3.28 B and C), the major catalytic residues R358 and H396, together with less prominent catalytic residues K361 and S395, are located only in the C-terminal subdomain (Figure 3.28 B). Structural comparisons were performed by Dr. Giedrė Tamulaitienė. Our sequence analysis revealed that almost all Lon-SAVED4 proteins, including CCaCalpL structural homologue SsCalpL, contain the CCaCalpL ring nuclease active site residues (Figures 3.28 E). Not surprisingly, in parallel with our study SsCalpL has been reported to possess cA_4 ring nuclease activity (Binder et al., 2024). This supports the hypothesis that all SAVED4 family proteins are self-limiting effectors that share a conserved catalytic mechanism for cA_4 hydrolysis.

In summary, since 2018, when the first dedicated cA_4-specific ring nuclease (Crn1 family) was identified (Athukoralage et al., 2018), our understanding of cA_n regulation in type III CRISPR-Cas systems has advanced significantly. This work contributed to this by elucidating the regulation of cA_6-based signaling by the self-limiting effectors of the Csm6 family and further expanded the repertoire of domains capable of cA_4 cleavage by identifying the SAVED domain of the CalpL protein as a new addition to this list.

CONCLUSIONS

1. StCsm6 and StCsm6' effectively reduce the levels of cA_6 and other oligoadenylate species (cA_n) produced by the StCsm complex in a heterologous *E. coli* host.
2. The CARF domain of StCsm6 functions as a cA_6-specific ring nuclease, while the HEPN domain exhibits non-specific RNase activity, degrading cA_6 and other cA_n species, with a preference for linear single-stranded RNA as a substrate.
3. cA_4 binding triggers the formation of CCaCalpL filament, which is essential for the recruitment and proteolysis of the substrate CCaCalpT-CalpS. Within the filament, CCaCalpT is cleaved by a neighboring CCaCalpL subunit.
4. The cleavage of CCaCalpT by CCaCalpL exposes the C-degron, targeting the cleavage product for further degradation by cellular proteases. This degradation results in the release of CCaCalpS, which subsequently binds to bacterial RNA polymerase.
5. The SAVED domain of CCaCalpL functions as a cA_4-specific ring nuclease degrading its activator, cA_4, via a sequential three-step mechanism, that generates two intermediates, A_4>p and A_4p.
6. The formation of CCaCalpL filaments, triggered by the binding of the activator cA_4 or its cleavage intermediates, is crucial for the protease and ring nuclease activities of the Lon-SAVED CalpL effector.

SANTRAUKA

SANTRUMPOS

[S] < [E]	substrato nepriteklius
[S] > [E]	substrato perteklius
6H	6 α-spiralių motyvas; pagal angl. *6 α-helices*
A1/A2/A4/A4	pirmasis/antrasis/trečiasis/ketvirtasis adenozinas
A_2>p/A_3>p/A_4>p /A_5>pA_6>p	linijinis di-/tri-/tetra-/penta-/heksaadenilatas su 2′-3′-cikliniu fosfatu
A_2p/A_4p	linijinis di-/tetraadenilatas su 3′-fosfatu
BLI	biosluoksnio interferometrija; pagal angl. *biolayer interferometry*
cA_3/cA_4/cA_5/cA_6	ciklinis tri-/tetra-/penta-/heksaadenilatas
CalpL	su CRISPR susijusi Lon proteazė; pagal angl. *CRISPR associated Lon protease*
cA_n	ciklinis(-iai) oligoadenilatas(-ai)
CARF	su CRISPR susijusi Rossman sankloda; pagal angl. *CRISPR-associated Rossman fold*
Cas	susijęs su CRISPR; pagal angl. *CRISPR-associated*
CCa	*Candidatu*s Cloacimonas acidaminovorans
CFU	kolonijas formuojantys vienetai; pagal angl. *colony forming units*
CRISPR	pagal angl. *clustered regularly interspaced short palindromic repeats*
Crn	su CRISPR susijusi žiedo nulеazė pagal angl. *CRISPR ring nuclease*
crRNR	CRISPR RNR
DUF	nežinomos funkcijos domenas; pagal angl. *domain of unknown function*
GFP	žaliai fluorescuojantis baltymas; pagal angl. *green fluorescent protein*
HEPN	aukštesniųjų eukariotų ir prokariotų nukleotidus surišantis; pagal angl. *higher eukaryotes and prokaryotes nucleotide-binding*
HPLC-MS	didelio našumo skysčių chromatografija – masių spetrometrija; pagal angl. *high-performance liquid chromatograph - mass spectrometry*
IPTG	izopropilo β-D-tiogalaktopiranozidas

krio-EM	kriogeninė elektroninė mikroskopija
NCBI	National Center for Biotechnology Information (biotechnologijos ir biomedicinos duomenų bazės)
OD_{600}	optinis tankis, nustatytas esant 600 nm bangos ilgio šviesai
PDB	Protein Data Bank (baltymų struktūrų duomenų bazė)
RNR pol	RNR polimerazė
s.v.	sąlyginiai vienetai
SAVED	su SMODS susijęs ir sulietas su įvairias efektoriais; pagal angl. *SMODS-associated and fused to various effector domains*
SEC-MALS	dydžio atskyrimo chromatografija su šviesos sklaida daugeliu kampų; pagal angl. *size exclusion chromatography – multi-angle light scattering*
St	*Streptococcus thermophilus*
STAS	sulfato transporteris ir anti-σ faktoriaus antagonistas; pagal angl. *Sulphate Transporter and AntiSigma factor antagonist*
To	*Thermococcus onnurineus*
Tt	*Thermus thermophilus*
UV	ultravioletinė
wt	laukinis tipas; pagal angl. *wild type*

ĮVADAS

Vykstant nuolatinėms ginklavimosi varžybos tarp bakterijų ir jas puolančių virusų, vadinamų bakteriofagais (arba fagais), bakterijos išvystė sudėtingas priešvirusines gynybos sistemas. Šios sistemos slopina virusų plitimą populiacijoje aptikdamos fago infekciją ląstelėje ir koordinuodamos efektorių atsaką. Prokariotuose paplitusios CRISPR-Cas sistemos suteikia ląstelėms adaptyvų imunitetą, nes pirminės infekcijos metu į CRISPR regioną įterpia svetimos nukleorūgšties fragmentą, kurio pagrindu susintetintos CRISPR RNR (crRNR) molekulės pakartotinės infekcijos metu nulemia užpuoliko atpažinimą (Nussenzweig and Marraffini, 2020). Nors visų tipų CRISPR-Cas sistemos (šiuo metu yra žinomi septyni tipai) svetimų nukleorūgščių atpažinimui ir tiesioginiam taikinio degradavimui naudoja interferencijos kompleksą suformuotą iš crRNR ir Cas baltymų, III tipo CRISPR-Cas sistemos turi papildomą gynybos liniją: ciklinių oligoadenilatų (cA_n, n=3–6)

pagrindu veikiantį signalinį kelią, kuris aktyvuoja pagalbinius efektorinius baltymus (Stella and Marraffini, 2024).

III tipo CRISPR-Cas interferencijos kompleksui (Csm III-A ir III-D potipiuose arba Cmr III-B potipyje) komplementarumo principu atpažinus svetimą RNR yra aktyvinamas didžiojo subvieneto Cas10 Palm domenas, kuris katalizuoja cA_n sintezę iš ATP (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Susintetintos cA_n molekulės jungiasi prie pagalbinių efektorių ir juos aktyvina. Dažniausiai pagalbiniai baltymai, susiję su III-A/B/D sistemomis, yra sudaryti iš sensorinio CARF (angl. *CRISPR-Cas associated Rossmann fold*) domeno sulieto su įvairiais efektoriniais domenais (Hoikkala et al., 2024; Makarova et al., 2020a). CARF domeną turintys efektoriniai baltymai pastaraisiais metais buvo plačiai tyrinėjami (Steens et al., 2022; Stella and Marraffini, 2024), tačiau kur kas mažiau žinoma apie sensorinį SAVED (angl. *SMODS-associated and fused to various effector domains*), o ne CARF, domeną turinčius kai kurių III tipo CRISPR-Cas sistemų efektorinius baltymus (Makarova et al., 2020a; Steens et al., 2022). SAVED domenai yra siejami su kita prokariotų gynybos sistema CBASS (angl. *cyclic oligonucleotide-based antiphage signaling system*), kuri remiasi efektorinių baltymų aktyvinimu per signalines molekules ir nulemia infekuotos ląstelės žūtį (Slavik and Kranzusch, 2023), bet šiuos domenus turinčių baltymų rolė III tipo CRISPR-Cas sistemos yra mažai ištyrinėta.

Signalinių cA_n molekulių gamyba yra išjungiama, kai Csm/Cmr kompleksas hidrolizuoja taikinio RNR (Kazlauskiene et al., 2017; Rouillon et al., 2018), tačiau tai neišjungia pagalbinių efektorių - juos toliau aktyvinti gali jau susintetintos cA_n molekulės. Siekiant išvengti per didelio efektorių aktyvavimo ir pačios ląstelės pažeidimo, yra būtina papildoma šio signalinio kelio kontrolė. Signalines cA_n molekules geba specifiškai degraduoti specializuoti baltymai, CRISPR žiedo nukleazės (Crn - *angl. CRISPR ring nuclease*) (Athukoralage et al., 2018), bet jos randomos ne visose III tipo CRISPR-Cas sistemose. Crn baltymų šerdis yra CARF domenas, tad CARF domeną turintys efektoriai taip pat galėtų pasižymėti žiedo nukleaziniu aktyvumu.

Darbo objektai

Šio darbo objektai yra *Streptococcus thermophilus* DGCC8004 III-A tipo CRISPR-Cas pagalbiniai efektoriniai baltymai Csm6 ir Csm6′ (StCsm6 ir StCsm6′) bei *Candidatu*s Cloacimonas acidaminovorans str. Evry III-A tipo CRISPR-Cas sistemos pagalbinis trinaris efektorius CalpL-CalpT-CalpS (CCaCalpL-CalpT-CalpS), kurio CalpL narys turi SAVED sensorinį domenų, o CalpT ir CalpS yra anti-σ faktorius ir σ faktorius.

Darbo tikslai

(i) išaiškinti cA_6 signalinio kelio deaktyvacijos mechanizmą *S. thermophilus* III-A tipo CRISPR-Cas sistemoje,

(ii) nustatyti detalų III-A tipo CRISPR-Cas trinario efektoriaus CCaCalpL-CalpT-CalpS veikimo mechanizmą.

Uždaviniai

1. Įvertinti III tipo CRISPR-Cas sistemos vykdomą cA_6 ir kitų cA_n sintezę bakterijų ląstelėse, esant ir nesant StCsm6 bei StCsm6′ baltymams.
2. Charakterizuoti StCsm6 CARF ir HEPN domenų fermentines savybes, skaidant cA_6 ir kitus cA_n *in vitro*.
3. Nustatyti molekulinius reikalavimus CCaCalpL efektoriaus aktyvacijai ir jo sąveikai su substratu CCaCalpT-CalpS.
4. Išsiaiškinti σ faktoriaus CCaCalpS paleidimo iš CCaCalpT-CalpS komplekso mechanizmą.
5. Nustatyti CCaCalpL efektoriaus deaktyvavimo mechanizmą, tiriant jo gebėjimą hidrolizuoti cA_4.
6. Struktūriškai pagrįsti CCaCalpL efektoriaus aktyvacijos ir reguliacijos mechanizmus.

Mokslinis naujumas ir praktinė vertė

2017 m. III tipo CRISPR-Cas sistemose buvo atrastas ciklinius oligoadenilatų (cA_n) signalinis kelias, kurio pagalba yra aktyvinami įvairiūs pagalbiniai efektoriai, turintys sensorinį CARF ar SAVED domeną, taip sustiprinant CRISPR-Cas gynybą (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Nors buvo nustatyta, kad cA_n sintezė sustoja, kai užpuoliko RNR yra perkerpama (Kazlauskiene et al., 2017), tačiau tai, kas vyksta su jau susintetintais cA_n ir ar yra papildomicA_n aktyvuotų efektoriųreguliacijos mechanizmai, liko neaišku.

Šiuo metu nemažai žinoma apie CARF domeną turinčius efektorius, tokius kaip RNR nukleazės (Csm6 ir Csx1), DNR nukleazės, deaminazės, transliacijos inhibitoriai bei membranos integralumą pažeidžiantys baltymai (Baca et al., 2024a; Li et al., 2025; Mogila et al., 2023; Stella and Marraffini, 2024). Tačiau SAVED domeną turinčių efektorių molekuliniai mechanizmai iš esmės liko neištirti, tik keletas reikšmingesnių įžvalgų buvo gauta apie SAVED-CHAT efektorių (Steens et al., 2024).

Šio tyrimo tikslas buvo atskleisti pilną III tipo CRISPR-Cas sistemų reguliavimo mechanizmą, išanalizuojant CARF efektoriaus Csm6 deaktyvaciją ir SAVED efektoriaus CalpL reguliavimą. Pagrindiniai šio darbo

moksliniai rezultatai, kurie ženkliai praplėtė suvokimą apie III tipo CRISPR-Cas sistemų pagalbinių baltymų veikimą ir reguliaciją:

1. cA_6 signalinio kelio reguliavimas. Stebėdami cA_n gamybą *E. coli* ląstelėse, išreiškiančiose *S. thermophilus* III tipo CRISPR-Cas sistemą, nustatėme, kad StCsm6 ir StCsm6′ degraduoja savo aktyvatorių cA_6 bei kitus cA_n junginius. *In vitro* tyrimais atskleidėme, kad StCsm6 CARF domenas veikia kaip cA_6 specifinė žiedo nukleazė, o HEPN domenas pasižymi nespecifiniu ribonukleaziniu aktyvumu ir hidrolizuoja cA_6 bei kitus cA_n esant didelėms substrato koncentracijoms. Šiuo darbu parodėme, kad, analogiškai pavienėms žiedo nukleazėms, Csm6 CARF domenas veikia kaip efektoriaus ir visos CRISPR-Cas sistemos vidinis laikmatis. Be to, atskleidėme CARF ir HEPN domenų funkcijas ir tarpusavio ryšius cA_n signalinio kelio reguliacijoje.

2. SAVED domeną turinčio efektoriaus CalpL reguliavimas. Tirdami CalpL efektorių iš *Ca.* C. acidaminovorans, nustatėme, kad jis hidrolizuoja savo aktyvatorių cA_4, pasitelkdamas unikalų trijų pakopų mechanizmą ir suformuodamas du linijinius tarpinius junginius (A_4>p - linijinį tetraadenilatą su cikliniu 2′-3′ fosfatu ir A_4p - linijinį tetraadenilatą su 3′ fosfatu). Šis mechanizmas skiriasi nuo anksčiau žinomų žiedo nukleazių veikimo būdų ir yra pirmasis eksperimentinis SAVED domeno fermentinio aktyvumo pademonstravimas.

3. Trijų komponentų CCaCalpL-CalpT-CalpS sistemos veikimo mechanizmas. Atlikdami biocheminius, biofizikinius ir toksiškumo *E. coli* ląstelėse tyrimus, nustatėme, kad cA_4 aktyvuojamas CCaCalpL efektorius, sudarytas iš Lon proteazės ir SAVED domeno, specifiškai proteolizuoja anti-σ faktorių CCaCalpT, esantį CCaCalpT-CalpS anti-σ/σ faktoriaus poroje. Taip pat atradome, kad cA_4 arba jo hidrolizės produktų (A_4>p ir A_4p) prisijungimas inicijuoja CalpL filamento formavimąsi, kuris yra būtinas CCaCalpT-CalpS substrato surišimui. CCaCalpT perkirpimas suformuoja C-degroną, kuris nukreipia šį baltymą tolimesniam degradavimui ląstelės proteazėmis. Tai leidžia iš komplekso išlaisvinti σ-faktorių CCaCalpS, kuris toliau jungiasi su RNR polimeraze ir nulemia *E. coli* ląstelių augimo ribojimą. Remdamiesi struktūriniais tyrimais išaiškinome, kodėl filamento susidarymas yra būtinas tiek proteazės, tiek žiedo nukleazės funkcijoms.

Disertacijoje pateikiami rezultatai atskleidžia ne tik III tipo CRISPR-Cas sistemų ir jų pagalbinių efektorių veikimo mechanizmo kompleksiškumą, bet ir kaip griežtai apsaugos sistemos yra reguliuojamos. Šie duomenys prisideda prie augančio supratimo, kad III tipo CRISPR-Cas sistemos, naudojančios cA_n signalinį kelią, ląstelėje veikia ne per abortyvinės infekcijos mechanizmą. Aprašyti darbai yra svarbūs ir praktiniam pritaikymui, siekiant tikslingai

panaudoti bakteriofagus kovai su antibiotikams atspariomis bakterijomis, kurios kelia vis didesnę grėsmę žmonių sveikatai bei maisto ir žemės ūkio sistemų tvarumui.

Ginamieji teiginiai:

1. StCsm6 ir StCsm6′ degraduoja cA_6 bei kitus ciklinius oligoadenilatus (cA_n), kuriuos *E. coli* ląstelėse sintetina StCsm kompleksas.
2. StCsm6 CARF domenas veikia kaip cA_6 specifiška žiedo nuklezė, o HEPN domenas pasižymi nespecifiniu RNR nukleaziniu aktyvumu ir hidrolizuoja tiek cA_6, tiek kitus cA_n.
3. cA_4 prisijungimas inicijuoja CCaCalpL filamento susiformavimą, kuris leidžia prisijungti CCaCalpT-CalpS substratui ir kaimyniniam CCaCalpL subvienetui filamente vykdyti CCaCalpT proteolizę.
4. CCaCalpT perkirpimas suformuoja C-degroną, kuris nukreipia baltymą tolimesnei ląstelės proteazių vykdomai degradacijai. Dėl CCaCalpT degradacijos CCaCalpS yra išlaisvinamas ir rišasi su RNR polimeraze.
5. CCaCalpL SAVED domenas yra cA_4 specifiška žiedo nuklezė, kuri savo aktyvatorių degraduoja pasitelkdama trijų nuoseklių pakopų hidrolizės mechanizmą.
6. CCaCalpL filamento susidarymas, kurį inicijuoja cA_4 arba jo hidrolizės tarpiniai produktai, yra būtinas tiek efektoriaus proteazės, tiek žiedo nukleazės aktyvumams.

METODAI

Šiam darbui atlikti buvo pasitelkiami įvairūs kompiuteriniai, biocheminiai ir mikrobiologiniai metodai.

Homologijos paieškos buvo atliekamos naudojant HHPred (Zimmermann et al., 2018) ir DaliSearch (Holm et al., 2023). Homologinių baltymų sekų palyginiai buvo atliekami naudojant MMseqs2 (Steinegger and Söding, 2017), MUSCLE (Edgar, 2004) ir EBI EMBOSS Needle (Madeira et al., 2024). Baltymų ir jų kompleksų struktūrų modeliavimas buvo atliekamas AlphaFold2 (Evans et al., 2022; Jumper et al., 2021; Mirdita et al., 2022) ir AlphaFold3 (Abramson et al., 2024).

Darbe naudotos plazmidės buvo sukonstruotos Gibson assembly ir kryptingos mutagenezės metodais. Visų plazmidžių sekos buvo patvirtintos atliekant Sanger arba visos plazmidės sekoskaitą.

Ląstelėse gaminamų cA_n nustatymui bakterijų lizatai buvo paruošti iš *E. coli* ląstelių, transformuotų plazmidėmis koduojančiomis *S. thermophilus*

CRISPR-Cas sistemos variantus ir taikinio geną. Genų raiška indukuota IPTG ir ląstelės surinktos centrifuguojant, suardytos ir atlikta metabolitų analizė HPLC-MS metodu. Bakteriofagų sukeliamos cA_n sintezės įvertinimui panaudotas MS2 bakteriofagas ir prieš jį nukreipta *S. thermophilus* III tipo CRISPR-Cas sistema *E. coli* NovaBlue (DE3) kamiene.

Baltymai ir Csm kompleksai buvo išgryninti iš *E. coli* ląstelių naudojant gimiringumo chromatografijos ir kai kuriais atvejais dydžių chromatografijos metodus. Atskirtų StCsm6 CARF ir HEPN domenų oligomerinės būsenos nustatytos naudojant dydžių chromatografijos metodą ir lyginant su žinomų dydžių standartais. CCaCalpL-CalpT-CalpS efektoriaus komponentų oligomerinės būsenos buvo įvertintos SEC-MALS metodu.

RNR substratai RNR nukleazinio aktyvumo tyrimams bei taikinio RNR Csm kompleksų aktyvacijai *in vitro* buvo paruošti *in vitro* transkripcijos metodu, naudojant atitinkamas DNR su T7 promotoriumi.

StCsm ir CCaCsm kompleksų *in vitro* gaminamų cA_n įvertinimui kompleksai buvo inkubuojamai su specifinėmis taikinio RNR ir ATP reakcijos mišiniuose esant reikiamų metalo jonų. Susidarę cA_n produktai toliau buvo analizuojamai HPLC-MS metodu. StCsm ir CCaCsm kompleksai analogiškai buvo panaudojami ir radioaktyviai žymėtų cA_6 ir cA_4 sintezei, reakcijos mišinius papildant α-^{32}P-ATP. Prieš naudojant tolimesniuose tyrimuose, tiksliniai sintezės produktai buvo išskiriami iš gelio.

Žiedo nukleazės reakcijos buvo atliekamos naudojant sintetinius cA_3, cA_4, cA_6, A_4>p, A_4p, A_6>p bei StCsm susintetintą cA_n mišinį ir analizuojant reakcijos produktus HPLC-MS arba naudojant radioaktyviai žymėtus cA_4 ir cA_6 ir analizuojant autoradiografiškai vizualizuotus reakcijos produktus denatūruojančiame gelyje. Kinetiniams reakcijų parametrams (k_{obs} ir k_{kat}) nustatyti buvo pritaikomas eksponentinis substrato nykimo modelis arba tiesinės regresijos analizė. CCaCalpL vykdomai cA_4 hidrolizei aprašyti buvo taikomas nuoseklaus skaidymo modelis, aprašantis cA_4 virtimą tarpiniais ir galutiniais produktais, o greičio konstantos nustatytos atliekant netiesinės regresijos analizę. RNR hidrolizės reakcijų produktai buvo analizuojami denatūruojančiame gelyje. Konkurenciniame RNR karpymo eksperimente į reakciją buvo pridedamas įvairus cA_6 kiekis (0-500 μM), o reakcijų kinetika buvo nustatyta pritaikant substrato nykimo eksponentinio modelio kreivę.

Proteazės reakcijos buvo atliekamos inkubuojant CCaCalpL ir CCaCalpT-CalpS (arba tik CCaCalpT) 37°C temperatūroje 120 minučių reakcijos mišinyje esant cA_3, cA_4, A_4>p, A_4p arba A_2>p ir A_2p. Reakcijų produktai buvo analizuojamai denatūruojančios baltymų elektroforezės (SDS-PAGE) metodu.

Duomenys CCaCalpL susirišusio su cA_4, A_4>p ir A_4p struktūrų nustatymui buvo surinkti naudojant „Glacios“ krio-transmisijos elektroninį mikroskopą (Thermo Fisher Scientific). Vaizdai buvo apdoroti „CryoSPARC“ (v.4.2.1) (Punjani et al. 2017; Punjani et al. 2020). Atominis modelis paruoštas naudojant Coot (v.0.9.8.1) (Emsley et al., 2010), remiantis AlphaFold2 sugeneruotu modeliu ir toliau patobulintas naudojant „phenix.real_space_refine“ (Liebschner et al., 2019). Struktūriniams duomenims atvaizduoti buvo naudojama „ChimeraX“ (v.1.5) (Meng et al., 2023).

Baltymų sąveikos tyrimams buvo pasitelkiamas BLI metodas, naudojant Octet K2 sistemą (Sartorius). Baltymų sąveikos tyrimai buvo vykdomi ant biosensoriaus imobilizuojant CCaCalpT-CalpS arba CCaCalpS, per CCaCalpS C-gale esančią gryninimo žymę ir perkeliant į šulinėlį su besirišančiu baltymu (CCaCalpL ir cA_4 mišiniu arba *E. coli* RNR pol.). Papildomai CCaCalpT-CalpS sąveika po kirpimo CCaCalpL buvo tiriama magnetinėmis dalelėmis išgaudant baltymus, kurie sąveikauja su per C-galo gryninimo žymę imobilizuotu CCaCalpS.

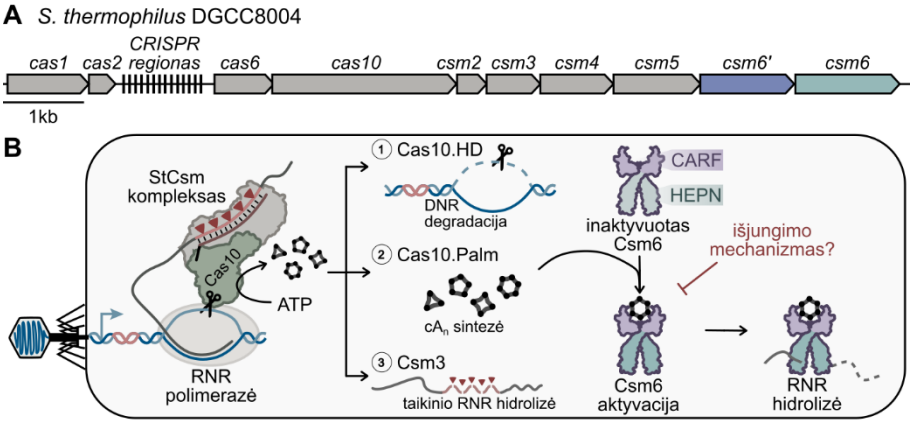
CCaCalpT-CalpS komplekso komponentų ir produktų susidarančių po CCaCalpL kirpimo toksiškumas *E. coli* ląstelėms buvo įvertintas auginant *E. coli* ląsteles, išreiškiančias pavienius komponentus ir komponentų kombinaciją, ant agarizuotos mitybinės terpės su induktoriumi arba skystoje terpėje su induktoriumi ir matuojant OD_{600} 16 valandų kas 10 minučių. Degrono įvertinimui reporterinėje sistemoje GFP degradacijos eksperimentai buvo atliekami tiek auginant ląsteles ant agarizuotos terpės, tiek skystoje terpėje 12 valandų sekant ląstelių augimą ir GFP fluorescenciją.

REZULTATAI

cA_6 signalinio kelio reguliacija

2017 m. mano kolegos, dirbantys dr. Gintauto Tamulaičio vadovaujamoje mokslinėje grupėje, atrado ciklinių oligoadenilatų (cA_n) signalinį kelią, dalyvaujantį bakterijų apsaugoje nuo jas puolančių virusų (Kazlauskiene et al., 2017). Šis atradimas buvo padarytas tiriant modelinę III-A tipo CRISPR-Cas sistemą iš *Streptococcus thermophilus* (St) DGCC8004, kurią sudaro adaptacijos genai (*cas1* ir *cas2*), CRISPR regionas, *cas6* genas, koduojantis už crRNR brendimą atsakingą baltymą, bei interferencijos kompleksą formuojančių Cas baltymų genai (*cas10, csm2, csm3, csm4* ir *csm5*) (1 pav. A) (Tamulaitis et al., 2014). Šalia taip pat yra koduojami du pagalbiniai efektoriai - StCsm6 ir StCsm6′, sudaryti iš CARF ir HEPN domenų.

Bakteriofago infekcijos metu Csm interferencijos kompleksas susiriša su užpuoliko RNR ir tai paruošia Cas10 subvienetą ciklinių oligoadenilatų sintezei (cA_n, kur n = 3–6), kuri yra nutraukiama StCsm komplekso Csm3 subvienetams perkirpus ir paleidus surištą taikinio RNR (1 pav. B). cA_n yra signalinės molekulės, kurios aktyvuoja pagalbinius efektorius. *S. thermophilus* sistemoje tai yra StCsm6 ir StCsm6′ – nespecifinės RNR nukleazės, kurias aktyvina cA_6 (Kazlauskiene et al., 2017). Tačiau homologas iš *Thermus thermophilus* (TtCsm6) yra aktyvinamas cA_4, o ne cA_6 (Kazlauskiene et al., 2017).

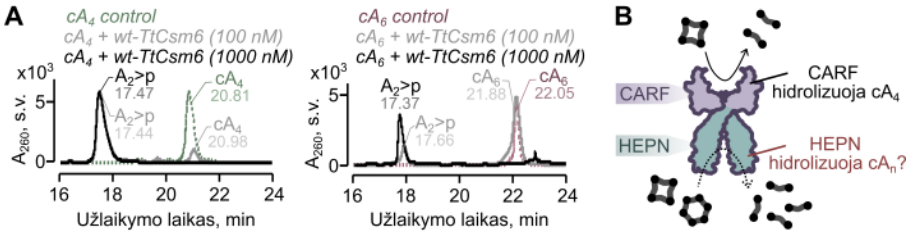


1 pav. *S. thermophilus* III-A tipo CRISPR-Cas Sistema. (**A**) III-A tipo CRISPR-Cas sistemos regionas *S. thermophilus* kamieno DGCC8004 genome (GenBank ID: KM222358.1). (**B**) Sistemos veikimo mechanizmas. StCsm interferencijos kompleksui atpažinus ir surišus bakteriofago transkriptą yra aktyvinamas didysis komplekso subvienetas Cas10. Cas10 HD nukleazinis domenas pradeda degraduoti viengrandinę DNR (1), išstumtą transkripcijos burbule, o Palm domenas iš ATP ima sintetinti cA_n signalines molekules (2). Pagalbinių efektorių StCsm6 ir StCsm6′ sensorinis CARF domenas suriša cA_6 ir tai aktyvina HEPN domeno RNR nukleazinį aktyvumą. Aktyvintas HEPN degraduoja RNR. Cas10 aktyvumai yra išjungiami StCsm komplekso surištą taikinio RNR perkirpus Csm3 subvienetams (3), tačiau nėra aišku, ar/kaip jau susintetintos cA_6 molekulės yra pašalinamos.

In vitro StCsm kompleksas iš ATP sintetina įvairių cA_n mišinį, kuriame dominuoja cA_3. Tačiau StCsm6 RNR nukleazę aktyvina tik nedideliais kiekiais sintetinamas cA_6 (Kazlauskiene et al., 2017). Šie rezultatai iškėlė keletą svarbių klausimų: (i) kokie cA_n sintetinami bakterijų ląstelėje; (ii) ar III tipo CRISPR-Cas sistema veikia kaip abortyvinės gynybos mechanizmas, sukeliantis ląstelės žūtį, o gal cA_n yra suskaidomi po jų sintezės nutraukimo; (iii) jei cA_n yra degraduojami, kaip tai yra reguliuojama?

Lūžis tyrimuose įvyko 2018 m., kai prof. M. White′o vadovaujama mokslininkų grupė surado, kad kai kurios III tipo CRISPR-Cas sistemos koduoja specializuotus CARF domeno baltymus, vadinamus CRISPR žiedo

nukleazėmis (Crn) ir specifiškai hidrolizuojančius cA_4 (Athukoralage et al., 2018). Todėl iškėlėme hipotezę, kad CARF-HEPN sandaros baltymai StCsm6, StCsm6′ ir TtCsm6 taip pat galėtų degraduoti cA_6 arba cA_4, pasitelkdami savo CARF domenus.

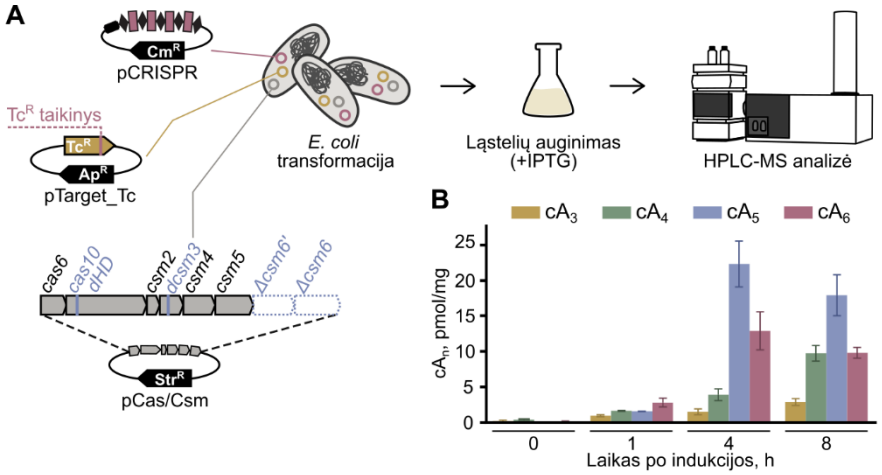


2 pav. TtCsm6 žiedo nukleazinis aktyvumas. (**A**) TtCsm6 hidrolizės reakcijų su cA_4 ir cA_6 analizė HPLC-MS metodu. Reakcijos mišinyje buvo 20 µM cA_4 arba cA_6 ir 0, 100 arba 1000 nM laukinio tipo (wt-TtCsm6), inkubuota 120 min 37°C temperatūroje. (**B**) Schematinis siūlomo Csm6 žiedinės nukleazės aktyvumo atvaizdavimas. HEPN domeno dalyvavimas cA_n skaidyme nėra nustatytas.

Siekdami patikrinti šią hipotezę, pirmiausia patikrinome, ar *TtCsm6* gali hidrolizuoti savo aktyvatorių cA_4 (2 pav. A). Šiuos eksperimentus atliko dr. Irmantas Mogila. Aukštos gebos skysčių chromatografijos ir masių spektrometrijos (HPLC-MS; *ang. high-performance liquid chromatography - mass spectrometry*) metodu nustatėme, kad per 2 valandas *TtCsm6* pavertė cA_4 į linijinį diadenilatą su cikliniu-2′,3′ fosfatu (A_2>p). HPLC-MS analizę atliko Audronė Rukšėnaitė. Ši reakcija atitiko Crn1 žiedo nukleazės vykdomą reakciją (Athukoralage et al., 2018) ir patvirtino, kad TtCsm6 taip pat gali veikti kaip žiedo nukleazė. Nepriklausomi tyrimai dar labiau sustiprino šiuos pastebėjimus, kai buvo pademonstruota, kad tiek TtCsm6, tiek *T. onnurineus* Csm6 (ToCsm6) specifiškai hidrolizuoja cA_4 (Athukoralage et al., 2019; Jia et al., 2019b).

Pastebėjome, kad esant aukštai baltymo koncentracijai, TtCsm6 hidrolizuoja ne tik cA_4, bet ir cA_6, nors cA_6 molekulė yra per didelė, kad tilptų į CARF domeno surišimo kišenę. Todėl iškėlėme hipotezę, jog Csm6 ribonukleazės gali pasitelkti HEPN domeną visų III tipo CRISPR-Cas Csm komplekso sintetinamų cA_n hidrolizei (2 pav. B). Šią hipotezę paremia ir tai, kad ToCsm6 kristalinėje struktūroje cA_4 yra surištas tiek CARF, tiek prie HEPN domenų (Jia et al., 2019b). Norėdami išsamiau ištirti cA_n signalinio kelio reguliaciją, toliau tyrėme *S. thermophilus* III-A tipo CRISPR-Cas sistemo cA_6 signalinio kelio deaktyvavimo mechanizmą bei StCsm6/StCsm6′ CARF ir HEPN domenų vaidmenį cA_n hidrolizėje.

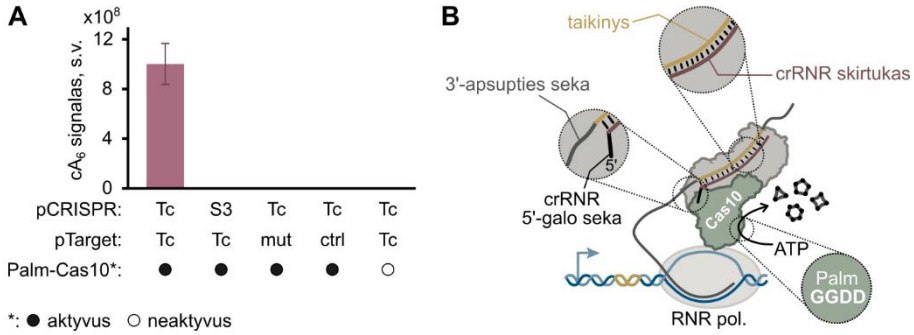
Pirmiausia, pasitelkę HPLC-MS metodą, nustatėme, kad StCsm interferencijos kompleksą išreiškiančiose *E. coli* ląstelėse cA_n pasiskirstymas skiriasi nuo stebėto *in vitro*. Metabolitų išskyrimą iš ląstelių ir jų analizę HPLC-MS metodu atliko dr. Jesper F. Havelund ir dr. Nils J. Færgeman. Ląstelėse daugiausiai buvo gaminama cA_5 ir cA_6 (3 pav.), o *in vitro* StCsm daugiausia gamina cA_3, šiek tiek cA_4 bei cA_5 ir tik nedidelį kiekį cA_6, kuris aktyvina StCsm6 (Kazlauskiene et al., 2017).



