https://doi.org/10.15388/vu.thesis.720
https://orcid.org/0000-0003-0082-3018

VILNIUS UNIVERSITY

Irmantas Mogila

Mechanistic Studies of Type III CRISPR-Cas Activity Regulation and Characterization of Associated Translation-Inhibiting Effectors

DOCTORAL DISSERTATION

Natural sciences,
Biochemistry (N 004)

VILNIUS 2025

The dissertation was prepared between 2018 and 2024 at Vilnius University, Life Sciences Center, Institute of Biotechnology.
The research was supported by Research Council of Lithuania.

This dissertation is defended on an external basis.

Academic consultants:
Dr. Gintautas Tamulaitis (Vilnius University, Natural Sciences, Biochemistry – N 004)
Prof. Dr. Virginijus Šikšnys (Vilnius University, Natural Sciences, Biochemistry – N 004)
Dr. Giedrė Tamulaitienė (Vilnius University, Natural Sciences, Biochemistry – N 004)

Dissertation defense panel:
Chairman – Prof. Dr. Saulius Serva (Vilnius University, Natural Sciences, Biochemistry – N 004).
Members:
Prof. Dr. Vasili Hauryliuk (Lund University, Sweden, Natural Sciences, Biochemistry – N 004),
Dr. Patrick Pausch (Vilnius University, Natural Sciences, Biochemistry – N 004),
Prof. Dr. Edita Sužiedelienė (Vilnius University, Natural Sciences, Biochemistry – N 004),
Assoc. Prof. Dr. Marijonas Tutkus (Vilnius University, Natural Sciences, Biophysics – N 011).

The dissertation shall be defended at a public meeting of the dissertation defence panel at 10:00 on 29th of January, 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.720
https://orcid.org/0000-0003-0082-3018

VILNIAUS UNIVERSITETAS

Irmantas Mogila

III tipo CRISPR-Cas aktyvumo reguliacijos mechanizmo tyrimai ir susijusių transliaciją slopinančių efektorių charakterizavimas

DAKTARO DISERTACIJA

Gamtos mokslai,
Biochemija (N 004)

VILNIUS 2025

Disertacija rengta 2018-2024 metais Vilniaus universitete, Gyvybės mokslų centre, Biotechnologijos institute.
Mokslinius tyrimus rėmė Lietuvos mokslo taryba.

Disertacija ginama eksternu.

Moksliniai konsultantai:
dr. Gintautas Tamulaitis (Vilniaus universitetas, gamtos mokslai, biochemija – N 004)
prof. dr. Virginijus Šikšnys (Vilniaus universitetas, gamtos mokslai, biochemija – N 004)
dr. Giedrė Tamulaitienė (Vilniaus universitetas, gamtos mokslai, biochemija – N 004)

Disertacijos gynimo taryba:
Pirmininkas – prof. dr. Saulius Serva (Vilniaus universitetas, gamtos mokslai, biochemija - N004)
Nariai:
prof. dr. Vasili Hauryliuk (Lundo universitetas, Švedija, gamtos mokslai, biochemija – N004),
dr. Patrick Pausch (Vilniaus universitetas, gamtos mokslai, biochemija – N004),
prof. dr. Edita Sužiedelienė (Vilniaus universitetas, gamtos mokslai, biochemija – N004),
doc. dr. Marijonas Tutkus (Vilniaus universitetas, gamtos mokslai, biofizika – N 011).

Disertacija ginama viešame gynimo tarybos posėdyje 2025 m. sausio mėn. 29 d. 10:00 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

A_2>p/A_4>p	linear di-/tetra-adenosine phosphate with 2′,3′-cyclic phosphate
Acr	anti-CRISPR
AMP	adenosine 5′-monophosphate
Ap	ampicillin
ATP	adenosine 5′-triphosphate
Av	*Allochromatium vinosum*
BLI	bio-layer interferometry
bp	base pair
BSA	bovine serum albumin
Ca	*Caldilinea aerophila*
cA_3/cA_4/cA_6	cyclic tri-/tetra-/hexa-adenosine phosphate
Cami	CRISPR-Cas associated mRNA interferase
CARF	CRISPR associated Rossmann fold
Cas	CRISPR associated
Cb	carbenicillin
CCa	*Candidatus* Cloacimonas acidaminovorans
CFU	colony forming unit
Ch	*Caldicellulosiruptor hydrothermalis*
Cm	chloramphenicol
cOA, cOAs	cyclic oligoadenylate, cyclic oligoadenylates
CRISPR	clustered regularly interspaced short palindromic repeats
Crn1	CRISPR associated ring nuclease 1
Cryo-EM	cryogenic sample electron microscopy
CTD	C-terminal domain
DLU	densitometry light unit
DNase	deoxyribonuclease
dRI	differential refractive index
DTT	dithiothreitol
Ec	*Escherichia coli*
Ei	*Enterococcus italicus*
EMSA	electrophoretic mobility shift assay
ESI	electrospray ionization
FCCS	fluorescence cross-correlation spectroscopy
FCS	fluorescence correlation spectroscopy
FTP	file transfer protocol
HMM	hidden Markov model

HPLC	high-performance liquid chromatography
HTH, wHTH	helix-turn-helix, winged helix-turn-helix
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ipTM	pTM of interfaces
kb	1000 bp
k_{cat}	catalytic constant
K_d	dissociation constant
Kn	kanamycin
MALS	multi-angle light scattering
Mp	*Marinitoga piezophila*
mRNA	messenger RNA
MS	mass spectrometry
MSA	multiple sequence alignment
MW	molecular weight
NCBI	National Center for Biotechnology Information
Nh	*Nitrosococcus halophilus*
NMR	nuclear magnetic resonance
OD_{600}	optical density at a wavelength of 600 nm in 1 cm light path
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PFS	protospacer flanking sequence
pLDDT	predicted local distance difference test (score)
PMSF	phenylmethylsulfonyl fluoride
PNK	polynucleotide kinase
pTM	predicted template modelling score
px	pixel
RIP	ribosome inactivating protein
RMSD	root mean square deviation
RNase	ribonuclease
rRNA	ribosomal RNA
SAM	S-adenosyl-L-methionine
SAVED	SMODS-associated and fused to various effector domains
SD	standard deviation
SDS	sodium dodecyl sulphate
Se	*Staphylococcus epidermidis*
SEC	size exclusion chromatography
Si	*Sulfolobus islandicus*

SMODS	secondary messenger oligonucleotide or dinucleotide synthetase
Ss	*Saccharolobus solfataricus*
ssDNA	single-stranded DNA
Str	streptomycin
TA	toxin-antitoxin
TEV	tobacco etch virus
TM	transmembrane
tmRNA	transfer-messenger RNA
To	*Thermococcus onnurineus*
tRNA	transfer RNA
Ts	*Treponema succinifaciens*
Vc	*Vibrio cholerae*
WT	wild type

INTRODUCTION

The battle between bacteria and the viruses that infect them called bacteriophages or phages is perhaps the longest-running clash between two of the world's largest "armies" in the history of life on Earth. The fierce coevolution has resulted in prokaryotes evolving defence systems to defend themselves against phages that both outnumber them and have higher rates of variability (Mayo-Muñoz et al., 2023).

CRISPR-Cas systems are widespread in the prokaryotic defence arsenal, providing adaptive protection against infecting nucleic acids (Makarova et al., 2020b). These systems insert short fragments of foreign genetic material into the CRISPR region of the genome (Barrangou et al., 2007). The genomic memory of previous infections is utilized by producing individual short CRISPR RNA (crRNA) molecules from CRISPR region transcript. crRNA form effector complexes together with CRISPR-associated (Cas) proteins. During subsequent infections crRNA guides effector complexes to foreign nucleic acids which are specifically degraded by Cas nucleases (Wang et al., 2022). CRISPR-Cas systems are characterised by a wide variety of effector complexes which destroy foreign DNA, RNA or both (Makarova et al., 2020b).

The Csm and Cmr complexes of type III CRISPR-Cas systems are activated upon recognition of foreign transcripts during infection and destroy both the foreign RNA and the DNA that is being transcribed (Tamulaitis et al., 2017). The DNase activity of complex signature protein Cas10 is upregulated upon target RNA binding to the complex. Yet, while the RNA is destroyed within seconds, the DNase activity of Cas10 persists for tens of minutes (Kazlauskiene et al., 2016). These results were explained by a model suggesting that prolonged retention of RNA cleavage products could keep Cas10 in an activated state. Arousing curiosity, this model has yet to be verified.

Recently, it has been shown that activated Cas10 additionally produces cyclic oligoadenylates (cOAs) from ATP which in turn activate ancillary effectors (Kazlauskiene et al., 2017; Niewoehner et al., 2017). This polymerase-cyclase activity seems to be regulated in a similar fashion as the DNase activity (Kazlauskiene et al., 2017). Produced cOAs activate diverse CRISPR ancillary proteins of which the most widespread are dimeric proteins consisting of cOA-binding CARF domain (forming nine major groups CARF1 to CARF9) and an effector domain (Makarova et al., 2020a). Structures and functions of several effector domains have been elucidated: the RNase activity of CARF-fused HEPN domains (Garcia-Doval et al., 2020; Smalakyte et al.,

2020), the DNase and RNase activities of restriction endonuclease-like effectors (McMahon et al., 2020; Rostøl et al., 2021; Zhu et al., 2021) and membrane depolarizing CARF-TM (Baca et al., 2024). However, the molecular mechanisms of most other effector domains of CARF proteins remain to be elucidated and can only be inferred by their sequence similarity to known proteins.

A group of these effector domains shows homology to so-called mRNA interferases of the RelE superfamily (Makarova et al., 2020a). Standalone RelE is a well characterized toxin from type II toxin-antitoxin (TA) systems and it is normally neutralized by antitoxin RelB (Li et al., 2009). However, under nutrient deprivation, in presence of antibiotic agents and especially during phage infection the RelB:RelE ratio may drop (Song and Wood, 2020). Free RelE toxin binds to ribosomal A-site during translation elongation and degrades the mRNA which results in ribosome stalling and translation inhibition (Christensen and Gerdes, 2003; Neubauer et al., 2009).

The presence of RelE domain in the context of dimeric CARF proteins is both unusual, since this toxin acts as a single domain protein in TA systems, and intriguing, because of the signal sensing function of CARF domain which could add a layer of regulation. Based on sequence similarities CARF-RelE proteins are termed here as CRISPR-Cas-associated mRNA interferases 1 (Cami1).

The objects of this doctoral thesis are *Streptococcus thermophilus* DGCC8004 type III-A CRISPR-Cas complex Csm (StCsm) and four Cami1 proteins associated to type III CRISPR-Cas in different hosts. **The goals** are to provide details into StCsm temporal regulation mechanism and to determine Cami1 functions in CRISPR-Cas defence. To achieve these goals, the following **objectives** were undertaken:

1. To measure the retention time of terminal products upon RNA cleavage by StCsm and compare it to the duration of Cas10 DNase activity.
2. To analyse phylogenetic relationships between CARF7 group proteins.
3. To identify Cami1 activator(s).
4. To determine the enzymatic activity of Cami1 effector domain.
5. To demonstrate Cami1 activity in live bacteria.
6. To provide details into Cami1 regulation and activity mechanism by structural approaches.

Scientific novelty and practical value. Csm and Cmr complexes of CRISPR-Cas type III-A and III-B systems, respectively, are well characterized both structurally and functionally with major findings being cross-

verified among different homologs in published studies (Molina et al., 2020; Stella and Marraffini, 2023; van Beljouw et al., 2022). This study in part aimed to supplement the established model of Csm/Cmr effector complex regulation. While it is generally accepted that these complexes are activated by target RNA binding resulting in the upregulation of DNase and polymerase-cyclase activity of Cas10 subunit, its downregulation is less understood. Here, fluorescence correlation spectroscopy (FCS) assays showed that despite rapid RNA shredding, the terminal RNA cleavage products remain bound to the complex for the mean duration of over an hour. Upon activation the DNase activity of Cas10 also persists for a similar duration suggesting a functional association between these two processes. This proposition is supported by previous observations that terminal RNA ends participate in Cas10 regulation by allowing self vs. non-self discrimination in Csm/Cmr complexes (Li et al., 2024). Corroboration of these findings allows to supplement type III CRISPR-Cas regulation model by explaining how the effector complexes are inactivated after infection subsides. When corroborated, these findings allow to explain how the type III CRISPR-Cas effector complexes are inactivated after infection subsides.

CARF protein sequence analysis provided in this thesis exemplifies how the evolutionary trajectories of different prokaryotic defence systems converge to yield new modes of molecular warfare. Phylogenetic sequence analysis shows that a mRNA interferase toxin of TA systems was co-opted by type III CRISPR-Cas signalling system on multiple occasions yielding a polyphyletic group of Cami1 ancillary effectors. Biochemical and *in vivo* toxicity assays showed that Cami1 inhibits the protein synthesis in a strictly regulated fashion. CARF domain fusion to the toxic RelE domain provides the latter with a molecular switch which is activated by cA_4 produced by Csm/Cmr complex when it detects phage RNA in the cell. Ring nuclease activity of CARF domain allows to cleave the activator.