3 pav. *E. coli* ląstelėse StCsm komplekso sintetinamų cA_n analizė. (**A**) Eksperimento schema. *E. coli* ląstelės buvo transformuotos plazmidėmis: pCRISPR_Tc, pCas/Csm (dHD-Cas10, dCsm3, ΔCsm6′ΔCsm6) ir pTarget_Tc. Plazmidė pCRISPR_Tc kodavo CRISPR regioną skirtą atsparumo tetraciklinui geno transkriptui komplementarios crRNR produkcijai, o pTarget_Tc plazmidė kodavo tetraciklino atsparumo geną (Tc^R). Siekdami padidinti detektuojamų cA_n kiekį bei sumažinti galimą toksiškumą naudojome pCas/Csm (dHD-Cas10, dCsm3, ΔCsm6′ΔCsm6), kuri kodavo StCsm komplekso baltymus, tačiau jų visų, išskyrus Cas10 Palm domeno, aktyvūs centrai turėjo išveiklinančias mutacijas, o pagalbinius efektorius StCsm6 ir StCsm6′ koduojantys genai buvo pašalinti iš konstrukto. Iš surinktų ląstelių, kuriose StCsm interferencijos komplekso raiška buvo indukuota, lizatų išskirti metabolitai buvo analizuojami HPLC-MS metodu. (**B**) cA_n variantų pasiskirstymas prieš StCsm komplekso raiškos indukciją (0 h) ir praėjus 1, 4 ir 8 valandoms po indukcijos. cA_n kiekis įvertintas pmol miligrame ląstelių kultūros. Paklaidos žymi bent dviejų eksperimento pakartojimų standartinį nuokrypį.

Šią cA_n sintezės *E. coli* ląstelėse strategiją taip pat panaudojome patvirtinti, kad anksčiau *in vitro* nustatyti cA_n sintezės reikalavimai (Han et al., 2018; Kazlauskiene et al., 2017; Nasef et al., 2019; Niewoehner et al., 2017; Rouillon et al., 2018) išlieka nepakitę sintezei vykstant ląstelėse. Panaudoję įvairias pCas/Csm, pCRISPR ir pTarget variantų kombinacijas patvirtinome, kad cA_n sintezei svarbu, jog: (i) crRNR skirtuko seka turi būti komplementari

taikinio RNR sekai – kompleksas turintis crRNR su S3 skirtuku nebuvo aktyvinamas Tc^R geno transkripto, o Tc^R geno neturinti pTarget_ctrl neaktyvino cA_6 sintezės; (ii) taikinio RNR 3′-apsupties seka turi nebūti komplementari crRNR 5′-galo sekai – mutacijos Tc^R gene (pTarget_Tc^{mut}), įvedančios 3′-apsupties ir crRNR 5′-galo sekos komplementarumą, neaktyvino cA_6 sintezės; (iii) Cas10 turi turėti aktyvų Palm domeną su GGDD katalitiniu motyvu – Cas10 Palm domeno mutantas negebėjo sintetinti cA_6 (4 pav.).

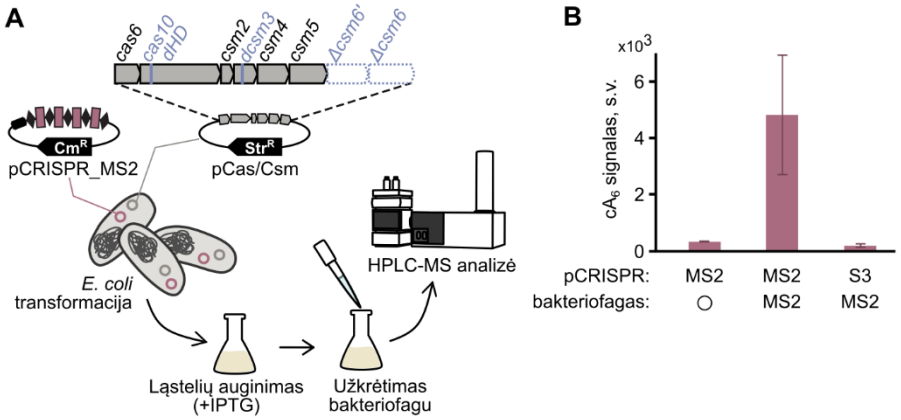


4 pav. III-A tipo CRISPR-Cas sistemos elementai reikalingi cA_6 sintezei vykti. (**A**) cA_6 sintezės įvertinimas *E. coli* ląstelėse transformuotose nurodytomis pTarget ir pCRISPR plazmidėmis bei išreiškiančiose StCsm kompleksą (dCsm3, dHD-Cas10, ΔCsm6′ΔCsm6) su nesugadintu Cas10 Palm domenu Palm-Cas10 (juodas apskritimas) arba GGDD katalitiniame motyve mutaciją turinčiu Palm domenu dPalm-Cas10 (baltas apskritimas). (**B**) Schema vaizduojanti cA_n sintezei reikalingus elementus: (i) crRNR skirtuko sekos ir taikinio RNR komplementarumas, (ii) komplementarumo nebuvimas tarp taikinio 3′-apsupties sekos ir crRNR 5′-galo sekos, (iii) aktyvus Cas10 Palm domenas su GGDD katalitiniu motyvu.

Pakeitę skirtuko seką pCRISPR plazmidėje, nukreipėme StCsm kompleksą prieš RNR genomą turintį bakteriofagą MS2 ir įvertinome cA_6 sintezę infekcijos metu (5 pav. A). Praėjus 16 valandų nuo StCsm komplekso išreiškiančių ląstelių kultūros užkrėtimo MS2 bakteriofagu buvo stebimas cA_6 kaupimasis (5 pav. B). Tad cA_6 sintezė yra aktyvinama StCsm kompleksui reaguojant tiek į plazmidines, tiek į virusines kilmės svetimas nukleorūgštis.

Nustatę, kad StCsm6 ir StCsm6′ aktyvatoriaus cA_6 sintezė efektyviai vyksta StCsm kompleksą išreiškiant *E. coli* ląstelėse, toliau patikrinome, kaip keičiasi cA_6 produkcija ląstelėms kartu su StCsm kompleksu išreiškiant pagalbinius efektorius StCsm6 ir StCsm6′. Lyginant su aukščiau aprašytu eksperimentu (3 pav. B), *E. coli* ląstelių, išreiškiančių StCsm6 ir StCsm6′, cA_6 kiekis buvo ženkliai žemesnis (1 lentelė). Įdomu, kad ne tik cA_6, bet ir visų kitų cA_n kiekiai buvo paveikti StCsm6 ir StCsm6′ raiškos (1 lentelė). Tai leidžia manyti, kad *S. thermophilus* III-A tipo CRISPR-Cas pagalbiniai

efektoriai yra atsakingi už juos aktyvianančio cA_6 ir visų kitų StCsm komplekso sintetinamų cA_n degradavimą.



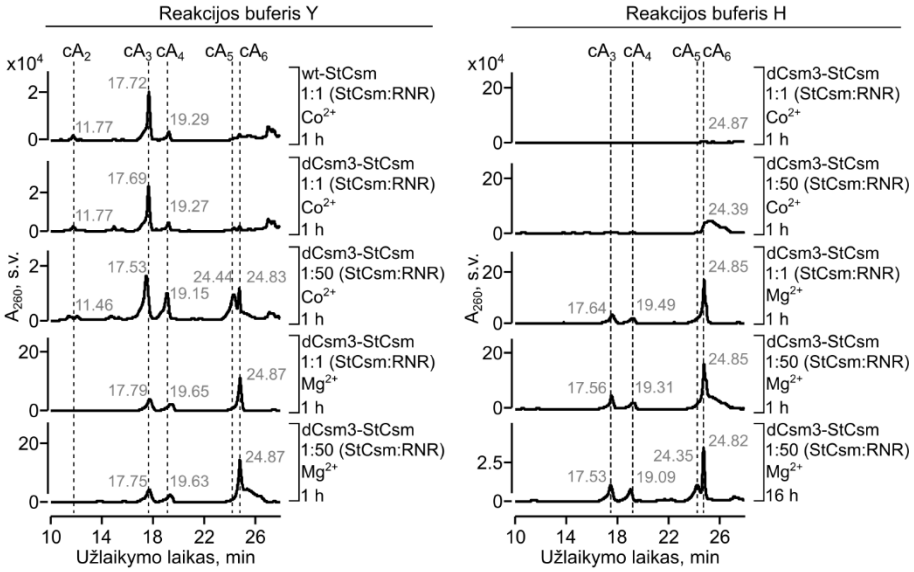
5 pav. cA_6 sintezė bakteriofago infekcijos metu. (**A**) Eksperimento schema. *E. coli* ląstelės buvo transformuotos plazmidėmis: pCRISPR_MS2 ir pCas/Csm (dHD-Cas10, dCsm3, ΔCsm6′ΔCsm6). Po indukcijos ląstelės buvo užkrėstos MS2 bakteriofagu. Iš po 16 valandų auginimo surinktų ląstelių lizatų išskirti metabolitai buvo analizuojami HPLC-MS metodu. (**B**) cA_6 sintezės įvertinimas *E. coli* ląstelėse, transformuotose nurodyta pCRISPR plazmide, išreiškiančiomis StCsm kompleksą (dCsm3, dHD-Cas10, ΔCsm6′ΔCsm6) ir užkrėstomis MS2 bakteriofagu arba neinfekuotomis (baltas apskritimas).

1 lentelė. cA_n kiekis nustatytas *E. coli* ląstelėse, kai ląstelėse nebuvo (Δ) arba buvo vykdoma StCsm6 ir StCsm6′ raiška. Pateikiamos bent dviejų eksperimento pakartojimų verčių vidurkis ir standartinis nuokrypis; n.d.- matavimo metu signalas nedetektuotas.

cA_n	laikas, h	cA_n kiekis, pmol/mg	
		ΔStCsm6′ΔStCsm6	StCsm6′, StCsm6
cA_3	0	0.24 ± *0.15*	0.000 ± *0.001*
	1	0.94 ± *0.20*	0.005 ± *0.007*
	4	1.51 ± *0.56*	0.002 ± *0.003*
	8	2.87 ± *0.68*	0.017 ± *0.007*
cA_4	0	0.39 ± *0.10*	n.d.
	1	1.62 ± *0.13*	0.003 ± *0.005*
	4	3.88 ± *1.18*	n.d.
	8	9.78 ± *1.57*	n.d.
cA_5	0	0.04 ± *0.03*	n.d.
	1	1.58 ± *0.004*	n.d.
	4	22.33 ± *4.58*	n.d.
	8	17.89 ± *4.05*	n.d.
cA_6	0	0.12 ± *0.07*	0.126 ± *0.178*
	1	2.80 ± *0.88*	0.005 ± *0.007*
	4	12.91 ± *3.80*	n.d.
	8	9.79 ± *1.07*	n.d.

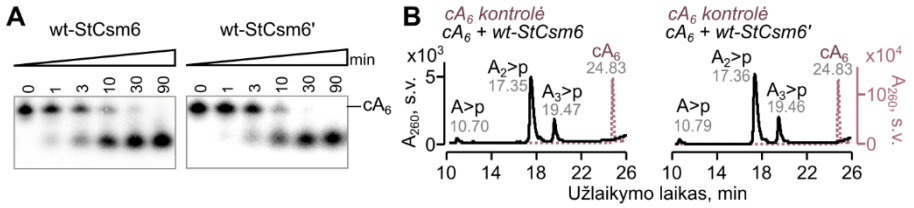
StCsm6 ir StCsm6′ vykdoma cA_6 hidrolizė in vitro

Siekdami išsamiai *in vitro* detalizuoti StCsm6 ir StCsm6′ vykdomą cA_6 hidrolizę, siekėme pasigaminti lengvai detektuojamą substratą – radioaktyviai žymėtą cA_6 arba/ir kitus žymėtus cA_n. Tam iš naujo įvertinome ir optimizavome cA_6 sintezės sąlygas. Atlikta sintezės reakcijų produktų HPLC-MS analizė parodė, kad naudojant taikinio RNR perteklių ir Mg^{2+} kaip kofaktorių sintetinamų cA_n mišinyje ženkliai padidėjo susintetinamo cA_6 dalis (6 pav.). Audronė Rukšėnaitė atliko šią ir visas toliau disertacijoje minimas HPLC-MS analizes. Šiomis optimizuotomis sąlygomis atlikdami reakciją ir papildomai į ją pridėję α-^{32}P-ATP, pasigaminome radioaktyviai žymėtą cA_6, kurį naudojome tolimesniuose tyrimuose.



6 pav. cA_6 sintezės sąlygų optimizavimas. StCsm komplekso cA_n sintezės HPLC-MS analizės rezultatai. Reakcijos buvo atliktos naudojant wt arba dCsm3-StCsm kompleksą esant StCsm:taikinio RNR moliniam santykiui 1:1 arba 1:50 nurodytame reakcijos buferyje su nurodytu metalo jono kofaktoriumi. Reakcijos buvo inkubuojamos 1 arba 16 valandų.

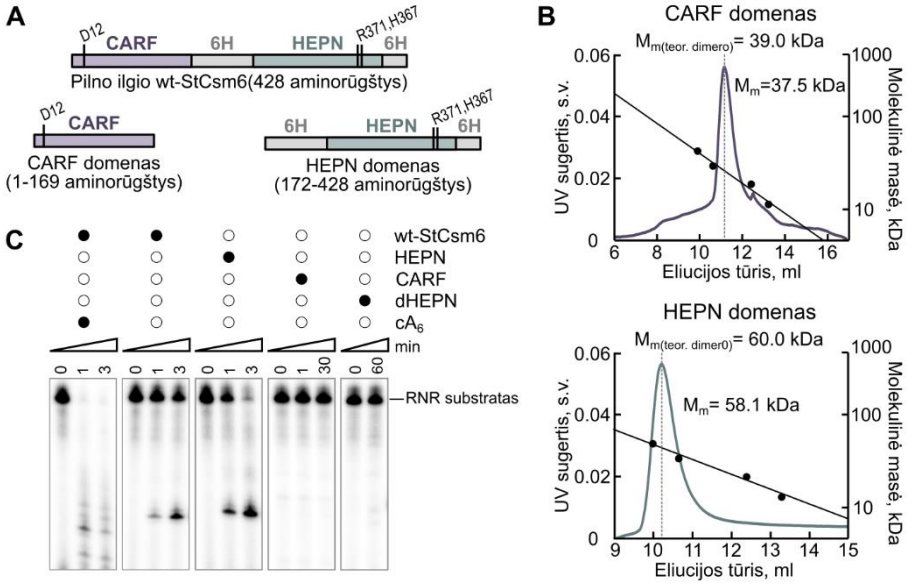
Išgrynintus StCsm6 ir StCsm6′ baltymus inkubavome su radioaktyviai pažymėtu cA_6 ir nustatėme, kad abu baltymai geba hidrolizuoti cA_6 ir reakcijai nėra reikalingi metalo jonai (7 pav. A). O pasitelkiant HPLC-MS analizė identifikavome reakcijos produktus: 2′,3′-ciklinį AMP (A>p), linijinį diadenilatą su 2′,3′-cikliniu fosfatu (A_2>p) ir linijinį triadenilatą su 2′,3′-cikliniu fosfatu (A_3>p) (7 pav. B).



7 pav. StCsm6 ir StCsm6′ vykdoma cA_6 hidrolizė. (**A**) StCsm6 ir StCsm6′ vykdoma radioaktyviai žymėto cA_6 hidrolizė. Reakcijos produktai analizuoti denatūruojančiame poliakrilamidiniame gelyje. (**B**) cA_6 karpymo reakcijos produktų analizė HPLC-MS metodu. Rausva spalva žymi cA_6 kontrolinę reakciją, t. y., reakcijos mišinį be baltymo.

Kadangi tiek StCsm6, tiek StCsm6′ baltymai pasižymėjo tokiu pačiu cA_6 karpymo aktyvumu ir yra artimi struktūriniai homologai, turintys ~34% identiškų amino rūgščių (Kazlauskiene et al., 2017; Tamulaitis et al., 2014), paprastumo dėlei toliau tyrinėjome tik vieną iš homologų – StCsm6.

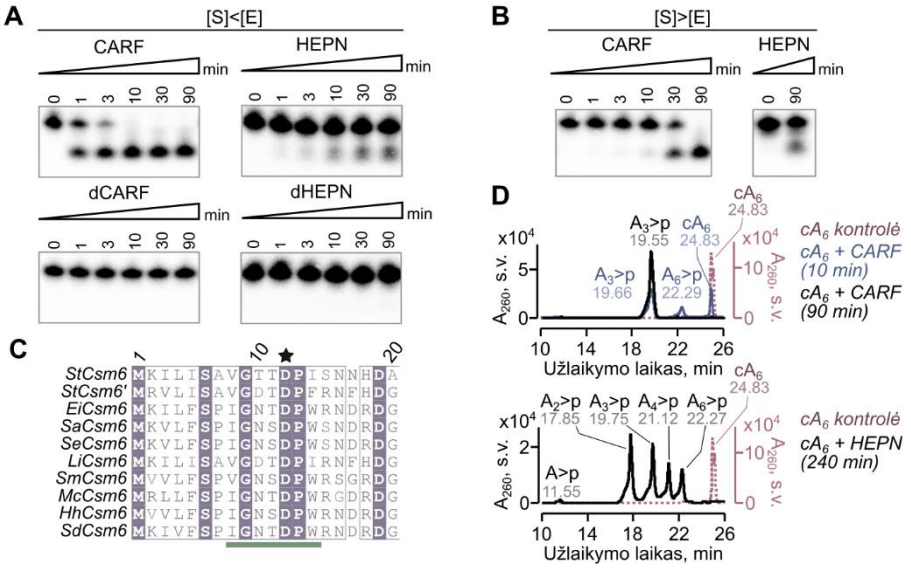
Atskirtų CARF ir HEPN domenų aktyvumo tyrimai



8 pav. Atskirtų StCsm6 CARF ir HEPN domenų biocheminė analizė. (**A**) Wt-StCsm6 ir jo CARF (1–169 aminorūgštys) bei HEPN (172–428 aminorūgštys) domenų schema. Aktyvumui svarbios aminorūgštys pažymėtos. (**B**) CARF ir HEPN domenų oligomerinės būsenos nustatymas dydžių atskirties chromatografijos metodu. Eksperimentinės molekulinės masės buvo apskaičiuotos remiantis dydžio standartais BSA (67 kDa), ovalbuminas (43 kDa), chimotripsinogenas A (25 kDa) ir ribonukleazė A (13,7 kDa). (**C**) Izoliuotų CARF ir HEPN domenų RNR nukleazinio aktyvumo palyginimas su neaktyvuotu ir cA_6 aktyvuotu wt-StCsm6. Reakcijose buvo naudojama 10 nM α-^{32}P žymėto RNR substrato, 10 nM baltymo ir, jei nurodyta, 10 nM cA_6. Reakcijos produktai analizuoti denatūruojančiame gelyje. dHEPN

(izoliuotam HEPN domenui su R371Air H367A mutacijomis) buvo naudojama 10 μM baltymo koncentracija.

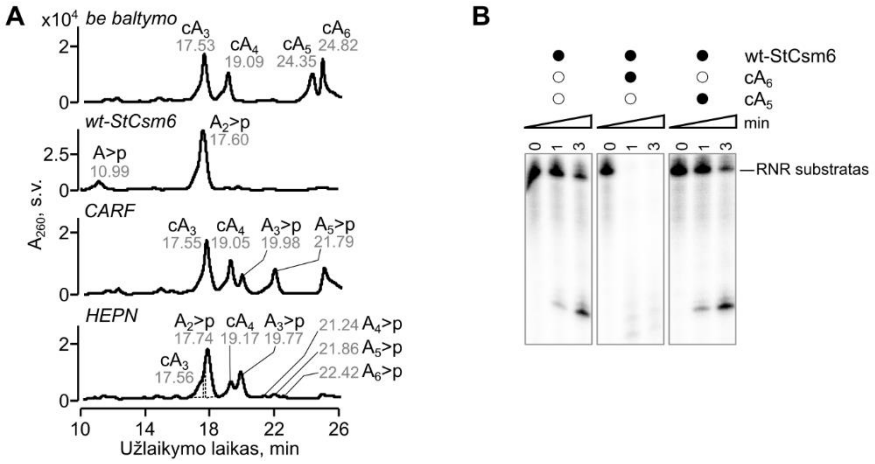
StCsm yra sudarytas iš CARF ir HEPN domenų, sujungtų α-spiraliniu 6H domenu. Norėdami identifikuoti, kuris domenas yra atsakingas už cA_6 hidrolizę, mes sukonstravome ir išgryninome atskirus StCsm6 CARF (1–169 aminorūgštys) ir HEPN (172–428 aminorūgštys) domenus (8 pav. A), kurie, kaip ir pilno ilgio StCsm6 (Kazlauskiene et al., 2017), tirpale formavo dimerus (8 pav. B). Domenų atskyrimą ir oligomerinės būsenos įvertinimą atliko Augustė Rimaitė. Izoliuotas HEPN domenas išlaikė nežymų RNR nukleazinį aktyvumą, panašų į neaktyvinto wt-StCsm6 aktyvumą, o atskirtas CARF domenas nepasižymėjo RNR nukleaziniu aktyvumu (8 pav. C).



9 pav. Atskirtų StCsm6 CARF ir HEPN domenų vykdoma cA_6 hidrolizė. (**A**) Atskirtų laukinio tipo CARF ir HEPN domenų bei jų mutantinių variantų dCARF (su D12A mutacija) ir dHEPN (su R371A, H376A mutacijomis) vykdoma cA_6 hidrolizė substrato nepritekliaus sąlygomis ([S]<[E]). Reakcijose buvo naudojama 50 nM žymėto cA_6 ir 1 μM CARF/dCARF variantų arba 10 μM HEPN/dHEPN variantų. Reakcijos produktai analizuoti denatūruojančiame gelyje. (**B**) Atskirtų laukinio tipo CARF ir HEPN domenų vykdoma cA_6 hidrolizė substrato pertekliaus sąlygomis ([S]>[E]). Reakcijose buvo naudojama 50 nM žymėto cA_6, 10 μM nežymėto cA_6 ir 1 μM CARF arba 10 μM HEPN domeno. Reakcijos produktai analizuoti denatūruojančiame gelyje. (**C**) Csm6 homologų CARF domeno fragmento (1-20 aminorūgštys) sekų palyginys. Žaliai pabrauktos aminorūgštys, esančios kilpoje, kuri, spėjama, yra netoli cA_6 surišimo vietos (8-14 aminorūgštys). Konservatyvus glutamatas pažymėtas žvaigždute. (**D**) cA_6 karpymo CARF ir HEPN domenais produktų analizė HPLC-MS metodu. Hidrolizės reakcijoms buvo naudojama 20 μM cA_6 ir 1 μM CARF arba 10 μM HEPN.

Toliau patikrinome, ar atskiri CARF ir HEPN domenai geba degraduoti radioaktyviai žymėtą cA_6 (9 pav. A ir B). Substrato nepritekliaus sąlygomis ([S]<[E]) CARF domenas hidrolizavo cA_6 efektyviau nei HEPN (9 pav. A). Šis tik silpnas HEPN aktyvumas gali būti nulemtas to, kad be prilieto CARF domeno HEPN negali būti alosteriškai aktyvinamas ir pasižymi tik nedideliu likutiniu RNR nukleaziniu aktyvumu (8 pav. C). Remdamiesi Csm6 homologų sekų palyginiu (9 pav. C) bei lygindami StCsm6 struktūrinį modelį su cA_4 karpančio *T. onnurineus* Csm6-cA_4 struktūra (PDB ID: 6O6V), identifikavome, kad konservatyvi aminorūgštis D12 galėtų būti svarbi CARF domeno žiedo nukleazės aktyvumui. Tai patvirtinome įvesdami D12A mutaciją CARF domene (dCARF variantas) – šis variantas visiškai neteko gebėjimo hidrolizuoti cA_6 (9 pav. A). Be to, HEPN domene įvedus mutacijas R371A ir H376A (dHEPN variantas), kurios, kaip žinoma, išjungia RNR karpymą (Kazlauskiene et al., 2017), buvo išjungta ir cA_6 hidrolizė (9 pav. A ir 8 pav. C). Vadinasi, tas pats HEPN domeno aktyvusis centras yra atsakingas tiek už linijinės RNR, tiek už cA_6 hidrolizę.

Pasitelkę HPLC-MS analizę nustatėme, kad CARF domenas hidrolizuoja cA_6 iki galutinio A_3>p produkto (9 pav. D). CARF dimere kiekvieno iš CARF domenų aktyvusis centras hidrolizuoja cA_6, tačiau tai įvyksta nekoordinuotai, todėl reakcijos pradžioje stebimas tarpinio junginio A_6>p susiformavimas. Kitaip nei CARF, HEPN domenas cA_6 karpo iki įvairaus ilgio linijinių produktų su 2′-3′-cikliniu fosfatu: A_6>p, A_4>p, A_3>p, A_2>p ir A>p (9 pav. D). Tad tiek CARF, tiek HEPN domenai hidrolizuoja cA_6, tačiau hidrolizės produktai skiriasi.



10 pav. Atskirtų CARF ir HEPN domenų vykdoma cA_n mišinio hidrolizė (**A**) cA_n hidrolizės reakcijų analizė HPLC-MS metodu. Reakcijose buvo ~30 µM cA_n, 10 nM wt-StCsm6 arba 1 µM CARF arba 10 µM HEPN domeno. Reakcijos buvo inkubuojamos 16 valandų arba, wt-StCsm atveju, 90 minučių. (**B**) Wt-StCsm6 RNR

hidrolizės reakcijos su cA_6 arba cA_5 signalinėmis molekulėmis. Reakcijose buvo naudojama 10 nM baltymo, 10 nM radioaktyviai žymėtos RNR ir 100 nM cA_6 arba cA_5. Mėginiai analizuoti elektroforezės denatūruojančiame gelyje metodu.

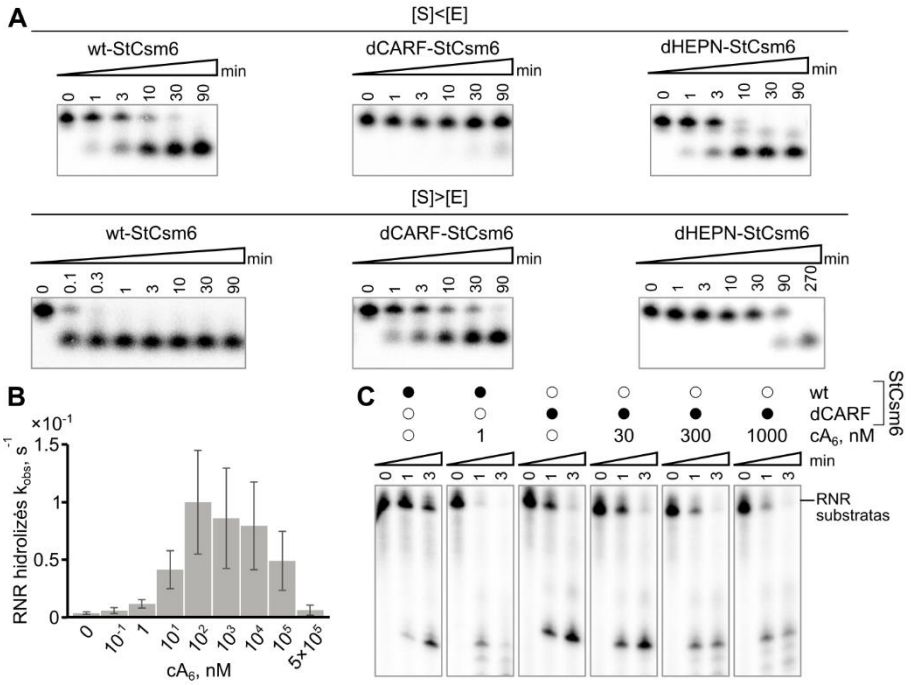
Toliau patikrinome, ar wt-StCsm6 ir atskirti jo domenai geba hidrolizuoti kitus StCsm komplekso gaminamus cA_n. Pilno ilgio StCsm6 hidrolizavo cA_n mišinį iki A_2>p ir A>p galutinių produktų, HEPN domenas linearizavo visų dydžių cA_n, o CARF domenas specifiškai hidrolizavo cA_6 iki A_3>p (10 pav. A). CARF domenas taip pat gebėjo linearizuoti cA_5 iki A_5>p, tačiau cA_5 wt-StCsm6 RNR nukleazinio aktyvumo nestimuliuoja (10 pav. B), todėl cA_5 hidrolizė, tikėtina, nėra specifinis reguliacijos mechanizmas. Tad StCsm6 CARF domenas yra žiedo nuklеazė, specifiška cA_6, o HEPN domenas degraduoja cA_6 ir kitus cA_n, pasitelkdamas savo RNR nukleazinį aktyvumą.

CARF ir HEPN domenų vaidmuo pilno ilgio StCsm6 baltyme

Siekdami įvertinti tikėtinos alosterinės sąveikos tarp CARF ir HEPN domenų pilno ilgio StCsm6 baltyme svarbą cA_6 hidrolizei, sukonstravome ir išgryninome pilno ilgio StCsm6 baltymus su išveiklintu CARF domenu (dCARF-StCsm6) arba su išveiklintu HEPN domenu (dHEPN-StCsm6). Substrato nepritekliaus sąlygomis ([S]<[E]) dHEPN-StCsm6 hidrolizavo cA_6 panašiu greičiu kaip ir wt-StCsm6 (atitinkamai k_{obs}= 0.27 ± 0.10 min^{-1} ir k_{obs}= 0.19 ± 0.04 min^{-1}) ir tik apytiksliai du kartus lėčiau nei pavienis CARF domenas (k_{obs}= 0.47 ± 0.18 min^{-1}) (11 pav. A ir 9 pav. A). Be to, dCARF-StCsm6, taip pat kaip pavienis HEPN domenas, beveik nehidrolizavo cA_6 (11 pav. A ir 9 pav. A). Tai rodo, kad esant mažai substrato koncentracijai, cA_6 yra surišamas ir hidrolizuojamas CARF domene.

Aktyvinto HEPN domeno giminingumą cA_6 įvertinome atlikdami konkurencinį RNR karpymo eksperimentą. Tik labai didelė cA_6 koncentracija (>100 μM lyginant su 10 nM RNR koncentracija) ėmė slopinti RNR karpymą (11 pav. B). Vadinasi, HEPN domenas pasižymi nedideliu giminiškumu cA_6, o pagrindinis aktyvinto HEPN domeno substratas yra linijinė RNR.

Substrato pertekliaus sąlygomis ([S]>[E]) dHEPN-StCsm6, kaip ir pavienis CARF domenas, hidrolizavo cA_6 gerokai lėčiau nei wt-StCsm6 ar dCARF-StCsm6 (11 pav. A ir 9 pav. B). Tai reiškia, kad alosteriškai aktyvintas HEPN domenas gerokai efektyviau hidrolizuoja cA_6 nei CARF domenas. O hidrolizės efektyvumo skirtumas tarp dCARF-StCsm6 ir wt-StCsm6 galėtų būti nulemtas D12A mutacijos, kuri taip pat sumažina CARF domeno giminingumą cA_6 aktyvatoriui – dCARF-StCsm6 RNR nukleazei aktyvinti reikia ženkliai didesnės cA_6 koncentracijos nei wt-StCsm6 atveju (11 pav. C).



11 pav. Pilno ilgio StCsm6 vykdoma cA_6 hidrolizė. (**A**) cA_6 hidrolizės reakcijos substrato nepritekliaus ([S]<[E]) ir substrato pertekliaus ([S]>[E]) sąlygomis. Substrato nepritekliaus sąlygomis atliktose reakcijose buvo naudojama 50 nM radioaktyviai žymėto cA_6 ir 1 µM nurodyto StCsm6 varianto. Substrato pertekliaus sąlygomis atliktose reakcijose papildomai pridėta 10 µM nežymėto cA_6. (**B**) Konkurencinis RNR karpymo eksperimentas. Buvo vertinamas RNR hidrolizės greitis, reakcijos mišinyje esant 10 nM radioaktyviai žymėtos RNR, 10 nM wt-StCsm6 ir 0-500 µM cA_6. Pateikiami trijų eksperimentų RNR hidrolizė greičio konstantų (k_{obs}) verčių vidurkiai su standartiniai nuokrypiais. (**C**) dCARF-StCsm6 vykdomos RNR hidrolizės priklausomybė nuo cA_6 koncentracijos. Reakcijose buvo naudojama 10 nM baltymo, 10 nM radioaktyviai žymėtos RNR ir 0-1000 nM cA_6.

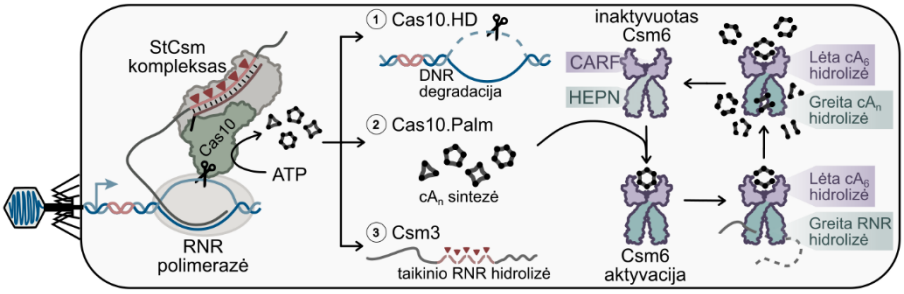
Apibendrinant, StCsm6 CARF domenas yra lėta, bet cA_6 specifiška žiedo nukleazė, o alosteriškai aktyvintas HEPN domenas yra efektyvi RNR nukleazė, kuri pirmiausia degraduoja linijinę RNR ir tik esant aukštai cA_6 koncentracijai hidrolizuoja šią signalinę molekulę.

III-A tipo CRISPR-Cas sistemos reguliacijos mechanizmas

Įvertinę StCsm komplekso vykdomą cA_n sintezę *E. coli* ląstelėse bei atlikę StCsm6 žiedo nukleazinio aktyvumo analizę *in vitro*, galėjome pasiūlyti papildytą *S. thermophilus* III-A tipo CRISPR-Cas veikimo mechanizmą su cA_6 signalinio kelio reguliacijos etapu (12 pav.).

Svetimą transkriptą surišęs StCsm interferencijos kompleksas ima sintetinti įvairaus dydžio cA_n. StCsm6 sensorinis CARF domenas suriša cA_6

aktyvatorių, o tai aktyvina HEPN domeno RNR nukleazinį aktyvumą ir yra inicijuojama linijinės RNR degradacija. Kadangi CARF pasižymi žiedo nukleaziniu aktyvumu, specifišku cA_6, jis lėtai hidrolizuoja surištą aktyvatorių iki A_3>p. Aktyvatoriaus perkirpimas veikia kaip laikmačio mechanizmas, išjungiantis HEPN domeną. Prie CARF domeno prisijungus naujai cA_6 molekulei, HEPN domenas vėl aktyvinamas. Kai dėl HEPN domeno RNR nukleazinio aktyvumo sumažėja lokali linijinės RNR koncentracija, o cA_n koncentracija padidėja, HEPN domenas ima efektyviai degraduoti cA_n. StCsm interferencijos komplekso Csm3 subvienetams perkerpus taikinio RNR ir paleidus kirpimo produktus, cA_n sintezė yra išjungiama, o ląstelėje susikaupę cA_n yra toliau degraduojami aktyvinto StCsm6, taip visiškai išjungiant signalinį kelią. Dėl to infekciją įveikusi ląstelė gali išvengti žūties ir atsigauti.