Structural studies of Cami1 proteins revealed possible new functionality of the translation machinery itself. We observed that Cami1 proteins are recruited to the ribosomes via ribosomal stalk protein bL12. The interaction between bL12 and wHTH domain of Cami1 upon its activation proved to be critical for mRNA interferase activity. Normally the stalk proteins interact with the G domain of translation factors and closely related ribosome protecting proteins to increase their local concentration (Liljas and Sanyal, 2018). To our knowledge this is the first observation of a prokaryotic ribosome toxin hijacking the ribosome stalk to facilitate binding to the ribosome.

Taken together, the results outlined here highlight stringent regulation mechanism of prokaryotic defence systems – an important aspect to be accounted for in the emerging era of antibiotic resistance of bacteria and their control using phage treatment in medicine, agriculture and food industries.

The major findings presented for defence in this thesis:

1. *S. thermophilus* Csm complex retains terminal RNA cleavage products and keeps Cas10 DNase activated for extended period of time after rapid RNA cleavage.
2. Cami1 proteins were formed upon CARF and RelE-like domain fusion events during evolution.
3. Cami1 CARF domain binds and cleaves cA_4.
4. Upon activation by cA_4 Cami1 proteins cleave ribosome-bound mRNA.
5. Activated Cami1 cleaves mRNAs during translation in live *Escherichia coli* bacteria.
6. wHTH domain of activated Cami1 proteins binds the C-terminal domain of ribosome stalk protein bL12.

1. LITERATURE OVERVIEW

It is generally accepted that the most abundant forms of life in the biosphere are prokaryotic viruses, also known as bacteriophages or simply phages (Parikka et al., 2017). Phages replicate and spread by infecting prokaryotic hosts, often causing their death (Ofir and Sorek, 2018). Phages outnumber their hosts in most biomes. For example, there are at least ten times more phages than bacteria in the most biomass-rich ocean ecosystems (Parikka et al., 2017).

Both the enormous scale of the struggle between bacteria and phages and the evidence that many cellular life forms are inevitably linked to the viruses that infect them suggest that this kind of coevolution between hosts and their parasites plays a significant role in the history of life on Earth (Koonin, 2016). The fierce coevolution results in a molecular arms race that has added a variety of systems to the prokaryotic arsenal of defence which are well documented (Dy et al., 2014; Labrie et al., 2010). It is noted that genes encoding defence systems are often clustered in genomic regions termed defence islands.

Recent systematic analysis of prokaryotic genes associated with genomic defence islands resulted in a great expansion of putative defence system list with more than hundred new systems (Doron et al., 2018; Gao et al., 2020; Millman et al., 2022; Payne et al., 2024; Vassallo et al., 2022). The antiviral activity for most of these systems is demonstrated *in vivo* by infecting model bacteria expressing candidate defence genes with various phages. Although *in vivo* experiments show varying degrees of protection, the molecular mechanism of defence can only be inferred from sequence or structural model comparisons of candidate systems and known proteins and only if characterized homologs can be found. Therefore, more thorough *in vitro* studies are required to decipher the underlying defence mechanisms. These results are promising, as the natural self vs. non-self recognition of defence systems, especially at the nucleic acid level, gives them great potential for applications in biotechnology and medicine.

1.1. CRISPR-Cas

CRISPR-Cas defence systems are detected in approximately 40% of bacterial and 85% of archaeal genomes (Makarova et al., 2020b). These protection systems are adaptive and preserve the records of past infections in the genomic memory of prokaryotes. This information can be passed on from

generation to generation and used by RNA-guided CRISPR-Cas nucleases against foreign nucleic acids during secondary infections (Fig. 1.1).

The elements of the CRISPR-Cas system are encoded by the CRISPR-Cas locus, which consists of the CRISPR region and the adjacent Cas protein genes. The CRISPR region is characterized by a series of identical direct repeats of approximately 20-50 bp in length, specific to each CRISPR-Cas system, which are separated by unique sequences of similar length termed "spacers" (Mojica et al., 2000; Jansen et al., 2002). A subset of spacers correspond to sequences of plasmids and phage genomes from which they likely originated; these are termed "protospacers" (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). The 5′-end of the CRISPR region contains leader sequence in which the transcription of the CRISPR region is initiated (Pul et al., 2010). The CRISPR transcript is matured to form individual crRNA (CRISPR RNA) molecules. It is the crRNA that determines the specificity of CRISPR-Cas system effectors for foreign nucleic acids (Brouns et al., 2008; Hale et al., 2008).

Cas proteins are involved in all stages of CRISPR-Cas system protection. The CRISPR-Cas system acquires its memory of the invader during adaptation, during which a short fragment (protospacer) of the foreign nucleic acid is taken up and inserted into the CRISPR region (Fig. 1.1A) (Lee and Sashital, 2022). A complex of conserved Cas1 and Cas2 proteins is responsible for the integration of the foreign DNA into the CRISPR region, although other host or Cas proteins may be additionally required in various CRISRP-Cas systems. A new spacer is inserted after the leader sequence as duplication of the first repeat occurs (Nuñez et al., 2015). This acquisition of spacers from previously unencountered attackers is referred to as "naive" spacers acquisition (Yosef et al., 2012). This is a rare event, since elevated activity of the adaptation complex could lead to spacer acquisition from own genome and thus autoimmunity (Wimmer and Beisel, 2019). According to the current model, naive spacer acquisition is assisted by host proteins such as RecBCD complex which processively degrades DNA from an exposed end after double strand break (Levy et al., 2015). Also, other defence systems targeting nucleic acids could be provide DNA fragments to be processed by Cas1-Cas2 (Lee and Sashital, 2022). A more efficient "primed" adaptation is also possible, in which spacers are acquired to the CRISPR region from already known attackers (Amitai and Sorek, 2016; Sternberg et al., 2016). Primed adaptation reduces the likelihood of viruses escaping the CRISPR-Cas system by introducing a protective mutation in the genomic promoter sequence. In some CRISPR-Cas systems, the Cas1 protein is found fused to the reverse transcriptase domain, suggesting

that in such systems the spacer could be inserted from RNA molecules, for example, viral transcripts (Silas et al., 2016).

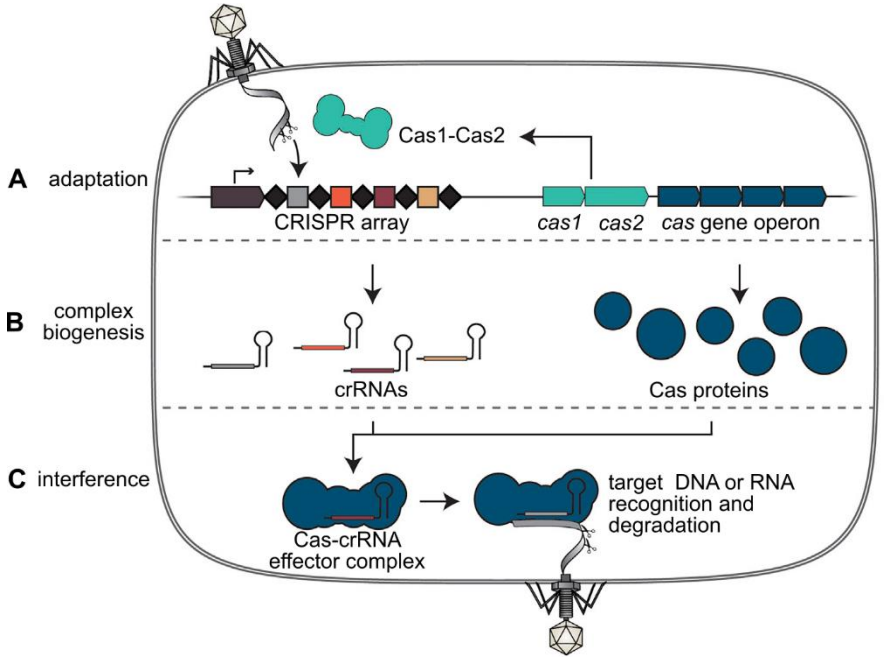


Fig. 1.1. General mechanism of CRISPR-Cas defence. (**A**) In the first phase, adaptation, a spacer is selected and integrated by the Cas1-Cas2 complex into the CRISPR array. (**B**) During complex biogenesis the CRISPR region is being transcribed and processed and matured crRNAs form effector complexes with Cas proteins. (**C**) Upon reinfection Cas effector complexes seek out and destroy foreign DNA or its transcripts in a sequence-specific manner guided by crRNA. Adapted from (Watson et al., 2021).

The CRISPR region is transcribed and processed to individual crRNA molecules, which determine the specificity of the CRISPR-Cas systems (Fig. 1.1B). In general, the CRISPR region, starting from the promoters in AT-rich part of the leader sequence, is transcribed into a long CRISPR transcript (Pul et al., 2010). The repeat sequences of CRISPR transcript are cut by cellular RNase III or Cas proteins such as Cas6 resulting in individual crRNA molecules of ~60-80 nt in length (reviewed in (Charpentier et al., 2015)). Some systems utilize trans-activating crRNAs (tracrRNAs) which are complementary to crRNA repeat sequences and form duplexes prone to cleavage in that region (Deltcheva et al., 2011). Depending on the CRISPR-Cas system, these RNAs function as the final form of crRNAs or they can be further shortened by cellular ribonucleases before being incorporated into effector complexes (Chou-Zheng and Hatoum-Aslan, 2022; Walker et al., 2017).

Effector complexes formed from mature crRNA and Cas proteins recognize the crRNA-complementary nucleic acids and destroy them during secondary infections (Fig 1.1C). These complexes are composed of either a single Cas protein or a multisubunit complex bound to crRNA guide. They are very diversified among different CRISPR-Cas types and subtypes and can act on foreign DNA and/or RNA both directly and via ancillary effectors.

1.2. Classification of CRISPR-Cas

Due to extensive diversification, there are no genes universally shared by all loci of CRISPR-Cas systems. Therefore, their classification is multipronged and based on phylogenies of repeat sequences, Cas locus organization and Cas protein families (Makarova et al., 2020b) (Fig. 1.2). These systems are divided into two classes and seven types including a candidate type VII CRISPR-Cas system (Altae-Tran et al., 2023). Class 1 consists of type I, III, IV and VII systems, which are characterized by defence effector complexes formed of multiple protein subunits. Class 2 includes type II, V and VI systems which utilize effector complexes formed of a single multidomain Cas protein bound to crRNA. Each type is further divided into multiple subtypes based on ancillary effector recruitment, *cas* gene composition and their phylogeny.

Core elements of CRISPR-Cas include a CRISPR region, an adaptation module composed of Cas1 and Cas2 with associated proteins or fused domains, RNases processing CRISPR transcript (such as Cas6 in class 1 systems) and interference module proteins (Fig. 1.2). Yet these elements can be dispensable and thus absent in some of the CRISPR-Cas loci. For example, plasmid-encoded type IV systems often lack the CRISPR region and adaptation module, but their effector complexes can be loaded with crRNA guides from type I loci encoded in host chromosome (Pinilla-Redondo et al., 2020). Great variation among interference module proteins forming effector complexes is also observed. For example, Cas7 family proteins that bind crRNA-target duplex and form the backbone of all class 1 effector complexes (Jackson and Wiedenheft, 2015) are very diverged. While type I and IV complexes show affinity to DNA, type III and VI complexes target RNA. Moreso, Cas7 family protein Csm3 of subtypes III-A (Tamulaitis et al., 2014) and III-B (Staals et al., 2013) systems, respectively, cleave bound RNAs, while Cas7 proteins of I and IV type systems are inactive.

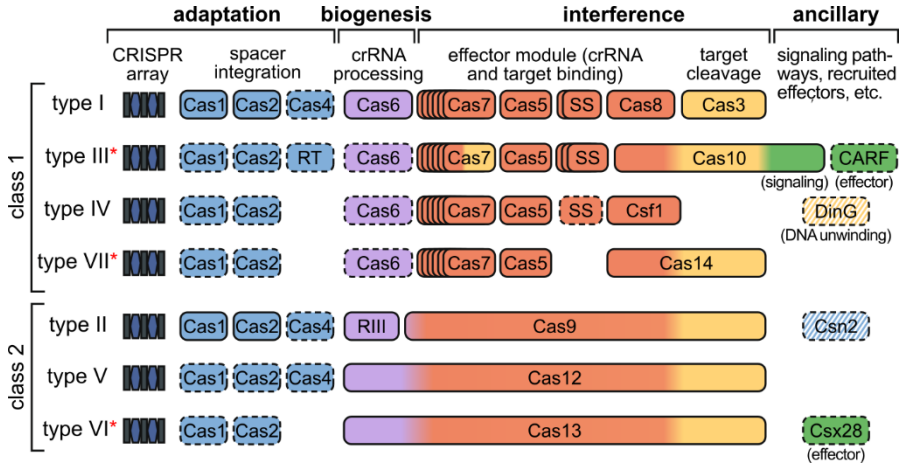


Fig. 1.2. Two classes and seven types of CRISPR-Cas systems. Rounded rectangles of different colours depict proteins or protein domains which participate in different states of CRISPR-Cas biogenesis and defence including the signalling pathway elements unique to type III systems. Names correspond to current nomenclature of protein families. Systems in which effector complexes are targeting RNA are marked with red asterisks. Proteins which are dispensable or absent in some subtypes or variants are marked with a dashed outline. SS – small subunit protein of Cas11 family, RT – reverse transcriptase, RIII – RNase III. Adapted from (Makarova et al., 2020b).