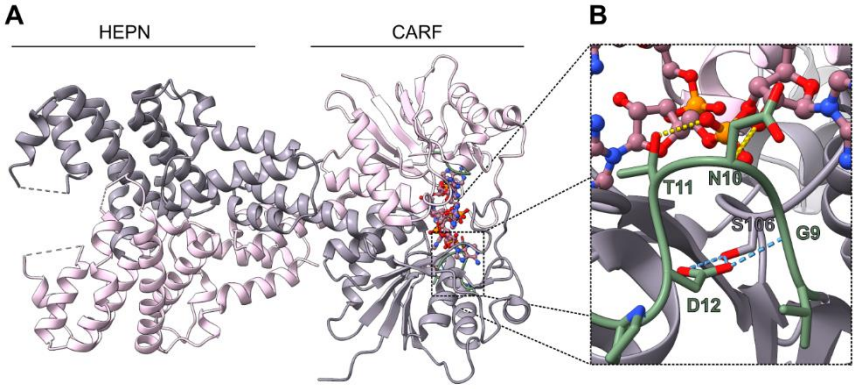
12 pav. Papildytas *S. thermophilus* III-A tipo CRISPR-Cas sistemos veikimo mechanizmas. StCsm interferencijos komplekso susirišimas su užpuoliko transkriptu įjungia Cas10 subvieneto fermentinius aktyvumus: HD nukleazės domenas ima degraduoti transkripcijos burbule išstumtą viengrandinę DNR (1), o Palm domenas ima iš ATP sintetinti cA_n (2). Pagalbinio efektoriaus StCsm6 CARF domenas suriša cA_6 ir tai inicijuoja nespecifinę RNR hidrolizę aktyvuotu HEPN RNR nukleazės domenu. StCsm6 aktyvavimą reguliuoja CARF domenas, kuris veikia kaip cA6 specifiška žiedo nukleazė, lėtai paverčianti cA_6 į linijinį A_3>p. Be to, StCsm komplekso Csm3 subvieneto atliekama tikslinės RNR hidrolizė (3) išjungia cA_n sintezę, o susikaupusius cA_n hidrolizuoja aktyvuotas StCsm6 HEPN domenas. Signalas yra visiškai išjungiamas ir ląstelė gali atsigauti.

Struktūriniais duomenimis paremtos įžvalgos į cA_6 hidrolizės mechanizmą

Lygiagrečiai su mūsų cA_6 signalinio kelio reguliacijos tyrimu, buvo publikuotas straipsnis, aprašantis *Enterococcus italicus* Csm6 (EiCsm6) kristalinę struktūrą ir cA_6 hidrolizę (Garcia-Doval et al., 2020), o vėliau struktūriškai ir biochemiškai buvo charakterizuotas ir StCsm6′ (McQuarrie et al., 2023). Csm6 homologų tyrimai patvirtino mūsų rezultatus, kad tiek CARF, tiek HEPN domenai dalyvauja cA_6 karpyme, o esant žemai cA_6

koncentracijai, pagrindinį vaidmenį atlieka CARF domeno žiedo nukleazinis aktyvumas (Garcia-Doval et al., 2020; McQuarrie et al., 2023).

Remiantis publikuotomis struktūromis, A_6 hidrolizę nulemia steriniai veiksniai, kurie priverčia ligandą įgyti konformaciją, leidžiančią ribozės 2′-OH vykdyti nukleofilinę ataką į hidrolizuojamą fosfatą (Garcia-Doval et al., 2020; McQuarrie et al., 2023). Nors konservatyvi D12 aminorūgštis nesąveikauja tiesiogiai su cA_6, ji gali stabilizuoti katalitinę kilpą, turinčią G-X-[T/S]-DP motyvą, per vandenilinius ryšius su G9 amidu arba konservatyvaus serino (S106 StCsm6′; S112 EiCsm6; S105 StCsm6) šonine grandine (13 pav.). Tai paaiškintų D12A mutacijos poveikį CARF domeno aktyvumui (9 ir 11 pav.).



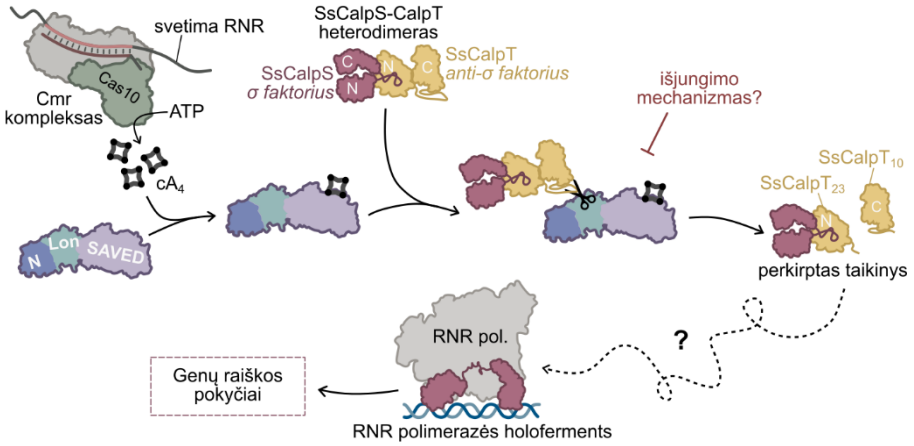
13 pav. Katalitinės kilpos stabilizavimas. (**A**) cA_6 aktyvatorių surišusio StCsm6′ struktūra PDB ID: 8PE3 (McQuarrie et al., 2023). StCsm6′ dimerą formuojantys monomerai nuspalvinti rožine ir violetine. (**B**) Priartinta katalitinė kilpa (žalia) ir jos formuojami kontaktai. T11 šoninės grandinės OH grupė sudaro vandenilinį ryšį su hidrolizuojamuoju fosfatu, o N10 amidas sudaro ryšį su ribozės 2′-OH (geltonos punktyrinės linijos). D12 šoninė grandinė galėtų sudaryti katalitinės kilpos padėtį stabilizuojančius vandenilinius ryšius su G9 pagrindinės grandinės atomais arba su S106 šonine grupe (mėlynos punktyrinės linijos).

Trinario efektoriaus mechanizmo tyrimai

Pastaraisiais metais buvo apibūdinta daugybė naujų III tipo CRISPR-Cas CARF efektorių (Steens et al., 2022; Stella and Marraffini, 2024), o dar daugiau jų identifikuota kompiuteriniais metodais (Altae-Tran et al., 2023; Hoikkala et al., 2024; Makarova et al., 2020a), tačiau geriausiai ištirtais efektoriais išlieka CARF-nukleazės, o SAVED domeną turinčių efektorių molekuliniai mechanizmai tebėra menkai suprasti. Todėl antruoju

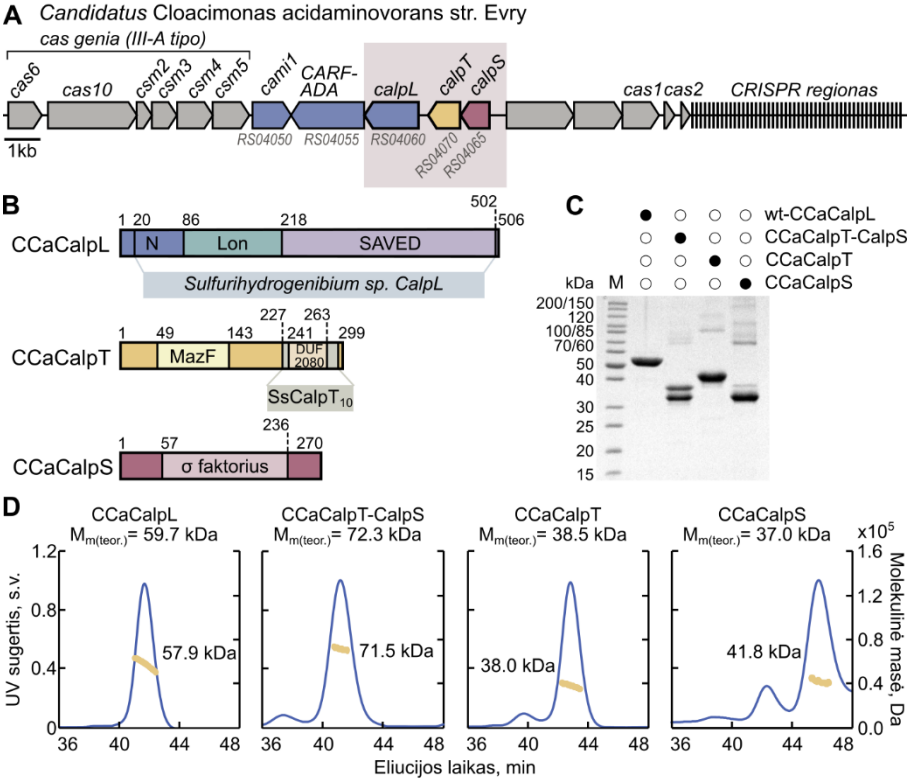
doktorantūros studijų tyrimo objektu pasirinkome Lon-SAVED efektorių, kuris, kartu su dviem tame pačiame operone koduojamais baltymais, yra randamas *Candidatus* Cloacimonas acidaminovorans (CCa) III-A tipo CRISPR-Cas sistemoje.

Mums bevykdant šiuos tyrimus Rouillon ir kt. apibūdino homologinį efektorių iš *Sulfurihydrogenibium* spp. YO3AOP1 (Ss) (Rouillon et al., 2023). Jie nustatė, kad Lon-SAVED baltymas, pavadintas CalpL (angl. *CRISPR-Cas associated Lon protease*), kartu su CalpT ir CalpS sudaro trinarį efektorių. Be to, parodė, kad cA_4 aktyvinta SsCalpL proteazė perkerpa anti-σ faktorių SsCalpT ir taip išlaisvina σ faktorių SsCalpS, kad jis sąveikautų su RNR polimeraze (RNR pol) ir keistų genų raišką (Rouillon et al., 2023) (14 pav.). Tačiau išliko neatsakyti svarbūs klausimai: koks yra CalpT perkirpimo CalpL proteaze mechanizmas, kokia yra įvykių seka, nulemianti CalpS paleidimą po CalpT perkirpimo, ir kaip šis trinaris efektorius yra reguliuojamas. Norėdami atsakyti į šiuos klausimus, mes tęsėme išsamių CCaCalpL-CalpT-CalpS efektoriaus charakterizavimą, ypatingą dėmesį skirdami reguliacijos mechanizmui.



14 pav. *Sulfurihydrogenibium* spp. III-B tipo CRISPR-Cas trinario pagalbinio efektoriaus CalpL-CalpT-CalpS veikimo mechanizmas. Užpuoliko RNR surišęs III-B tipo interferencijos kompleksas sintetina cA_4, kuriuos atpažįsta ir suriša SsCalpL baltymo SAVED domenas. cA_4 surišimas aktyvina SsCalpL Lon proteazės domeną, kuris proteolizuoja anti-σ faktorių SsCalpT. Dėl SsCalpT perkirpimo dar nenustatytu būdu išlaisvinamas σ faktorius SsCalpS, kuris toliau prisijungia prie RNR polimerazės ir gali keisti genų raišką.

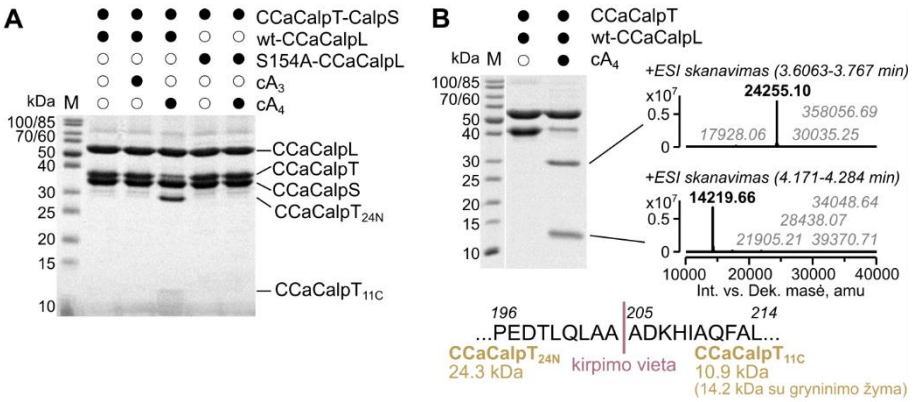
Pirmiausiai atlikę homologinių regionų paiešką patvirtinome, kad Ca. *C. acidaminovorans* šalia III-A tipo CRISPR-Cas sistemos koduojamas trinaris CalpL-CalpT-CalpS efektorius yra homologiškas randamam *Sulfurihydrogenibium* spp. YO3AOP1 (15 pav. A ir B) (Rouillon et al., 2023).



15 pav. *Ca.* C. acidaminovorans koduoja trinarį pagalbinį efektorių. (**A**) *Ca.* C. acidaminovorans III-A tipo CRISPR-Cas operono (NCBI ID: NC_020449.1) schema. Mėlyna spalva pažymėti pagalbinius efektorius koduojantys genai. Rožiniame fone – *calpL-calpT-calpS* trinario efektoriaus genai. (**B**) CCaCalpL, CCaCalpT ir CCaCalpS baltymų sandaros schema su homologinėmis sritimis, nustatytomis naudojant HHpred. (**C**) Išgrynintų baltymų SDS-PAGE elektroforezė: wt-CCaCalpL, CCaCalpT-CalpS (CalpS su C-gryninimo žyme), CCaCalpT (su C-gryninimo žyme) ir CCaCalpS (su C-gryninimo žyme). M – baltymų molekulinės masės standartas. (**D**) Išgrynintų baltymų SEC-MALS analizė: CCaCalpL (C-gryninimo žymė pašalinta TEV proteaze), CCaCalpT-CalpS (C-gryninimo žyme ant CCaCalpS), CCaCalpT (su C-gryninimo žyme) ir CCaCalpS (su C-gryninimo žyme).

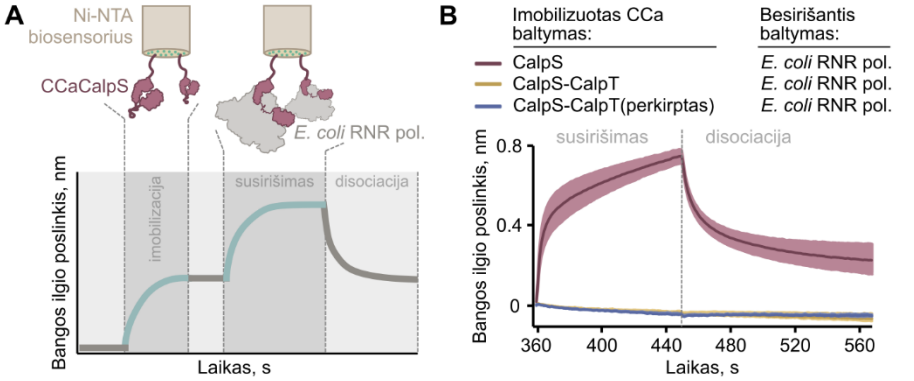
Išgrynintų trinario efektoriaus komponentų SEC-MALS analizė atskleidė, kad CCaCalpL, CCaCalpT ir CCaCalpS tirpale yra monomerai, o kartu

išreikšti CCaCalpT ir CCaCalpS baltymai sudaro stabilų heterodimerą (15 pav. D).



16 pav. CCaCalpL aktyvatoriaus identifikavimas ir CCaCalpT kirpimo vietos nustatymas. (**A**) CCaCalpL proteazinio aktyvumo analizė SDS-PAGE metodu. Reakcijose buvo naudojama 5 µM CCaCalpL (wt arba Lon proteazės aktyvaus centro S154A mutanto), 5 µM CCaCalpT-CalpS ir 0 arba 25 µM cA_4 arba cA_3. Reakcijos buvo inkubuotos 120 min. (**B**) CCaCalpT kirpimo vietos nustatymas. Proteazinės reakcijos produktai buvo analizuojami SDS-PAGE metodu ir tiksli masė įvertinama HPLC-MS metodu. Apačioje schematiškai pavaizduota kirpimo vieta CCaCalpT baltymo sekoje.

Toliau nustatėme, kad tik cA_4 veikia kaip CCaCalpL proteazės aktyvatorius ir aktyvinta CCaCalpL perskelia CCaCalpT, esantį CCaCalpT-CalpS heterodimere, į du proteolizės fragmentus (16 pav. A). Taip pat nustatėme, kad CCaCalpT baltymas yra kerpamas alaninais praturtiname motyve ^{202}LAAA, tarp A204 ir A205 aminorūgščių. Po kirpimo susidaro 24.3 kDa N-galinis ($CCaCalpT_{24N}$) ir 10.9 kDa C-galinis ($CCaCalpT_{11C}$) fragmentai (16 pav. B).

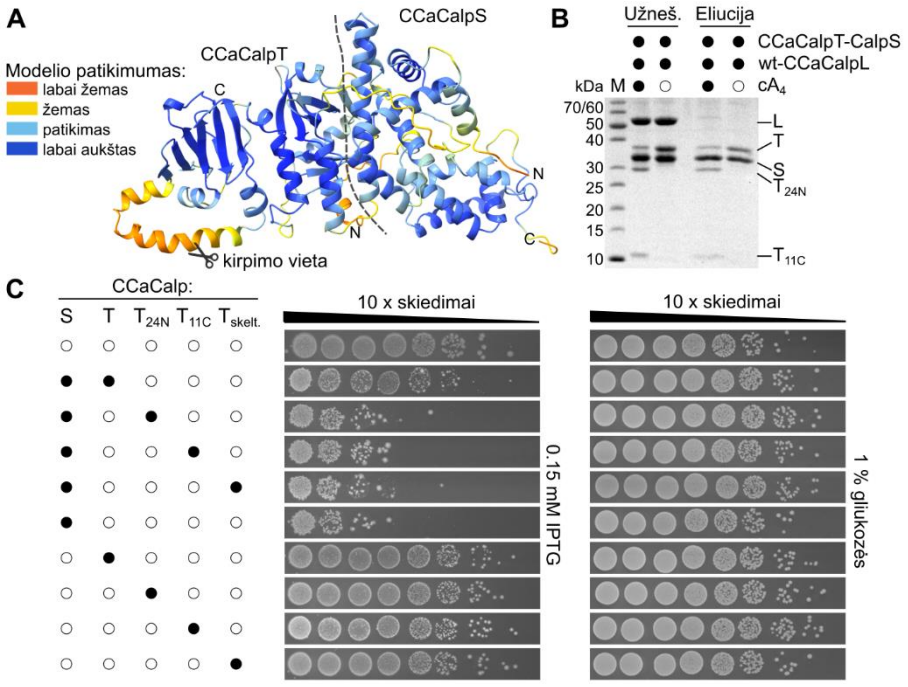


17 pav. σ faktorių primenančios CCaCalpS baltymo savybės. (**A**) BLI eksperimento schema. CCaCalpS, CCaCalpT-CalpS arba perkirptas CCaCalpT-

CalpS buvo imobilizuoti ant Ni-NTA biosensoriaus (imobilizacija), ir sensorius patalpinamas į šulinėlį su *E. coli* RNR pol (susirišimas). Susirišimą indikuoja stebimas bangos ilgio poslinkis. (**B**) CCaCalpS sąveikos su *E. coli* RNR pol BLI eksperimento asociacijos ir disociacijos kreivės. Pateikiamos vidutinės trijų eksperimentų reikšmės su standartiniais nuokrypiais.

Kadangi CCaCalpS yra panašus į σ faktorių, panaudoję biosluoksnio interferometrijos (BLI, angl. *biolayer interferometry*) metodą pademonstravome, kad ant biosensoriaus imobilizuotas pavienis CCaCalpS geba sąveikauti su *E. coli* RNR pol (17 pav. A ir B), o CCaCalpT veikia kaip anti-σ faktorius ir blokuoja CCaCalpS sąveiką su RNR pol. Tačiau CCaCalpL vykdoma CCaCalpT proteolizė neišlaisvino CCaCalpS susirišimui su RNR pol (17 pav. C), vadinasi, vien CCaCalpT kirpimo CCaCalpL proteaze neužtenka, kad CCaCalpS būtų paleidžiamas iš CCaCalpT-CalpS komplekso.

CCaCalpS paleidimo mechanizmas

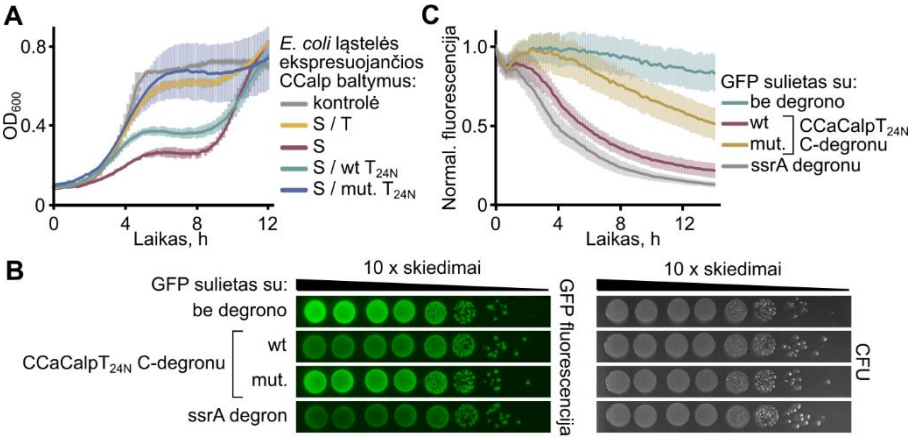


18 pav. CCaCalpS paleidimo tyrimai. (**A**) CCaCalpT-CalpS heterodimero AlphaFold2 modelis, nuspalvintas pagal modelio patikimumo įverčius: labai žemas (*pLDDT* <50); žemas (70> *pLDDT* >50); patikimas (90> *pLDDT* >70); labai aukštas (*pLDDT* >90). Punktyrinė linija žymi sąveikos paviršių. (**B**) CCaCalpT-CalpS heterodimero disociacijos po CCaCalpT perkirpimo CCaCalpL proteaze analizė. Su CCaCalpS susirišę baltymai analizuoti SDS-PAGE metodu. M – baltymų molekulinės masės standartas. (**C**) *E. coli* ląstelių, išreiškiančių CCaCalpS ir pilno ilgio CCaCalpT,

$CCaCalpT_{24N}$, $CCaCalpT_{11C}$ arba perkirpto CCaCalpT kombinacijas, serijiniai skiedimai.

Pagal CCaCalpT-CalpS heterodimero AlphaFold2 modelį, CCaCalpT su CCaCalpS sąveikauja per CCaCalpT N-galinį domeną, o CCaCalpT kirpimo vieta yra nutolusi nuo baltymų sąveikos paviršiaus (18 pav. A). Išanalizavę kokie baltymai lieka asocijuoti su CCaCalpS po CCaCalpL katalizuojamos reakcijos, nustatėme, kad CCaCalpS liko susirišęs su abiem CCaCalpT fragmentais (18 pav. B). Tai reiškia, kad ląstelėse $CCaCalpT_{24N}$ turi būti papildomai apdorojamas, kad paleistų CCaCalpS.

Tam patikrinti *E. coli* ląstelėse išreiškėme įvairias CCaCalpS ir CCaCalpT variantų kombinacijas (18 pav. C). Pavienio CCaCalpS raiška buvo toksiška *E. coli* ląstelėms, o šį toksiškumą gerokai sumažino pilno ilgio CCaCalpT raiška, bet ne jo kirpimo produktų (18 pav. C), patvirtinant, kad tik pilno ilgio CCaCalpT veika kaip anti-σ faktorius.



19 pav. $CCaCalpT_{24N}$ degrono tyrimai. (**A**) *E. coli* ląstelių, išreiškiančių CCaCalpS ir laukinio tipo (wt) $CCaCalpT_{24N}$ arba A204D mutaciją turintį $CCaCalpT_{24N}$. Pateikiamos trijų eksperimentų vidutinės reikšmės su standartiniais nuokrypiais. (**B**) GFP fluorescencijos slopinimo eksperimentas. *E. coli* ląstelės, transformuotos plazmidėmis, koduojančiomis GFP, sulietą su skirtingais degronais (PETLQLAA – wt $CCaCalpT_{24N}$ C-degronas, PETLQLAD – mutantinis $CCaCalpT_{24N}$ C-degronas, ENYALAA – dalinis ssrA C-degrono fragmentas). Wt-$CCaCalpT_{24N}$ C-degronas sumažino GFP fluorescenciją (kairėje), tačiau neturėjo įtakos kolonijas formuojančių vienetų (CFU) skaičiui (dešinėje). (**C**) Su degronais (kaip B dalyje) sulieto GFP degradacijos tyrimas *E. coli* ląstelėse laikui bėgant. GFP fluorescencijos signalas buvo normalizuotas pagal optinį tankį. Pateikiamos trijų eksperimentų vidutinės reikšmės su standartiniais nuokrypiais.

Priešingai nei *in vitro* (17 pav. B), *E. coli* ląstelėse CCaCalpT kirpimo produktų ko-raiška neturėjo jokios įtakos CCaCalpS toksiškumui (18 pav. C), kas reiškia, kad ląstelėse CCaCalpS yra paleidžiamas po CCaCalpT

perkirpimo. Šį paleidimą nulemia ant $CCaCalpT_{24N}$ fragmento naujai susiformavusiame C-gale eksponuojama degrono seka, kuri nukreiptų jį tolimesnei nuo ATP priklausomų proteazių degradacijai. Tai patvirtinome stebėdami CCaCalpS toksiškumą *E. coli* ląstelėms, išreiškiančioms $CCaCalpT_{24N}$ su laukinio tipo arba mutaciją turinčiu C-degronu (19 pav. A), bei prilieję $CCaCalpT_{24N}$ C-degrono seką prie reporterinio žaliai fluorescuojančio baltymo (GFP, angl. *Green fluorescent protein*) (19 pav. B ir C). Kaip teigiamą kontrolę prie GFP priliejome ir dalinį, gerai characterizuotą SsrA degroną (ENYALAA) (Fei et al., 2020).

Taigi, po CCaCalpL proteazės vykdomo CCaCalpT kirpimo susidarantis $CCaCalpT_{24N}$ fragmentas turi C-degroną, kuris nukreipia jį tolimesniam degradavimui, taip išlaisvindamas *E. coli* ląstelėms toksišką CCaCalpS.

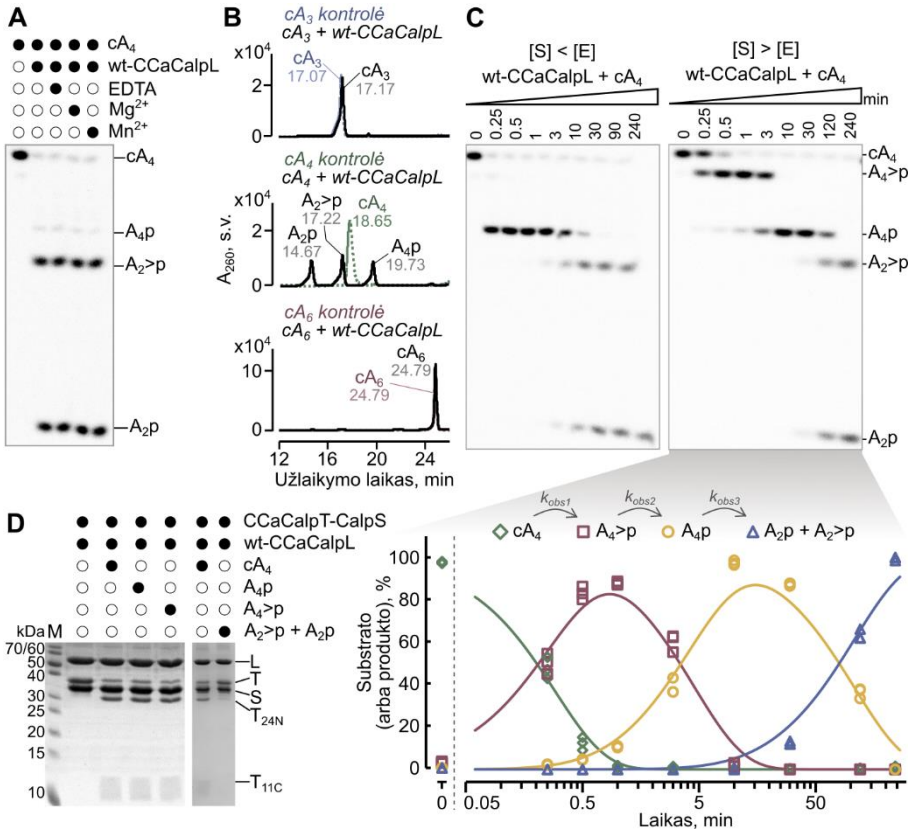
Reguliacijos mechanizmas – CCaCalpL pasižymi žiedo nukleaziniu aktyvumu

Toliau siekėme patikrinti, ar CCaCalpL geba reguliuoti save hidrolizuodamas aktyvuojančias signalines molekules kaip StCsm6. Nustatėme, kad CCaCalpL pasižymi žiedo nukleaziniu aktyvumu – baltymas hidrolizavo radioaktyviai žymėtą cA_4, o šiai reakcijai nebuvo reikalingi metalo jonai (20 pav. A). CCaCalpL specifiškai hidrolizuoja tik cA_4, susidarant trims kirpimo produktams: A_4p (linijinis tetraadenilatas su 3′-fosfatu), A_2>p (linijinis diadenilatas su 2′,3′-cikliniu fosfatu) ir A_2p (linijinis diadenilatas su 3′-fosfatu) (20 pav. B).

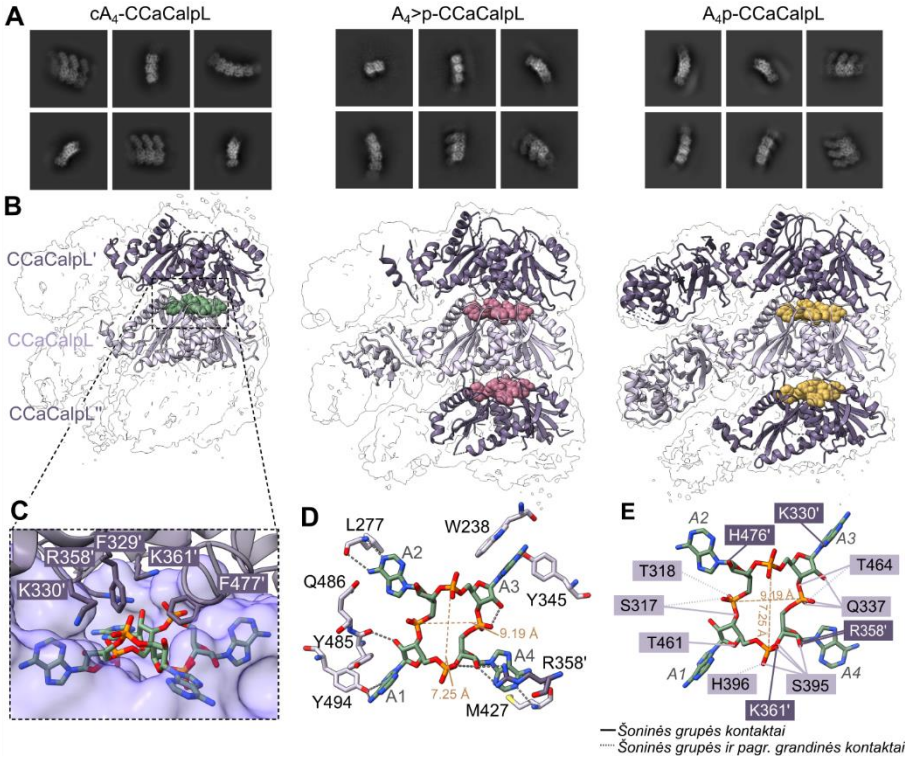
Substrato pertekliaus sąlygomis ([S] > [E]) cA_4 pirmiausia buvo verčiamas į pirmąjį tarpinį reakcijos produktą A_4>p (linijinį tetraadenilatą su 2′,3′-cikliniu fosfatu), o tuomet A_4>p buvo konvertuojamas į antrą linijinį tarpinį produktą A_4p, kuris galiausiai buvo perskeliamas į du linijinius dinukleotidus A_2>p ir A_2p (20 pav. C). Tarpinių produktų kaupimasis reakcijos eigoje rodo, kad (i) po kiekvieno kirpimo etapo tarpiniai reakcijos produktai disocijuoja iš CCaCalpL aktyvaus centro ir (ii) kiekviename vėlesniame reakcijos etape CCaCalpL žiedo nukleazės reakcijos greitis mažėja daugiau nei 10 kartų (20 pav. C).

Substrato nepritekliaus atveju ([S] < [E], 50 nM cA_4 ir 1 μM CCaCalpL) reakcija vyko pagal tą patį mechanizmą: cA_4 buvo palyginti greitai (per keletą sekundžių) konvertuojamas į A_4>p ir A_4p, o A_4p palyginti lėtai (per dešimtis minučių) konvertuojamas į A_2>p ir A_2p (20 pav. C).

Be to, nustatėme, kad tiek A_4>p, tiek A_4p stimuliuoja CCaCalpL proteazės aktyvumą, o galutiniai reakcijos produktai A_2>p ir A_2p neaktyvina CCaCalpL (20 pav. D).



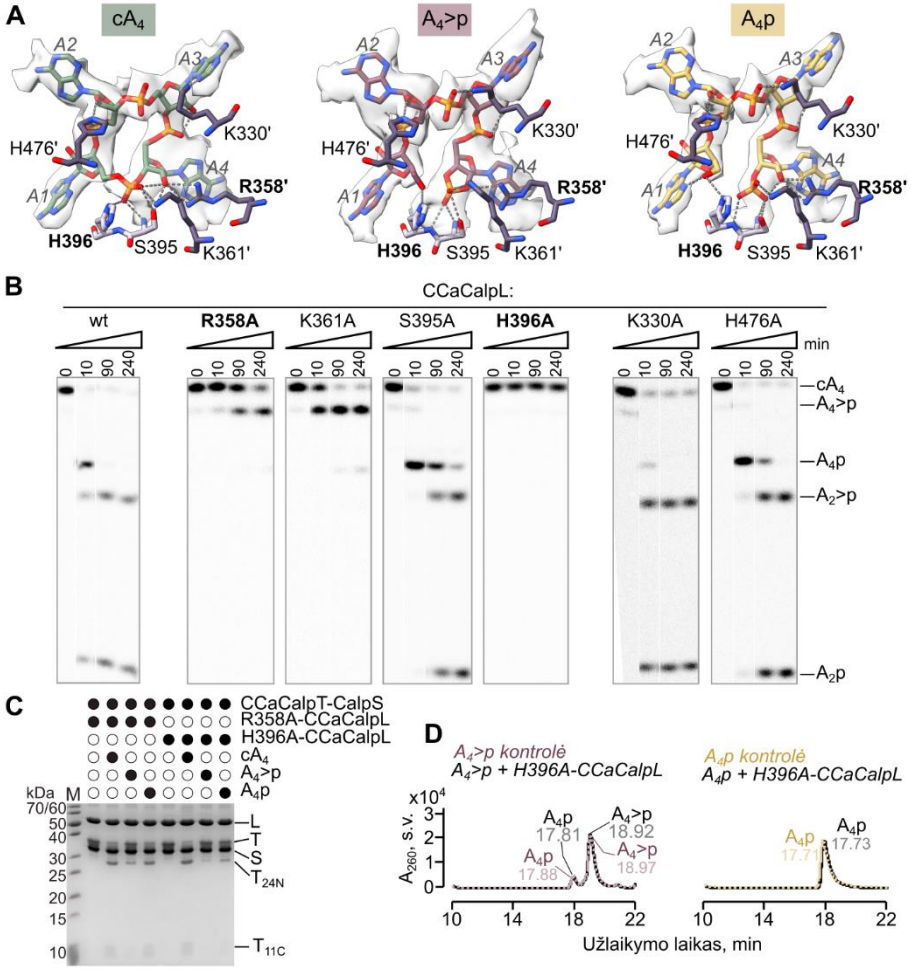
20 pav. CCaCalpL žiedo nukleazinio aktyvumo tyrimai. (**A**) Radioaktyviai žymėto cA_4 karpymas wt-CCaCalpL baltymu reakcijos mišinyje su metalo jonais arba be jų. Reakcija, kurioje buvo 1 µM CCaCalpL, 50 nM radioaktyviai žymėto cA_4 ir 1 mM EDTA, 10 mM $Mg(CH_3COO)_2$ arba 1 mM $MnCl_2$, buvo inkubuota 90 minučių. (**B**) cA_3, cA_4 ir cA_6 karpymo wt-CCacalpL baltymu reakcijų HPLC-MS analizė. Punktyrinės linijos žymi kontrolines reakcijas be baltymo (cA_3 – mėlyna, cA_4 – žalia, cA_6 – raudona). (**C**) cA_4 karpymas substrato nepritekliaus ([S] < [E]) (kairėje) ir substrato pertekliaus ([S] > [E]) sąlygomis (dešinėje). Reakcijos mišiniuose buvo 1 µM wt-CCaCalpL ir 50 nM radioaktyviai žymėto cA_4, reakcijoms su substrato pertekliumi buvo papildomai įdėta 10 µM nežymėto cA_4. Produktai buvo patvirtinti atliekant HPLC-MS analizę. Apačioje substrato cA_4 ir tarpinių kirpimo produktų kiekio priklausomybė nuo laiko. Pateikiami trijų eksperiemntų duomenys. Duomenims pritaikius eksponentinio kitimo modelį (vientisos linijos) buvo įvertintos reakcijų greičio konstantos (k_{obs}): $k_{obs1} = 34 \pm 2$ min^{-1}, $k_{obs2} = 2.2 \pm 0.1$ min^{-1} ir $k_{obs3} = 0.092 \pm 0.003$ min^{-1}. (**D**) Wt-CCaCalpL proteazinė reakcija aktyvacijai naudojant cA_4, A_4>p, A4p arba A_2>p ir A_2p. Reakcijos mišinyje buvo 5 µM CCaCalpL, 5 µM CCaCalpT-CalpS ir 0 arba 25 µM nurodytų adenilatų. Reakcijos buvo vykdomos 120 min ir analizuotos SDS-PAGE metodu. M – baltymų molekulinės masės standartas.



21 pav. CCaCalpL struktūriniai tyrimai. (**A**) Wt-CCaCalpL, susirišusio su cA_4, A_4>p arba A_4p, krio-EM analizė, 2D klasės. (**B**) Wt-CCaCalpL susirišusio su cA_4 (žalias), A_4>p (raudonas) ir A_4p (geltonas) modeliai ir elektronų tankio žemėlapiai. Gretimi subvientai filamentuose nuspalvinti skirtingais violetinės atspalviais. (**C**) Sąveikos tarp gretimų SAVED domenų ir tarp jų surišto cA_4. Šviesiai violetinė - pirminis CCaCalpL subvienetas, tamsiai violetinė – antrinis CCaCalpL′ subvienetas. (**D**) Kontaktai tarp baltymo ir cA_4 adenino bazių. (**E**) Kontaktai tarp baltymo ir cA_4 fosfodiesterinio karkaso.

Siekdami nustatyti detalų cA_4 karpymo mechanizmą atlikome wt-CCaCalpL, susirišusio su cA_4, A_4>p arba A_4p, struktūrinius tyrimus naudodami kriogeninės elektroninės mikroskopijos (krio-EM) metodą. Krio-EM eksperimentus atliko dr. Giedrius Sasnauskas ir dr. Giedrė Tamulaitienė. Wt-CCaClpL, susirišęs tiek su cA_4, tiek su A_4>p, tiek su A_4p, formavo trumpas filamentines struktūras (21 pav. A). Baltymo, susirišusio su cA_4, atveju gavome 2,97 Å skiriamosios gebos krio-EM žemėlapį, susirišusio su A_4>p - 2,81 Å, o susirišusio su A_4p – 2,75 Å. Gauti elektroniniai tankiai leido rekonstruoti modelius dviejų ir trijų gretimų CCaCalpL subvienetų SAVED domenų bei dviejų pilnų CCaCalpL subvienetų ir SAVED domeno iš trečiojo CCaCalpL subvieneto (21 pav. B). Visais trejais atvejais CCaCalpL

monomerai filamente oligomerizuojasi „galva - uodega“ principu, o aktyvatorius yra įsiterpęs tarp gretimų SAVED domenų (21 pav. B). Pirminis CCaCalpL subvienetas suriša aktyvatorių gilios kišenės viduje, o gretimas (antrinis) subvienetas CCaCalpL′ su aktyvatoriumi sąveikauja per iškilų paviršiaus motyvą priešingoje domeno pusėje (21 pav. C). Surišto cA_4 adenino bazės ir fosfodiesterinis karkasas sudaro sąveikas su dvejais SAVED domenais (21 pav. D ir E) - aktyvatoriaus surišimas SAVED-SAVED sąveikos paviršiuje nulemia CCaCalpL oligomerizaciją.



22 pav. Už cA_4 hidrolizę atsakingos CCaCalpL aminorūgštys. (**A**) Spėjamos katalitinės aminorūgštys ir jų sąveika su cA_4, A_4>p ir A_4p. (**B**) Žymėto cA_4 hidrolizė naudojant potencialių katalitinių aminorūgščių mutantinius ir wt CCaCalpL variantus. Reakcijose buvo 1 µM CCaCalpL ir 50 nM žymėto cA_4. wt CCaCalpL gelis pakartotinai panaudotas iš 20 pav. C. (**C**) R358A ir H396A-CCaCalpL proteazės reakcija, naudojant cA_4, A_4>p arba A_4p kaip aktyvatorius. Reakcijose buvo 5 µM CCaCalpL, 5 µM CCaCalpT-CalpS ir 0 arba 25 µM cA_4, A_4>p ar A_4p. Reakcijos

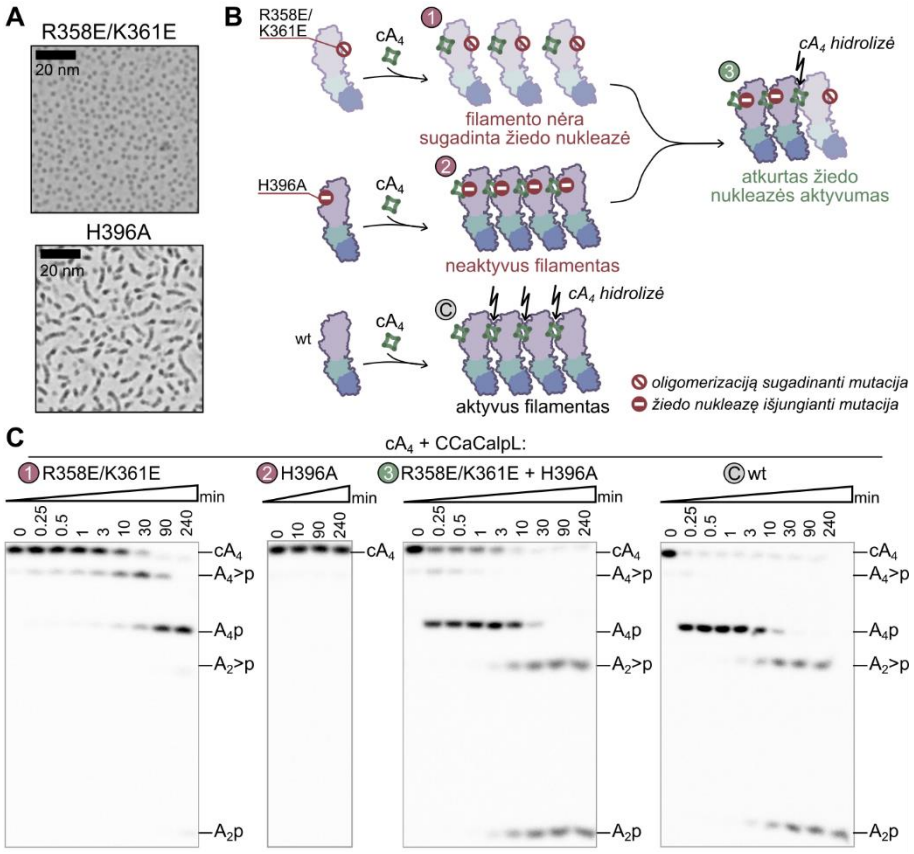
inkubuotos 120 min ir analizuotos SDS-PAGE. (**D**) A_4>p ir A_4p hidrolizės reakcijų HPLC-MS analizė naudojant H396A-CCaCalpL.

Remiantis nustatytomis CCaCalpL struktūromis, S395 ir H396 iš pirminio CCaCalpL subvieneto ir R358′ ir K361′ iš antrinio CCaCalpL′ subvieneto išsidėsto netoli kerpamojo fosfodiesterinio ryšio, jungiančio ketvirtąjį (A4) ir pirmąjį (A1) adenozinus, todėl galėtų sudaryti žiedo nukleazės aktyvųjį centrą (22 pav. A). R358A ir H396A mutacijos turėjo didžiausią įtaką cA_4 hidrolizei: R358A stipriai susilpnino cA_4 hidrolizę, o H396 visiškai panaikino žiedo nukleazinį aktyvumą (22 pav. B). Įdomu, kad abu mutantai silpniau rišo tarpinius reakcijos produktus (A_4>p ir A_4p) – jie prasčiau stimuliavo proteazinį aktyvumą (22 pav. C). Tad H396 ir R358′ galėtų dalyvauti koordinuojant 5′-OH ir 2′,3′-ciklinį fosfatą / 3′-fosfatą A_4>p ir A_4p tarpinių reakcijos produktų galuose. Tad galime manyti, kad H396 ir R358′ yra katalitinės aminorūgštys atsakingos už du pirmuosius žiedo nukleazės reakcijos etapus: cA_4 hidrolizę iki A_4>p ir už A_4>p pavertimą į A_4p.

Tačiau fosfodiesterinio ryšio tarp A2 ir A3 perkirpimas paverčiant A_4p į galutinius reakcijos produktus A_2>p ir A_2p yra gerokai neaiškesnis. CCaCalpL struktūrose vyrauja viena A_4>p ir A_4p surišimo padėtis, kurioje molekulių galai orientuoti link katalitinių H396 ir R358′ liekanų. Už A2-A3 ryšio hidrolizę galėtų būti atsakingos H476′ ir K330′, tačiau jų mutantiniai variantai vykdė cA_4 hidrolizę iki galutinių diadeniltų (22 pav. B). Alternatyviai, tiek A1-A4, tiek A2-A3 fosfodiesteriniai ryšiai galėtų būti hidrolizuojami tame pačiame aktyviajame centre, jei A_4p gali būti surišamas dviem skirtingomis padėtimis. Tai paaiškintų lėtą A_4p hidrolizę ir H396A-CCaCalpL nesugebėjimą linearizuoti cA_4, tačiau šiuo metu negalime nei patvirtinti, nei paneigti antrojo aktyviojo centro egzistavimo.

Filamento svarba CCaCalpL aktyvumams

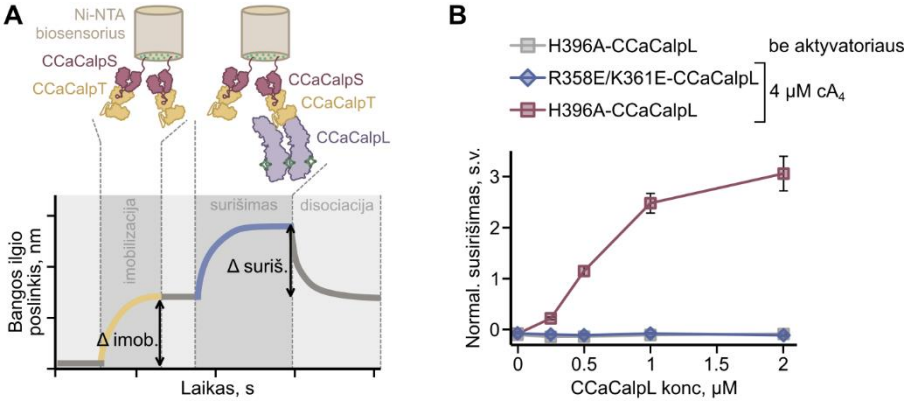
CCaCalpL filamente už žiedo nukleazinį aktyvumą yra atsakingos aminorūgštys H396 ir R358′ iš gretimų SAVED domenų (22 pav.), tad nenuostabu, kad filamento formuoti negebantis R358E/K361E-CCaCalpL variantas prastai karpė ir cA_4 (>100 kartų lėčiau nei wt-CCaCalpL) (23 pav.). Sumaišę R358E/K361E-CCaCalpL su žiedo nukleaziniu aktyvumu nepasižyminčiu H395A-CCaCalpL variantu, atkūrėme SAVED-SAVED sąveikos paviršių ir žiedo nukleazės aktyvųjį centrą - toks baltymų mišinys hidrolizavo cA_4 panašiai efektyviai kaip ir wt-CCaCalpL (23 pav. C). Vadinasi, CCaCalpL oligomerizacija yra svarbi žiedo nukleazės aktyvumui.



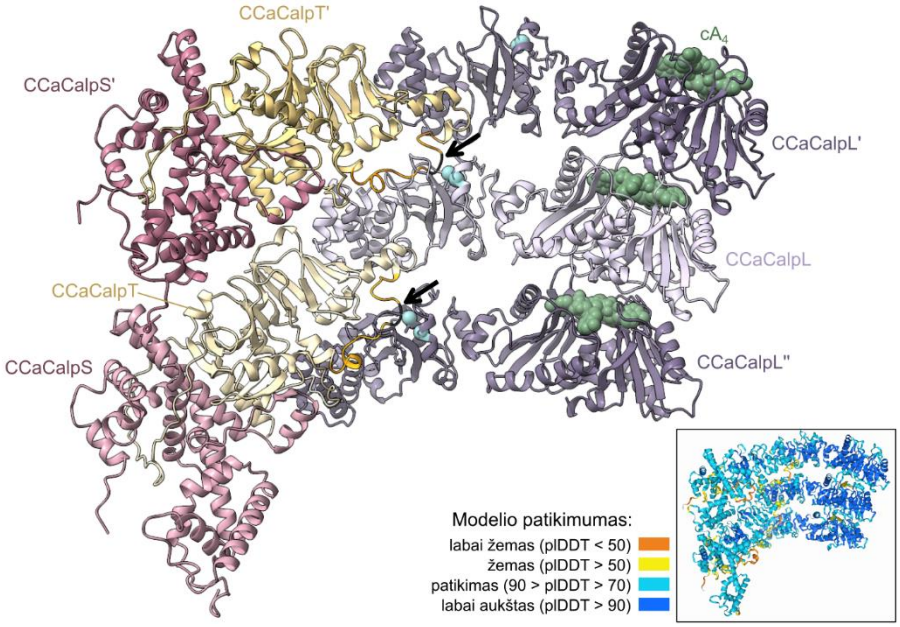
23 pav. CCaCalpL filamento svarbos žiedo nukleaziniam aktyvumui tyrimai. (**A**) R358E/K361E-CCaCalpL ir H396A-CCaCalpL sumažinto foninio triukšmo krio-EM mikrografijos, pavyzdžio mišinyje esant cA_4. (**B**) cA_4 karpymui reikalingo SAVED-SAVED sąveikos paviršiaus atkūrimo schema. Sumaišius žiedo nukleazės aktyvumo neturintį H396A-CCaCalpL variantą (2) su filamentų formuoti negebančiu R358E/K361E-CCaCalpL (1), turėtų būti formuojami dimerai ar R358E/K361E-CCaCalpL užsibaigiantys trumpi filamentukai, atkuriantys SAVED-SAVED sąveikos paviršių ir atstatantys žiedo nukleazės aktyvumą (3). (**C**) Žiedo nukleazinio aktyvumo atstatymo tyrimas. Reakcijų numeracija pagal schemą B dalyje. cA_4 hidrolizės reakcijos buvo atliekamos substrato nepritekliaus sąlygomis, naudojant 1 μM CCaCalpL ir 50 nM cA_4. Siekiant patogesnio palyginimo, cA_4 hidrolizės, naudojant H396A-CCaCalpL ir wt-CCaCalpL baltymus, gelių vaizdai iš 22 pav. B ir 20 pav. C panaudoti dar kartą.

Taip pat patikrinome, ar aktyvatoriaus prisijungimo sukelta CCaCalpL oligomerizacija yra svarbi proteazės taikinio surišimui. BLI metodu įvertinome ant biosensoriaus įmobilizuotos CCaCalpT-CalpS susirišimą su CCaCalpL (24 pav. A). Iš tirtų variantų, susirišimas su taikiniu buvo stebimas tik filamentus formuojančio H396A-CCaCalpL atveju, kai susirišimo

mišinyje buvo cA_4 (23 pav. A ir 24 pav. B). Tai rodo, kad CCaCalpL oligomerizacija yra būtina sąveikai su proteazės taikiniu CCaCalpT-CalpS.



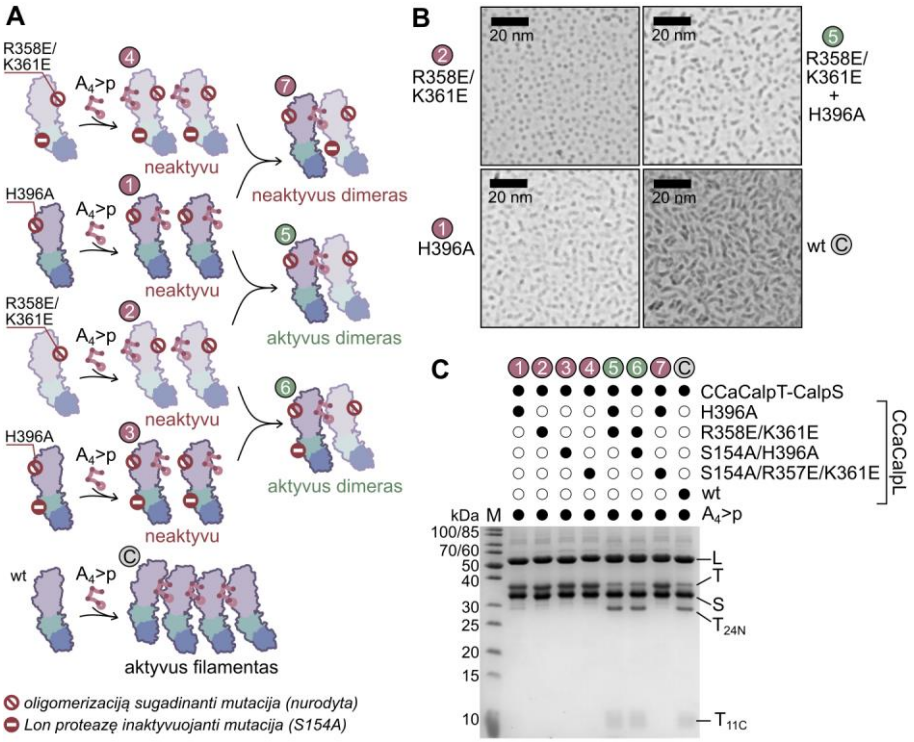
24 pav. CCaCalpL filamento svarbos proteazės taikinio surišimui tyrimai. (**A**) BLI eksperimento schema. CCaCalpT-CalpS imobilizuojamas ant Ni-NTA biosensoriaus (imobilizacijos žingsnis) ir stebima, ar prie jo prisiriša CCaCalpL, esant įvairioms pastarojo koncentracijoms (surišimo žingsnis). (**B**) Susirišimo priklausomybė nuo CCaCalpL koncentracijos. Normalizuotas susirišimas yra lygus bangos ilgio pokytis susirišimo žingsnyje (Δ suriš. / bangos ilgio poslinkio imobilizacijos žingsnyje (Δ imob.). Pateikiamos trijų eksperimentų vidutinės vertės su standartiniais nuokrypiais.



25 pav. CCaCalpL ir CCaCalpT sąveikos filamente modelis. Trimerinio CCaCalpL filamento (CCaCalpL′-CCaCalpL-CCaCalpL″), susirišusio su dviem CCaCalpT-CalpS heterodimerais ([CCaCalpT′-CalpS′]-[CCaCalpT-CalpS]),

AlphaFold3 modelis. AAAA RNR molekulės (žalios sferos) buvo sumodeliuotos kaip aktyvatoriaus atitikmuo. CCaCalpS pavaizduoti raudonos spalvos atspalviais, CCaCalpT – geltonos spalvos atspalviais, CCaCalpL – violetinės spalvos atspalviais.

Remiantis aktyvinto trijų CCaCalpL subvienetų ir dviejų CCaCalpT-CalpS komplekso kopijų filamento AplhaFold3 modeliu (25 pav.), CCaCalpT′ sudaro aukšto patikimumo sąveikos paviršių (960 Å²) su CCaCalpL′. CCaCalpT′ kilpa (194–210 aminorūgštys) įsiterpia į šalia esančio CCaCalpL Lon proteazės domeną ir CCaCalpT kirpimo vieta atsiduria šalia CCaCalpL Lon proteazės aktyviojo centro.

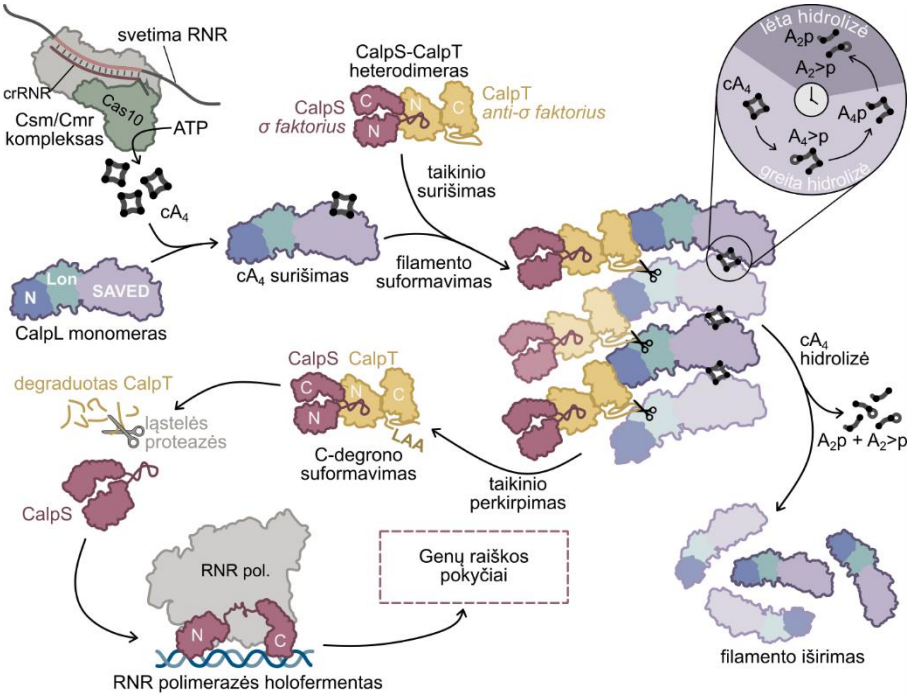


26 pav. CCaCalpL oligomerizacijos svarbos proteaziniam aktyvumui tyrimas. (**A**) A_4>p aktyvinamų CCaCalpL dimerų atkūrimo schema. (**B**) Wt, R358E/K361E, H396A CCaCalpL variantų bei R358E/K361E-CCaCalpL sumaišyto su H396A-CCaCalpL sumažinto triukšmo krio-EM mikrografijos, pavyzdžio mišinyje esant A_4>p. (**C**) Proteazės reakcijos. Numeravimas pagal A dalies schemą. Reakcijose buvo naudojama 5 μM CCaCalpL, 5 μM CCaCalpT-CalpS ir 25 μM A_4>p.

Remiantis modeliu, reikia mažiausiai dviejų CCaCalpL subvienetų, kad CCaCalpT būtų perkerpamas. Šį pastebėjimą patvirtinome iš dviejų filamentų neformuojančių ir dėl to neaktyvių CCaCalpL variantų (R358E/K361E-CCaCalpL ir H396A-CCaCalpL esant A_4>p) atkurdami aktyvų dimerą (26 pav.). Be to, panaudoję R358E/K360E ir H396A CCaCalpL variantus

turinčius papildomą S154A mutaciją Lon proteazės aktyviajame centre patvirtinome CCaCalpL ir CCaCalpT-CalpS išsidėstymą filamente, nuspėtą AlphaFold3 modelyje (25 pav.) – proteazės aktyvumą atkūrė H396A/S154A ir R358E/K361E mutantinių baltymų kombinacija (26 pav.). Tad galime teigti, kad CCaCalpL oligomerizacija yra būtina sąveikai su proteazės taikiniu, proteaziniam ir žiedo nukleazės aktyvumams.

Papildytas CalpL-CalpT-CalpS trinario efektoriaus veikimo mechanizmas



27 pav. Papildytas III tipo CRISPR-Cas apsaugos, veikiančios per trinarį CalpL-CalpT-CalpS efektorių, mechanizmas. Su užpuoliko RNR susirišęs Csm/Cmr interferencijos kompleksas sintetina cA_4, kurį specifiškai atpažįsta ir suriša CalpL SAVED domenas. Su cA_4 susirišęs CalpL oligomerizuojasi ir susiriša su CalpL proteazės taikiniu - CalpT-CalpS kompleksu. Suformuotame filamente CalpL perkerpa CalpT ir ant fragmento suformuojama C-degrono seka. Ši seka nukreipia CalpT fragmentą tolimesnei degradacijai, dėl ko suyra kompleksas ir išlaisvinamas CalpS, kuris toliau sąveikauja su RNR polimeraze ir keičia genų raišką.

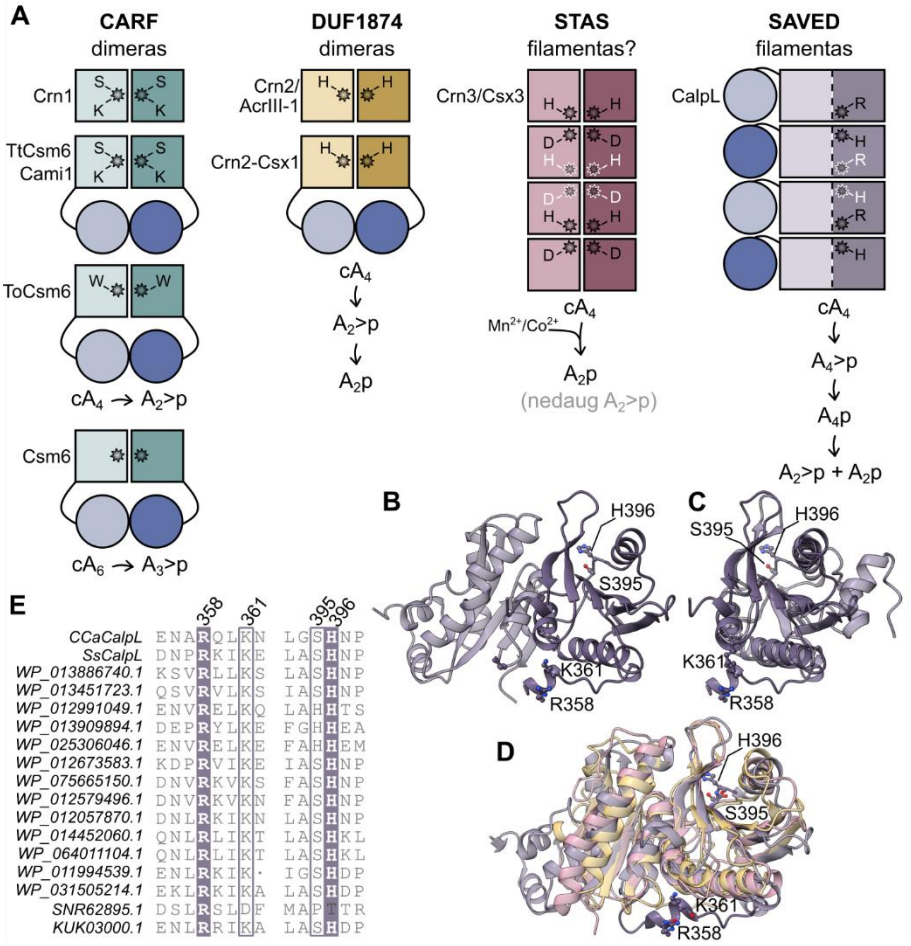
Palyginus su šio skyriaus pradžioje pateiktu CalpL-CalpT-CalpS veikimo mechanizmu (14 pav.), čia aprašytas išsamus trinario efektoriaus iš *Ca.* C. acidaminovorans charakterizavimas atskleidė trūkstamas mechanizmo detales. 27 paveikslas iliustruoja šio darbo rezultatais papildytą CalpL-CalpT-

CalpS trinario efektoriaus veikimo mechanizmą. III tipo CRISPR-Cas sistemos interferencijos kompleksui atpažinus ir surišus užpuoliko transkriptą, pradedamos gaminti cA_4 signalinės molekulės, kurios specifiškai rišasi su CalpL baltymo SAVED domenu. Šis susirišimas inicijuoja CalpL filamento formavimąsi ir CalpT-CalpS komplekso prisijungimą. Filamente vienas CalpL subvienetas suriša CalpT, o šalia esančio CalpL subvieneto Lon proteazės aktyvusis centras vykdo proteolizę. Perkirpto CalpT fragmentai lieka susirišę su CalpS, tačiau ant $CalpT_{24N}$ fragmento suformuotas C-degronas nukreipia jį tolimesniam degradavimui ląstelės proteazėmis, taip iš komplekso paleidžiant CalpS. Tuomet iš komplekso išlaisvintas CalpS gali susirišti su RNR polimeraze, suformuojant holofermentą, ir moduliuoti transkripciją. Be to, CalpL turi reguliacinį mechanizmą: CalpL SAVED domenas veikia kaip žiedo nukleazė, kuri per tris nuoseklius etapus hidrolizuoja cA_4. Ši aktyvatoriaus hidrolizė veikia kaip fermento „laikmatis“, ribojantis efektoriaus aktyvacijos trukmę.

Baigiamieji pastabėjimai: Išplėsta žiedo nukleazių įvairovė

Šioje disertacijoje aprašyti cA_6 ir cA_4 reguliacijos ir CalpL-CalpT-CalpS trinario efektoriaus charakterizavimo darbai leido papildyti žinomų žiedo nukleaziniu aktyvumu pasižyminčių baltymų grupę (28 pav. A).

Šiame darbe pristatyti StCsm6 žiedo nukleazinio aktyvumo tyrimai papildė žiedo nukleazių grupę cA_6 specifiškais CARF domenais (28 pav. A). Mūsų atlikti StCsm6 biocheminiai tyrimai bei kitų grupių atlikti StCsm6′ ir EiCsm6 struktūriniai tyrimai (Garcia-Doval et al., 2020; McQuarrie et al., 2023) nuosekliai pademonstravo, kad cA_6 specifinio Csm6 CARF domenas katalizuoja cA_6 hidrolizę priversdamas substratą užimti katalitiškai palankią konformaciją. Kadangi ne tik šoninės grandinės, bet ir pagrindinės aminorūgščių grandinės atomai, esantys vadinamojoje katalitinėje kilpoje, sąveikauja su cA_6, sunku tiksliai nustatyti katalitines aminorūgštis. Tačiau kilpoje esantis treoninas yra geras katalitinės amino rūgšties kandidatas, nes jo šoninės grandinės atomai sąveikauja su hidrolizuojamuoju fosfatu (Garcia-Doval et al., 2020; McQuarrie et al., 2023). Be to, mes pademonstravome, kad StCsm6 HEPN domeno nespecifinis RNR nukleazinis aktyvumas prisideda prie cA_6 ir kitų cA_n degradacijos esant aukštai signalinių molekulių koncentracijai, o tai dar labiau padidina cA_n hidrolizuoti gebančių baltymų įvairovę.



28 pav. Išplėsta žiedo nukleazių įvairovė. (**A**) CARF, DUF1874, STAS ir SAVED šeimų žiedo nukleazės. Žiedo nukleaziniai domenai pavaizduoti kvadratais, apskritimais pažymėti įvairūs efektoriniai domenai. Aktyvieji centrai pavaizduoti žvaigždutėmis su nurodytomis katalitinėmis amino rūgštimis. Punktyrinė linija žymi subdomenų ribas. Kiekvienos žiedo nukleazės tipo oligomerinė būsena įvardinta. (**B**) CCaCalpL SAVED domeno struktūrinis modelis, nuspalvintas pagal subdomenus: N-subdomenas (202–352 aminorūgštys) šviesiai violetinis, C-subdomenas (357–504 aminorūgštys) tamsiai violetinis. Nurodytos aktyviojo centro aminorūgštys. (**C**) CCaCalpL SAVED domeno N-subdomeno ir C-subdomeno superpozicija (spalvos kaip A dalyje). (**D**) CCaCalpL SAVED domeno (violetinis) palyginimas su CBASS susijusių efektorių SAVED domenais: Cap5 (rožinis, PDB ID: 7RWK) (Fatma et al., 2021) ir Cap4 (rusvas, PDB ID: 7YIB) (Lowey et al., 2020). Paryškinta α-spiralė, kurioje yra CCaCalpL žiedo nukleazės aktyviojo centro liekanos. Šios spiralės neturi nei Cap5 nei Cap4. CCaCalpL SAVED aktyviojo centro aminorūgštys nurodytos kaip A dalyje. (**E**) Konservatyvios žiedo nukleazės aktyviojo centro aminorūgštys Lon-SAVED4 baltymuose.

Be to, šioje disertacijoje aprašomi CCaCalpL-CalpT-CalpS trinario efektoriaus reguliacijos mechanizmo tyrimai cA_4 specifiškų žiedo nukleazių

gretas papildė ir SAVED domenu (28 pav. A). CCaCalpL kartu su kitais Lon-SAVED bei TM-SAVED efektoriais priklauso SAVED4 grupei, kurios atstovai dažnai randami šalia III tipo CRISPR-Cas sistemų (Makarova et al., 2020a). Kitaip nei CBASS sistemų efektoriuose aptinkami SAVED domenai, kurie specifiškai sąveikauja su asimetriniais (nelyginio nukleotidų skaičiaus) aktyvatoriais (Fatma et al., 2021; Hogrel et al., 2022; Lowey et al., 2020), CCaCalpL specifiškai sąveikauja su simetriniu (lyginio nukleotidų skaičiaus) aktyvatoriumi cA_4. CCaCalpL SAVED domeną sudarančių dviejų subdomenų palyginimas juos perklojant rodo, kad SAVED pseudodimeras yra simetriškas (28 pav. B ir C), tačiau pagrindinės katalitinės aminorūgštys R358 ir H396, kartu su mažiau išreikštomis K361 ir S395, yra išsidėsčiusios tik C-galiniame SAVED subdomene (28 pav. B). Struktūrų palyginimą atliko dr. Giedrė Tamulaitienė. Lon-SAVED4 baltymų sekų analizė parodė, kad beveik visi Lon-SAVED4 baltymai, įskaitant ir CCaCalpL struktūrinį homologą SsCalpL, kurio žiedo nukleazins aktyvumas jau buvo patvirtintas (Binder et al., 2024), turi konservatyvias žiedo nukleazės aktyviojo centro aminorūgštis (28 pav. E). Taigi, tikėtina, kad visi SAVED4 grupės baltymai pasižymi konservatyviu savireguliaciniu mechanizmu.

Apibendrinant, nuo 2018 m., kai pirmą kartą buvo identifikuotos cA_4 specifinės žiedo nukleazės (Athukoralage et al., 2018), supratimas apie cA_n reguliaciją III tipo CRISPR-Cas sistemose tapo žymiai gilesnis ir detalesnis, prie ko itin reikšmingai prisidėjo šis darbas.

IŠVADOS

1. StCsm6 ir StCsm6′ efektyviai sumažina cA_6 ir kitų oligoadenilatų (cA_n), kuriuos sintetina StCsm kompleksas, kiekį heterologiniame *E. coli* šeimininke.
2. StCsm6 CARF domenas yra cA_6 specifiška žiedo nukleazė, o HEPN domenas pasižymi nespecifiniu ribonukleaziniu aktyvumu, hidrolizuojančiu tiek cA_6, tiek kitus cA_n, tačiau pirmenybę teikiančiu linijinei viengrandinei RNR.
3. cA_4 prisijungimas inicijuoja CCaCalpL filamento susidarymą, kuris yra būtinas substrato CCaCalpT-CalpS prijungimui ir proteolizei. Filamente CCaCalpT proteolizę vykdo kaimyninio CCaCalpL subvieneto Lon proteazinis domenas.
4. CCaCalpL vykdoma CCaCalpT proteolizė suformuoja C-degroną, kuris nukreipia susidariusį fragmentą tolimesnei degradacijai ląstelės

proteazėmis. Šios degradacijos metu iš komplekso išlaisvinamas CCaCalpS , kuris vėliau jungiasi prie bakterinės RNR polimerazės.