A recent exhaustive survey of rare CRISPR-Cas systems discovered variants in which nucleolytic or other enzymatic domains are fused to interference module proteins marking CRISPR-Cas diversity (Altae-Tran et al., 2023). This study also unveiled a novel type VII CRISPR-Cas system which forms a multisubunit complex from Cas5 and several copies of Cas7 proteins bound to crRNA (Yang et al., 2024). Upon target RNA binding this complex recruits Cas14 which acts as an RNase to cleave the bound RNA (Altae-Tran et al., 2023; Yang et al., 2024).

Adding to the differentiation and versatility of CRISPR-Cas systems various ancillary genes are often associated with loci of certain systems (Shmakov et al., 2018) (Fig. 1.2).

1.3. Type III CRISPR-Cas

Type III CRISPR-Cas include six subtypes III-A to III-F (Fig. 1.3). These systems are present in both bacteria and archaea are thought to represent evolutionary ancestors of other CRISPR-Cas types (Makarova et al., 2020b).

Many type III systems lack CRISPR region and accompanying adaptation module genes (Makarova et al., 2020b). It is demonstrated that type III-B complexes from *Pyrococcus furiosus* (Majumdar et al., 2015) and

Marinomonas mediterranea (Silas et al., 2017a) use crRNAs originating from CRISPR regions adjacent to type I systems co-occurring in host genome. Cas6 protein which cleaves CRIPSR transcript into individual crRNAs is also dispensable in some type III systems (Makarova et al., 2020b). Phylogenetic analysis of type III systems which have CRISPR-*cas1-cas2* module present in the same loci suggest that the association between adaptation and interference modules is mostly sporadic in the evolutionary timescales. The phylogenies of Cas1 and signature type III Cas proteins do not match and mutually homologous Cas1-Cas2 dyads can adopted by different subtypes (Silas et al., 2017b). ~8% of bacterial type III systems encode Cas1 proteins fused to reverse transcriptase enabling them to directly obtain spacers from foreign RNA (Silas et al., 2016). Moreso, some of these proteins are additionally fused to a domain resembling Cas6 as exemplified by *M. mediterranea* Cas6-RT-Cas1 protein which performs in both adaptation and crRNA biogenesis stages (Mohr et al., 2018). RNA recording by RT-Cas1 into minimal CRISPR region was utilized to track fluctuations in gut bacteria transcriptome under different diet regimes or inflation in the host organism (Choi et al., 2022).

Effector complexes of type III CRISPR-Cas are formed of several subunits which belong to Cas10 (large subunit), Cas5, Cas7 and Cas11 (small subunit) protein families (Makarova et al., 2020b) (Fig. 1.3). Type III-E systems are unique since these effector complexes consist of Cas11 fused to four copies of Cas7 proteins forming one large effector protein (Cas11-7 or gRAMP) mimicking class 2 systems (Özcan et al., 2021; van Beljouw et al., 2021). These systems are devoid of Cas10 protein. Conversely, Cas10 is stably associated with other type III systems and forms monophyletic groups for each of the subtype (Wiegand et al., 2023). Sequence analysis of Cas10 homologs show that two enzymatically active domains are common for this protein family: a N-terminal DNase HD domain and polymerase-cyclase Palm domain forming the bilobed core of the protein (Cocozaki et al., 2012; Zhu and Ye, 2012a). Yet some homologs are truncated and devoid of HD (most variants of type III-D, some of III-B) while Palm domain is mostly inactivated in type III-C and III-F subtypes (Fig. 1.3) (Wiegand et al., 2023). A similar protein Cas10d is signature of type I-D systems in which it also acts as the large subunit of the effector complex. Yet, Cas10d is formed of a diverged HD domain fused to inactivated Palm core (Lin et al., 2020; McBride et al., 2020) and is believed to be the evolutionary intermediate between Cas10 and the typical large subunit Cas8 of type I systems (Makarova et al., 2020b). C-terminal domains of type III Cas10 proteins and Cas10d are structurally similar to small subunits Cas11 with *cas10d* gene giving rise to both large and small subunits

utilizing alternative translation start sites (McBride et al., 2020) whereas in type III systems small subunits are encoded separately with the exception of III-E subtype.

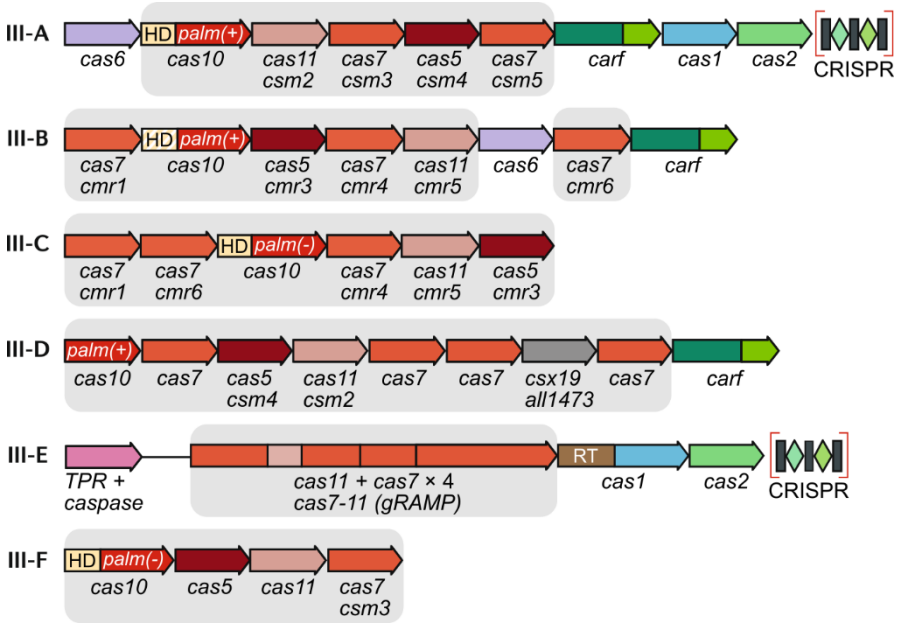


Fig. 1.3. Subtypes of type III CRISPR-Cas. Consensus CRISPR-Cas loci of subtypes III-A to III-F are presented with homologous genes color-coded and identified by a family name (upper) and conventional name (lower). Most III-B, III-C, III-D and III-F systems lack adjacent CRISPR region and adaptation module which are not shown. Genes encoding proteins of the interference complex are indicated with background shade. The most common variants of *cas10* gene in different subtypes are depicted by indicating the presence or absence of ssDNase HD domain and the putative activity of polymerase-cyclase Palm domain with "(+)" and "(-)" indicating intact or mutated active site, respectively, as presented in (Wiegand et al., 2023). RT – reverse transcriptase, TPR – tetratricopeptide repeat. The figure is adapted from (Makarova et al., 2020b).

Type III-A, III-B and III-D systems on average have more than two ancillary genes present it their loci with most of them related to CARF (CRISPR-associated Rossmann fold) superfamily (Makarova et al., 2020a, 2020b). These proteins are not included in the effector complex. Instead, CARF and related SAVED domains are often fused to various toxic domains such as RNases, DNases, transmembrane helices, and others (Makarova et al., 2020a). During infection these proteins provide a secondary line of defence when activated via specific signalling pathway which will be covered in detail further

(Kazlauskiene et al., 2017; Makarova et al., 2020a; Niewoehner et al., 2017; Stella and Marraffini, 2023).

1.3.1. Biogenesis and structure of effector complexes

Effector complex biogenesis and structure is best studied for type III-A and III-B CRISPR-Cas systems which form effector complexes called Csm and Cmr, respectively. These complexes are formed of several protein subunits bound to a crRNA molecule. While most type III-A systems encode crRNAs in the CRISPR regions adjacent to *cas* gene locus, this association is exceptional in type III-B loci (Makarova et al., 2020b).

Type III crRNA undergoes a two-step maturation process (Fig. 1.4A). First the CRISPR transcript is cleaved by Cas6 proteins at the stem loop secondary structures formed due to partial self-complementarity of repeat sequences (Carte et al., 2008; Hatoum-Aslan et al., 2014). Cas6 cleavage yields ~70 nt individual crRNAs with remnants of repeat sequences flanking the spacer sequence. The 3′-repeat flank is further trimmed by host 3′-5′ exonucleases (Hale et al., 2008; Hatoum-Aslan et al., 2011). In *Staphylococcus epidermidis* III-A system crRNA is trimmed by endogenous PNPase (Walker et al., 2017) and RNase R (Chou-Zheng and Hatoum-Aslan, 2022) which are associated with RNA degradosomes regulating mRNA decay (Tejada-Arranz et al., 2020). crRNA also undergoes trimming when type III-A effector complexes from *S. epidermidis*, *Thermococcus onnurineus*, *Lactococcus lactis*, *Streptococcus thermophilus* or *Mycobacterium tuberculosis* are expressed in heterologous *Escherichia coli* host (Hatoum-Aslan et al., 2013; Jia et al., 2019b; Sridhara et al., 2022; Tamulaitis et al., 2014; Zhang et al., 2024). A recent report presents another candidate maturation RNase of NYN family which putatively trims crRNA in *Bacteroides fragilis* (Chi and White, 2024). The matured crRNA of type III effector complexes is ~40 nt in size with 8 nt of 5′-tag part of repeat origin and the guiding part originating from variable spacers.

It is thought that crRNA undergoes 3′-end trimming during the assembly of effector complex (Chou-Zheng and Hatoum-Aslan, 2022; Mogila et al., 2019) (Fig. 1.4A). While not investigated directly, it is supported by several observations. In *S. epidermidis* type III-A Csm complex (SeCsm) final crRNA length is determined by the number of backbone Csm3 subunits each covering 6 nt of spacer sequence (Hatoum-Aslan et al., 2013). Moreso, the genetic deletion of Csm5 subunit, which is normally present in WT complexes, disrupts crRNA maturation entirely yielding complexes with intermediate 71 nt

crRNA (Hatoum-Aslan et al., 2014). Similar results were also observed for *S. thermophilus* Csm (StCsm) complex in which Csm5 omission yields mutant complexes with Cas6 bound to intermediate 72 nt crRNA (Mogila et al., 2019). Moreso, the Csm5 subunit is shown to be critical for PNPase and RNase R recruitment for crRNA 3′-trimming in SeCsm (Chou-Zheng and Hatoum-Aslan, 2022; Walker et al., 2017).

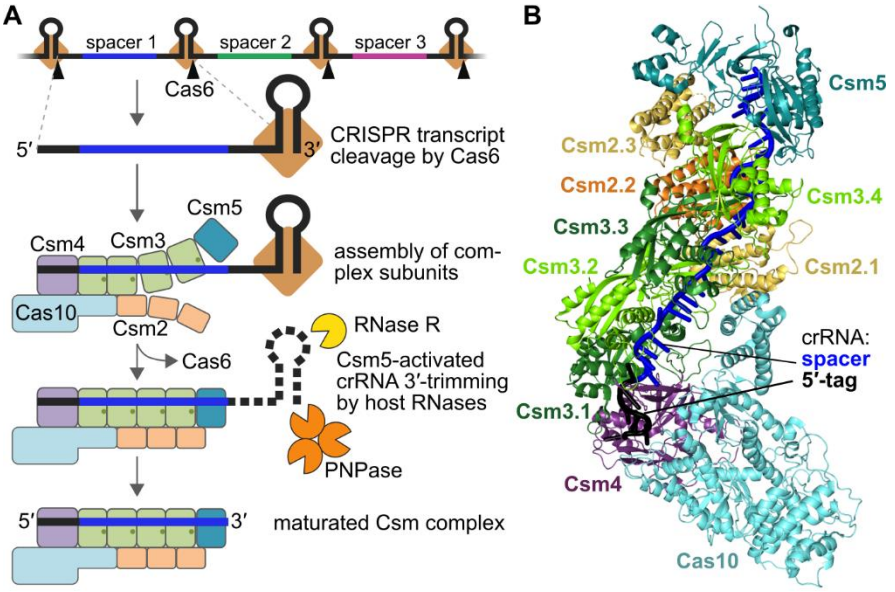


Fig. 1.4. Biogenesis and structure of type III CRISPR-Cas effector complex. (**A**) Scheme showing biogenesis of crRNA and Csm complex assembly. CRISPR transcript is cleaved by Cas6 which is retained at short hairpin structures formed by palindromic repeat sequences. Csm subunit assembly and 3′-trimming by non-Cas RNases occurs simultaneously yielding maturated complex. Figure was adapted from (Chou-Zheng and Hatoum-Aslan, 2022; Mogila et al., 2019). (**B**) Cryo-EM structure of *L. lactis* Csm (LlCsm) complex (PDB: 6XN3).

The III-A and III-B system-specific effector complexes consist of proteins homologous between these subtypes, Cas10 (or Csm1/Cmr2 for type III-A/III-B), Csm2/Cmr5 of Cas11 family, Csm3/Cmr4 of Cas7 family, Csm4/Cmr3 of Cas5 family, Csm5/Cmr1-Cmr6 of Cas7 family, respectively, and the crRNAs bound to these proteins (Tamulaitis et al., 2017). Although the ribonuclease Cas6 is involved in the biogenesis of crRNAs in these complexes, this protein is not found in WT Csm and Cmr complexes.