5. CCaCalpL SAVED domenas veikia kaip cA_4 specifiška žiedo nukleazė, kuri hidrolizuoja savo aktyvatorių cA_4 trimis nuosekliais etapais, susidarant dviem tarpiniams junginiams – A_4>p ir A_4p.
6. CCaCalpL filamentų susidarymas, kurį inicijuoja aktyvatoriaus cA_4 arba jo hidrolizės tarpinių produktų prisijungimas, yra būtinas CCaCalpL efektoriaus Lon-SAVED proteazės ir žiedo nukleazės funkcijoms.

REFERENCES

1. Ablasser, A., Chen, Z.J., 2019. cGAS in action: Expanding roles in immunity and inflammation. Science 363, eaat8657. https://doi.org/10.1126/science.aat8657
2. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w
3. Ades, S.E., 2008. Regulation by destruction: design of the σE envelope stress response. Curr. Opin. Microbiol. 11, 535–540. https://doi.org/10.1016/j.mib.2008.10.004
4. Ahlawat, S., Morrison, D.A., 2009. ClpXP degrades SsrA-tagged proteins in Streptococcus pneumoniae. J. Bacteriol. 191, 2894–2898. https://doi.org/10.1128/jb.01715-08
5. Altae-Tran, H., Kannan, S., Suberski, A.J., Mears, K.S., Demircioglu, F.E., Moeller, L., Kocalar, S., Oshiro, R., Makarova, K.S., Macrae, R.K., Koonin, E.V., Zhang, F., 2023. Uncovering the functional diversity of rare CRISPR-Cas systems with deep terascale clustering. Science 382, eadi1910. https://doi.org/10.1126/science.adi1910
6. Athukoralage, J.S., Graham, S., Grüschow, S., Rouillon, C., White, M.F., 2019. A Type III CRISPR Ancillary Ribonuclease Degrades Its Cyclic Oligoadenylate Activator. J. Mol. Biol. 431, 2894–2899. https://doi.org/10.1016/j.jmb.2019.04.041
7. Athukoralage, J.S., Graham, S., Rouillon, C., Grüschow, S., Czekster, C.M., White, M.F., 2020a. The dynamic interplay of host and viral enzymes in type III CRISPR-mediated cyclic nucleotide signalling. eLife 9, e55852. https://doi.org/10.7554/eLife.55852
8. Athukoralage, J.S., McMahon, S.A., Zhang, C., Grüschow, S., Graham, S., Krupovic, M., Whitaker, R.J., Gloster, T.M., White, M.F., 2020b. An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. Nature 577, 572–575. https://doi.org/10.1038/s41586-019-1909-5
9. Athukoralage, J.S., McQuarrie, S., Grüschow, S., Graham, S., Gloster, T.M., White, M.F., 2020c. Tetramerisation of the CRISPR ring nuclease Crn3/Csx3 facilitates cyclic oligoadenylate cleavage. eLife 9, e57627. https://doi.org/10.7554/eLife.57627
10. Athukoralage, J.S., Rouillon, C., Graham, S., Grüschow, S., White, M.F., 2018. Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. Nature 562, 277–280. https://doi.org/10.1038/s41586-018-0557-5

11. Athukoralage, J.S., White, M.F., 2022. Cyclic Nucleotide Signaling in Phage Defense and Counter-Defense. Annu. Rev. Virol. https://doi.org/10.1146/annurev-virology-100120-010228
12. Athukoralage, J.S., White, M.F., 2021. Cyclic oligoadenylate signaling and regulation by ring nucleases during type III CRISPR defense. RNA 27, 855–867. https://doi.org/10.1261/rna.078739.121
13. Baca, C.F., Majumder, P., Hickling, J.H., Ye, L., Teplova, M., Brady, S.F., Patel, D.J., Marraffini, L.A., 2024a. The CRISPR-associated adenosine deaminase Cad1 converts ATP to ITP to provide antiviral immunity. Cell 187, 7183-7195.e24. https://doi.org/10.1016/j.cell.2024.10.002
14. Baca, C.F., Yu, Y., Rostøl, J.T., Majumder, P., Patel, D.J., Marraffini, L.A., 2024b. The CRISPR effector Cam1 mediates membrane depolarization for phage defence. Nature 625, 797–804. https://doi.org/10.1038/s41586-023-06902-y
15. Baker, T.A., Sauer, R.T., 2006. ATP-dependent proteases of bacteria: recognition logic and operating principles. Trends Biochem. Sci. 31, 647–653. https://doi.org/10.1016/j.tibs.2006.10.006
16. Banh, D.V., Roberts, C.G., Morales-Amador, A., Berryhill, B.A., Chaudhry, W., Levin, B.R., Brady, S.F., Marraffini, L.A., 2023. Bacterial cGAS senses a viral RNA to initiate immunity. Nature 623, 1001–1008. https://doi.org/10.1038/s41586-023-06743-9
17. Barkow, S.R., Levchenko, I., Baker, T.A., Sauer, R.T., 2009. Polypeptide Translocation by the AAA+ ClpXP Protease Machine. Chem. Biol. 16, 605–612. https://doi.org/10.1016/j.chembiol.2009.05.007
18. Battesti, A., Hoskins, J.R., Tong, S., Milanesio, P., Mann, J.M., Kravats, A., Tsegaye, Y.M., Bougdour, A., Wickner, S., Gottesman, S., 2013. Anti-adaptors provide multiple modes for regulation of the RssB adaptor protein. Genes Dev. 27, 2722–2735. https://doi.org/10.1101/gad.229617.113
19. Baytshtok, V., Fei, X., Shih, T.-T., Grant, R.A., Santos, J.C., Baker, T.A., Sauer, R.T., 2021. Heat activates the AAA+ HslUV protease by melting an axial autoinhibitory plug. Cell Rep. 34, 108639. https://doi.org/10.1016/j.celrep.2020.108639
20. Benda, C., Ebert, J., Scheltema, R.A., Schiller, H.B., Baumgärtner, M., Bonneau, F., Mann, M., Conti, E., 2014. Structural Model of a CRISPR RNA-Silencing Complex Reveals the RNA-Target Cleavage Activity in Cmr4. Mol. Cell 56, 43–54. https://doi.org/10.1016/j.molcel.2014.09.002
21. Bernal-Bernal, D., Abellón-Ruiz, J., Iniesta, A.A., Pajares-Martínez, E., Bastida-Martínez, E., Fontes, M., Padmanabhan, S., Elías-Arnanz, M., 2018. Multifactorial control of the expression of a CRISPR-Cas system by an extracytoplasmic function σ/anti-σ pair and a global regulatory complex. Nucleic Acids Res. 46, 6726–6745. https://doi.org/10.1093/nar/gky475
22. Bernal-Bernal, D., Pantoja-Uceda, D., López-Alonso, J.P., López-Rojo, A., López-Ruiz, J.A., Galbis-Martínez, M., Ochoa-Lizarralde, B., Tascón, I., Elías-Arnanz, M., Ubarretxena-Belandia, I., Padmanabhan, S., 2024. Structural basis for regulation of a CBASS-CRISPR-Cas defense island by a transmembrane anti-σ factor and its ECF σ partner. Sci. Adv. 10, eadp1053. https://doi.org/10.1126/sciadv.adp1053
23. Binder, S.C., Schneberger, N., Schmitz, M., Engeser, M., Geyer, M., Rouillon, C., Hagelueken, G., 2024. The SAVED domain of the type III

CRISPR protease CalpL is a ring nuclease. Nucleic Acids Res. 52, 10520–10532. https://doi.org/10.1093/nar/gkae676
24. Bittner, L.-M., Arends, J., Narberhaus, F., 2017. When, how and why? Regulated proteolysis by the essential FtsH protease in Escherichia coli. Biol. Chem., Biological Chemistry 398, 625–635. https://doi.org/10.1515/hsz-2016-0302
25. Bokori-Brown, M., Martin, T.G., Naylor, C.E., Basak, A.K., Titball, R.W., Savva, C.G., 2016. Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat. Commun. 7, 11293. https://doi.org/10.1038/ncomms11293
26. Brenzinger, S., Airoldi, M., Ogunleye, A.J., Jugovic, K., Amstalden, M.K., Brochado, A.R., 2024. The Vibrio cholerae CBASS phage defence system modulates resistance and killing by antifolate antibiotics. Nat. Microbiol. 9, 251–262. https://doi.org/10.1038/s41564-023-01556-y
27. Brown, S., Gauvin, C.C., Charbonneau, A.A., Burman, N., Lawrence, C.M., 2020. Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type III CRISPR–Cas that degrades the second messenger cA4. J. Biol. Chem. 295, 14963–14972. https://doi.org/10.1074/jbc.RA120.014099
28. Buckling, A., Rainey, P.B.., 2002. Antagonistic coevolution between a bacterium and a bacteriophage. Proc. R. Soc. Lond. B Biol. Sci. 269, 931–936. https://doi.org/10.1098/rspb.2001.1945
29. Burroughs, A.M., Zhang, D., Schäffer, D.E., Iyer, L.M., Aravind, L., 2015. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. Nucleic Acids Res. 43, 10633–10654. https://doi.org/10.1093/nar/gkv1267
30. Butala, M., Žgur-Bertok, D., Busby, S.J.W., 2008. The bacterial LexA transcriptional repressor. Cell. Mol. Life Sci. 66, 82. https://doi.org/10.1007/s00018-008-8378-6
31. Cao, Xueli, Xiao, Y., Huiting, E., Cao, Xujun, Li, D., Ren, J., Fedorova, I., Wang, H., Guan, L., Wang, Y., Li, L., Bondy-Denomy, J., Feng, Y., 2024. Phage anti-CBASS protein simultaneously sequesters cyclic trinucleotides and dinucleotides. Mol. Cell 84, 375-385.e7. https://doi.org/10.1016/j.molcel.2023.11.026
32. Carte, J., Christopher, R.T., Smith, J.T., Olson, S., Barrangou, R., Moineau, S., Glover III, C.V.C., Graveley, B.R., Terns, R.M., Terns, M.P., 2014. The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in treptococcus thermophilus. Mol. Microbiol. 93, 98–112. https://doi.org/10.1111/mmi.12644
33. Čepaitė, R., Klein, N., Mikšys, A., Camara-Wilpert, S., Ragožius, V., Benz, F., Skorupskaitė, A., Becker, H., Žvejytė, G., Steube, N., Hochberg, G.K.A., Randau, L., Pinilla-Redondo, R., Malinauskaitė, L., Pausch, P., 2024. Structural variation of types IV-A1- and IV-A3-mediated CRISPR interference. Nat. Commun. 15, 9306. https://doi.org/10.1038/s41467-024-53778-1
34. Charbonneau, A.A., Eckert, D.M., Gauvin, C.C., Lintner, N.G., Lawrence, C.M., 2021. Cyclic Tetra-Adenylate (cA4) Recognition by Csa3; Implications for an Integrated Class 1 CRISPR-Cas Immune Response in Saccharolobus solfataricus. Biomolecules 11. https://doi.org/10.3390/biom11121852

35. Charpentier, E., Richter, H., van der Oost, J., White, M.F., 2015. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol. Rev. 39, 428–441. https://doi.org/10.1093/femsre/fuv023
36. Chi, H., Hoikkala, V., Grüschow, S., Graham, S., Shirran, S., White, M.F., 2023. Antiviral type III CRISPR signalling via conjugation of ATP and SAM. Nature 622, 826–833. https://doi.org/10.1038/s41586-023-06620-5
37. Chi, H., White, M.F., 2024. RNA processing by the CRISPR-associated NYN ribonuclease. Biochem. J. 481, 793–804. https://doi.org/10.1042/BCJ20240151
38. Chou-Zheng, L., Hatoum-Aslan, A., 2022. Critical roles for 'housekeeping' nucleases in type III CRISPR-Cas immunity. eLife 11, e81897. https://doi.org/10.7554/eLife.81897
39. Chylinski, K., Makarova, K.S., Charpentier, E., Koonin, E.V., 2014. Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Res. 42, 6091–6105. https://doi.org/10.1093/nar/gku241
40. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., Sorek, R., 2019. Cyclic GMP–AMP signalling protects bacteria against viral infection. Nature 574, 691–695. https://doi.org/10.1038/s41586-019-1605-5
41. Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173, 33–38. https://doi.org/10.1016/0378-1119(95)00685-0
42. Cui, N., Zhang, J.-T., Li, Z., Liu, X.-Y., Wang, C., Huang, H., Jia, N., 2022. Structural basis for the non-self RNA-activated protease activity of the type III-E CRISPR nuclease-protease Craspase. Nat. Commun. 13, 7549. https://doi.org/10.1038/s41467-022-35275-5
43. Cui, N., Zhang, J.-T., Liu, Yongrui, Liu, Yanhong, Liu, X.-Y., Wang, C., Huang, H., Jia, N., 2023. Type IV-A CRISPR-Csf complex: Assembly, dsDNA targeting, and CasDinG recruitment. Mol. Cell 83, 2493-2508.e5. https://doi.org/10.1016/j.molcel.2023.05.036
44. Dartigalongue, C., Missiakas, D., Raina, S., 2001. Characterization of the Escherichia coliçE Regulon*. J. Biol. Chem. 276, 20866–20875. https://doi.org/10.1074/jbc.M100464200
45. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., Sorek, R., 2018. Systematic discovery of antiphage defense systems in the microbial pangenome. Science 359, eaar4120. https://doi.org/10.1126/science.aar4120
46. Du, L., Zhang, D., Luo, Z., Lin, Z., 2023. Molecular basis of stepwise cyclic tetra-adenylate cleavage by the type III CRISPR ring nuclease Crn1/Sso2081. Nucleic Acids Res. 51, 2485–2495. https://doi.org/10.1093/nar/gkad101
47. Du, L., Zhu, Q., Lin, Z., 2024. Molecular mechanism of allosteric activation of the CRISPR ribonuclease Csm6 by cyclic tetra-adenylate. EMBO J. 43, 304–315. https://doi.org/10.1038/s44318-023-00017-w
48. Duncan-Lowey, B., McNamara-Bordewick, N.K., Tal, N., Sorek, R., Kranzusch, P.J., 2021. Effector-mediated membrane disruption controls cell death in CBASS antiphage defense. Mol. Cell 81, 5039-5051.e5. https://doi.org/10.1016/j.molcel.2021.10.020

49. Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340
50. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Biol. Crystallogr. 66, 486–501. https://doi.org/10.1107/S0907444910007493
51. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J., Hassabis, D., 2022. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021.10.04.463034. https://doi.org/10.1101/2021.10.04.463034
52. Fang, C., Li, L., Shen, L., Shi, J., Wang, S., Feng, Y., Zhang, Y., 2019. Structures and mechanism of transcription initiation by bacterial ECF factors. Nucleic Acids Res. 47, 7094–7104. https://doi.org/10.1093/nar/gkz470
53. Fatma, S., Chakravarti, A., Zeng, X., Huang, R.H., 2021. Molecular mechanisms of the CdnG-Cap5 antiphage defense system employing 3′,2′-cGAMP as the second messenger. Nat. Commun. 12, 6381. https://doi.org/10.1038/s41467-021-26738-2
54. Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., Sauer, R.T., 2020. Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife 9, e61496. https://doi.org/10.7554/eLife.61496
55. Flynn, J.M., Levchenko, I., Sauer, R.T., Baker, T.A., 2004. Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. Genes Dev. 18, 2292–2301. https://doi.org/10.1101/gad.1240104
56. Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., Baker, T.A., 2001. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. 98, 10584–10589. https://doi.org/10.1073/pnas.191375298
57. Foster, K., Grüschow, S., Bailey, S., White, M.F., 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. https://doi.org/10.1093/nar/gkaa176
58. Foster, K., Kalter ,Joshua, Woodside ,Walter, Terns ,Rebecca 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 Biol. 16, 449–460. https://doi.org/10.1080/15476286.2018.1493334
59. Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. 3, 722–732. https://doi.org/10.1038/nrmicro1235
60. Fuchs, R.T., Curcuru, J.L., Mabuchi, M., Noireterre, A., Weigele, P.R., Sun, Z., Robb, G.B., 2022. Characterization of Cme and Yme thermostable Cas12a orthologs. Commun. Biol. 5, 325. https://doi.org/10.1038/s42003-022-03275-2
61. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., Zhang, F., 2020. Diverse

enzymatic activities mediate antiviral immunity in prokaryotes. Science 369, 1077–1084. https://doi.org/10.1126/science.aba0372
62. Gao, L.A., Wilkinson, M.E., Strecker, J., Makarova, K.S., Macrae, R.K., Koonin, E.V., Zhang, F., 2022. Prokaryotic innate immunity through pattern recognition of conserved viral proteins. Science 377, eabm4096. https://doi.org/10.1126/science.abm4096
63. Garcia-Doval, C., Schwede, F., Berk, C., Rostøl, J.T., Niewoehner, O., Tejero, O., Hall, J., Marraffini, L.A., Jinek, M., 2020. Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. Nat. Commun. 11, 1596. https://doi.org/10.1038/s41467-020-15334-5
64. Ghanbarpour, A., Cohen, S.E., Fei, X., Kinman, L.F., Bell, T.A., Zhang, J.J., Baker, T.A., Davis, J.H., Sauer, R.T., 2023. A closed translocation channel in the substrate-free AAA+ ClpXP protease diminishes rogue degradation. Nat. Commun. 14, 7281. https://doi.org/10.1038/s41467-023-43145-x
65. Gogarten, J.P., Townsend, J.P., 2005. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. 3, 679–687. https://doi.org/10.1038/nrmicro1204
66. Goswami, H.N., Ahmadizadeh, F., Wang, B., Addo-Yobo, D., Zhao, Y., Whittington, A.C., He, H., Terns, M.P., Li, H., 2024. Molecular basis for cA6 synthesis by a type III-A CRISPR–Cas enzyme and its conversion to cA4 production. Nucleic Acids Res. 52, 10619–10629. https://doi.org/10.1093/nar/gkae603
67. Grüschow, S., Adamson, C.S., White, M.F., 2021. Specificity and sensitivity of an RNA targeting type III CRISPR complex coupled with a NucC endonuclease effector. Nucleic Acids Res. 49, 13122–13134. https://doi.org/10.1093/nar/gkab1190
68. Grüschow, S., Athukoralage, J.S., Graham, S., Hoogeboom, T., White, M.F., 2019. Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. Nucleic Acids Res. 47, 9259–9270. https://doi.org/10.1093/nar/gkz676
69. Grüschow, S., McQuarrie, S., Ackermann, K., McMahon, S., Bode, B.E., Gloster, T.M., White, M.F., 2024. CRISPR antiphage defence mediated by the cyclic nucleotide-binding membrane protein Csx23. Nucleic Acids Res. 52, 2761–2775. https://doi.org/10.1093/nar/gkae167
70. Guo, M., Zhang, K., Zhu, Y., Pintilie, G.D., Guan, X., Li, S., Schmid, M.F., Ma, Z., Chiu, W., Huang, Z., 2019. Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry. Cell Res. 29, 305–312. https://doi.org/10.1038/s41422-019-0151-x
71. Gur, E., Biran, D., Ron, E.Z., 2011. Regulated proteolysis in Gram-negative bacteria — how and when? Nat. Rev. Microbiol. 9, 839–848. https://doi.org/10.1038/nrmicro2669
72. Gur, E., Sauer, R.T., 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. Proc. Natl. Acad. Sci. 106, 18503–18508. https://doi.org/10.1073/pnas.0910392106
73. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., Montoya, G., She, Q., 2018. A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative

substrate binding. Nucleic Acids Res. 46, 10319–10330. https://doi.org/10.1093/nar/gky844
74. Hatoum-Aslan Asma, Maniv Inbal, Samai Poulami, Marraffini Luciano A., 2014. Genetic Characterization of Antiplasmid Immunity through a Type III-A CRISPR-Cas System. J. Bacteriol. 196, 310–317. https://doi.org/10.1128/jb.01130-13
75. Heinrich, J., Wiegert, T., 2009. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. Spec. Issue Proteolysis Prokaryotes Protein Qual. Control Regul. Princ. 160, 696–703. https://doi.org/10.1016/j.resmic.2009.08.019
76. Herman, C., Thévenet, D., D'Ari, R., Bouloc, P., 1995. Degradation of sigma 32, the heat shock regulator in Escherichia coli, is governed by HflB. Proc. Natl. Acad. Sci. 92, 3516–3520. https://doi.org/10.1073/pnas.92.8.3516
77. Hobbs, S.J., Kranzusch, P.J., 2024. Nucleotide Immune Signaling in CBASS, Pycsar, Thoeris, and CRISPR Antiphage Defense. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev-micro-041222-024843
78. Hobbs, S.J., Wein, T., Lu, A., Morehouse, B.R., Schnabel, J., Leavitt, A., Yirmiya, E., Sorek, R., Kranzusch, P.J., 2022. Phage anti-CBASS and anti-Pycsar nucleases subvert bacterial immunity. Nature 605, 522–526. https://doi.org/10.1038/s41586-022-04716-y
79. Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E., Doudna, J.A., 2014. CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. Proc. Natl. Acad. Sci. 111, 6618–6623. https://doi.org/10.1073/pnas.1405079111
80. Hogrel, G., Guild, A., Graham, S., Rickman, H., Grüschow, S., Bertrand, Q., Spagnolo, L., White, M.F., 2022. Cyclic nucleotide-induced helical structure activates a TIR immune effector. Nature 608, 808–812. https://doi.org/10.1038/s41586-022-05070-9
81. Hoikkala, V., Graham, S., White, M.F., 2024. Bioinformatic analysis of type III CRISPR systems reveals key properties and new effector families. Nucleic Acids Res. 52, 7129–7141. https://doi.org/10.1093/nar/gkae462
82. Holm, L., Laiho, A., Törönen, P., Salgado, M., 2023. DALI shines a light on remote homologs: One hundred discoveries. Protein Sci. 32, e4519. https://doi.org/10.1002/pro.4519
83. Hu, C., van Beljouw, S.P.B., Nam, K.H., Schuler, G., Ding, F., Cui, Y., Rodríguez-Molina, A., Haagsma, A.C., Valk, M., Pabst, M., Brouns, S.J.J., Ke, A., 2022. Craspase is a CRISPR RNA-guided, RNA-activated protease. Science 377, 1278–1285. https://doi.org/10.1126/science.add5064
84. Huiting, E., Cao, X., Ren, J., Athukoralage, J.S., Luo, Z., Silas, S., An, N., Carion, H., Zhou, Y., Fraser, J.S., Feng, Y., Bondy-Denomy, J., 2023. Bacteriophages inhibit and evade cGAS-like immune function in bacteria. Cell 186, 864-876.e21. https://doi.org/10.1016/j.cell.2022.12.041
85. Huo, Y., Li, T., Wang, N., Dong, Q., Wang, X., Jiang, T., 2018. Cryo-EM structure of Type III-A CRISPR effector complex. Cell Res. 28, 1195–1197. https://doi.org/10.1038/s41422-018-0115-6
86. Irmisch, P., Mogila, I., Samatanga, B., Tamulaitis, G., Seidel, R., 2024. Retention of the RNA ends provides the molecular memory for maintaining

the activation of the Csm complex. Nucleic Acids Res. 52, 3896–3910. https://doi.org/10.1093/nar/gkae080
87. Jenal, U., Fuchs, T., 1998. An essential protease involved in bacterial cell-cycle control. EMBO J. 17, 5658–5669. https://doi.org/10.1093/emboj/17.19.5658
88. Jenson, J.M., Li, T., Du, F., Ea, C.-K., Chen, Z.J., 2023. Ubiquitin-like conjugation by bacterial cGAS enhances anti-phage defence. Nature 616, 326–331. https://doi.org/10.1038/s41586-023-05862-7
89. Jia, N., Jones, R., Sukenick, G., Patel, D.J., 2019a. Second Messenger cA4 Formation within the Composite Csm1 Palm Pocket of Type III-A CRISPR-Cas Csm Complex and Its Release Path. Mol. Cell 75, 933-943.e6. https://doi.org/10.1016/j.molcel.2019.06.013
90. Jia, N., Jones, R., Yang, G., Ouerfelli, O., Patel, D.J., 2019b. CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. Mol. Cell 75, 944-956.e6. https://doi.org/10.1016/j.molcel.2019.06.014
91. Jia, N., Mo, C.Y., Wang, C., Eng, E.T., Marraffini, L.A., Patel, D.J., 2019c. Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. Mol. Cell 73, 264-277.e5. https://doi.org/10.1016/j.molcel.2018.11.007
92. Johnson, A.G., Mayer, M.L., Schaefer, S.L., McNamara-Bordewick, N.K., Hummer, G., Kranzusch, P.J., 2024. Structure and assembly of a bacterial gasdermin pore. Nature 628, 657–663. https://doi.org/10.1038/s41586-024-07216-3
93. Johnson, A.G., Wein, T., Mayer, M.L., Duncan-Lowey, B., Yirmiya, E., Oppenheimer-Shaanan, Y., Amitai, G., Sorek, R., Kranzusch, P.J., 2022. Bacterial gasdermins reveal an ancient mechanism of cell death. Science 375, 221–225. https://doi.org/10.1126/science.abj8432
94. Johnson, K.A., Garrett, S.C., Noble-Molnar, C., Elgarhi, H.A., Woodside, W.T., Cooper, C., Zhang, X., Olson, S., Catchpole, R.J., Graveley, B.R., Terns, M.P., 2024. Selective degradation of phage RNAs by the Csm6 ribonuclease provides robust type III CRISPR immunity in Streptococcus thermophilus. Nucleic Acids Res. 52, 12549–12564. https://doi.org/10.1093/nar/gkae856
95. Jonas, K., Liu, J., Chien, P., Laub, M.T., 2013. Proteotoxic Stress Induces a Cell-Cycle Arrest by Stimulating Lon to Degrade the Replication Initiator DnaA. Cell 154, 623–636. https://doi.org/10.1016/j.cell.2013.06.034
96. Julien, O., Wells, J.A., 2017. Caspases and their substrates. Cell Death Differ. 24, 1380–1389. https://doi.org/10.1038/cdd.2017.44
97. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2
98. Jungfer, K., Moravčík, Š., Garcia-Doval, C., Knörlein, A., Hall, J., Jinek, M., 2025. Mechanistic determinants and dynamics of cA6 synthesis in type

III CRISPR-Cas effector complexes. Nucleic Acids Res. 53, gkae1277. https://doi.org/10.1093/nar/gkae1277
99. Kato, K., Okazaki, S., Schmitt-Ulms, C., Jiang, K., Zhou, W., Ishikawa, J., Isayama, Y., Adachi, S., Nishizawa, T., Makarova, K.S., Koonin, E.V., Abudayyeh, O.O., Gootenberg, J.S., Nishimasu, H., 2022a. RNA-triggered protein cleavage and cell growth arrest by the type III-E CRISPR nuclease-protease. Science 378, 882–889. https://doi.org/10.1126/science.add7347
100. Kato, K., Zhou, W., Okazaki, S., Isayama, Y., Nishizawa, T., Gootenberg, J.S., Abudayyeh, O.O., Nishimasu, H., 2022b. Structure and engineering of the type III-E CRISPR-Cas7-11 effector complex. Cell 185, 2324-2337.e16. https://doi.org/10.1016/j.cell.2022.05.003
101. Kaur, G., Burroughs, A.M., Iyer, L.M., Aravind, L., 2020. Highly regulated, diversifying NTP-dependent biological conflict systems with implications for the emergence of multicellularity. eLife 9, e52696. https://doi.org/10.7554/eLife.52696
102. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., Siksnys, V., 2017. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science 357, 605–609. https://doi.org/10.1126/science.aao0100
103. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č., Siksnys, V., 2016. Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. Mol. Cell 62, 295–306. https://doi.org/10.1016/j.molcel.2016.03.024
104. Keiler, K.C., Waller, P.R.H., Sauer, R.T., 1996. Role of a Peptide Tagging System in Degradation of Proteins Synthesized from Damaged Messenger RNA. Science 271, 990–993. https://doi.org/10.1126/science.271.5251.990
105. Kelleher, P., Ortiz Charneco, G., Kampff, Z., Diaz-Garrido, N., Bottacini, F., McDonnell, B., Lugli, G.A., Ventura, M., Fomenkov, A., Quénée, P., Kulakauskas, S., de Waal, P., van Peij, N.N.M.E., Cambillau, C., Roberts, R.J., van Sinderen, D., Mahony, J., 2024. Phage defence loci of Streptococcus thermophilus—tip of the anti-phage iceberg? Nucleic Acids Res. 52, 11853–11869. https://doi.org/10.1093/nar/gkae814
106. Kimura, S., Suzuki, T., Chen, M., Kato, K., Yu, J., Nakamura, A., Tanaka, I., Yao, M., 2016. Template-dependent nucleotide addition in the reverse (3′-5′) direction by Thg1-like protein. Sci. Adv. 2, e1501397. https://doi.org/10.1126/sciadv.1501397
107. Klimecka, M.M., Antosiewicz, A., Izert, M.A., Szybowska, P.E., Twardowski, P.K., Delaunay, C., Górna, M.W., 2021. A Uniform Benchmark for Testing SsrA-Derived Degrons in the Escherichia coli ClpXP Degradation Pathway. Molecules 26. https://doi.org/10.3390/molecules26195936
108. Ko, T.-P., Wang, Y.-C., Yang, C.-S., Hou, M.-H., Chen, C.-J., Chiu, Y.-F., Chen, Y., 2022. Crystal structure and functional implication of bacterial STING. Nat. Commun. 13, 26. https://doi.org/10.1038/s41467-021-26583-3
109. Koonin, E.V., Makarova, K.S., 2022. Evolutionary plasticity and functional versatility of CRISPR systems. PLOS Biol. 20, e3001481. https://doi.org/10.1371/journal.pbio.3001481
110. Krüger, L., Gaskell-Mew, L., Graham, S., Shirran, S., Hertel, R., White, M.F., 2024. Reversible conjugation of a CBASS nucleotide cyclase

regulates bacterial immune response to phage infection. Nat. Microbiol. 9, 1579–1592. https://doi.org/10.1038/s41564-024-01670-5
111. Kuroda, A., Nomura, K., Ohtomo, R., Kato, J., Ikeda, T., Takiguchi, N., Ohtake, H., Kornberg, A., 2001. Role of Inorganic Polyphosphate in Promoting Ribosomal Protein Degradation by the Lon Protease in E. coli. Science 293, 705–708. https://doi.org/10.1126/science.1061315
112. Langklotz, S., Baumann, U., Narberhaus, F., 2012. Structure and function of the bacterial AAA protease FtsH. Biochim. Biophys. Acta BBA - Mol. Cell Res. 1823, 40–48. https://doi.org/10.1016/j.bbamcr.2011.08.015
113. 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., Mekalanos, J.J., Kranzusch, P.J., Jain, M., Pogliano, J., Corbett, K.D., 2020. Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity. Mol. Cell 77, 723-733.e6. https://doi.org/10.1016/j.molcel.2019.12.010
114. Leavitt, A., Yirmiya, E., Amitai, G., Lu, A., Garb, J., Herbst, E., Morehouse, B.R., Hobbs, S.J., Antine, S.P., Sun, Z.-Y.J., Kranzusch, P.J., Sorek, R., 2022. Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. Nature 611, 326–331. https://doi.org/10.1038/s41586-022-05375-9
115. Ledvina, H.E., Ye, Q., Gu, Y., Sullivan, A.E., Quan, Y., Lau, R.K., Zhou, H., Corbett, K.D., Whiteley, A.T., 2023. An E1–E2 fusion protein primes antiviral immune signalling in bacteria. Nature 616, 319–325. https://doi.org/10.1038/s41586-022-05647-4
116. Lee, H., Sashital, D.G., 2022. Creating memories: molecular mechanisms of CRISPR adaptation. Trends Biochem. Sci. 47, 464–476. https://doi.org/10.1016/j.tibs.2022.02.004
117. Levchenko, I., Grant, R.A., Flynn, J.M., Sauer, R.T., Baker, T.A., 2005. Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. Nat. Struct. Mol. Biol. 12, 520–525. https://doi.org/10.1038/nsmb934
118. Li, X., Han, J., Yang, J., Zhang, H., 2024. The structural biology of type III CRISPR-Cas systems. J. Struct. Biol. 216, 108070. https://doi.org/10.1016/j.jsb.2024.108070
119. Li, Y., Li, Z., Yan, P., Hua, C., Kong, J., Wu, W., Cui, Y., Duan, Y., Li, S., Li, G., Ji, S., Chen, Y., Zhao, Y., Yang, P., Hu, C., Lu, M., Chen, M., Xiao, Y., 2025. Antiviral signaling of a type III CRISPR-associated deaminase. Science 387, eadr0393. https://doi.org/10.1126/science.adr0393
120. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Biol. Crystallogr. 75, 861–877. https://doi.org/10.1107/S2059798319011471
121. Liu, X., Zhang, L., Wang, H., Xiu, Y., Huang, L., Gao, Z., Li, N., Li, F., Xiong, W., Gao, T., Zhang, Y., Yang, M., Feng, Y., 2022. Target RNA activates the protease activity of Craspase to confer antiviral defense. Mol. Cell 82, 4503-4518.e8. https://doi.org/10.1016/j.molcel.2022.10.007
122. Lowey, B., Whiteley, A.T., Keszei, A.F.A., Morehouse, B.R., Mathews, I.T., Antine, S.P., Cabrera, V.J., Kashin, D., Niemann, P., Jain, M., Schwede, F., Mekalanos, J.J., Shao, S., Lee, A.S.Y., Kranzusch, P.J., 2020. CBASS Immunity Uses CARF-Related Effectors to Sense 3′–5′- and 2′–5′-