Both Csm and Cmr complexes are well characterized by structural studies (Guo et al., 2019; Huo et al., 2018; Jia et al., 2019b; Osawa et al., 2015; Sofos et al., 2020; Sridhara et al., 2022; Taylor et al., 2015; You et al., 2019;

Zhang et al., 2024) (Fig. 1.4B). The 8 nt crRNA 5′-tag is bound by the Csm4 (only type III-A subunits are listed onwards) protein in the Csm complexes. Csm4 interacts with the large subunit Cas10 with crRNA 5′-tag located near this interaction surface. The crRNA spacer sequence is bound by a filament consisting of several copies of Csm3 subunits, covered by a parallel filament formed by Csm2 small subunits in one-to-one fashion. Cas10-promixal Csm3 subunit is covered by C-terminal domain of Cas10. Each Csm3 subunit binds six nucleotides of crRNA spacer sequence. The Cas10-distal 3′-end of the crRNA is bound by the Csm5 subunit in Csm complexes. The Cas10-distal part of Cmr complexes is different and formed by a heterodimer of Cmr1 and Cmr6 proteins bound to the 3′-end of the crRNA (Sofos et al., 2020).

It is observed that the number of Csm3/Cmr4 subunits differ in Csm/Cmr complexes of different hosts and could range from two in *Thermococcus onnurineus* Csm (Jia et al., 2019b) to six in *Mycobacter tuberculosis* Csm (Zhang et al., 2024). The length of bound crRNA varies accordingly with 27 nt in ToCsm and 46 nt in MtCsm structures. This variation can be explained by dynamic crRNA maturation mechanism involving both complex assembly and 3′-end trimming (Fig. 1.4A).

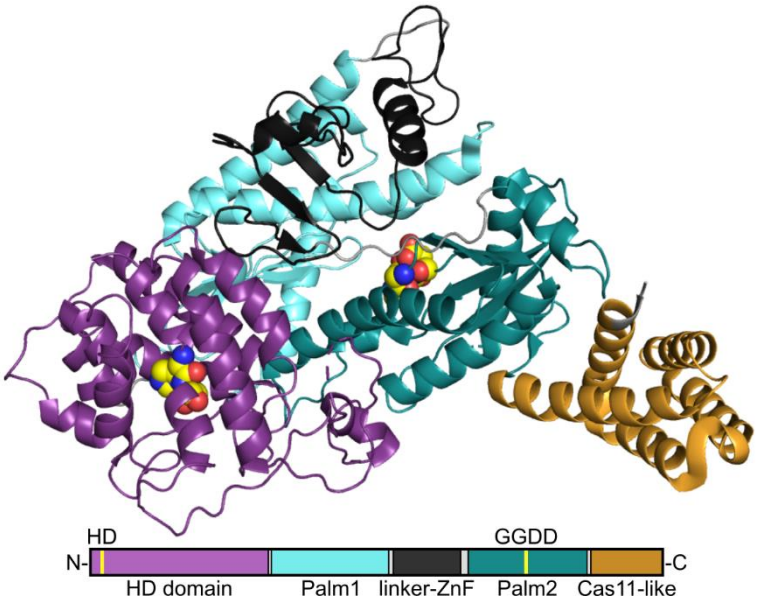


Fig. 1.5. Cas10 structure. Cryo-EM Cas10 protein structure of LlCsm complex (PDB: 6XN3). Cas10 protein domains are coloured according to the protein chain scheme depicted below. Conserved HD and GGDD residues are shown as spheres. ZnF – zinc finger structure.

Structural studies of the large subunit Cas10 have revealed that it is composed of five distinct domains: N-terminal HD, polymerase-like Palm1 and Palm2 with zinc-finger insertion and the C-terminal domain adopting Cas11 small subunit fold (Cocozaki et al., 2012; Jung et al., 2015; Venclovas, 2016) (Fig. 1.5). The conserved Palm domains form a pseudosymmetric dimer and feature a β-loop-β structure and GGDD motif conserved in non-Cas nucleotide cyclases (Cocozaki et al., 2012; Jung et al., 2015; Zhu and Ye, 2012b). C-terminal Cas11-like domain is the first member of Csm2/Cmr5 filament in Csm/Cmr complexes (Osawa et al., 2015; You et al., 2019).

1.3.2.Defence mechanism

Type III-A and III-B effector complexes stand out from other CRISPR-Cas effectors because they destroy both the DNA and RNA during infection (Tamulaitis et al., 2017; van Beljouw et al., 2022). (Fig. 1.6A). Both Csm and Cmr complexes specifically bind to the ssRNA target which is complementary to the spacer sequence of crRNA. In Csm complexes the Csm5 subunit binds RNA molecules in non-specific fashion (Mogila et al., 2019) while in Cmr complexes the Cmr6 and Cmr1 subunits are critical for RNA binding (Hale et al., 2014). These interactions allow type III complex to scan potential targets for complementarity with the seed region located at the Cas10-distal part of crRNA (Steens et al., 2021; You et al., 2019).

Upon target RNA binding HD and Palm domains of Cas10 subunit are activated (Fig. 1.6A). HD domain has a DNase activity specific for single-stranded DNA (Elmore et al., 2016; Estrella et al., 2016; Kazlauskiene et al., 2016; T. Y. Liu et al., 2017; Samai et al., 2015). According to the current model, during the foreign gene expression Cas10 subunit degrades the DNA strand displaced by RNA polymerase with the RNA transcript targeted by the Csm/Cmr complex acting as a tether (Samai et al., 2015; Tamulaitis et al., 2017). Alternative ssDNA targets could be single-stranded intermediates of DNA replication in some plasmids and phages or R-loops (Liu et al., 2019). Palm domains possess polymerase-cyclase activity and produce cyclic oligoadenylates (cOA) from ATP when triggered (Kazlauskiene et al., 2017; Niewoehner et al., 2017). These unique signalling molecules in turn activate associated CARF and SAVED family proteins causing infected cell dormancy or death as overviewed further in Chapter 1.4. Reinforcement by cOA signalling pathway is critical when the production of target RNA is low or delayed during infection (Jiang et al., 2016; Rostøl and Marraffini, 2019).

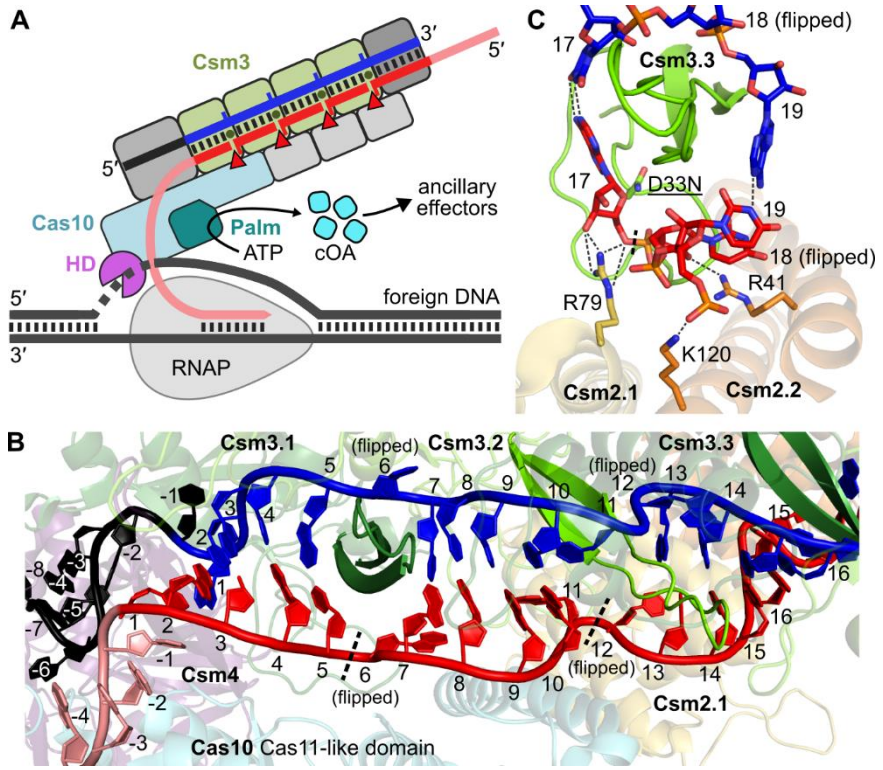


Fig. 1.6. Defensive mechanism of type III CRISPR-Cas. (**A**) Csm complex activation upon binding crRNA-complementary foreign transcript. HD domain of Cas10 cleaves DNA strand displaced by RNA polymerase while Palm domain converts ATP to cOA for ancillary effector activation. RNA target is cleaved by Csm3 proteins in 6 nt increments (red arrowheads). Figure was prepared according to (van Beljouw et al., 2022). RNAP – RNA polymerase. (**B**) Fragment of crRNA:target duplex in LlCsm complex (PDB: 6XN3). crRNA (5′-tag – black, spacer – blue) and substrate (salmon and red) nucleotides are numerated with positive indices corresponding to variable spacer and its cognate target sequence. Base pairs flipped by Csm3 "thumb" structure are indicated. Scissile phosphodiester bonds are marked with dashed lines. (**C**) Composite active centre of StCsm formed by Csm3 and Csm2 proteins (PDB: 6IG0). Mutated D33N active site residue of Csm3 is indicated and underlined. Scissile P-5′O bond between 17th and 18th nucleotides in the target is marked with a dashed line.

The RNA is cleaved at several sites 6 nt apart by Csm3/Cmr4 subunits (Staals et al., 2014, 2013; Tamulaitis et al., 2014) (Fig. 1.6A). The crRNA:target hybrid is distorted and unwound by the Csm3/Cmr4 "thumb" structure at every sixth nucleotide (Fig. 1.6B) placing the scissile phosphodiester bond near the composite RNase active centre formed by Csm3/Cmr4 and Csm2/Cmr5 (Guo et al., 2019; Jia et al., 2019b; Mogila et al., 2019; Osawa et al., 2015; Sofos et al., 2020; You et al., 2019; Zhang et al., 2024) (Fig 1.6C).

Cmr4 of PfCmr complex hydrolyses P-5′O bond leaving products with 5′-OH and 3′-phosphate or 2′3′-cyclic phosphate ends (Hale et al., 2009). Divalent metal cofactors are required for all three (DNase, polymerase-cyclase, RNase) activities. Upon RNA cleavage and dissociation Cas10 returns to deactivated state (Kazlauskiene et al., 2016, 2017).

Type III complexes have been observed to tolerate multi-nucleotide mismatches between target and crRNA allowing to suppress immunity escape due to mutations in the target sequence (Maniv et al., 2016; Pyenson et al., 2017). Yet the Cas10-proximal part of the target is found to be more susceptible and single nucleotide mutations in this region do not permit Cas10 activation (Guo et al., 2019; Steens et al., 2021).

In addition to Csm3 and Cas10, housekeeping RNase J2, RNase R and PNPase are shown to assist *S. epidermidis* type III-A system in destroying foreign nucleic acids (Chou-Zheng and Hatoum-Aslan, 2019, 2022).

1.3.3. Regulation of effector complexes

It has been observed that in some CRISPR-Cas systems, antisense transcription of the CRISPR region occurs to form crRNA complementary transcripts (Lillestøl et al., 2006). Yet, type III CRISPR-Cas does not cause self-immunity when antisense transcripts are produced *in vivo* in *S. epidermidis* (Marraffini and Sontheimer, 2010). Later studies revealed that Cas10 subunits of type III-A and III-B effectors are inhibited when the targeted transcript is complementary to the 8 nt 5′-tag sequence of the crRNA despite a perfect match with the spacer part (Elmore et al., 2016; Kazlauskiene et al., 2016; Li et al., 2024; T. Y. Liu et al., 2017; Wang et al., 2018) (Fig. 1.7A). 5′-tag of crRNA is derived from the repeat sequence specific to each CRISPR region (Section 1.3.1). This means that upstream transcripts of the CRISPR region are fully complementary to the crRNA 5′-tag sequence and thus inhibit Cas10 activity while viral transcripts that are not complementary to 5′-tag activate Cas10 (Tamulaitis et al., 2017). Nevertheless, antisense transcripts are readily degraded by effector complex RNases (Estrella et al., 2016; Kazlauskiene et al., 2016). The 3′-part of targeted RNA which is adjacent to the sequence targeted by the spacer and for which the complementarity with 5′-tag is checked in the complex is termed protospacer flanking sequence (PFS).

Structural studies of type III-A StCsm complexes bound to either cognate (PFS not matching 5′-tag) or non-cognate (PFS matching 5′-tag) substrates allows to compare activated and repressed states and provide insights into how activating signal is transferred to Cas10 upon RNA binding (You et al., 2019).

Between Csm4 and Cas10 proteins a four nucleotide (-5 to -2) stretch of 5′-tag is exposed to the surface (Fig. 1.7B). If the respective part of PFS is complementary and hybridizes to this 4 nt stretch, the Cas10 is not activated while the RNA is degraded. Differently, when PFS of the substrate is not complementary to 5′-tag it binds the positively charged groove between linker-ZnF and Palm1 parts of Cas10 (Fig. 1.7C, see also Fig. 1.5). PFS-Cas10 interaction drives subtle conformation changes and activation of HD and Palm domains while the mutations in this Cas10 groove disrupt both activities (You et al., 2019). Such activation mode of Csm complex is also supported by single molecule fluorescence studies which show that 3′-flanking part of a cognate substrate is arranged in the vicinity of Cas10 Palm domain while non-cognate substrate does not interact with it (Wang et al., 2018).

Structural studies investigating the regulation of type III-B Cmr complexes also show that Cas10 activity depends on PFS complementarity to 5′-tag of crRNA (Sofos et al., 2020). Differently from StCsm complex, in which the binding of cognate or non-cognate substrates causes distinct conformational rearrangements of regulatory linker-ZnF motif of Cas10 protein (You et al., 2019), the regulatory function in *S. islandicus* Cmr-β complex is attributed to an extended loop of Cmr3 protein (which is a homolog of Csm4) named the stalk loop (Sofos et al., 2020). Stalk loop interacts with the ZnF and Palm2 domains of Cas10 protein and is located between 5′-tag of crRNA and PFS of an RNA substrate. It adopts a rigid "retracted" conformation upon complex binding to an inhibitory non-cognate substrate and becomes flexible when bound to a cognate substrate. This causes subsequent conformational changes in HD and Palm domains permitting ssDNase and polymerase-cyclase activities (Sofos et al., 2020). ATP and ssDNA binding can stabilize the stalk loop in the "extended" conformation (Sofos et al., 2020). The PFS part of the substrate is not resolved in *S. islandicus* Cmr-β structural studies suggesting that it is more flexible compared to StCsm as Cmr-β does not have a positively charged groove between ZnF and Palm1 domains (Sofos et al., 2020).

As overviewed in section 1.3.2, type III effector activity is temporally controlled by RNA substrate degradation by Csm3/Cmr4. However, *in vitro* biochemical studies of StCsm show discrepancy between RNA hydrolysis rate and the duration of Cas10 DNase and polymerase-cyclase activity: while cognate RNA is degraded in seconds, the Cas10 remains activated for tens of minutes (Kazlauskiene et al., 2017, 2016). It is speculated that Cas10 remains in an activated state due to the slow dissociation of RNA cleavage products (Kazlauskiene et al., 2016). If this is the case, it does not hinder the rapid

RNase turnover by StCsm when RNA is provided in excess (Kazlauskiene et al., 2016). To corroborate these observations, the temporal control mechanism of the type III CRISPR-Cas defence needs to be addressed directly.

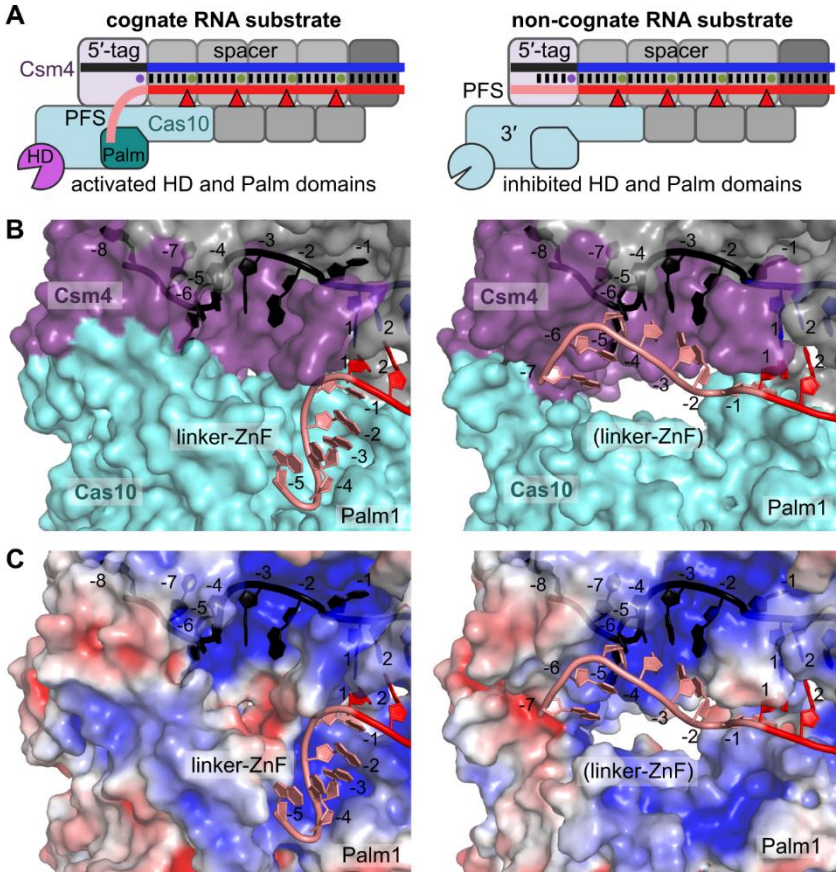


Fig. 1.7. Self vs. non-self RNA discrimination in Csm complex. (**A**) Scheme depicting Csm complex regulation by either cognate or non-cognate substrate binding. PFS (salmon) of cognate substrate is not complementary to 5′-tag (black) of crRNA and such substrate binding activates Csm whereas PFS of non-cognate substrate, such as antisense CRISPR transcript, hybridization with 5′-tag inhibits Csm. RNA is cleaved by Csm3 in both cases. (**B**) Structures of StCsm complexes bound to cognate (left; PDB: 6IG0) or non-cognate (right; PDB: 6IFL) RNA substrates. crRNA and RNA substrates are coloured as in (A), and nucleotides are numbered as in Fig. 1.6B. Csm4 and Cas10 proteins with Cas10 domains are indicated; linker-ZnF structure is not fully resolved in non-cognate complex structure suggesting flexibility. (**C**) Structures of (B) with protein surfaces coloured according to electrostatic potential (positive – blue, negative – red).

1.3.4. Anti-CRISPR

Anti-CRISPR (acr) proteins are natural inhibitors of CRISPR-Cas systems. Found present in phage genomes, these small proteins disrupt CRISPR-Cas effector complex formation, target binding or cleavage often by mimicking nucleic acids (Davidson et al., 2020).

While most research has been done on acrs which suppress DNA-targeting CRISPR-Cas, there are known three families of acrs which interfere with type III CRISPR-Cas defence. AcrIIIB1 and AcrIIIB2 proteins were found in the genomes of phages infecting archaea of *Sulfolobacea* family and during infection they bind to and deactivate type III-B Cmr complex allowing phage to propagate (Bhoobalan-Chitty et al., 2019; Lin et al., 2023; Liu et al., 2023). AcrIIIB1 binds to Cmr complex and disrupts cOA production and subsequent CARF protein activation (Bhoobalan-Chitty et al., 2019). AcrIIIB2 binds Cmr4 protein in Cmr complex and traps RNA cleavage products thus blocking their dissociation and complex reactivation under excess of target transcripts (Lin et al., 2023; Liu et al., 2023).

While no acrs directly targeting type III-A systems are described, a universal inhibitor of cOA-producing type III systems AcrIII-1 is characterized to cleave cOA signalling molecules (Athukoralage et al., 2020). cOA signalling pathway is important when type III CRISPR-Cas targets late or lowly expressed phage genes (Jiang et al., 2016; Rostøl and Marraffini, 2019) thus efficient ring nuclease activity of AcrIII-1 helps phages to circumvent the defence in such cases.

1.4. Ancillary CARF effectors

As presented previously, effector complexes of three CRISPR-Cas subtypes III-A, III-B and III-D utilize Cas10 proteins with active Palm domains (Fig. 1.3). When a complex of this group binds to a foreign transcript during infection, its Palm domain starts to produce cOA signalling molecules from ATP (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Cas10 proteins produce either three-, four- or six-membered cOAs (cA_3, cA_4 or cA_6) in which AMP moieties are connected by canonical 3′-5′ phosphodiester bonds (Grüschow et al., 2021; Kazlauskiene et al., 2017; Rouillon et al., 2018). The size of produced cOA depends on the number of adenine base binding pockets conserved in Palm2 domain (Goswami et al., 2024).

cOAs activate ancillary effector proteins which mostly come from CARF or related SAVED families (Makarova et al., 2014, 2020a). These proteins are

often encoded in the vicinity of type III CRISPR-Cas loci. CARF and SAVED domains in these proteins mostly are fused to other effector domains such as RNases, DNases, proteases, transmembrane helices and others (Makarova et al., 2020a). Ancillary proteins are expressed in inactive form and upon specific cOA binding to CARF or SAVED domain the fused effector is activated. Due to the nature of effector domains, they are toxic to the host cell. For example, non-specific RNases and DNases degrade, respectively, transcriptome and genome of the cell (Jia et al., 2019a; Rostøl et al., 2021; Smalakyte et al., 2020) (Fig. 1.8) while transmembrane segments cause membrane puncturing (Baca et al., 2024; Tak et al., 2023). Such effects cause dormancy or even death of the infected cell reducing viral spread in the population (Stella and Marraffini, 2023). The notion that cOA signalling provides immunity at population level is supported by several observations of CRISPR-Cas systems themselves (including type III) being upregulated in high cell density conditions via well-known quorum sensing pathways (Høyland-Kroghsbo et al., 2017; Patterson et al., 2016).

Cyclic oligonucleotides are also utilized by a different defence system CBASS (cyclic-oligonucleotide-based antiphage signalling system). Two core proteins of CBASS are a CD-NTase (cGAS/DncV-like nucleotidyltransferase) which senses infection and produces cyclic oligonucleotides and an effector protein causing infected cell dormancy or death upon signal binding (Millman et al., 2020). CBASS CD-NTases synthesize diverse cyclic di- and trinucleotides such as cyclic cGAMP, cyclic UMP-AMP, cyclic UMP-UMP and cyclic AMP-AMP-GMP (Whiteley et al., 2019). In 30% of CBASS systems these signalling molecules are sensed by CBASS-specific SAVED proteins fused to toxic effectors (Burroughs et al., 2015).

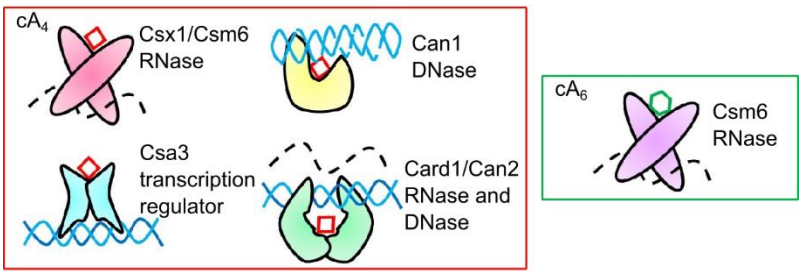


Fig. 1.8. Examples of CARF nucleases regulated by different cOAs. CARF proteins are often found fused to effector domains with RNase and DNase activities or DNA binding HTH domains. Different CARF domains correspond to cOAs of different ring size with cA_4 and cA_6 being the most common activators. Adapted from (Athukoralage and White, 2021). See also Table 1.1.

1.4.1.Classification of CARF effectors

CARF domains adopt a common tertiary structure termed Rossmann fold composed of a beta sheet of parallel strands sandwiched by alpha helices (Makarova et al., 2014). CARF domains form symmetrical dimers whereas the related SAVED domains are monomeric with double CARF-like structures in a single chain that arranges in pseudosymmetric fashion (Lowey et al., 2020; Makarova et al., 2020a). CARF proteins are found in bacteria and archaea while SAVED proteins are mostly bacterial (Makarova et al., 2020a).

A recent classification of CARF (and SAVED) proteins placed them in different groups (10 major and 13 minor of CARF proteins) based on full length sequence alignments (Makarova et al., 2020a) (Table 1.1). Despite such an approach CARF proteins with different fused effectors were often found in the same group while for some effectors, such as HEPN RNase, CARF-HEPN fusions were spread over several groups. Such distribution suggests a dynamic evolution of CARF proteins diversified by multiple fusions with effector domains (Makarova et al., 2020a).

A diverged CARF domain is also present in a widespread transcription regulator RtcR which presumably responds to tRNA fragments and controls the expression of RNA repair genes unrelated to CRISPR-Cas (Chen and Wolin, 2023; Kotta-Loizou et al., 2022).

Beyond CARF and SAVED effectors there are ancillary proteins related to type III CRISPR-Cas which also can be activated by signalling molecules synthesized by Cas10 protein. One such example is trimeric NucC DNase which binds cA_3 at the interface of subunits and forms an active hexamer (Grüschow et al., 2021). A recent discovery of type III system producing not cOA, but SAM-AMP signalling molecules which active transmembrane effectors (Chi et al., 2023) exemplify the diversity and versatility of type III systems and their ancillary effectors.

Table 1.1. Major groups of non-redundant CARF proteins. Table is based on CARF protein classification data presented in (Makarova et al., 2020a).