Linked Cyclic Oligonucleotide Signals and Protect Bacteria from Phage Infection. Cell 182, 38-49.e17. https://doi.org/10.1016/j.cell.2020.05.019
123. Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A.R.N., Lopez, R., Butcher, S., 2024. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. Nucleic Acids Res. 52, W521–W525. https://doi.org/10.1093/nar/gkae241
124. Mahmoud, S.A., Chien, P., 2018. Regulated Proteolysis in Bacteria. Annu. Rev. Biochem. https://doi.org/10.1146/annurev-biochem-062917-012848
125. Makarova, K.S., Timinskas, A., Wolf, Y.I., Gussow, A.B., Siksnys, V., Venclovas, Č., Koonin, E.V., 2020a. Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antivirus defense. Nucleic Acids Res. 48, 8828–8847. https://doi.org/10.1093/nar/gkaa635
126. Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Scott, D., Shah, S.A., Siksnys, V., Terns, M.P., Venclovas, Č., White, M.F., Yakunin, A.F., Yan, W., Zhang, F., Garrett, R.A., Backofen, R., van der Oost, J., Barrangou, R., Koonin, E.V., 2020b. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. 18, 67–83. https://doi.org/10.1038/s41579-019-0299-x
127. Manik, M.K., Shi, Y., Li, S., Zaydman, M.A., Damaraju, N., Eastman, S., Smith, T.G., Gu, W., Masic, V., Mosaiab, T., Weagley, J.S., Hancock, S.J., Vasquez, E., Hartley-Tassell, L., Kargios, N., Maruta, N., Lim, B.Y.J., Burdett, H., Landsberg, M.J., Schembri, M.A., Prokes, I., Song, L., Grant, M., DiAntonio, A., Nanson, J.D., Guo, M., Milbrandt, J., Ve, T., Kobe, B., 2022. Cyclic ADP ribose isomers: Production, chemical structures, and immune signaling. Science 377, eadc8969. https://doi.org/10.1126/science.adc8969
128. Mascher, T., 2023. Past, Present, and Future of Extracytoplasmic Function σ Factors: Distribution and Regulatory Diversity of the Third Pillar of Bacterial Signal Transduction. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev-micro-032221-024032
129. Mayo-Muñoz, D., Pinilla-Redondo, R., Birkholz, N., Fineran, P.C., 2023. A host of armor: Prokaryotic immune strategies against mobile genetic elements. Cell Rep. 42, 112672. https://doi.org/10.1016/j.celrep.2023.112672
130. Mayo-Muñoz, D., Smith, L.M., Garcia-Doval, C., Malone, L.M., Harding, K.R., Jackson, S.A., Hampton, H.G., Fagerlund, R.D., Gumy, L.F., Fineran, P.C., 2022. Type III CRISPR-Cas provides resistance against nucleus-forming jumbo phages via abortive infection. Mol. Cell 82, 4471-4486.e9. https://doi.org/10.1016/j.molcel.2022.10.028
131. McMahon, S.A., Zhu, W., Graham, S., Rambo, R., White, M.F., Gloster, T.M., 2020. Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. Nat. Commun. 11, 500. https://doi.org/10.1038/s41467-019-14222-x
132. McQuarrie, S., Athukoralage, J.S., McMahon, S.A., Graham, S., Ackermann, K., Bode, B.E., White, M.F., Gloster, T.M., 2023. Activation of Csm6 ribonuclease by cyclic nucleotide binding: in an emergency, twist

to open. Nucleic Acids Res. 51, 10590–10605. https://doi.org/10.1093/nar/gkad739
133. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., Ferrin, T.E., 2023. UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. 32, e4792. https://doi.org/10.1002/pro.4792
134. Micevski, D., Zammit, J.E., Truscott, K.N., Dougan, D.A., 2015. Anti-adaptors use distinct modes of binding to inhibit the RssB-dependent turnover of RpoS (σS) by ClpXP. Front. Mol. Biosci. 2.
135. Millman, A., Melamed, S., Amitai, G., Sorek, R., 2020. Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. Nat. Microbiol. 5, 1608–1615. https://doi.org/10.1038/s41564-020-0777-y
136. Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., Garb, J., Bechon, N., Brandis, A., Lopatina, A., Ofir, G., Hochhauser, D., Stokar-Avihail, A., Tal, N., Sharir, S., Voichek, M., Erez, Z., Ferrer, J.L.M., Dar, D., Kacen, A., Amitai, G., Sorek, R., 2022. An expanded arsenal of immune systems that protect bacteria from phages. Cell Host Microbe 30, 1556-1569.e5. https://doi.org/10.1016/j.chom.2022.09.017
137. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M., 2022. ColabFold: making protein folding accessible to all. Nat. Methods 19, 679–682. https://doi.org/10.1038/s41592-022-01488-1
138. Mogila, I., Kazlauskiene, M., Valinskyte, S., Tamulaitiene, G., Tamulaitis, G., Siksnys, V., 2019. Genetic Dissection of the Type III-A CRISPR-Cas System Csm Complex Reveals Roles of Individual Subunits. Cell Rep. 26, 2753-2765.e4. https://doi.org/10.1016/j.celrep.2019.02.029
139. Mogila, I., Tamulaitiene, G., Keda, K., Timinskas, A., Ruksenaite, A., Sasnauskas, G., Venclovas, Č., Siksnys, V., Tamulaitis, G., 2023. Ribosomal stalk-captured CARF-RelE ribonuclease inhibits translation following CRISPR signaling. Science 382, 1036–1041. https://doi.org/10.1126/science.adj2107
140. Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., Soria, E., 2005. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J. Mol. Evol. 60, 174–182. https://doi.org/10.1007/s00239-004-0046-3
141. Molina, R., Garcia-Martin, R., López-Méndez, B., Jensen, A.L.G., Ciges-Tomas, J.R., Marchena-Hurtado, J., Stella, S., Montoya, G., 2022. Molecular basis of cyclic tetra-oligoadenylate processing by small standalone CRISPR-Cas ring nucleases. Nucleic Acids Res. 50, 11199–11213. https://doi.org/10.1093/nar/gkac923
142. Molina, R., Jensen, A.L.G., Marchena-Hurtado, J., López-Méndez, B., Stella, S., Montoya, G., 2021. Structural basis of cyclic oligoadenylate degradation by ancillary Type III CRISPR-Cas ring nucleases. Nucleic Acids Res. 49, 12577–12590. https://doi.org/10.1093/nar/gkab1130
143. Molina, R., Sofos, N., Montoya, G., 2020. Structural basis of CRISPR-Cas Type III prokaryotic defence systems. Curr. Opin. Struct. Biol. 65, 119–129. https://doi.org/10.1016/j.sbi.2020.06.010
144. Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., López-Méndez, B., She, Q., Montoya, G., 2019. Structure of Csx1-cOA4

complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. Nat. Commun. 10, 4302. https://doi.org/10.1038/s41467-019-12244-z
145. Morehouse, B.R., Govande, A.A., Millman, A., Keszei, A.F.A., Lowey, B., Ofir, G., Shao, S., Sorek, R., Kranzusch, P.J., 2020. STING cyclic dinucleotide sensing originated in bacteria. Nature 586, 429–433. https://doi.org/10.1038/s41586-020-2719-5
146. Morehouse, B.R., Yip, M.C.J., Keszei, A.F.A., McNamara-Bordewick, N.K., Shao, S., Kranzusch, P.J., 2022. Cryo-EM structure of an active bacterial TIR–STING filament complex. Nature 608, 803–807. https://doi.org/10.1038/s41586-022-04999-1
147. Murtazalieva, K., Mu, A., Petrovskaya, A., Finn, R.D., 2024. The growing repertoire of phage anti-defence systems. Trends Microbiol. 32, 1212–1228. https://doi.org/10.1016/j.tim.2024.05.005
148. Nasef, M., Muffly, M.C., Beckman, A.B., Rowe, S.J., Walker, F.C., Hatoum-Aslan, A., Dunkle, J.A., 2019. Regulation of cyclic oligoadenylate synthesis by the Staphylococcus epidermidis Cas10-Csm complex. RNA 25, 948–962. https://doi.org/10.1261/rna.070417.119
149. Niewoehner, O., Garcia-Doval, C., Rostøl, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., Jinek, M., 2017. Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature 548, 543–548. https://doi.org/10.1038/nature23467
150. Nussenzweig, P.M., Marraffini, L.A., 2020. Molecular Mechanisms of CRISPR-Cas Immunity in Bacteria. Annu. Rev. Genet. https://doi.org/10.1146/annurev-genet-022120-112523
151. Ofir, G., Herbst, E., Baroz, M., Cohen, D., Millman, A., Doron, S., Tal, N., Malheiro, D.B.A., Malitsky, S., Amitai, G., Sorek, R., 2021. Antiviral activity of bacterial TIR domains via immune signalling molecules. Nature 600, 116–120. https://doi.org/10.1038/s41586-021-04098-7
152. Osawa, T., Inanaga, H., Sato, C., Numata, T., 2015. Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. Mol. Cell 58, 418–430. https://doi.org/10.1016/j.molcel.2015.03.018
153. Österberg, S., Peso-Santos, T. del, Shingler, V., 2011. Regulation of Alternative Sigma Factor Use. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev.micro.112408.134219
154. Özcan, A., Krajeski, R., Ioannidi, E., Lee, B., Gardner, A., Makarova, K.S., Koonin, E.V., Abudayyeh, O.O., Gootenberg, J.S., 2021. Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. Nature 597, 720–725. https://doi.org/10.1038/s41586-021-03886-5
155. Patel, D.J., Yu, Y., Jia, N., 2022. Bacterial origins of cyclic nucleotide-activated antiviral immune signaling. Mol. Cell 82, 4591–4610. https://doi.org/10.1016/j.molcel.2022.11.006
156. Philipps-Wiemann, P., 2018. Chapter 12 - Proteases—general aspects, in: Nunes, C.S., Kumar, V. (Eds.), Enzymes in Human and Animal Nutrition. Academic Press, pp. 257–266. https://doi.org/10.1016/B978-0-12-805419-2.00012-5
157. Pinilla-Redondo, R., Russel, J., Mayo-Muñoz, D., Shah, S.A., Garrett, R.A., Nesme, J., Madsen, J.S., Fineran, P.C., Sørensen, S.J., 2022. CRISPR-Cas systems are widespread accessory elements across bacterial and archaeal plasmids. Nucleic Acids Res. 50, 4315–4328. https://doi.org/10.1093/nar/gkab859

158. Pruteanu, M., Baker, T.A., 2009. Proteolysis in the SOS response and metal homeostasis in Escherichia coli. Spec. Issue Proteolysis Prokaryotes Protein Qual. Control Regul. Princ. 160, 677–683. https://doi.org/10.1016/j.resmic.2009.08.012
159. Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296. https://doi.org/10.1038/nmeth.4169
160. Punjani, A., Zhang, H., Fleet, D.J., 2020. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat. Methods 17, 1214–1221. https://doi.org/10.1038/s41592-020-00990-8
161. Rankin, D.J., Rocha, E.P.C., Brown, S.P., 2011. What traits are carried on mobile genetic elements, and why? Heredity 106, 1–10. https://doi.org/10.1038/hdy.2010.24
162. Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., Finn, R.D., 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 46, D624–D632. https://doi.org/10.1093/nar/gkx1134
163. Rechkoblit, O., Sciaky, D., Kreitler, D.F., Buku, A., Kottur, J., Aggarwal, A.K., 2024. Activation of CBASS Cap5 endonuclease immune effector by cyclic nucleotides. Nat. Struct. Mol. Biol. 31, 767–776. https://doi.org/10.1038/s41594-024-01220-x
164. Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324. https://doi.org/10.1093/nar/gku316
165. Rosenthal, P.B., Henderson, R., 2003. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. J. Mol. Biol. 333, 721–745. https://doi.org/10.1016/j.jmb.2003.07.013
166. Rostøl, J.T., Marraffini, L.A., 2019. Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR–Cas immunity. Nat. Microbiol. 4, 656–662. https://doi.org/10.1038/s41564-018-0353-x
167. Rostøl, J.T., Xie, W., Kuryavyi, V., Maguin, P., Kao, K., Froom, R., Patel, D.J., Marraffini, L.A., 2021. The Card1 nuclease provides defence during type III CRISPR immunity. Nature 590, 624–629. https://doi.org/10.1038/s41586-021-03206-x
168. Rouillon, C., Athukoralage, J.S., Graham, S., Grüschow, S., White, M.F., 2018. Control of cyclic oligoadenylate synthesis in a type III CRISPR system. eLife 7, e36734. https://doi.org/10.7554/eLife.36734
169. Rouillon, C., Schneberger, N., Chi, H., Blumenstock, K., Da Vela, S., Ackermann, K., Moecking, J., Peter, M.F., Boenigk, W., Seifert, R., Bode, B.E., Schmid-Burgk, J.L., Svergun, D., Geyer, M., White, M.F., Hagelueken, G., 2023. Antiviral signalling by a cyclic nucleotide activated CRISPR protease. Nature 614, 168–174. https://doi.org/10.1038/s41586-022-05571-7
170. Rousset, F., Osterman, I., Scherf, T., Falkovich, A.H., Leavitt, A., Amitai, G., Shir, S., Malitsky, S., Itkin, M., Savidor, A., Sorek, R., 2025. TIR

signaling activates caspase-like immunity in bacteria. Science 387, 510–516. https://doi.org/10.1126/science.adu2262
171. Rousset, F., Yirmiya, E., Nesher, S., Brandis, A., Mehlman, T., Itkin, M., Malitsky, S., Millman, A., Melamed, S., Sorek, R., 2023. A conserved family of immune effectors cleaves cellular ATP upon viral infection. Cell 186, 3619-3631.e13. https://doi.org/10.1016/j.cell.2023.07.020
172. Sabonis, D., Avraham, C., Lu, A., Herbst, E., Silanskas, A., Leavitt, A., Yirmiya, E., Zaremba, M., Amitai, G., Kranzusch, P.J., Sorek, R., Tamulaitiene, G., 2024. TIR domains produce histidine-ADPR conjugates as immune signaling molecules in bacteria. bioRxiv 2024.01.03.573942. https://doi.org/10.1101/2024.01.03.573942
173. Sajed, T., Marcu, A., Ramirez, M., Pon, A., Guo, A.C., Knox, C., Wilson, M., Grant, J.R., Djoumbou, Y., Wishart, D.S., 2016. ECMDB 2.0: A richer resource for understanding the biochemistry of E. coli. Nucleic Acids Res. 44, D495–D501. https://doi.org/10.1093/nar/gkv1060
174. Samolygo, A., Athukoralage, J.S., Graham, S., 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. https://doi.org/10.1093/nar/gkaa298
175. Schwartz, E.A., Bravo, J.P.K., Ahsan, M., Macias, L.A., McCafferty, C.L., Dangerfield, T.L., Walker, J.N., Brodbelt, J.S., Palermo, G., Fineran, P.C., Fagerlund, R.D., Taylor, D.W., 2024. RNA targeting and cleavage by the type III-Dv CRISPR effector complex. Nat. Commun. 15, 3324. https://doi.org/10.1038/s41467-024-47506-y
176. Severin, G.B., Ramliden, M.S., Ford, K.C., Van Alst, A.J., Sanath-Kumar, R., Decker, K.A., Hsueh, B.Y., Chen, G., Yoon, S.H., Demey, L.M., 2023. Activation of a Vibrio cholerae CBASS anti-phage system by quorum sensing and folate depletion. MBio 14, e00875-23. https://doi.org/10.1128/mbio.00875-23
177. Severin, G.B., Ramliden, M.S., Hawver, L.A., Wang, K., Pell, M.E., Kieninger, A.-K., Khataokar, A., O'Hara, B.J., Behrmann, L.V., Neiditch, M.B., Benning, C., Waters, C.M., Ng, W.-L., 2018. Direct activation of a phospholipase by cyclic GMP-AMP in El Tor Vibrio cholerae. Proc. Natl. Acad. Sci. 115, E6048–E6055. https://doi.org/10.1073/pnas.1801233115
178. Sheppard, N.F., Glover, C.V.C., Terns, R.M., Terns, M.P., 2016. The CRISPR-associated Csx1 protein of Pyrococcus furiosus is an adenosine-specific endoribonuclease. RNA 22, 216–224. https://doi.org/10.1261/rna.039842.113
179. Shi, M., Zhang, P., Vora, S.M., Wu, H., 2020. Higher-order assemblies in innate immune and inflammatory signaling: A general principle in cell biology. Curr. Opin. Cell Biol. 63, 194–203. https://doi.org/10.1016/j.ceb.2020.03.002
180. Skerker, J.M., Laub, M.T., 2004. Cell-cycle progression and the generation of asymmetry in Caulobacter crescentus. Nat. Rev. Microbiol. 2, 325–337. https://doi.org/10.1038/nrmicro864
181. Slavik, K.M., Kranzusch, P.J., 2023. CBASS to cGAS-STING: The Origins and Mechanisms of Nucleotide Second Messenger Immune Signaling. Annu. Rev. Virol. https://doi.org/10.1146/annurev-virology-111821-115636

182. Sofos, N., Feng, M., Stella, S., Pape, T., Fuglsang, A., Lin, J., Huang, Q., Li, Y., She, Q., Montoya, G., 2020. Structures of the Cmr-β Complex Reveal the Regulation of the Immunity Mechanism of Type III-B CRISPR-Cas. Mol. Cell 79, 741-757.e7. https://doi.org/10.1016/j.molcel.2020.07.008
183. Staals, R.H.J., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varossieau, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P.J., Urlaub, H., Heck, A.J.R., Nogales, E., Doudna, J.A., Shinkai, A., van der Oost, J., 2014. RNA Targeting by the Type III-A CRISPR-Cas Csm Complex of Thermus thermophilus. Mol. Cell 56, 518–530. https://doi.org/10.1016/j.molcel.2014.10.005
184. Steens, J.A., Bravo, J.P.K., Salazar, C.R.P., Yildiz, C., Amieiro, A.M., Köstlbacher, S., Prinsen, S.H.P., Andres, A.S., Patinios, C., Bardis, A., Barendregt, A., Scheltema, R.A., Ettema, T.J.G., van der Oost, J., Taylor, D.W., Staals, R.H.J., 2024. Type III-B CRISPR-Cas cascade of proteolytic cleavages. Science 383, 512–519. https://doi.org/10.1126/science.adk0378
185. Steens, J.A., Salazar, C.R.P., Staals, R.H.J., 2022. The diverse arsenal of type III CRISPR–Cas-associated CARF and SAVED effectors. Biochem. Soc. Trans. 50, 1353–1364. https://doi.org/10.1042/BST20220289
186. Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. 35, 1026–1028. https://doi.org/10.1038/nbt.3988
187. Stella, G., Marraffini, L., 2024. Type III CRISPR-Cas: beyond the Cas10 effector complex. Trends Biochem. Sci. 49, 28–37. https://doi.org/10.1016/j.tibs.2023.10.006
188. Strecker, J., Demircioglu, F.E., Li, D., Faure, G., Wilkinson, M.E., Gootenberg, J.S., Abudayyeh, O.O., Nishimasu, H., Macrae, R.K., Zhang, F., 2022. RNA-activated protein cleavage with a CRISPR-associated endopeptidase. Science 378, 874–881. https://doi.org/10.1126/science.add7450
189. Sundar, S., McGinness, K.E., Baker, T.A., Sauer, R.T., 2010. Multiple Sequence Signals Direct Recognition and Degradation of Protein Substrates by the AAA+ Protease HslUV. J. Mol. Biol. 403, 420–429. https://doi.org/10.1016/j.jmb.2010.09.008
190. Swarts, D.C., Jinek, M., 2019. Mechanistic Insights into the cis- and trans-Acting DNase Activities of Cas12a. Mol. Cell 73, 589-600.e4. https://doi.org/10.1016/j.molcel.2018.11.021
191. Tal, N., Morehouse, B.R., Millman, A., Stokar-Avihail, A., Avraham, C., Fedorenko, T., Yirmiya, E., Herbst, E., Brandis, A., Mehlman, T., Oppenheimer-Shaanan, Y., Keszei, A.F.A., Shao, S., Amitai, G., Kranzusch, P.J., Sorek, R., 2021. Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. Cell 184, 5728-5739.e16. https://doi.org/10.1016/j.cell.2021.09.031
192. Tamulaitiene, G., Sabonis, D., Sasnauskas, G., Ruksenaite, A., Silanskas, A., Avraham, C., Ofir, G., Sorek, R., Zaremba, M., Siksnys, V., 2024. Activation of Thoeris antiviral system via SIR2 effector filament assembly. Nature 627, 431–436. https://doi.org/10.1038/s41586-024-07092-x
193. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, Č., Nwokeoji, A.O., Dickman, M.J., Horvath, P., Siksnys, V., 2014. Programmable RNA

Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. Mol. Cell 56, 506–517. https://doi.org/10.1016/j.molcel.2014.09.027
194. Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., Lyumkis, D., 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods 14, 793–796. https://doi.org/10.1038/nmeth.4347
195. Tatsuta, T., Joo, D.M., Calendar, R., Akiyama, Y., Ogura, T., 2000. Evidence for an active role of the DnaK chaperone system in the degradation of σ32. FEBS Lett. 478, 271–275. https://doi.org/10.1016/S0014-5793(00)01869-X
196. Tesson, F., Bernheim, A., 2023. Synergy and regulation of antiphage systems: toward the existence of a bacterial immune system? Curr. Opin. Microbiol. 71, 102238. https://doi.org/10.1016/j.mib.2022.102238
197. Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., Bernheim, A., 2022. Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat. Commun. 13, 2561. https://doi.org/10.1038/s41467-022-30269-9
198. Urbaitis, T., Gasiunas, G., Young, J.K., Hou, Z., Paulraj, S., Godliauskaite, E., Juskeviciene, M.M., Stitilyte, M., Jasnauskaite, M., Mabuchi, M., Robb, G.B., Siksnys, V., 2022. A new family of CRISPR-type V nucleases with C-rich PAM recognition. EMBO Rep. 23, e55481. https://doi.org/10.15252/embr.202255481
199. van Beljouw, S.P.B., Haagsma, A.C., Rodríguez-Molina, A., van den Berg, D.F., Vink, J.N.A., Brouns, S.J.J., 2021. The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. Science 373, 1349–1353. https://doi.org/10.1126/science.abk2718
200. van Beljouw, S.P.B., Sanders, J., Rodríguez-Molina, A., Brouns, S.J.J., 2023. RNA-targeting CRISPR–Cas systems. Nat. Rev. Microbiol. 21, 21–34. https://doi.org/10.1038/s41579-022-00793-y
201. van den Berg, D.F., Costa, A.R., Esser, J.Q., Stanciu, I., Geissler, J.Q., Zoumaro-Djayoon, A.D., Haas, P.-J., Brouns, S.J.J., 2024. Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity highlight shared evolutionary roots of viral defenses. Cell Host Microbe 32, 1427-1443.e8. https://doi.org/10.1016/j.chom.2024.07.007
202. Van Opdenbosch, N., Lamkanfi, M., 2019. Caspases in Cell Death, Inflammation, and Disease. Immunity 50, 1352–1364. https://doi.org/10.1016/j.immuni.2019.05.020
203. Vassallo, C.N., Doering, C.R., Littlehale, M.L., Teodoro, G.I.C., Laub, M.T., 2022. A functional selection reveals previously undetected anti-phage defence systems in the E. coli pangenome. Nat. Microbiol. 7, 1568–1579. https://doi.org/10.1038/s41564-022-01219-4
204. Velisetty, P., Chakrapani, S., 2012. Desensitization Mechanism in Prokaryotic Ligand-gated Ion Channel. J. Biol. Chem. 287, 18467–18477. https://doi.org/10.1074/jbc.M112.348045
205. Wah, D.A., Levchenko, I., Baker, T.A., Sauer, R.T., 2002. Characterization of a Specificity Factor for an AAA+ ATPase: Assembly of SspB Dimers with ssrA-Tagged Proteins and the ClpX Hexamer. Chem. Biol. 9, 1237–1245. https://doi.org/10.1016/S1074-5521(02)00268-5

206. Walker, F.C., Chou-Zheng, L., Dunkle, J.A., Hatoum-Aslan, A., 2017. Molecular determinants for CRISPR RNA maturation in the Cas10–Csm complex and roles for non-Cas nucleases. Nucleic Acids Res. 45, 2112–2123. https://doi.org/10.1093/nar/gkw891
207. Walsh, N.P., Alba, B.M., Bose, B., Gross, C.A., Sauer, R.T., 2003. OMP Peptide Signals Initiate the Envelope-Stress Response by Activating DegS Protease via Relief of Inhibition Mediated by Its PDZ Domain. Cell 113, 61–71. https://doi.org/10.1016/S0092-8674(03)00203-4
208. Whiteley, A.T., Eaglesham, J.B., de Oliveira Mann, C.C., Morehouse, B.R., Lowey, B., Nieminen, E.A., Danilchanka, O., King, D.S., Lee, A.S.Y., Mekalanos, J.J., Kranzusch, P.J., 2019. Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature 567, 194–199. https://doi.org/10.1038/s41586-019-0953-5
209. Xia, P., Dutta, A., Gupta, K., Batish, M., Parashar, V., 2022. Structural basis of cyclic oligoadenylate binding to the transcription factor Csa3 outlines cross talk between type III and type I CRISPR systems. J. Biol. Chem. 298, 101591. https://doi.org/10.1016/j.jbc.2022.101591
210. Yan, W.X., Hunnewell, P., Alfonse, L.E., Carte, J.M., Keston-Smith, E., Sothiselvam, S., Garrity, A.J., Chong, S., Makarova, K.S., Koonin, E.V., Cheng, D.R., Scott, D.A., 2019. Functionally diverse type V CRISPR-Cas systems. Science 363, 88–91. https://doi.org/10.1126/science.aav7271
211. Yang, H., Patel, D.J., 2022. A type III-E CRISPR Craspase exhibiting RNase and protease activities. Cell Res. 32, 1044–1046. https://doi.org/10.1038/s41422-022-00739-2
212. Yang, J., Li, X., He, Q., Wang, X., Tang, J., Wang, T., Zhang, Y., Yu, F., Zhang, S., Liu, Z., Zhang, L., Liao, F., Yin, H., Zhao, H., Deng, Z., Zhang, H., 2024. Structural basis for the activity of the type VII CRISPR–Cas system. Nature 633, 465–472. https://doi.org/10.1038/s41586-024-07815-0
213. Yang, W., 2011. Nucleases: diversity of structure, function and mechanism. Q. Rev. Biophys. 44, 1–93. https://doi.org/10.1017/S0033583510000181
214. Ye, Qiaozhen, Lau, R.K., Mathews, I.T., Birkholz, E.A., Watrous, J.D., Azimi, C.S., Pogliano, J., Jain, M., Corbett, K.D., 2020. HORMA Domain Proteins and a Trip13-like ATPase Regulate Bacterial cGAS-like Enzymes to Mediate Bacteriophage Immunity. Mol. Cell 77, 709-722.e7. https://doi.org/10.1016/j.molcel.2019.12.009
215. Ye, Qing, Zhao, X., Liu, J., Zeng, Z., Zhang, Z., Liu, T., Li, Y., Han, W., Peng, N., 2020. CRISPR-Associated Factor Csa3b Regulates CRISPR Adaptation and Cmr-Mediated RNA Interference in Sulfolobus islandicus. Front. Microbiol. 11.
216. Yirmiya, E., Leavitt, A., Lu, A., Ragucci, A.E., Avraham, C., Osterman, I., Garb, J., Antine, S.P., Mooney, S.E., Hobbs, S.J., Kranzusch, P.J., Amitai, G., Sorek, R., 2024. Phages overcome bacterial immunity via diverse anti-defence proteins. Nature 625, 352–359. https://doi.org/10.1038/s41586-023-06869-w
217. Yoshioka, K., 2002. KyPlot — A User-oriented Tool for Statistical Data Analysis and Visualization. Comput. Stat. 17, 425–437. https://doi.org/10.1007/s001800200117

218. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., Wang, Y., 2019. Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. Cell 176, 239-253.e16. https://doi.org/10.1016/j.cell.2018.10.052
219. Zhang, D., Du, L., Gao, H., Yuan, C., Lin, Z., 2024. Structural insight into the Csx1–Crn2 fusion self-limiting ribonuclease of type III CRISPR system. Nucleic Acids Res. 52, 8419–8430. https://doi.org/10.1093/nar/gkae569
220. Zhao, R., Yang, Y., Zheng, F., Zeng, Z., Feng, W., Jin, X., Wang, J., Yang, K., Liang, Y.X., She, Q., Han, W., 2020. A Membrane-Associated DHH-DHHA1 Nuclease Degrades Type III CRISPR Second Messenger. Cell Rep. 32, 108133. https://doi.org/10.1016/j.celrep.2020.108133
221. Zhou, Y., Gottesman, S., Hoskins, J.R., Maurizi, M.R., Wickner, S., 2001. The RssB response regulator directly targets ςS for degradation by ClpXP. Genes Dev. 15, 627–637. https://doi.org/10.1101/gad.864401
222. Zhu, D., Wang, Lijun, Shang, G., Liu, X., Zhu, J., Lu, D., Wang, Lei, Kan, B., Zhang, J., Xiang, Y., 2014. Structural Biochemistry of a Vibrio cholerae Dinucleotide Cyclase Reveals Cyclase Activity Regulation by Folates. Mol. Cell 55, 931–937. https://doi.org/10.1016/j.molcel.2014.08.001
223. Zhu, W., McQuarrie, S., Grüschow, S., McMahon, S.A., Graham, S., Gloster, T.M., White, M.F., 2021. The CRISPR ancillary effector Can2 is a dual-specificity nuclease potentiating type III CRISPR defence. Nucleic Acids Res. 49, 2777–2789. https://doi.org/10.1093/nar/gkab073
224. Zhuang, J., Xie, L., Zheng, L., 2022. A Glimpse of Programmed Cell Death Among Bacteria, Animals, and Plants. Front. Cell Dev. Biol. 9. https://doi.org/10.3389/fcell.2021.790117
225. Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., Alva, V., 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol. 430, 2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007

REPRODUCTION LICENCES

https://doi.org/10.1093/nar/gkaa635.

APPENDICES

Appendix 1. Plasmids used in the study of cA_6 signaling regulation.

Name	Backbone / antibiotic resistance	Details	Source
pCas/Csm	pCDFDuet-1 / Str^R	Plasmid encoding *S. thermophilus* DGCC8004 genes: *cas6*, *cas10*, *csm2*, *csm3*, *csm4*, *csm5*, *csm6′*, *csm6*. Used for expression of wt-StCsm complex.	(Tamulaitis et al., 2014)
pCas/Csm_dCsm3	pCDFDuet-1 / Str^R	Plasmid encoding *S. thermophilus* DGCC8004 genes: *cas6*, *cas10*, *csm2*, (D33A)*csm3*, *csm4*, *csm5*, *csm6′*, *csm6*. Used for expression of RNase deficient StCsm complex in the presence of StCsm6 and StCsm6′.	(Tamulaitis et al., 2014)
pCas/Csm_dCsm3_dHD Cas10	pCDFDuet-1 / Str^R	Plasmid encoding *S. thermophilus* DGCC8004 genes: *cas6*, (D16A)*cas10*, *csm2*, (D33A)*csm3*, *csm4*, *csm5*, *csm6′*, *csm6*. Used for expression of RNase and DNase deficient StCsm complex in the presence of StCsm6 and StCsm6′.	Provided by dr. M. Kazlauskiene
pCas/Csm_dCsm3_dHD Cas10_ ΔCsm6′ΔCsm6	pCDFDuet-1 / Str^R	Plasmid encoding *S. thermophilus* DGCC8004 genes: *cas6*, (D16A)*cas10*, *csm2*, (D33A)*csm3*, *csm4*, *csm5*, *csm6′*, *csm6*. Used for expression of RNase and DNase deficient StCsm complex in the absence of StCsm6 and StCsm6′.	Provided by dr. M. Kazlauskiene
pCas/Csm_dCsm3_dHD Cas10_dPalmCas10_ΔCsm6′ ΔCsm6	pCDFDuet-1 / Str^R	Plasmid encoding *S. thermophilus* DGCC8004 genes: *cas6*, (D16A, D575A, D576A)*cas10*, *csm2*, (D33A)*csm3*, *csm4*, *csm5*, *csm6′*, *csm6*. Used for expression of RNase, DNase and synthethase deficient StCsm complex in the absence of StCsm6 and StCsm6′.	Provided by dr. M. Kazlauskiene
pCRISPR_Tc	pACYCDuet-1 / Cm^R	Plasmid encoding synthetic *S. thermophilus* DGCC8004 CRISPR locus composed of the leader sequence, (repeat, spacer)$_4$ and terminal repeat. Tc spacer sequence targets the transcript of gene coding tetracycline resistance protein (Tc^R). Used for expression of crRNA targeting Tc^R transcript.	Provided by dr. M. Kazlauskiene

pCRISPR_S3	pACYCDuet-1 / Cm^R	Plasmid encoding synthetic *S. thermophilus* DGCC8004 CRISPR locus composed of the leader sequence, (repeat, spacer)$_4$ and terminal repeat. Spacer sequence is obtained from native *S. thermophilus* DGCC8004 CRISPR locus. Used for expression of StCsm complex with crRNA containing S3 spacer.	(Tamulaitis et al., 2014)
pCRISPR_MS2	pACYCDuet-1 / Cm^R	Plasmid encoding minimal synthetic *S. thermophilus* DGCC8004 CRISPR locus composed of the leader sequence, repeat, spacer and terminal repeat. MS2 pacer sequence targets the *rep* transcript of MS2 phage. Used for expression of crRNA targeting MS2 phage.	Provided by dr. M. Kazlauskiene
pTarget_Tc (Tc^R, Ap^R)	pBR322 / Tc^R, Ap^R	Plasmid containing gene coding tetracycline resistance protein Tc^R. Used for constitutive Tc target RNA production in cells.	Thermo Fisher Scientific
pTarget_Tc^{mut}	pBR322 / Tc^R, Ap^R	Plasmid containing gene encoding mutated version of Tc^R. Mutations in Tc target site result in the 5′-GTTTCCGT3′ 3′-flanking sequence, making it complementary to the 5′-handle (5′ACGGAAAC-3′) of crRNA bound in StCsm complex.	Provided by dr. M. Kazlauskiene
pTarget_ctrl	pUC18 / Ap^R	Plasmid used as control to pTarget_Tc plasmid. No gene coding tetracycline resistance protein.	Thermo Fisher Scientific
pCsm2N-Tag	pETDuet-1 / Ap^R	Plasmid encoding *S. thermophilus* DGCC8004 *csm2* fused with N-terminal StrepII-tag. Used for expression of N-terminal StrepII-tagged Csm2, for StCsm complex purification.	(Tamulaitis et al., 2014)
pCsm6	pBAD24 / Ap^R	Plasmid encoding *S. thermophilus* DGCC8004 *csm6* fused with N-terminal His_6-StrepII-His_6 tag. Used for expression of tagged wt-StCsm6.	(Kazlauskiene et al., 2017)
pCsm6′	pBAD24 / Ap^R	Plasmid encoding *S. thermophilus* DGCC8004 *csm6′* fused with N-terminal His_6-StrepII-His_6 tag. Used for expression of tagged wt-StCsm6′.	(Kazlauskiene et al., 2017)
pdCARF-Csm6	pBAD24 / Ap^R	Plasmid encoding *S. thermophilus* DGCC8004 (D12A)*csm6* fused with N-terminal His_6-StrepII-His_6 tag. Used for expression of tagged CARF ring nuclease activity-deficient StCsm6.	This study
pdHEPN-Csm6	pBAD24 / Ap^R	Plasmid encoding *S. thermophilus* DGCC8004 (R371A, H376A) *csm6* fused with N-terminal His_6-StrepII-His_6 tag. Used for expression of tagged RNase activity-deficient StCsm6.	(Kazlauskiene et al., 2017)

pCARF	pBAD24 / Ap^R	Plasmid encoding 1-169 codons of *csm6* from *S. thermophilus* DGCC8004 fused with N-terminal TEV protease-cleavable MBP and His_{10} tag. Used for expression of tagged CARF domain of StCsm6.	This study
pHEPN	pBAD24 / Ap^R	Plasmid encoding 172-428 codons of *csm6* from *S. thermophilus* DGCC8004 fused with N-terminal TEV protease-cleavable His_6 tag. Used for expression of tagged HEPN domain of StCsm6.	This study
pdCARF	pBAD24 / Ap^R	Plasmid encoding 1-169 codons of (D12A)*csm6* from *S. thermophilus* DGCC8004 fused with N-terminal TEV protease-cleavable MBP and His_{10} tag. Used for expression of tagged ring nuclease activity-deficient CARF domain of StCsm6.	This study
pdHEPN	pBAD24 / Ap^R	Plasmid encoding 172-428 codons of (R371A, H376A)*csm6* from *S. thermophilus* DGCC8004 fused with N-terminal TEV protease-cleavable His_6 tag. Used for expression of tagged RNase activity-deficient HEPN domain of StCsm6.	This study
pUC18_S3/1	pUC18_S3/1 / Ap^R	Plasmid for generation of S3/1 and NS RNA substrates	(Tamulaitis et al., 2014)

Appendix 2. Plasmids used in the study of CCaCalpL-CalpT-CalpS tripartite effector system.