Group	Size (%, excluding RtcR)	Main effector domains (activities)	CRISPR-Cas link	Representative proteins
CARF1	329 (14%)	HEPN (RNase)	III-A	TtCsx1 (Csm6) (Niewoehner and Jinek, 2016), StCsm6 (Smalakyte et al., 2020), EiCsm6 (Garcia-Doval et al., 2020)

CARF2	239 (10%)	HEPN (RNase), mCpol (polymerase)	III	MpCsx1 (Samolygo et al., 2020)
CARF3	436 (19%)	wHTH (DNA binding)	I	SiCsa3 (T. Liu et al., 2017)
CARF4	400 (18%)	PD-(D/E)XK (nuclease), AAA (ATPase)	partial	TtCan1 (McMahon et al., 2020), TsCan2 (Card1) (Rostøl et al., 2021)
CARF5	128 (6%)	ADA (deaminase)	III	-
CARF6	167 (7%)	Serine/threonine protein kinase	no	-
CARF7	154 (7%)	RelE (RNase), HD (phosphodiesterase), CYTH (cyclase/ phosphatase), wHTH, DUF2103	III	SsCrn1 (Athukoralage et al., 2018), SiCrn1 (Molina et al., 2021)
CARF8	50 (2%)	transmembrane	III	NsCam1 (Baca et al., 2024)
CARF9	183 (8%)	HEPN (RNase)	III	ToCsm6 (Jia et al., 2019a)
RtcR	1925	AAA (ATPase)	no	EcRtcR (Kotta-Loizou et al., 2022)

1.4.2.CARF ring nuclease activity

As discussed previously, toxic activities of CARF proteins reduce the growth of infected cell or even kill it benefiting the population. The latter outcome would be expected in systems without active cOA degradation module. However, some CARF-HEPN RNases have evolved CARF domains with ring nuclease activity allowing the cleavage of cOAs. Such property diminishes toxic activity upon foreign nucleic acid destruction and cOA-producing Cas10 protein deactivation. Ring nuclease activity was observed for Csm6 proteins of CARF1 and CARF9 groups (Jia et al., 2019a; Athukoralage et al., 2020; Garcia-Doval et al., 2020; Smalakyte et al., 2020). Additionally, CARF-wHTH proteins (Crn1) of CARF7 (Athukoralage et al., 2018; Molina et al., 2021) and standalone CARF protein of minor CARF_m13 clade (Molina et al., 2022) also cleave cognate cA_4 activators. It is proposed that such internal ring nucleases reduce intracellular cOA concentration after infection ensuring temporal control of non-ring nuclease CARF or SAVED proteins (Athukoralage et al., 2018).

The cOA cleavage reaction occurs in two steps. First the cOA ring is opened forming a linear oligoadenylate with 5′-OH and 2′3′-cyclic phosphate

ends (A_n>p) which is then cut in half producing two identical products with the same ends. Ring nucleases which are activated by cA_6 produce intermediate A_6>p and A_3>p final products (Athukoralage et al., 2018; Garcia-Doval et al., 2020; Molina et al., 2021; Smalakyte et al., 2020) while cA_4-cleaving CARF domains generate intermediate A_4>p which is cut to two A_2>p as final products (Athukoralage et al., 2018, 2019; Jia et al., 2019a; Molina et al., 2021). *In vitro* studies of cOA cleavage kinetics by CARF domains indicate that activators are cleaved in time range from tens of minutes to hours suggesting a timeframe of ancillary effector activation during infection and type III CRISPR-Cas signalling (Athukoralage et al., 2019, 2018; Jia et al., 2019a; Smalakyte et al., 2020). As mentioned in section 1.3.4, cOAs are also cleaved by non-CARF anti-CRISPR ring nuclease AcrIII-1. The viral AcrIII-1 ring nuclease from *Sulfolobus islandicus* rod-shaped virus 1 shows sub-minute degradation of cA_4 under single turnover conditions which explains phage resistance to type III CRISPR-Cas defence when AcrIII-1 is present in the cell (Athukoralage et al., 2020). The reaction mechanism of AcrIII-1 is different from CARF since cA_4 is cleaved to A_2>p and A_2p having linear 3′-phosphate (Athukoralage et al., 2020).

1.4.3.CARF7 group proteins

Sequence analysis of CARF7 group proteins shows that most of them adopt a core of fused CARF and wHTH domains (Makarova et al., 2020a) as exemplified by characterized archaeal Crn1 ring nucleases (Athukoralage et al., 2018; Molina et al., 2021). CARF7 protein genes are strongly linked to type III CRISPR-Cas loci which often have additional ancillary effectors encoded in the vicinity.

~30% of CARF7 group proteins are two domain CARF-wHTH fusions. However, the most common (~50% of sequences) architecture in CARF7 group is CARF-wHTH-RelE with RelE domain fused to the C-terminus of CARF-wHTH core. RelE domain corresponds to well-characterized mRNA interferase toxins of numerous type II toxin-antitoxin systems (see section 1.5.1). ~15% of sequences exhibit uncharacterized DUF2103 as putative effector domain. A minor fraction of CARF7 group (~5%) is composed of (CYTH-)CARF-wHTH-HD proteins with an optional CYTH triphosphate tunnel metalloenzyme domain (Iyer and Aravind, 2002) fused to the N-terminal and putative phosphodiesterase HD domain (Aravind and Koonin, 1998) fused to the C-terminal end of CARF-wHTH core. The latter proteins are mostly restricted to the thermophilic bacteria of the order *Thermotogales*.

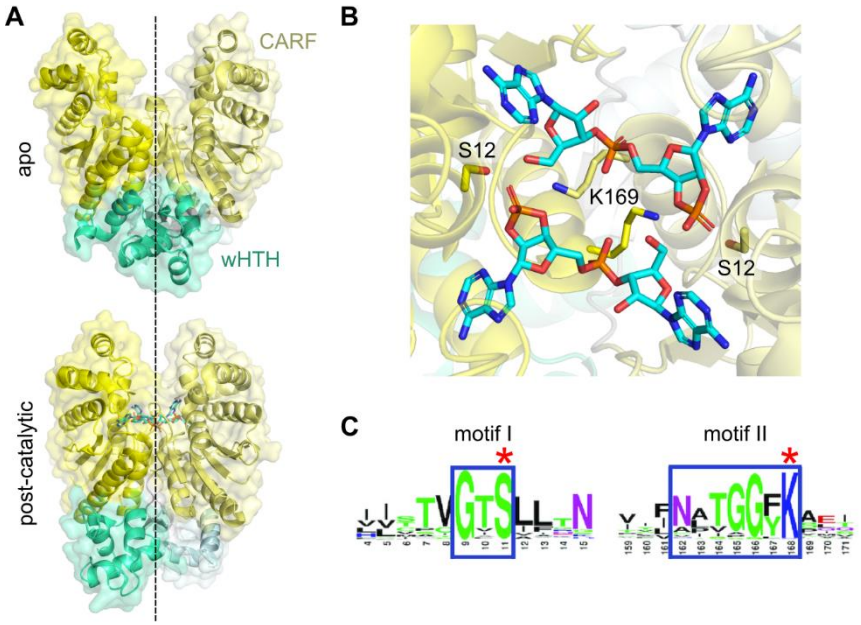


Fig. 1.9. Structure of *S. islandicus* Crn1 protein of CARF7 group. (**A**) Crystal structures of apo SiCrn1 (PDB: 7PQ2) and post-catalytic complex of SiCrn1 bound to reaction products A_2>p (PDB: 7PQ3) showing CARF domains closure and movement of wHTH domains. (**B**) 2 A_2>p bound in the CARF domain in post-catalytic complex with product ends located near conserved active site residues S12 and K169. (**C**) Conserved ring nuclease motifs of CARF7 group proteins, adapted from (Makarova et al., 2020a).

Structural characterization of *S. islandicus* Crn1 protein has shown that CARF-wHTH core is flexible (Molina et al., 2021) and wHTH domain undergoes a cork-screw motion during cognate cA_4 binding and/or cleavage (Fig. 1.9A). The serine and lysine residues critical for ring nuclease activity in Crn1 (Fig 1.9B) are conserved in the whole CARF7 group (Fig. 1.9C) suggesting that CARF7 proteins with fused effector domains could autoinhibit themselves (Makarova et al., 2020a; Molina et al., 2021).

wHTH domains usually act as DNA binding moieties in various transcriptional regulators (Gajiwala and Burley, 2000). Such function is experimentally confirmed for Csa3 proteins (CARF3 group) which regulate CRISPR-Cas and other genes expression by DNA binding via wHTH domains (T. Liu et al., 2017). However, Csa3 affinity to promoters in the DNA is not affected by cognate cA_4 binding (Xia et al., 2022). Despite substantial structural similarity to transcriptional regulators (Molina et al., 2021), no DNA binding or other functional properties have been attributed to wHTH domains of CARF7 group proteins.

1.5. Toxin-antitoxin systems

The effector domains (HEPN, RelE) of some CARF proteins correspond to toxins from widespread prokaryotic toxin-antitoxin (TA) systems (Jurėnas et al., 2022). TA systems, as the name suggests, are composed of toxins and antitoxins which are often found coupled in the same operon. Toxins usually are diverse proteins which impair DNA replication, degrade or modify RNA, inhibit translation, disrupt cell envelope or cause metabolic stress. Antitoxins are either proteins or RNAs which neutralize toxins at protein or mRNA level and regulate TA expression. According to the mode of toxin neutralization, TA systems are classified into eight types (I to VIII) (Jurėnas et al., 2022). To this date the most abundant and diverse are type II TA systems in which the toxin is neutralized by direct antitoxin protein binding often involving the covering of toxin active site (Jurėnas and Van Melderen, 2020).

While the targets and molecular mechanisms of numerous toxins are well studied, the biological functions for most of them have remained enigmatic. Naturally, functions of type II TA systems are linked to toxin activation which is caused by dysregulation of toxin:antitoxin protein ratio. This is clearly exemplified by first described plasmid-borne TA system CcdAB allowing F plasmid stabilization in *E. coli* population (Ogura and Hiraga, 1983). If the plasmid encoding CcdAB is lost in one of the daughter cells during cell fission, the antitoxin CcdA is readily degraded by Lon protease (even in the presence of its binding partner CcdB (Van Melderen et al., 1994)) and not replenished leading to the accumulation of free CcdB toxin which inhibits DNA gyrase and disrupts genome replication (Bernard and Couturier, 1992). This mechanism, termed post-segregational killing or addiction, is also adopted by other TA systems encoded on plasmids or mobile chromosomal regions such as integrons (Harms et al., 2018).

Yet, type II TA systems are often encoded in the genetically stable parts of the genome and most bacterial species harbours tens of them with substantial variation between strains (Akarsu et al., 2019). Differently from strict post-segregational killing mechanism, these systems are regulated more subtly at transcription level. Type II antitoxins have DNA-binding domains in addition to toxin-binding motifs enabling them to regulate their own transcription (De Bruyn et al., 2021; Lee and Lee, 2023). Under normal conditions antitoxins and toxins form complexes which block TA operon transcription by binding to operator regions. TA expression is alleviated due reduced affinity of toxin:antitoxin complex to DNA when additional toxin subunits bind to it (Grabe et al., 2024; Overgaard et al., 2008). It was observed that during

various abiotic stresses resulting in translation inhibition or increased protease activity on antitoxins, TA transcription levels rise significantly while no toxins are being freed from toxin-rich TA complexes (LeRoux et al., 2020). In type II TA systems the translation of toxins is less pronounced than antitoxins (Li et al., 2014) which helps to explain why toxins remain neutralized during substantial changes in operon transcription levels. Taken together these results cast doubt on previous reports stating that toxins unleashed in response to stress could put cells into dormancy or persistence allowing them to survive (Pizzolato-Cezar et al., 2023).

An another scenario leading to dysregulation of toxin:antitoxin ratio and actual toxin activation could unfold upon phage infection during which host transcription-translation machinery is hijacked to produce viral proteins (Song and Wood, 2020; LeRoux and Laub, 2022; Kelly et al., 2023). While the exact mechanisms of toxin unleash are mostly obscure, the evidences of chromosomal TA functioning as antiphage systems are accumulating during recent years (LeRoux and Laub, 2022; Kelly et al., 2023). One well-studied instance is RnlAB type II TA protecting *E. coli* K12 against T4 infection during which the endoribonuclease activity of toxin RnlA is unleashed (Koga et al., 2011). Interestingly, RnlAB protects only against T4 mutant lacking *dmd* gene (Koga et al., 2011). Later studies explained this observation by showing that T4 Dmd protein functions as an alternative antitoxin to RnlA and other related toxin LsoB from LsoAB system (Otsuka and Yonesaki, 2012). Antiphage protection was also observed for type I *hok*/*sok* (Pecota and Wood, 1996), type II toxSAS (Dedrick et al., 2017; Kurata et al., 2021; Zhang et al., 2022), type III ToxIN (Fineran et al., 2009; Guegler and Laub, 2021), type IV DarTG (LeRoux et al., 2022) and other TA systems.

Recently, a detailed toxin activation mechanism was described for type II TA-like antiphage system PARIS. In PARIS AriA antitoxin sequesters AriB toxin which inhibits translation by cleaving $tRNA^{Lys}$ when freed (Burman et al., 2024; Deep et al., 2024). It was previously noticed, that PARIS defend against phages which encode a counter-defence Ocr protein inhibiting restriction-modification defence systems (Rousset et al., 2022). Structural studies revealed that Ocr binding by AriA distorts its structure freeing AriB which inhibits translation during the infection (Burman et al., 2024; Deep et al., 2024). Interestingly, some phages such as T5 subverts counter-counter-defence by PARIS by encoding a variant of $tRNA^{Lys}$ that is not cleaved by AriB (Burman et al., 2024).

These examples and the emergence of anti-TA counter-defence elements, such as phage T4 Dmd against RnlAB and phage T5 variant $tRNA^{Lys}$ against

PARIS, strengthen the notion that TA systems primarily act in genetic conflicts (Song and Wood, 2020; LeRoux and Laub, 2022; Kelly et al., 2023).