Name	Backbone / antibiotic resistance	Details	Source
pCCaCsm	pCDFDuet-1 / Str^R	Plasmid encoding *Ca.* C. acidaminovorans *cas10* (fused with TEV-cleavable His_6-StrepII-His_6 tag), *csm2*, *csm3*, *csm4* and *csm5* genes. Used for expression of tagged CCaCsm complex.	This study
pCCaCRISPR	pETDuet1 / Ap^R	Plsmid encoding minimal synthetic *Ca.* C. acidaminovorans CRISPR locus (repeat-spacer-repeat) and *cas6* gene. Used for expression of CCaCsm complex.	This study
pCCaCalpL-^{N}Tag	pBAD/HisA / Ap^R	Plasmid encoding codon optimized *calpL* variant from *Ca.* C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on N terminus. Used for expression of N-tagged CCaCalpL.	This study
pCCaCalpL-^{C}Tag	pBAD/HisA / Ap^R	Plasmid encoding codon optimized *calpL* variant from *Ca.* C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL.	This study
pSAVED-^{C}Tag	pBAD/HisA / Ap^R	Plasmid encoding 209-506 optimized codons of *calpL* (with I209A mutation) from *Ca.* C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged SAVED domain of CCaCalpL.	This study
p(S154A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^R	Plasmid encoding codon optimized (S154A)*calpL* variant from *Ca.* C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged protease deficient CCaCalpL.	This study
p(K330A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^R	pBAD/HisA derivative encoding codon optimized (K330A)*calpL* variant from *Ca.* C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying K330A mutation.	This study

p(R358A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	pBAD/HisA derivative encoding codon optimized (R358A)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying R58A mutation.	This study
p(K361A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	Plasmid encoding codon optimized (K361A)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying K361A mutation.	This study
p(S395A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	Plasmid encoding codon optimized (S395A)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying S395A mutation.	This study
p(H396A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	Plasmid encoding codon optimized (H396A)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying H396A mutation.	This study
p(H476A)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	Plasmid encoding codon optimized (H476A)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying H476A mutation.	This study
p(R358E-K361E)CCaCalpL-^{C}Tag	pBAD/HisA / Ap^{R}	Plasmid encoding codon optimized (R358E-K361E)*calpL* variant from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpL carrying R358E and K361E mutations.	This study
pET-CCaCalpS-^{C}Tag	pETDuet-1 / Ap^{R}	Plasmid encoding codon optimized *calpS* from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpS in the presence and absence of CCaCalpT.	This study
pCDF-CCaCalpT	pCDFDuet-1 / Str^{R}	Plasmid encoding codon optimized *calpT* from *Ca*. C. acidaminovorans under IPTG-inducible promoter. Used in combination with pET-CCaCalpS-^{C}Tag for CCaCalpT-CalpS co-expression.	This study
pET-CalpT-^{C}Tag	pETDuet-1 / Ap^{R}	Plasmid encoding codon optimized *calpT* from *Ca*. C. acidaminovorans fused with TEV-cleavable His_6-StrepII-His_6 tag on C terminus. Used for expression of C-tagged CCaCalpT.	This study

pCCaCalpS	pCDFDuet-1 / Str^R	Plasmid encoding codon optimized *calpS* from *Ca*. C. acidaminovorans. Used for the toxicity assay in *E. coli.*	This study
pCCaCalpT	pETDuet-1 / Ap^R	Plasmid encoding codon optimized full-length (1-299 codons) *calpT* from *Ca.* C. acidaminovorans. Used for the toxicity assay in *E. coli.*	This study
pCCaCalpT$_{24N}$	pETDuet-1 / Ap^R	Plasmid encoding 1-204 codons of codon optimized *calpT* from *Ca.* C. acidaminovorans. Used for expression of CCaCalpT$_{24N}$ with wt C-degron in the toxicity assay in *E. coli.*	This study
pCCaCalpT$_{11C}$	pETDuet-1 / Ap^R	Plasmid encoding 205-299 codons of codon optimized *calpT* from *Ca*. C. acidaminovorans. Used for expression of CCaCalpT$_{11C}$ in the toxicity assay in *E. coli.*	This study
pCCaCalpT$_{split}$	pETDuet-1 / Ap^R	Plasmid encoding 1-204 and 205-299 codons of codon optimized *calpT* from *Ca*. C. acidaminovorans. Used in the toxicity assay in *E. coli* for expression of CCaCalpT$_{24N}$ with wt C-degron and CCaCalpT$_{11C}$.	This study
pCCaCalpT$_{24N}$(mut)	pETDuet-1 / Ap^R	Plasmid encoding 1-204 codons of codon optimized (A204D)*calpT* variant from *Ca*. C. acidaminovorans. Used in the toxicity assay in *E. coli* for expression of CCaCalpT$_{24N}$ with the A204D mutation in C-degron.	This study
pGFP	pETDuet-1 / Ap^R	Plasmid encoding *mut1-GFP*. Used in GFP degradation assays for expression of GFP without the C-degron.	Provided by Konstanty Keda
pGFP-(wt)CCaCalpT$_{24N}$-degron	pETDuet-1 / Ap^R	Plasmid for expression of mut1-GFP fused with CCaCalpT$_{24N}$ C-degron sequence PETLQLAA at the C-terminus. Used in GFP degradation assays.	This study
pGFP-(mut)CCaCalpT$_{24N}$-degron	pETDuet-1 / Ap^R	Plasmid for expression of mut1-GFP fused with mutant version of CCaCalpT$_{24N}$ C-degron sequence PETLQLAD at the C-terminus. Used in GFP degradation assays.	This study
pGFP-ssrA-degron	pETDuet-1 / Ap^R	Plasmid for expression of mut1-GFP fused with partial ssrA C-degron sequence ENYALAA at the C-terminus. Used in GFP degradation assays.	This study
pCDFDuet-1	pCDFDuet-1 / Str^R	Control plasmid used in the toxicity assay in *E. coli*.	Novagen
pETDuet-1	pETDuet-1 / Ap^R	Control plasmid used in the toxicity assay in *E. coli*.	Novagen

Appendix 3. RNA used in the study

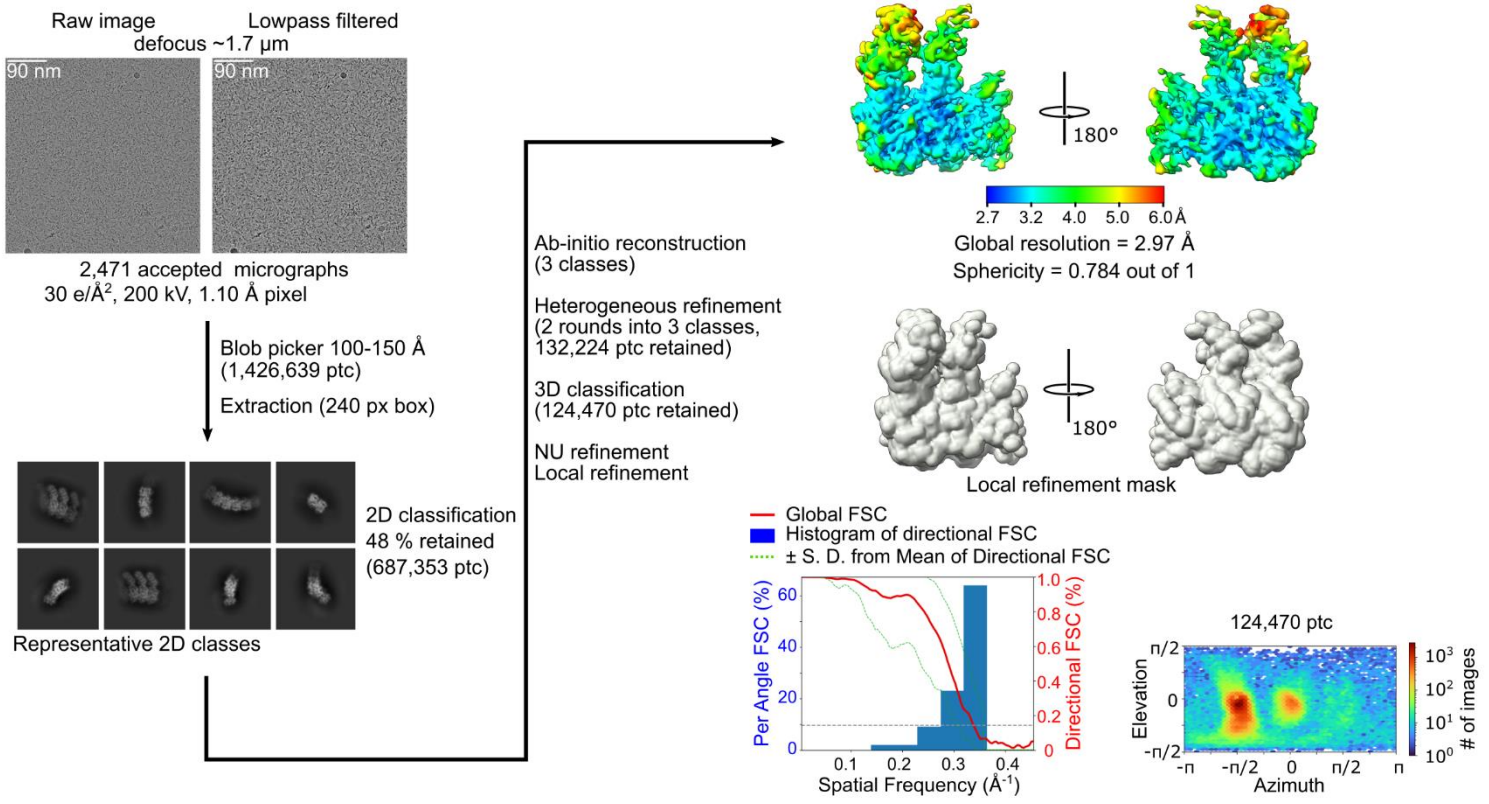
Name	Sequence 5′-3′/explanation	Usage	Source
S3/1	GGGCGGCAAAUUGAGGAGGUGGAAUAAGUGAACAGAAU UAAACAGUUACGAAAAACAAAAGGUGGCGA	Target RNA for StCsm complex to induce the production of cA_n *in vitro*.	(Tamulaitis et al., 2014)
CCaTarget	GGGUGAAGAGCAAUGAGCUCUCGAGGUGCGAUAUCGCUC UUCCCAGUGUA	Target RNA for CCaCsm complex to induce the production of cA_n *in vitro*.	This study
NS RNA	GGGCGGCAAAUUGAGGAUUUCGUAACUGUUUAAUUCUG UUCACUUAUUCCACCAACAAAAGGUGGCGA	Non-target RNA used as control in cA_n production *in vitro* experiments. Radiolabeled variant used as RNA substrate in StCsm6 RNA hydrolysis assays.	(Tamulaitis et al., 2014)
cA_3	3′-5′ linked cyclic AAA	Standard and control in ring nuclease reactions.	Biolog
cA_4	3′-5′ linked cyclic AAAA	Substrate in ring nuclease reactions, activator in CCaCalpL protease reactions and structural studies.	Biolog
cA_5	3′-5′ linked cyclic AAAAA	Activator in StCsm6 RNase reaction.	Dr. M. Kazlauskiene and Dr. A. Šilanskas
cA_6	3′-5′ linked cyclic AAAAAA	Substrate in ring nuclease reactions, activator in StCsm RNase reactions.	Biolog
A_4>p	3′-5′ linked linear AAAA with 2′-3′ cyclic phosphate	Substrate in ring nuclease reactions, activator in CCaCalpL protease reactions and structural studies.	ChemGenes
A_4p	3′-5′ linked linear AAAA with 3′ phosphate	Substrate in ring nuclease reactions, activator in CCaCalpL protease reactions and structural studies.	Metabion
A_6>p	3′-5′ linked linear AAAAAA with 2′-3′ cyclic phosphate	Substrate in ring nuclease reactions, activator in StCsm RNase reactions.	ChemGenes

Appendix 4. Cryo-EM data collection, refinement and validation statistics.

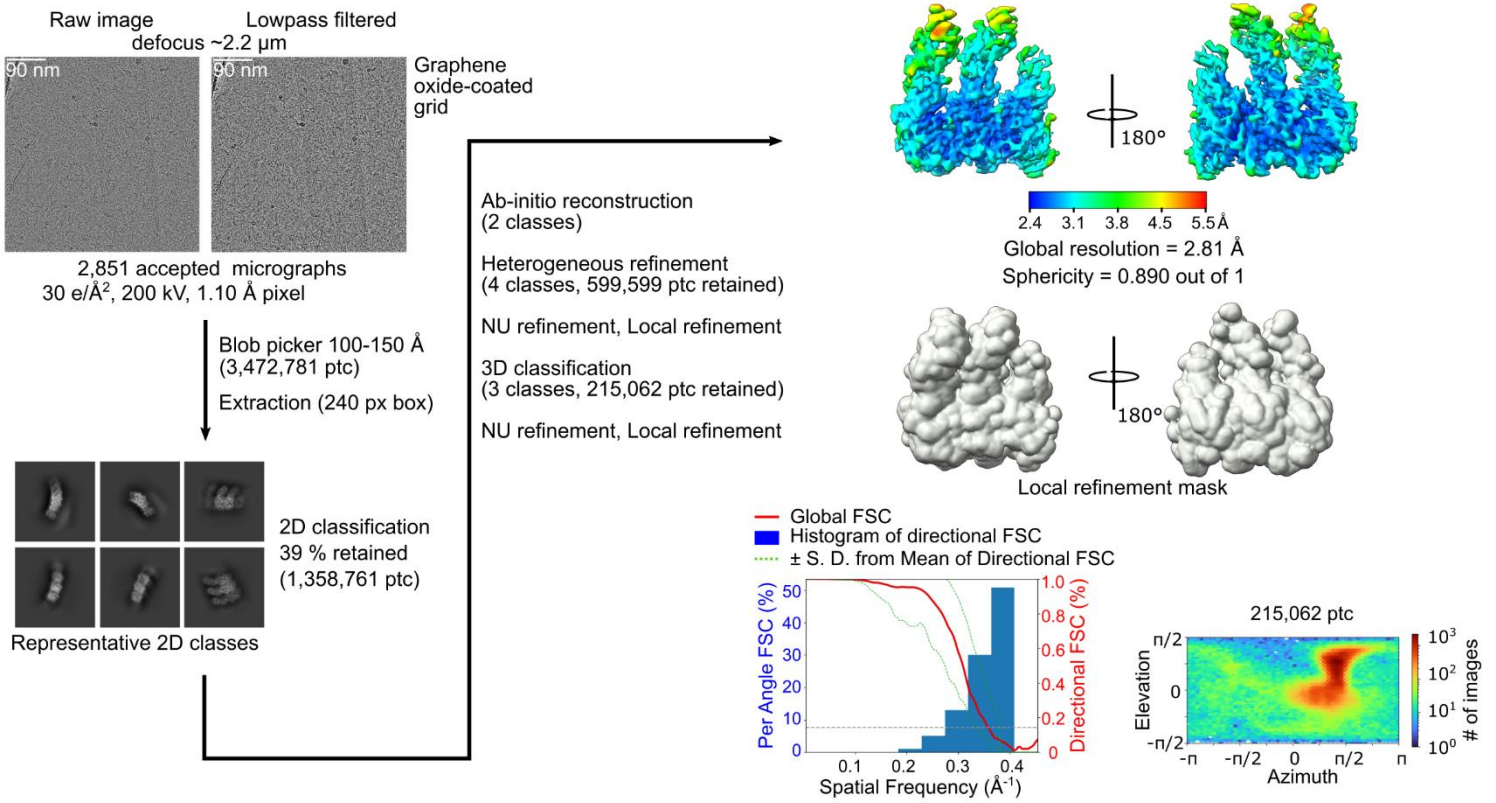
PDB **EMDB**	**cA$_4$-bound CCaCalpL** **9EYJ** **EMD-50055**	**A$_4$>p-bound CCaCalpL** **9EYK** **EMD-50056**	**A$_4$p-bound CCaCalpL** **9EYI** **EMD-50054**
Data collection and processing			
Microscope	Glacios (TFS)	Glacios (TFS)	Glacios (TFS)
Detector	Falcon III camera (TFS)	Falcon III camera (TFS)	Falcon III camera (TFS)
Magnification	92,000	92,000	92,000
Automation software	EPU	EPU	EPU
Voltage (kV)	200	200	200
Electron exposure (e-/Å^2)	29.7	29.7	29.7
Defocus range (μm)	0.5 to 2.5	0.5 to 2.5	0.5 to 2.5
Pixel size (Å)	1.1	1.1	1.1
Symmetry imposed	C1	C1	C1
Micrographs collected/used	2,546/2,471	3,017/2,851	3,469/3,127
Initial/final particle images (no.)	1,426,639/124,470	3,472,781/215,062	2,604,122/175,826
Map resolution (Å) (masked/unmasked); FSC threshold 0.143	2.97/3.18	2.81/2.94	2.75/2.92
Map sharpening B factor (Å^2)	105.8	108.5	119.7
3D FSC sphericity score	0.784	0.890	0.925
Refinement (Phenix)			
Initial model used	A$_4$p-bound CCaCalpL	A$_4$p-bound CCaCalpL	AlphaFold
Model resolution (Å) FSC threshold 0.5	3.38	3.29	3.00
Model composition			

Non-hydrogen atoms	4334	6639	9453
Protein residues	576	891	1248
Nucleotides	4	8	8
B factors (Å^2)			
Protein	72.07	80.57	83.48
Nucleotide	51.65	63.93	54.42
RMSDs			
Bond lengths (Å)	0.004	0.005	0.005
Bond angles (°)	0.587	0.750	0.633
Validation			
MolProbity score	2.22	2.28	1.99
Clashscore	13.32	13.61	10.21
Poor rotamers (%)	2.14	3.10	2.15
Ramachandran plot			
Favored (%)	95.16	96.13	96.72
Allowed (%)	4.84	3.87	3.28
Disallowed (%)	0	0	0
CCvolume/CCmask	0.71/0.72	0.73/0.73	0.81/0.81
Cb outliers (%)	0.18	0	0
CaBLAM outliers (%)	0.93	0.98	0.42

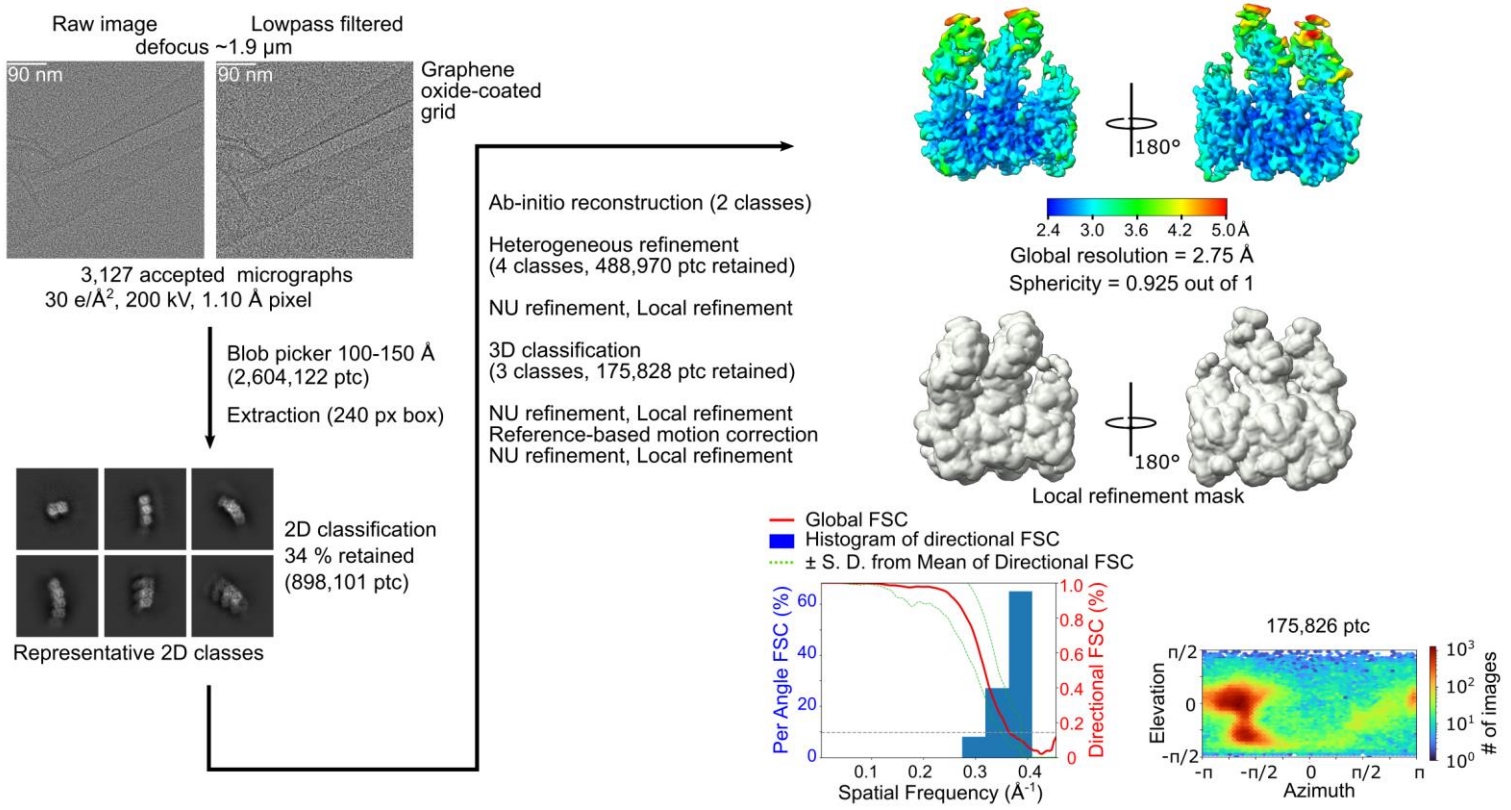
Appendix 5. Cryo-EM single particle analysis workflow of the cA$_4$-bound CCaCalpL sample



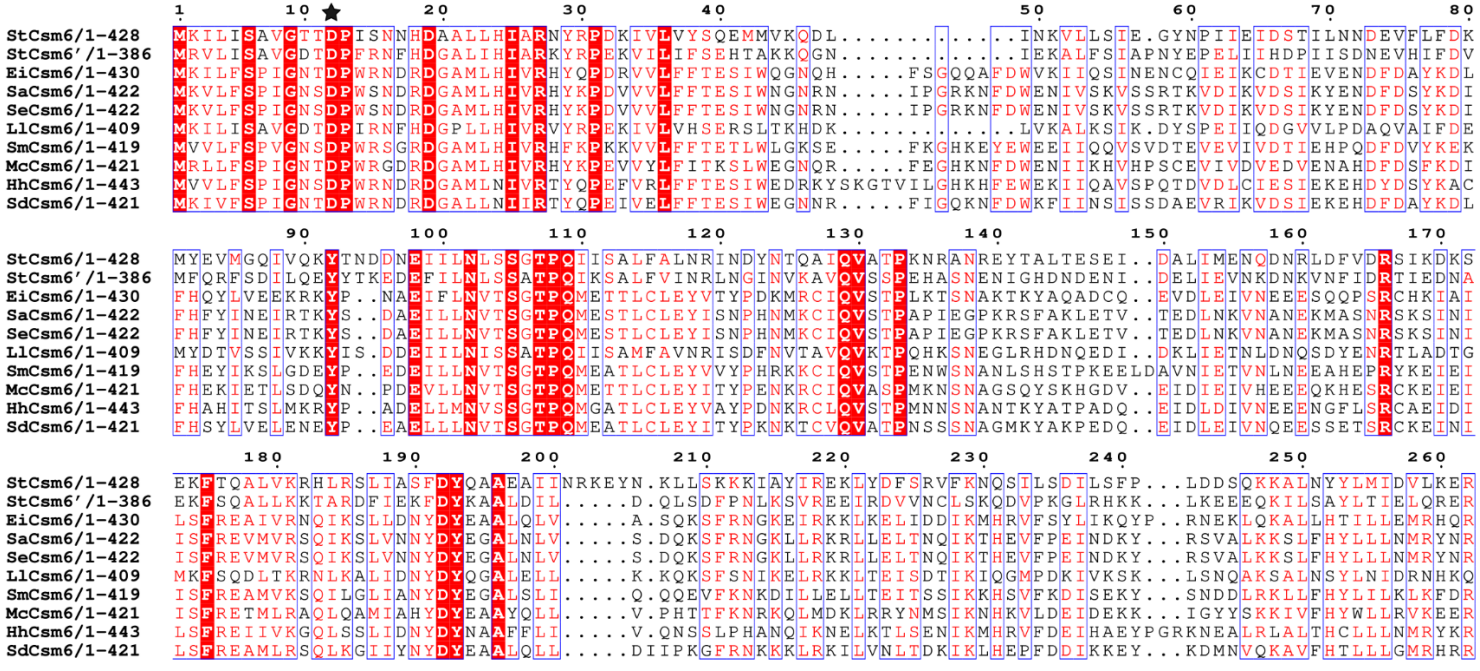
Appendix 6. Cryo-EM single particle analysis workflow of the A_4>p-bound CCaCalpL sample

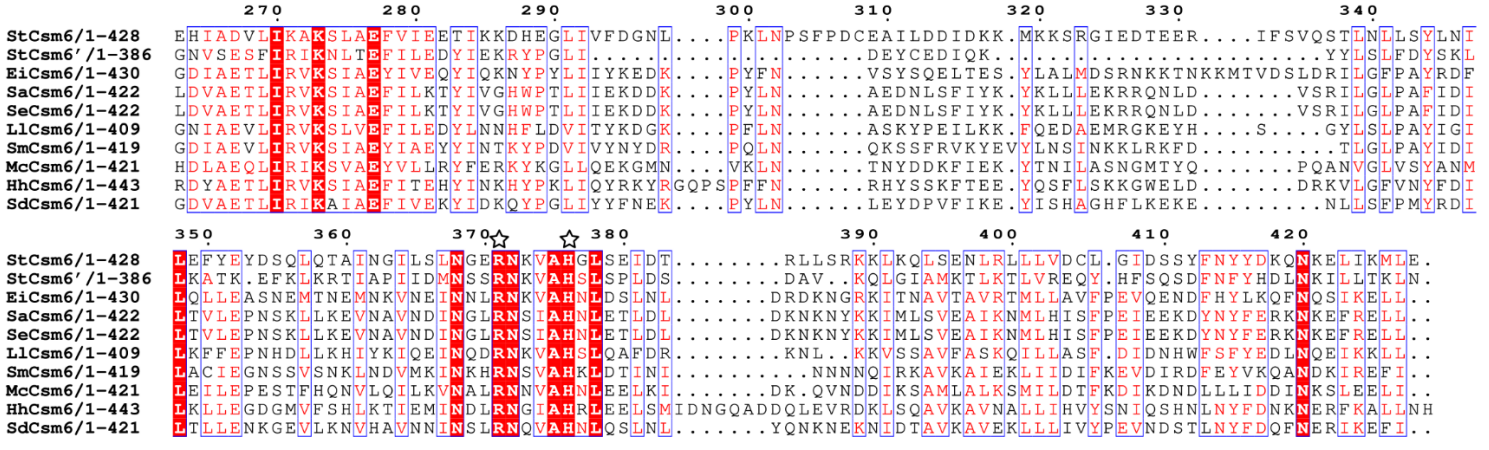


Appendix 7. Cryo-EM single particle analysis workflow of the A_4p-bound CCaCalpL sample

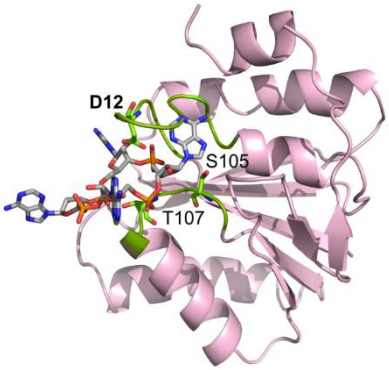


Appendix 8. Multiple sequence alignment of Csm6 homologues. The alignment was generated using MUSCLE and displayed using Espript 3.0. Black star indicates conserved aspartate important for ring nuclease activity of the CARF domain. White stars mark the active site residues of the HEPN RNase. StCsm6 - *Streptococcus thermophilus* Csm6 (WP_014621552.1), StCsm6′ *Streptococcus thermophilus* Csm6′ (WP_014621551.1), EiCsm6 - *Enterococcus italicus* Csm6 (WP_007208953.1), SaCsm6 - *Staphylococcus aureus* Csm6 (EZS13226.1), SeCsm6 - *Staphylococcus epidermidis* Csm6 (WP_002502662.1), LlCsm6 - *Lactococcus lactis* Csm6 (AGA14268.1), SmCsm6 - *Staphylococcus microti* Csm6 (WP_044359384.1), McCsm6 - *Macrococcoides caseolyticum* Csm6 (WP_133445588.1), HhCsm6 - *Halolactibacillus halophilus* Csm6 (WP_089831386.1), SdCsm6 *Staphylococcus debuckii* Csm6 (WP_123144061.1).

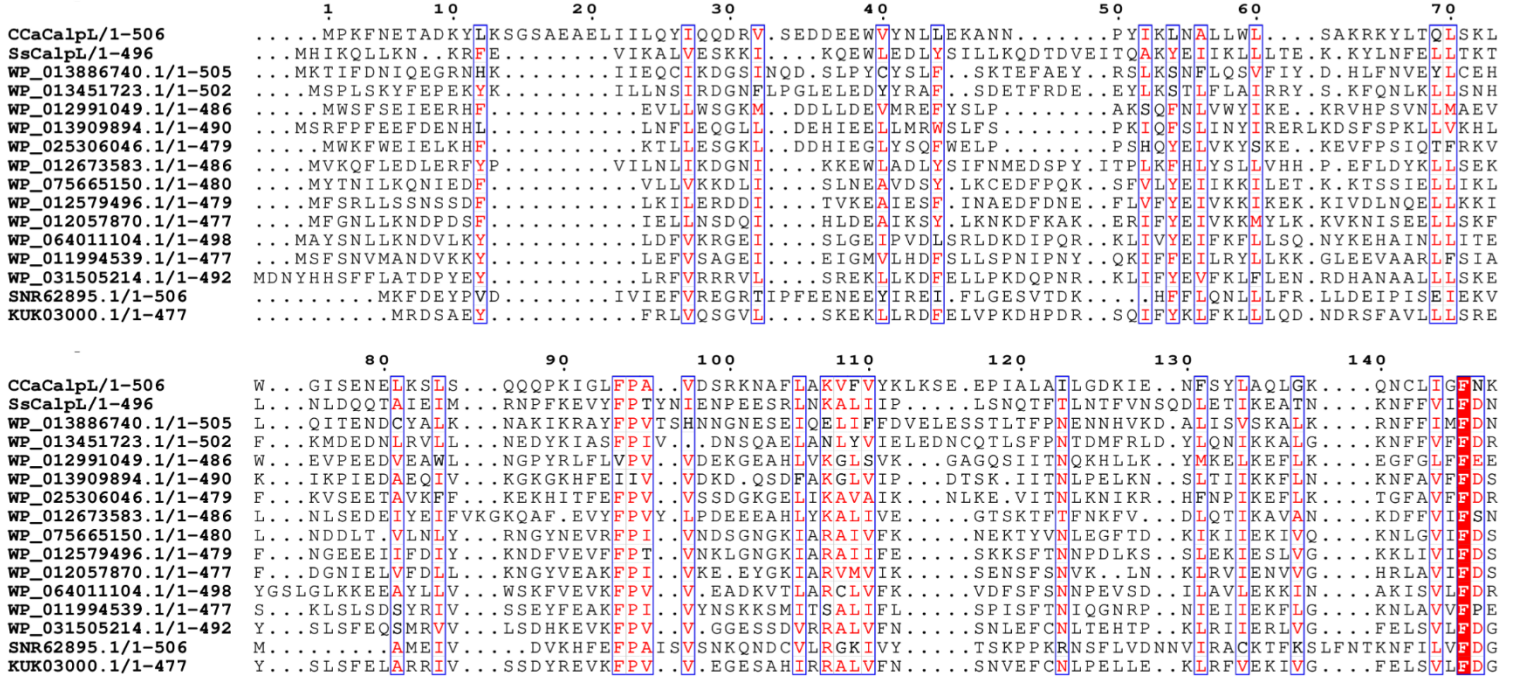


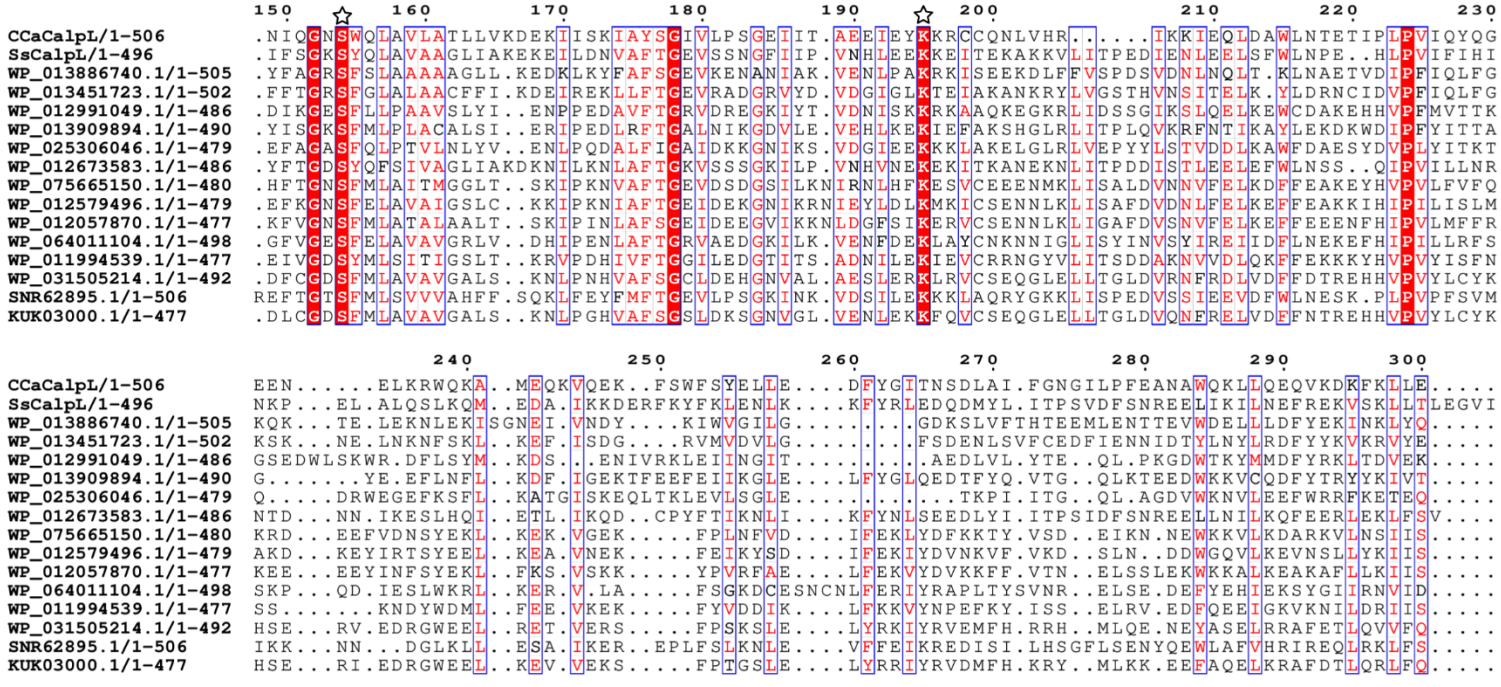

270 280 290 300 310 320 330 340
StCsm6/1-428 EHIADVLIKAKSLAEFVIEETIKKDHEGLIVFDGNL....PKLNPSFPDCEAILDDIDKK.MKKSRGIEDTEER....IFSVQSTLNLLSYLNI
StCsm6'/1-386 GNVSESFIRIKNLTEFILEDYIEKRYPGLI.....DEYCEDIQK.....YYLSLFDYSKL
EiCsm6/1-430 GDIAETLIRVKSIAEYIVEQYIQKNYPYLIIYKEDK....PYFN.....VSYSQELTES.YLALMDSRNKKTNKKMTVDSLDRILGFPAYRDF
SaCsm6/1-422 LDVAETLIRVKSIAEFILKTYIVGHWPTLIIEKDDK....PYLN.....AEDNLSFIYK.YKLLLEKRRQNLD.....VSRILGLPAFIDI
SeCsm6/1-422 LDVAETLIRVKSIAEFILKTYIVGHWPTLIIEKDDK....PYLN.....AEDNLSFIYK.YKLLLEKRRQNLD.....VSRILGLPAFIDI
LlCsm6/1-409 GNIAEVLIRVKSLVEFILEDYLNNHFLDVITYKDGK....PFLN.....ASKYPEILKK.FQEDAEMRGKEYH....S....GYLSLPAYIGI
SmCsm6/1-419 GDIAEVLIRVKSIAEYIAEYYINTKYPDVIVYNYDR....PQLN.....QKSSFRVKYEVYLNSINKKLRKFD.....TLGLPAYIDI
McCsm6/1-421 HDLAEQLIRIKSVAEYVLLRYFERKYKGLLQEKGMN....VKLN.....TNYDDKFIEK.YTNILASNGMTYQ.....PQANVGLVSYANM
HhCsm6/1-443 RDYAETLIRVKSIAEFITEHYINKHYPKLIQYRKYRGQPSPFFN.....RHYSSKFTEE.YQSFLSKKGWELD.....DRKVLGFVNYFDI
SdCsm6/1-421 GDVAETLIRIKAIAEFIVEKYIDKQYPGLIYYFNEK....PYLN.....LEYDPVFIKE.YISHAGHFLKEKE.....NLLSFPMYRDI
350 360 370 380 390 400 410 420
StCsm6/1-428 LEFYEYDSQLQTAINGILSLNGERNKVAHGLSEIDT.....RLLSRKKLKQLSENLRLLLVDCL.GIDSSYFNYYDKQNKELIKMLE.
StCsm6'/1-386 LKATK.EFKLKRTIAPIIDMNSSRNKVAHSLSPLDS.....DAV..KQLGIAMKTLKTLVREQY.HFSQSDFNFYHDLNKILLTKLN.
EiCsm6/1-430 LQLLEASNEMTNEMNKVNEINNLRNKVAHNLDSLNL.....DRDKNGRKITNAVTAVRTMLLAVFPEVQENDFHYLKQFNQSIKELL..
SaCsm6/1-422 LTVLEPNSKLLKEVNAVNDINGLRNSIAHNLETLDL.....DKNKNYKKIMLSVEAIKNMLHISFPEIEEKDYNYFERKNKEFRELL..
SeCsm6/1-422 LTVLEPNSKLLKEVNAVNDINGLRNSIAHNLETLDL.....DKNKNYKKIMLSVEAIKNMLHISFPEIEEKDYNYFERKNKEFRELL..
LlCsm6/1-409 LKFFEPNHDLLKHIYKIQEINQDRNKVAHSLQAFDR.....KNL..KKVSSAVFASKQILLASF.DIDNHWFSFYEDLNQEIKKLL..
SmCsm6/1-419 LACIEGNSSVSNKLNDVMKINKHRNSVAHKLDTINI.....NNNNQIRKAVKAIEKLIIDIFKEVDIRDFEYVKQANDKIREFI..
McCsm6/1-421 LEILEPESTFHQNVLQILKVNALRNNVAHNLEELKI.....DK.QVNDDIKSAMLALKSMILDTFKDIKDNDLLLIDDINKSLEELI..
HhCsm6/1-443 LKLLEGDGMVFSHLKTIEMINDLRNGIAHRLEELSMIDNGQADDQLEVRDKLSQAVKAVNALLIHVYSNIQSHNLNYFDNKNERFKALLNH
SdCsm6/1-421 LTLLENKGEVLKNVHAVNNINSLRNQVAHNLQSLNL.....YQNKNEKNIDTAVKAVEKLLLIVYPEVNDSTLNYFDQFNERIKEFI..