1.5.1.RelE-like toxins

Most toxins of type II TA systems target translation by either cleaving RNA (mRNA, tRNA or rRNA) or inactivating the enzymes driving protein synthesis such as aminoacyl-tRNA synthetases or translation factors (Jurėnas and Van Melderen, 2020). RelE is structurally and enzymatically well characterized type II toxin which targets mRNA exposed in the A (aminoacyl-tRNA accepting) site of the 70S ribosome (Pedersen et al., 2003; Neubauer et al., 2009) (Fig. 1.10A). This activity causes a two-fold toxic effect: it reduces the amount of mRNA available for translation and lowers the proportion of functional ribosome pool due to their stalling at the 3′-end of degraded mRNA (as discussed in detail in Chapter 1.6). Similar ribosome-dependent activity was also observed for other toxin families: YoeB (Zhang and Inouye, 2009), YafQ (Prysak et al., 2009) and HigB (Hurley and Woychik, 2009) which here are called RelE-like. These toxins show low (10–20%) sequence identity to RelE proteins and utilize different active site residues but adopt similar α/β sandwich fold (Fig. 1.10B). While RelE-like toxins are found only in prokaryotes, it was noted that *E. coli* RelE also could cleave mRNAs in stalled eukaryotic ribosomes (Andreev et al., 2008).

RelE-like toxins show low mRNA sequence specificity (Hurley et al., 2011; Goeders et al., 2013). Instead they rely on binding to the A-site of the ribosome in which they reorganize mRNA for cleavage (Neubauer et al., 2009; Feng et al., 2013; Maehigashi et al., 2015). Most of these toxins were shown to cleave mRNA after 2nd or 3rd nucleotide of the codon exposed in the A-site (Choi et al., 2017; Goeders et al., 2013) with some toxins showing flexibility (Hurley et al., 2011). EcRelE leaves 2′3′-cyclic phosphate at the 3′-end of ribosome-bound product while YafQ, YoeB and HigB yield linear 3′-phosphate end similarly to structurally related RNase T1 which is not a part of TA systems (Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016).

Further structural comparisons have shown that RelE-like toxins are related to ParE toxin family which are gyrase poisons (Anantharaman and Aravind, 2003). ParE proteins lack active site residues typical for RelE-like toxins but structures of *E. coli* ParDE and RelBE antitoxin-toxin pairs show similar modes of neutralization with antitoxins blocking the same surfaces of cognate toxins (Bøggild et al., 2012; Li et al., 2009). Network analysis of TA operons

has shown that some protein families of antitoxins are promiscuous and could neutralize both ParE and RelE-like antitoxins (Ernits et al., 2023).

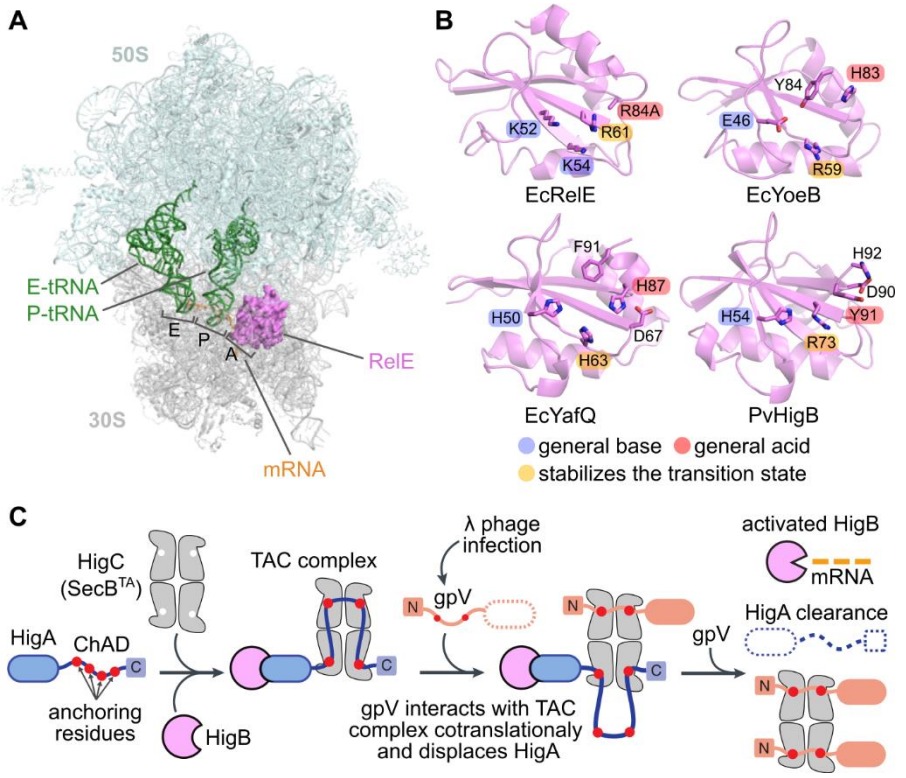


Fig. 1.10. Overview of RelE-like toxins. (**A**) Crystal structure of *in vitro* reconstituted ribosome-RelE complex. EcRelE (magenta) is bound to mRNA exposed in of *T. thermophilus* 70S ribosome (cyan and grey) (PDB: 4V7J). E (exit), P (peptidyl-tRNA) and A (aminoacyl-tRNA) sites of the ribosomes are indicated. (**B**) Spatially aligned RelE-like proteins with indicated active site residues (PDB: 4V7J (EcRelE), 4V8X (EcYoeB), 4ML2 (EcYafQ), 4ZSN (*P. vulagaris* HigB)). Adapted from (Schureck et al., 2016). (**C**) Defence mechanism of HigBAC tripartite toxin-antitoxin system utilizing SecB-like ($SecB^{TA}$) chaperone HigC. Adapted from (Mets et al., 2024).

Despite extensive computational, structural and biochemical studies, the biological functions of TA systems utilizing RelE-like toxins are debated. Yet the discourse on general functions of TA systems is leaning towards attributing a defence role to them (Song and Wood, 2020; LeRoux and Laub, 2022; Kelly et al., 2023) with new data from RelE-like toxin studies to support this notion. Recently, Mets et al. described a tripartite TA system HigBAC encoded by *E. coli* prophages which grants the bacterial host specific immunity

against phage λ infection (Mets et al., 2024). This system is composed of RelE-like HigB toxin, its cognate antitoxin HigA and a SecB-like chaperone HigC which binds chaperone addiction (ChAD) regions of C-terminal part of HigA thus stabilizing it and preventing from degradation (Fig 1.10C). Upon phage λ infection and viral protein expression, HigC binds N-terminal region of major tail protein gpV which outcompetes HigA. This causes HigA degradation and release of HigB toxin in the infected cell. These findings allowed to functionally pair HigC which acts as sensor for phage λ gpV protein expression to the effector HigB.

RelE-like toxins are also found in other systems involved in antiphage defence either as standalone proteins or parts of multidomain proteins. A recent study of genomic fragments from different *E. coli* strains describes two instances with putative RelE-like effectors among other novel defence systems (Vassallo et al., 2022). A singular protein termed PD-T4-7 composed of RelE-like and phage sheath domains provides robust immunity against T2, T4 and T6 phages while a tripartite system PD-λ-2 with HigB-like protein, peptidase and phage P4 beta family protein (Ghisotti et al., 1990; Rousset et al., 2022) defend bacteria against λ, T3 and other phages. Interestingly, the latter system shows functional similarity with TA systems, however here the HigB-like protein seems to function as an antitoxin neutralizing the toxic peptidase (Vassallo et al., 2022). Another single-protein system termed HEC07 composed of adenosylmethionine decarboxylase, inactivated thioredoxin and RelE-like domains provides defence against a group of T4-like phages (Payne et al., 2024). As mentioned in Chapter 1.4.3, RelE-like domains are also found fused to CARF proteins associated with type III CRISPR-Cas systems (Makarova et al., 2020a). These examples suggest that RelE-like protein domains are evolutionary promiscuous and were adapted in the context of prokaryotic defence.

The pace of defence system discovery is accelerating (Payne et al., 2022; Tesson et al., 2024) and the future unveiling of new systems employing RelE-like toxins should be not surprising. While the mRNA degradation activity and subsequent ribosome stalling caused by RelE-like toxins is well documented, the triggers and activation mechanisms for most TA or other antiphage defence systems utilizing them remain to be elucidated.

1.6. Ribosome stalling and rescue

Protein synthesis (translation) consists of three phases: initiation, elongation and termination followed by ribosome recycling (Rodnina, 2018). Each phase involves ribosomes decoding mRNA sequence to start the synthesis of a protein chain, add amino acids to it or stop the synthesis, respectively. Ribosomes scan the coding part of mRNA in three nucleotide increments termed codons each specific to either tRNAs, which transport amino acids to the ribosomes, or release factors (RFs) which induce translation termination. RFs are specific to certain stop codons and cause protein chain hydrolysis from tRNA and subsequent ribosome disassembly.

Rare codon stretches, mRNA secondary structures or certain amino acid composition of nascent peptides can cause temporary pausing of the ribosome which could be alleviated spontaneously or with assistance by other factors (Samatova et al., 2020). Physiological modifications of mRNA and tRNA could also influence the rate of translation due to ribosome pausing (Franco and Koutmou, 2022) while oxidative and alkylative damage to mRNA cause prolonged ribosome stalling at the damaged site (Simms et al., 2014; Thomas et al., 2020). mRNA hydrolysis, stop codon read-through or frameshifting could lead to a scenario under which ribosomes remain in the elongation phase until the 3′-end of mRNA. Since the translocation step of elongation puts the 3′-end of mRNA in the P site, the A-site remains devoid of mRNA preventing the recruitment of tRNAs or RFs (Keiler et al., 1996). This causes ribosome stalling preventing termination and ribosome recycling. The stalling can only be alleviated via specialized ribosome rescue pathways (Müller et al., 2021). The most widespread mechanism termed trans-translation utilizes a hybrid transfer-messenger RNA (tmRNA) which has both tRNA and mRNA properties (Fig. 1.11A). tmRNA is found in >99% of bacterial genomes (Hudson et al., 2014).

tmRNA is approximately 360 nt in length and adopts a looped structure (Fig. 1.11A). Its 5′- and 3′-terminal parts come together to form a tRNA-like domain (TLD) which could be loaded with alanine amino acid (Komine et al., 1994). TLD lacks an anticodon stem loop which is required for canonical tRNAs to bind specific mRNA codons. Instead, the space of this stem loop is occupied by SmpB protein stabilizing the tmRNA structure and assisting with ribosome binding (Karzai et al., 1999). The larger part of tmRNA chain is structured as helices and pseudoknots leaving a single-stranded stretch termed mRNA-like domain (MLD) (Fig. 1.11A). MLD is exposed to serve as a matrix for translation *in trans* (or trans-translation). It encodes a short degradation tag

to be attached to the nascent peptide of stalled ribosomes marking it for degradation (Keiler et al., 1996). Like a usual open reading frame, this coding sequence ends with canonical stop codon allowing proper translation termination.

tmRNA-SmpB complex moves through a stalled ribosome (①; further numbering follows Fig. 1.11B) in several steps during which it resumes the translation (Guyomar et al., 2021; Rae et al., 2019). Similarly to other canonical aminoacylated tRNAs, alanyl-tmRNA-SmpB complex is brought to the ribosome by elongation factor Tu (EF-Tu) (②). SmpB probes the empty mRNA entry channel with C-terminal helix and accommodates tmRNA in the A-site allowing for peptide transfer reaction from peptidyl-tRNA in the P site to alanyl-tmRNA (③). tmRNA-SmpB is then translocated to the P site by elongation factor G (EF-G) (④). During this the MLD passes into the A-site previously occupied by SmpB which flips to the opposite side of mRNA channel (⑤). The first codon of MLD is decoded by tRNA resuming the translation (⑥). This and subsequent peptide transfer (⑦) with tRNA translocation (⑧) steps are repeated until the stop codon of MLD is reached. Then RF is recruited to the A-site allowing tagged peptide release and ribosome disassembly (⑨). The tagged peptide is recognized and degraded by proteases (Keiler et al., 1996) while the ribosome subunits can be reused to translate other mRNAs.

Early studies showed that disruption of tmRNA encoding *ssrA* gene inhibits growth of *E. coli* (Oh and Apirion, 1991). The accumulation of unhydrolyzed peptidyl-tRNA molecules was observed in *E. coli* mutants devoid of main ribosome rescue pathways even when the bacteria were cultured in rich media (Ito et al., 2011). This signifies a prominent ribosome stalling without apparent stress (Ito et al., 2011). Yet, certain cellular states could drastically induce ribosome stalling, for example due to mRNA degradation under activation of RelE-like toxins. Direct *in vivo* experiments show that toxicity caused by RelE could be resuscitated by overproduction of tmRNA even when the MLD of tmRNA itself is targeted by RelE (Christensen and Gerdes, 2003). EcRelE was found to cleave A-site mRNA after second nucleotide of an exposed codon, effectively leaving 2 nt flanking past P-site (Hurley et al., 2011). This flanking end can be readily outcompeted by tmRNA-SmpB complex which tolerates flanks up to 9 nt (Ivanova et al., 2004).