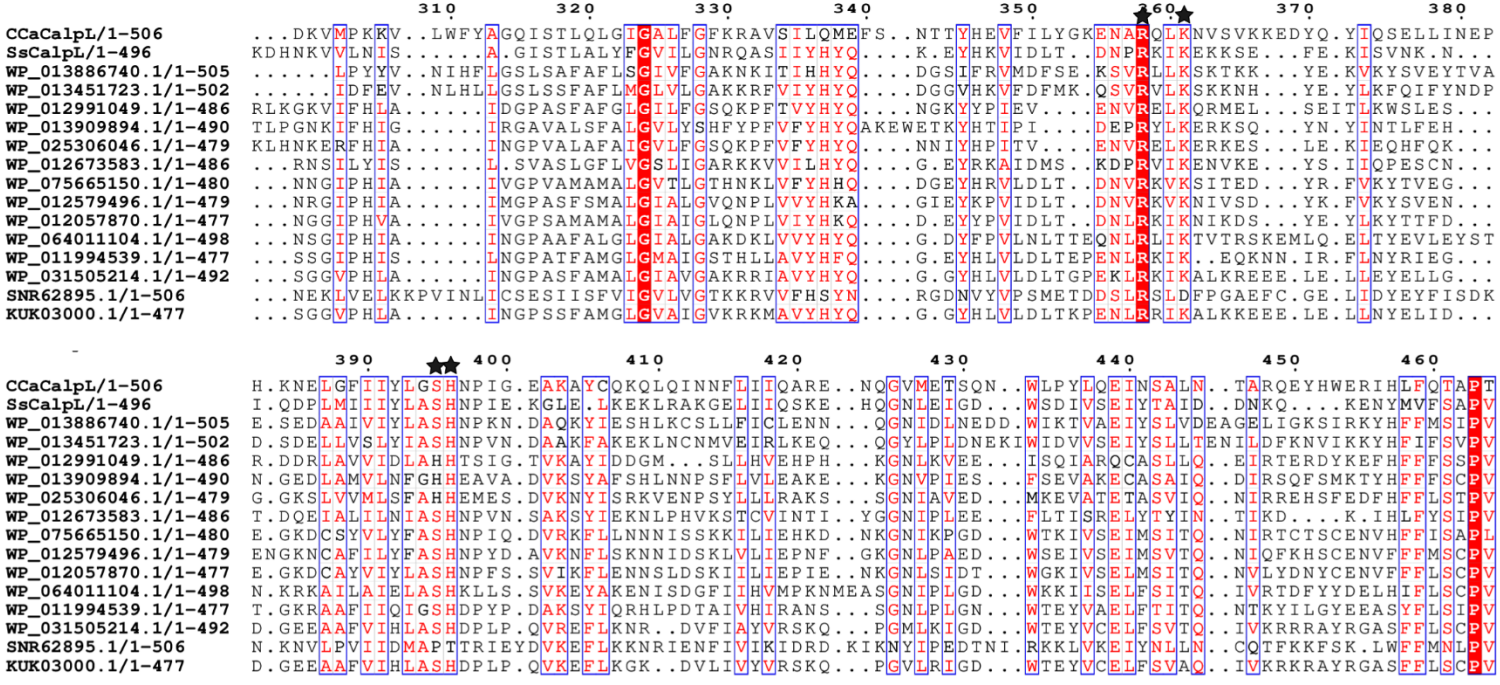
Appendix 9. Superimposition of StCsm6 structural model (Kazlauskiene et al., 2017) **and *T. onnurineus* Csm6-cA$_4$ structure (PDB ID: 6O6V)** (Jia et al., 2019b)**.**

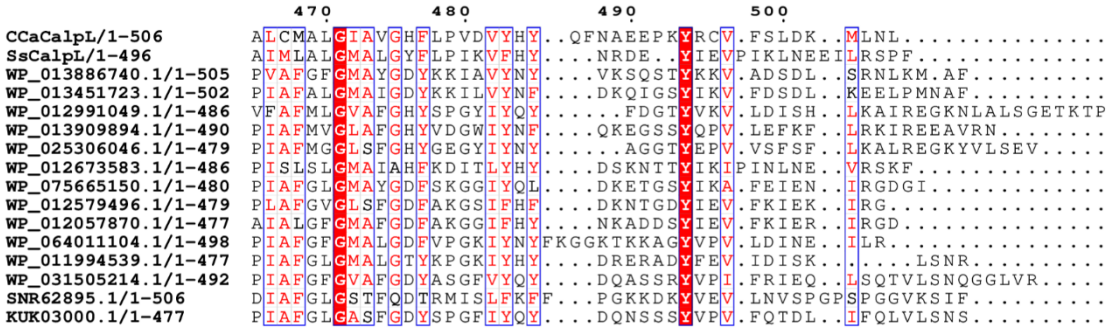


Appendix 10. Multiple sequence alignment of SAVED4-Lon homologues. The alignment is displayed using Espript 3.0. White stars indicate conserved Lon active site residues, black stars mark the active site residues of the SAVED ring nuclease. CCaCalpL – SAVED4-Lon architecture protein CalpL from *Candidatus* cloacimonas acidominovorans (WP_015424587.1), SsCalpL – SAVED4-Lon architecture protein from *Sulfurihydrogenibium* spp. YO3AOP1 (WP_012459369.1)

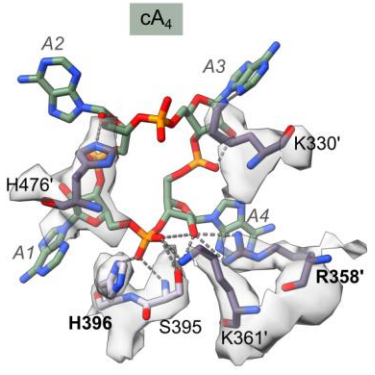



150 160 170 180 190 200 210 220 230
CCaCalpL/1-506 .NIQGNSWQLAVLATLLVKDEKIISKIAYSGIVLPSGEIIT.AEEIEYKKRCCQNLVHR.....IKKIEQLDAWLNTETIPLPVIQYQG
SsCalpL/1-496 .IFSGKSYQLAVAAGLIAKEKEILDNVAFTGEVSSNGFIIP.VNHLEEKKEITEKAKKVLITPEDIENLEELSFWLNPE..HLPVIFIHI
WP_013886740.1/1-505 .YFAGRSFSLAAAAAGLL.KEDKLKYFAFSGEVKENANIAK.VENLPAKRKISEEKDLFFVSPDSVDNLNQLT.KLNAETVDIPFIQLFG
WP_013451723.1/1-502 .FFTGRSFGLALAACFFI.KDEIREKLLFTGEVRADGRVYD.VDGIGLKTEIAKANKRYLVGSTHVNSITELK.YLDRNCIDVPFIQLFG
WP_012991049.1/1-486 .DIKGESFLLPAAVSLYI..ENPPEDAVFTGRVDREGKIYT.VDNISKKRKAAQKEGKRLIDSSGIKSLQELKEWCDAKEHHVPFMVTTK
WP_013909894.1/1-490 .YISGKSFMLPLACALSI..ERIPEDLRFTGALNIKGDVLE.VEHLKEKIEFAKSHGLRLITPLQVKRFNTIKAYLEKDKWDIPFYITTA
WP_025306046.1/1-479 .EFAGASFQLPTVLNLYV..ENLPQDALFIGAIDKKGNIKS.VDGIEEKKKLAKELGLRLVEPYYLSTVDDLKAWFDAESYDVPLYITKT
WP_012673583.1/1-486 .YFTGDSYQFSIVAGLIAKDKNILKNLAFTGKVSSSGKILP.VNHVNEKEKITKANEKNLITPDDISTLEELEFWLNSS..QIPVILLNR
WP_075665150.1/1-480 .HFTGNSFMLAITMGGLT..SKIPKNVAFTGEVDSDGSILKNIRNLHFKESVCEEENMKLISALDVNNVFELKDFFEAKEYHVPVLFVFQ
WP_012579496.1/1-479 .EFKGNSFELAVAIGSLC..KKIPKNIAFTGEIDEKGNIKRNIEYLDLKMKICSENNLKLISAFDVDNLFELKEFFEAKKIHIPILISLM
WP_012057870.1/1-477 .KFVGNSFMLATALAALT..SKIPINLAFSGEIDEEGVIKKNLDGFSIKERVCSENNLKLIGAFDVSNVFELKEFFEEENFHIPVLMFFR
WP_064011104.1/1-498 .GFVGESFELAVAVGRLV..DHIPENLAFTGRVAEDGKILK.VENFDEKLAYCNKNNIGLISYINVSYIREIIDFLNEKEFHIPILLRFS
WP_011994539.1/1-477 .EIVGDSYMLSITIGSLT..KRVPDHIVFTGGILEDGTITS.ADNILEKIEVCRRNGYVLITSDDAKNVVDLQKFFEKKKYHVPVYISFN
WP_031505214.1/1-492 .DFCGDSFMLAVAVGALS..KNLPNHVAFSGCLDEHGNVAL.AESLERKLRVCSEQGLELLTGLDVRNFRDLVDFFDTREHHVPVYLCYK
SNR62895.1/1-506 REFTGTSFMLSVVVAHFF.SQKLFEYFMFTGEVLPSGKINK.VDSILEKKKLAQRYGKKLISPEDVSSIEEVDFWLNESK.PLPVPFSVM
KUK03000.1/1-477 .DLCGDSFMLAVAVGALS..KNLPGHVAFSGSLDKSGNVGL.VENLEKKFQVCSEQGLELLTGLDVQNFRELVDFFNTREHHVPVYLCYK
240 250 260 270 280 290 300
CCaCalpL/1-506 EEN......ELKRWQKA..MEQKVQEK..FSWFSYELLE....DFYGITNSDLAI.FGNGILPFEANAWQKLLQEQVKDKFKLLE.....
SsCalpL/1-496 NKP...EL.ALQSLKQM..EDA.IKKDERFKYFKLENLK....KFYRLEDQDMYL.ITPSVDFSNREELIKILNEFREKVSKLLTLEGVI
WP_013886740.1/1-505 KQK...TE.LEKNLEKISGNEI.VNDY....KIWVGILG.....GDKSLVFTHTEEMLENTTEVWDELLLDFYEKINKLYQ.....
WP_013451723.1/1-502 KSK...NE.LNKNFSKL..KEF.ISDG....RVMVDVLG.....FSDENLSVFCEDFIENNIDTYLNYLRDFYYKVKRVYE.....
WP_012991049.1/1-486 GSEDWLSKWR.DFLSYM..KDS..ENIVRKLEIINGIT.....AEDLVL.YTE..QL.PKGDWTKYMMDFYRKLTDVEK.....
WP_013909894.1/1-490 G.......YE.EFLNFL..KDF.IGEKTFEEFEIIKGLE....LFYGLQEDTFYQ.VTG..QLKTEEDWKKVCQDFYTRYYKIVT.....
WP_025306046.1/1-479 Q.....DRWEGEFKSFL..KATGISKEQLTKLEVLSGLE.....TKPI.ITG..QL.AGDVWKNVLEEFWRRFKETEQ.....
WP_012673583.1/1-486 NTD...NN.IKESLHQI..ETL.IKQD..CPYFTIKNLI....KFYNLSEEDLYI.ITPSIDFSNREELLNILKQFEERLEKLFSV....
WP_075665150.1/1-480 KRD...EEFVDNSYEKL..KEK.VGEK.....FPLNFVD....IFEKLYDFKKTY.VSD..EIKN.NEWKKVLKDARKVLNSIIS.....
WP_012579496.1/1-479 AKD...KEYIRTSYEEL..KEA.VNEK.....FEIKYSD....IFEKIYDVNKVF.VKD..SLN..DDWGQVLKEVNSLLYKIIS.....
WP_012057870.1/1-477 KEE...EEYINFSYEKL..FKS.VSKK.....YPVRFAE....LFEKVYDVKKFF.VTN..ELSSLEKWKKALKEAKAFLLKIIS.....
WP_064011104.1/1-498 SKP...QD.IESLWKRL..KER.V.LA.....FSGKDCESNCNLFERIYRAPLTYSVNR..ELSE.DEFYEHIEKSYGIIRNVID.....
WP_011994539.1/1-477 SS.......KNDYWDML..FEE.VKEK.....FYVDDIK....LFKKVYNPEFKY.ISS..ELRV.EDFQEEIGKVKNILDRIIS.....
WP_031505214.1/1-492 HSE...RV.EDRGWEEL..RET.VERS.....FPSKSLE....LYRKIYRVEMFH.RRH..MLQE.NEYASELRRAFETLQVVFQ.....
SNR62895.1/1-506 IKK...NN...DGLKLL..ESA.IKER..EPLFSLKNLE....VFFEIKREDISI.LHSGFLSENYQEWLAFVHRIREQLRKLFS.....
KUK03000.1/1-477 HSE...RI.EDRGWEEL..KEV.VEKS.....FPTGSLE....LYRRIYRVDMFH.KRY..MLKK.EEFAQELKRAFDTLQRLFQ.....

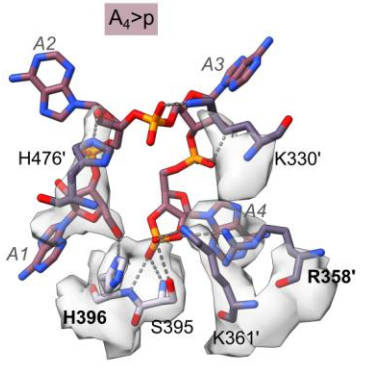

310 320 330 340 350 360 370 380
CCaCalpL/1-506 ...DKVMPKKV..LWFYAGQISTLQLGIGALFGFKRAVSILQMEFS..NTTYHEVFILYGKENARQLKNVSVKKEDYQ.YIQSELLINEP
SsCalpL/1-496 KDHNKVVLNIS.....A.GISTLALYFGVILGNRQASIIYHYQ....K.EYHKVIDLT..DNPRKIKEKKSE...FE.KISVNK.N...
WP_013886740.1/1-505 ......LPYYV..NIHFLGSLSAFAFLSGIVFGAKNKITIHHYQ....DGSIFRVMDFSE.KSVRLLKSKTKK...YE.KVKYSVEYTVA
WP_013451723.1/1-502 ......IDFEV..NLHLLGSLSSFAFLMGLVLGAKKRFVIYHYQ....DGGVHKVFDFMK.QSVRVLKSKKNH...YE.YLKFQIFYNDP
WP_012991049.1/1-486 RLKGKVIFHLA.....IDGPASFAFGLGILFGSQKPFTVYHYQ....NGKYYPIEV....ENVRELKQRMEL...SEITLKWSLES...
WP_013909894.1/1-490 TLPGNKIFHIG.....IRGAVALSFALGVLYSHFYPFVFYHYQAKEWETKYHTIPI....DEPRYLKERKSQ...YN.YINTLFEH...
WP_025306046.1/1-479 KLHNKERFHIA.....INGPVALAFAIGVLFGSQKPFVFYHYQ....NNIYHPITV....ENVRELKERKES...LE.KIEQHFQK...
WP_012673583.1/1-486 ...RNSILYIS.....L.SVASLGFLVGSLIGARKKVVILHYQ....G.EYRKAIDMS..KDPRVIKENVKE...YS.IIQPESCN...
WP_075665150.1/1-480 ...NNGIPHIA.....IVGPVAMAMALGVTLGTHNKLVFYHHQ....DGEYHRVLDLT..DNVRKVKSITED...YR.FVKYTVEG...
WP_012579496.1/1-479 ...NRGIPHIA.....IMGPASFSMALGIALGVQNPLVVYHKA....GIEYKPVIDLT..DNVRKVKNIVSD...YK.FVKYSVEN...
WP_012057870.1/1-477 ...NGGIPHVA.....IVGPSAMAMALGIAIGLQNPLVIYHKQ....D.EYYPVIDLT..DNLRKIKNIKDS...YE.YLKYTTFD...
WP_064011104.1/1-498 ...NSGIPHIA.....INGPAAFALGLGIALGAKDKLVVYHYQ....G.DYFPVLNLTTEQNIRLIKTVTRSKEMLQ.ELTYEVLEYST
WP_011994539.1/1-477 ...SSGIPHIS.....LNGPATFAMGLGMAIGSTHLLAVYHFQ....G.EYHLVLDLTEPENLRKIK..EQKNN.IR.FLNYRIEG...
WP_031505214.1/1-492 ...SGGVPHLA.....INGPASFAMALGIAVGAKRRIAVYHYQ....G.GYHLVLDLTGPEKLRKIKALKREEE.LE.LLEYELLG...
SNR62895.1/1-506 ...NEKLVELKKPVINLICSESIISFVIGVLVGTKKRVVFHSYN....RGDNVYVPSMETDDSLRSLDFPGAEFC.GE.LIDYEYFISDK
KUK03000.1/1-477 ...SGGVPHLA.....INGPSSFAMGLGVAIGVKRKMAVYHYQ....G.GYHLVLDLTKPENLRRIKALKKEEE.LE.LLNYELID...
390 400 410 420 430 440 450 460
CCaCalpL/1-506 H.KNELGFIIYLGSHNPIG.EAKAYCQKQLQINNFLIIQARE..NQGVMETSQN..WLPYLQEINSALN..TARQEYHWERIHLFQTAPT
SsCalpL/1-496 I.QDPLMIIIYLASHNPIE.KGLE.LKEKLRAKGELIIQSKE..HQGNLEIGD...WSDIVSEIYTAID..DNKQ....KENYMVFSAPV
WP_013886740.1/1-505 E.SEDAAIVIYLASHNPKN.DAQKYIESHLKCSLLFICLENN...QGNIDLNEDD.WIKTVAEIYSLVDEAGELIGKSIRKYHFFMSIPV
WP_013451723.1/1-502 D.SDELLVSLYIASHNPVN.DAAKFAKEKLNCNMVEIRLKEQ...QGYLPLDNEKIWIDVVSEIYSLLTENILDFKNVIKKYHFIFSVPV
WP_012991049.1/1-486 R.DDRLAVVIDLAHHTSIG.TVKAYIDDGM...SLLHVEHPH...KGNLKVEE...ISQIARQCASLLQ..EIRTERDYKEFHFFFSSPV
WP_013909894.1/1-490 N.GEDLAMVLNFGHHEAVA.DVKSYAFSHLNNPSFLVLEAKE...KGNVPIES...FSEVAKECASAIQ..DIRSQFSMKTYHFFFSCPV
WP_025306046.1/1-479 G.GKSLVVMLSFAHHEMES.DVKNYISRKVENPSYLLLRAKS...SGNIAVED...MKEVATETASVIQ..NIRREHSFEDFHFFLSTPV
WP_012673583.1/1-486 T.DQEIALILNIASHNPVN.SAKSYIEKNLPHVKSTCVINTI..YGGNIPLEE...FLTISRELYTYIN..TIKD....K.IHLFYSIPV
WP_075665150.1/1-480 E.GKDCSYVLYFASHNPIQ.DVRKFLLNNNISSKKILIEHKD..NKGNIKPGD...WTKIVSEIMSITQ..NIRTCTSCENVHFFISAPL
WP_012579496.1/1-479 ENGKNCAFILYFASHNPYD.AVKNFLSKNNIDSKLVLIEPNF..GKGNLPAED...WSEIVSEIMSVTQ..NIQFKHSCENVFFFMSCPV
WP_012057870.1/1-477 E.GKDCAYVIYLASHNPFS.SVIKFLENNSLDSKIILIEPIE..NKGNLSIDT...WGKIVSELMSITQ..NVLYDNYCENVFFFLSCPV
WP_064011104.1/1-498 N.KRKAILAIELASHKLLS.SVKEYAKENISDGFIIHVMPKNMEASGNIPLGD...WKKIISELFSITQ..IVRTDFYYDELHIFLSCPV
WP_011994539.1/1-477 T.GKRAAFIIQIGSHDPYP.DAKSYIQRHLPDTAIVHIRANS...SGNLPLGN...WTEYVAELFTITQ..NTKYILGYEEASYFLSIPV
WP_031505214.1/1-492 D.GEEAAFVIHLASHDPLP.QVREFLKNR..DVFIAYVRSKQ...PGMLKIGD...WTEYVCELFSVTQ..IVKRRRAYRGASFFLSCPV
SNR62895.1/1-506 N.KNVLPVIIDMAPTTRIEYDVKEFLKKNRIENFIVIKIDRD.KIKNYIPEDTNI.RKKLVKEIYNLLN..CQTFKKFSK.LWFFMNLPV
KUK03000.1/1-477 D.GEEAAFVIHLASHDPLP.QVKEFLKGK..DVLIVYVRSKQ...PGVLRIGD...WTEYVCELFSVAQ..IVKRKRAYRGASFFLSCPV


470
480
490
500
CCaCalpL/1-506 ALCMALGIAVGHFLPVDVYHY..QFNAEEPKYRCV.FSLDK..MLNL.....
SsCalpL/1-496 AIMLALGMALGYFLPIKVFHY....NRDE..YIEVPIKLNEEILRSPF.....
WP_013886740.1/1-505 PVAFGFGMAYGDYKKIAVYNY....VKSQSTYKKV.ADSDL..SRNLKM.AF.....
WP_013451723.1/1-502 PIAFALGMAIGDYKKILVYNF....DKQIGSYIKV.FDSDL..KEELPMNAF.....
WP_012991049.1/1-486 VFAFMLGVAFGHYSPGYIYQY.....FDGTYVKV.LDISH..LKAIREGKNLALSGETKTP
WP_013909894.1/1-490 PIAFMVGLAFGHYVDGWIYNF....QKEGSSYQPV.LEFKF..LRKIREEAVRN.....
WP_025306046.1/1-479 PIAFMGGLSFGHYGEGYIYNY.....AGGTYEPV.VSFSF..LKALREGKYVLSEV.....
WP_012673583.1/1-486 PISLSLGMAIAHFKDITLYHY....DSKNTTYIKIPINLNE..VRSKF.....
WP_075665150.1/1-480 PIAFGLGMAYGDFSKGGIYQL....DKETGSYIKA.FEIEN..IRGDGI.....
WP_012579496.1/1-479 PLAFGVGLSFGDFAKGSIFHF....DKNTGDYIEV.FKIEK..IRG.....
WP_012057870.1/1-477 AIALGFGMAFGDFAKGGIFHY....NKADDSYIEV.FKIER..IRGD.....
WP_064011104.1/1-498 PIAFGFGMALGDFVPGKIYNYFKGGKTKKAGYVPV.LDINE..ILR.....
WP_011994539.1/1-477 PIAFGLGMALGTYKPGKIYHY....DRERADYFEV.IDISK.....LSNR.....
WP_031505214.1/1-492 PIAFGFGVAFGDYASGFVYQY....DQASSRYVPI.FRIEQ..LSQTVLSNQGGLVR.....
SNR62895.1/1-506 DIAFGLGSTFQDTRMISLFKFF...PGKKDKYVEV.LNVSPGPSPGGVKSIF.....
KUK03000.1/1-477 PIAFGLGASFGDYSPGFIYQY....DQNSSSYVPV.FQTDL..IFQLVLSNS.....

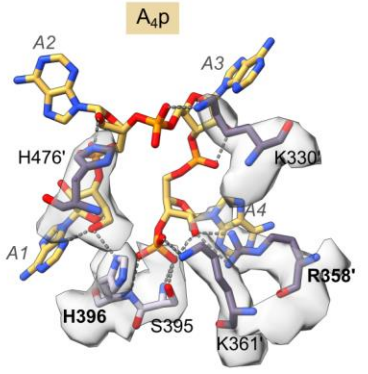
Appendix 11. Potential catalytic residues interacting with cA_4, A_4>p and A_4p with indicated bond lengths.



Bond between		Bond length
Amino acid atom	Nucleotide atom	
K330' Nζ	A3 N3	2.9 Å
S395 Oγ	A4 O1P	2.8 Å
S395 Oγ	A4 O2'	3.0 Å
S395 N	A4 O2P	3.1 Å
H396 Nδ1	A1 O2P	2.9 Å
R358' Nη1	A4 O2'	3.1 Å
R358' Nη2	A4 O2'	3.1 Å
K361' Nζ	A4 O1P	3.2 Å
H476' Nε	A2 O2'	2.7 Å



Bond between		Bond length
Amino acid atom	Nucleotide atom	
K330' Nζ	A3 O2P	2.5 Å
S395 Oγ	A4 O2C	3.2 Å
S395 N	A4 O1C	2.6 Å
H396 Nδ1	A1 O5'	2.7 Å
H396 N	A4 O1C	2.9 Å
R358' Nη1	A4 O2'	3.3 Å
K361' Nζ	A4 O2C	2.8 Å
H476' Nε	A2 O2'	2.5 Å



Bond between		Bond length
Amino acid atom	Nucleotide atom	
K330' Nζ	A3 O2P	2.8 Å
S395 Oγ	A4 O3P	2.3 Å
S395 N	A4 O3P	2.7 Å
H396 Nδ1	A1 O5'	3.0 Å
H396 N	A4 O1P	2.8 Å
R358' Nη1	A4 O2'	2.6 Å
R358' Nη2	A4 O2'	3.3 Å
K361' Nζ	A4 O2P	3.1 Å
H476' Nε	A2 O2'	2.7 Å

LIST OF PUBLICATIONS

This thesis is based on two original publications listed below.

1. Dalia Smalakyte*, Migle Kazlauskiene*, Jesper F. Havelund, Audronė Rukšėnaitė, Auguste Rimaite, Giedre Tamulaitiene, Nils J. Færgeman, Gintautas Tamulaitis, and Virginijus Siksnys. Type III-A CRISPR-associated protein Csm6 degrades cyclic hexa-adenylate activator using both CARF and HEPN domains. *Nucleic Acids Research* 48, no. 16 (September 18; 2020): 9204-9217.
DOI: 10.1093/nar/gkaa634

* – these authors are to be considered joined first authors.

My contributions: participated in experiment planning, performed cloning and protein purification together (with M. K. and A. Rimaite), prepared bacterial cultures for cA_n analysis (with M. K.), optimized cA_n synthesis reactions, prepared RNA and cA_n substrates, carried out cA_n hydrolysis, cA_6 and RNA competition experiments and contributed to preparation of the publication.

2. Dalia Smalakyte, Audrone Ruksenaite, Giedrius Sasnauskas, Giedre Tamulaitiene, and Gintautas Tamulaitis. Filament formation activates protease and ring nuclease activities of CRISPR Lon-SAVED. *Molecular Cell* 84, no. 21 (November 7; 2024): 4239-4255.
DOI: 10.1016/j.molcel.2024.09.002

My contributions: participated in study design and conceptualization, performed homology searches and structure predictions, carried out cloning and mutagenesis, purified and analyzed proteins, designed and performed biochemical assays and assays in E. coli cells, prepared samples for HPLC-MS and cryoEM, contributed to the writing of the manuscript and addressing reviewer comments.

CONFERENCE PRESENTATIONS

Oral presentations:

1. Dalia Smalakytė. III tipo CRISPR-Cas sistemų signalinio kelio reguliacija (Regulation of type III CRISPR-Cas signaling pathway). Bioateitis: gamtos ir gyvybės mokslų perspektyvos. Vilnius, Lithuania (virtual event). 2020 12 04. Award for the best presentation (1st place).

2. Dalia Smalakyte. Self-regulation of the type III CRISPR-Cas signaling pathway. FEBS3+ Conference of Estonian, Latvian and Lithuanian Biochemical Societies. Tallinn, Estonia. 2022 06 15 – 17.

3. Dalia Smalakytė, Audronė Rukšėnaitė, Giedrius Sasnauskas, Giedrė Tamulaitienė ir Gintautas Tamulaitis. Reguliuojama CRISPR-Cas protеazė (Regulated CRISPR-Cas protease). Bioateitis: gamtos ir gyvybės mokslų perspektyvos. Vilnius, Lithuania. 2024 11 21. Award for the best presentation (1st place).

Poster presentations:

1. Dalia Smalakyte, Migle Kazlauskiene, Jesper F Havelund, Audronė Rukšėnaitė, Auguste Rimaite, Giedre Tamulaitiene, Nils J. Færgeman, Gintautas Tamulaitis, Virginijus Siksnys. Streptococcus thermophilus Csm6 degrades its own activator using both CARF and HEPN domains. CRISPR 2021. Paris, France (virtual event) 2021 06 1 – 10.

2. Dalia Smalakyte, Audrone Ruksenaite, Giedrius Sasnauskas, Giedre Tamulaitiene, Gintautas Tamulaitis. Activation and regulation of type III CRISPR-Cas associated signaling cascade. Symposium on the Immune System of Bacteria. Boston, Massachusetts, USA. 2024 04 16 – 18. Presented by G. Tamulaitis.

3. Dalia Smalakyte, Audrone Ruksenaite, Giedrius Sasnauskas, Giedre Tamulaitiene, Gintautas Tamulaitis. Activation and regulation of type III CRISPR-Cas associated signaling cascade. International Conference for Students of Physics and Natural Sciences Open Readings 2024. Vilnius, Lithuania. 2024 04 23 – 26.

4. Dalia Smalakyte, Audrone Ruksenaite, Giedrius Sasnauskas, Giedre Tamulaitiene, Gintautas Tamulaitis. Filament structure activates and regulates CRISPR-Cas ‘protein scissors’. CRISPR 2025. Christchurch, New Zeland. 2025 02 16 – 20. Presented by G. Tamulaitis.

5. Dalia Smalakyte, Audronė Rukšėnaitė, Giedrius Sasnauskas, Giedre Tamulaitiene, Gintautas Tamulaitis. The Art of Destruction: Intrinsic control and sigma factor release in type III CRISPR-Cas tripartite effector. Symposium on the Immune System of Bacteria. Paris, France. 2025 04 08 – 10. Award for the best poster.

Presentations were delivered by the author unless otherwise noted.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Gintautas Tamulaitis, for introducing me to science and guiding me every step of the way. Your mentorship has been invaluable in shaping my scientific journey.

I am also immensely grateful to Prof. Dr. Virginijus Šikšnys for giving me the opportunity to join the lab, for sharing ideas, and for ensuring that our projects moved forward smoothly.

A special thank you goes to Dr. Miglė Kazlauskienė, who has been not only a great mentor but also a beloved friend. Despite being thousands of kilometers away, you have always been there to support and encourage me. I am equally grateful to Dr. Irmantas Mogila, my academic big brother, who not only convinced me to apply to this lab but also provided endless support, brainstorming sessions, and problem-solving discussions. Thank you!

The whole *GT team* also deserves thanks for supporting me and teaching me the value of teamwork. My deepest gratitude to my student, Marija Duchovskytė, for teaching me how to teach and for helping me discover the joy of passing down knowledge.

I want to sincerely thank all the co-authors of the publications that form the foundation of this thesis. The use of *us* throughout this work reflects the truly collaborative nature of our research, and I am deeply grateful for the collective effort that made these discoveries possible. I especially thank Dr. Miglė Kazlauskienė for laying the foundation, initiating, and guiding the StCsm6 research. I also extend my gratitude to Dr. Giedrius Sasnauskas and Dr. Giedrė Tamulaitienė, whose dedication and patience in performing cryo-EM studies were instrumental—they tirelessly prepared countless grids before we finally obtained the structures. Audronė Rukšėnaitė, our MS specialist, thank you for always finding the time to squeeze in our samples; your expertise in HPLC-MS analysis was invaluable. I am also grateful to Augustė Rimaitė for her work in purifying and determining the oligomeric state of the CARF and HEPN domains, as well as to Dr. Jesper F. Havelund and Dr. Nils J. Færgeman for their work on metabolite extraction and HPLC-MS analysis. Additionally, I appreciate Dr. Irmantas Mogila contributions to the TtCsm6 experiments.

I also extend my appreciation to the BNSTS members and staff for fostering such a warm and supportive lab atmosphere—I was lucky to meet so many brilliant people here! My special thanks go to the Lunch/Friday Kebab Club (Dr. Irmantas Mogila, Dr. Edvardas Golovinas, Jonas Juozapaitis, Antanas Vaitkus, Dr. Andrius Merkys, Algirdas Grybauskas, Dr. Giedrius

Sasnauskas, and others) for providing sustenance in both the physical and mental realms. Our discussions—often about the most random topics—were always a highlight. I also want to thank the past and present members of the V337 office (Dr. Arūnas Šilanskas, Dr. Greta Bigelytė, Rugilė Puteikienė) for their immeasurable moral support and patience when I was struggling and complaining. A special mention also goes to Dr. Arūnas Šilanskas, the guru of protein (and beyond) purification, for his patience and insight in answering all my questions. To Mantvyda Marija Juškevičienė and Donata Dakinevičienė, whom I met in the lab but whose friendship has flourished far beyond, thank you for being my safe harbor—always ready to listen, laugh, and share moments of respite.

I would like to thank Dr. Giedrė Tamulaitienė for reading and offering valuable advice on the initial manuscript of this dissertation. I am also grateful to Dr. Miglė Kazlauskienė for careful revision of the Lithuanian summary. Additionally, I sincerely thank the reviewers of this thesis, Dr. Patrick Pausch and Dr. Constantinos Patinios—your insightful comments were instrumental in improving the final version.

Finally, to my friends and family—thank you for your unwavering support, encouragement, love, and patience. And to my partner, Tomas—thank you for always being there for me and for tolerating my occasional over-obsession with work.

ABOUT THE AUTHOR

Dalia Smalakyė (born 1993 in Utena, Lithuania) earned her bachelor's degree with *cum laude* honors (2012-2016) and a master's degree (2016-2018) in biochemistry from Vilnius University. She began her research career in the Department of Protein-DNA Interactions at the Institute of Biotechnology (Vilnius University) during the final year of her undergraduate studies, focusing on CRISPR-Cas defense systems. Through this work, she gained extensive experience in biochemical, microbiological and structural methods.

Her research has been published in well-regarded scientific journals and presented at international conferences. During her doctoral studies, she received a scholarship for academic accomplishments from the Research Council of Lithuania and the Vilnius University Life Sciences Center Nominal Scholarship for excellent learning outcomes. She was also an active community member, serving as a PhD student representative at the Graduate School Council of the Life Sciences Center (2021-2022).

Vilniaus universiteto leidykla
Saulėtekio al. 9, III rūmai, LT-10222 Vilnius
El. p. info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Tiražas 15 egz.