When tmRNA is mutated or overwhelmed with non-stop ribosome complexes, an alternative rescue factor ArfA is produced in *E. coli*

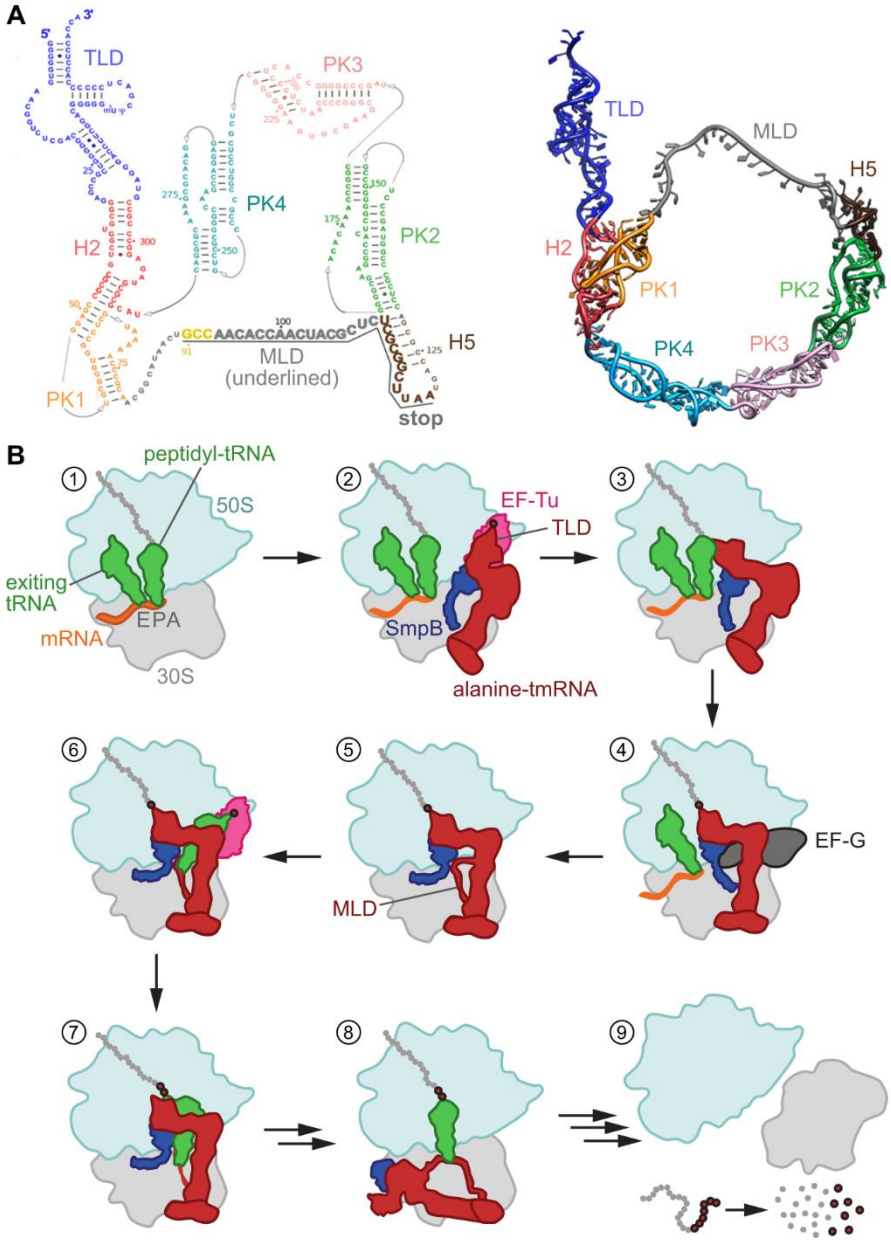


Fig. 1.11. tmRNA and trans-translation. (**A**) Composition (left) and structure (right; PDB: 3IYQ) of *T. thermophilus* tmRNA. Structural elements of tmRNA molecule are color-coded in both views. TLD – tRNA-like domain, MLD – mRNA-like domain, PK – pseudoknot, H – helix. UAA stop codon of MLD is indicated. Adapted from (Giudice et al., 2014). (**B**) Trans-translation ribosome rescue pathway. Steps from tmRNA delivery to stalled ribosome, translation of MLD to ribosome disassembly (①-⑨) are described in the main text accordingly. Adapted from (Rae et al., 2019).

(Chadani et al., 2011a; Garza-Sánchez et al., 2011). It recruits release factor 2 (RF2) to ribosomes with empty A-site where RF2 performs peptidyl-tRNA hydrolysis (Chadani et al., 2012; Shimizu, 2012) leading to ribosome recycling. A double *E. coli* mutant lacking both tmRNA and ArfA is non-viable signifying that sequestration of ribosomes due to spontaneous stalling is deleterious for the cell (Chadani et al., 2011a). Another *E. coli* alternative rescue factor termed ArfB (previously YaeJ) can cleave peptidyl-tRNA in non-stop ribosomes without assistance by RFs (Chadani et al., 2011b; Handa et al., 2011). The chromosomal copy of *arfB* gene does not resuscitate *E. coli* bacteria devoid of tmRNA and ArfA unless ArfB is overexpressed (Chadani et al., 2011b), yet chromosomal ArfB is sufficient to ensure survival of tmRNA and/or ArfA mutants of other species (Feaga et al., 2014; Moreno et al., 2022). Since ArfA and ArfB mediated rescue depends on peptidyl-tRNA hydrolysis, the nascent peptides are not additionally labelled for degradation as in trans-translation (Müller et al., 2021).

ArfB (but not tmRNA or ArfA) homologs are also found in the genomes of eukaryotes (Duarte et al., 2012). Best characterized example is an essential human ICT1 (immature colon carcinoma transcript-1) protein which is an integral part of mitochondrial ribosome (Amunts et al., 2015), but also functions in stalled mitoribosome rescue (Akabane et al., 2014; Richter et al., 2010). Interestingly, ICT1 was found to rescue stalled ribosomes of bacterium *Caulobacter crescentus* while *C. crescentus* ArfB prevented death of human cells with ICT1 knocked down suggesting mutually similar rescue mechanism (Feaga et al., 2016). ICT1 is found to be upregulated in various types of cancer and thus considered as a marker for tumour cells (Chicherin et al., 2021).

1.7. Ribosome stalk

The major drivers of elongation EF-Tu and EF-G together with initiation factor 2 (IF-2) and RF3 belong to a protein family called translational GTPases (trGTPases) which also accommodates other translation factors having a GTP-hydrolysing G domain (Maracci and Rodnina, 2016). trGTPases are activated upon binding to the large subunit when their G domain comes into contact with a universally conserved adenosine in sarcin-ricin loop (SRL) of 23S rRNA (Clementi and Polacek, 2010). GTP hydrolysis induces conformational changes related to trGTPase function. For example, upon GTP hydrolysis EF-Tu releases cargo aminoacyl-tRNA allowing it to accommodate into peptidyl

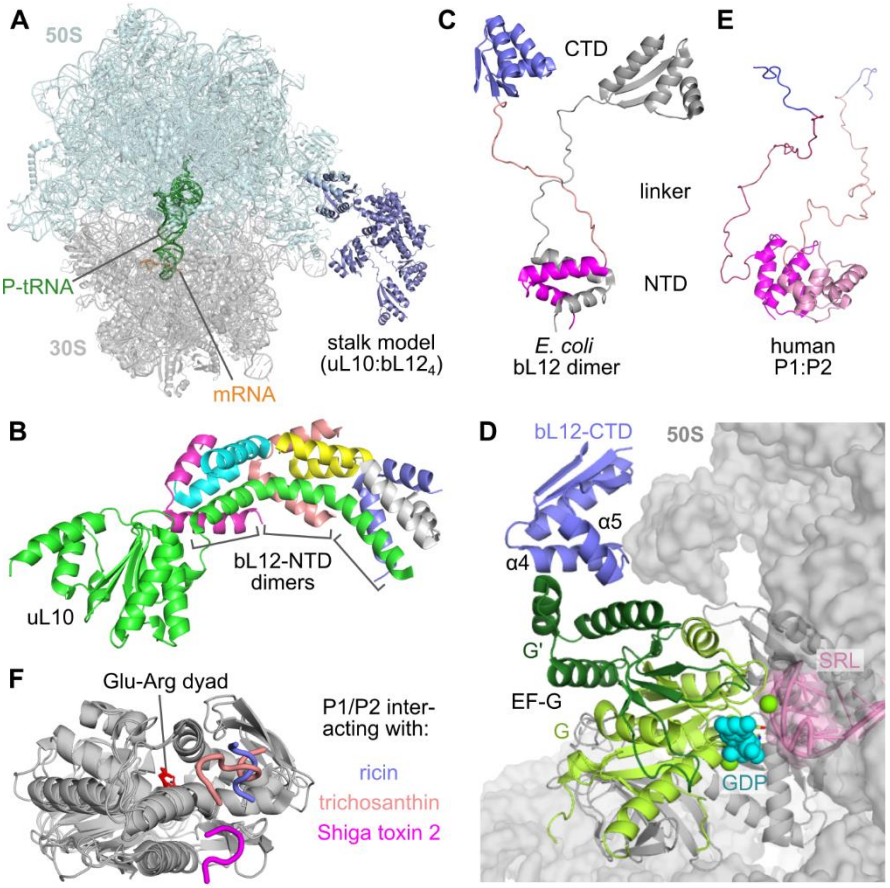


Fig. 1.12. Ribosome stalk structures. (**A**) Ribosome stalk model superimposed with cryoEM structure of *E. coli* 70S ribosome via uL10-NTD (PDB: 5MDZ). Stalk model composed of *E. coli* uL10:bL12$_4$ was prepared by the author using AlphaFold 2 (Jumper et al., 2021). (**B**) Crystal structure of *T. maritima* ribosome stalk core consisting of uL10:bL12-NTD$_6$ heptamer (PDB: 1ZAX). (**C**) NMR structure of full-length bL12 homodimer from *E. coli* (PDB: 1RQU). (**D**) EF-G:GDP bound to SRL of 50S ribosome subunit (PDB: 7N2C). G' extension (dark green) of EF-G G domain (light green) interacts with α4 and α5 helices of bL12-CTD (slate blue). (**E**) NMR structure of human P1:P2 heterodimer (PDB: 4BEH). (**F**) Interactions between RIPs and conserved C-terminal parts of P1/P2. Structures of three different RIPs (grey) are superimposed and three interacting peptides are coloured according to the bound RIP: ricin (slate blue; PDB: 5GU4), trichosanthin (salmon; PDB: 2JDL), Shiga toxin 2 (magenta; PDB: 6X6H). Conserved glutamic acid and arginine residues critical for RIP N-glycosylase activity are coloured in red.

transferase centre (Pape et al., 1998) and EF-G translocates ribosomes along mRNA (Rodnina et al., 1997).

trGTPases are recruited to the ribosomes by a structure called ribosome stalk (Liljas and Sanyal, 2018). Due to high mobility (Imai et al., 2020) the stalk structure has not been resolved in full in ribosome structural studies (Diaconu et al., 2005), however its position relative to the ribosome core can be estimated by protein structural modelling (Fig. 1.12A). Bacterial ribosome stalk consists of protein uL10 bound to four (as in *E. coli*), six (as in *Thermotoga maritima*, Fig. 1.12B) or eight (only in some cyanobacteria) copies of bL12 (Davydov et al., 2013). NTD of uL10 connects the stalk to 50S subunit and its C-terminal helix acts as a chassis to bind homodimers of bL12 proteins (Diaconu et al., 2005). bL12 consists of two dimerization helices at N-terminus connected to a globular CTD via flexible linker (Fig. 1.12C). CTD of bL12 binds to trGTPase G domain (Ge et al., 2018; Helena-Bueno et al., 2024) or its extension G' in case of EF-G (Fig. 1.12D) and RF3 (Pallesen et al., 2013) via two conserved helices (Helgstrand et al., 2007). These interactions are critical for trGTPase recruitment to the ribosome (Carlson et al., 2017; Ge et al., 2018). Ribosome stalk also binds other trGTPases such as BipA which facilitates 50S subunit maturation (Kumar et al., 2015) and TetM protein protecting ribosomes from antibiotic tetracycline (Arenz et al., 2015). Mitochondrial and chloroplast ribosomes also have bacterial-like stalk structures with bL12 homologs bL12m and bL12c (Ban et al., 2014) which recruit organelle-specific trGTPases (Koripella et al., 2020).

Stalk structure of archaeal and eukaryotic cytoplasmic ribosomes utilize P1/P2 proteins which do not share homology with bL12 (Naganuma et al., 2010). These proteins do not form globular CTDs and instead are disordered in eukaryotes (Fig. 1.12E) or adopt short helical secondary structures in archaea (Naganuma et al., 2010). C-terminal parts of these proteins bind hydrophobic pockets in the G/G' domains of corresponding trGTPases (Ito et al., 2014; Tanzawa et al., 2018) whereas bacterial bL12 utilizes charge complementarity (Helgstrand et al., 2007; Ge et al., 2018). Conversely, P1/P2 proteins can be hijacked by poisonous ribosome inactivating proteins (RIPs; Fig. 1.12F) such as ricin which then depurinates the G domain-activating adenosine in SRL making trGTPases unable to function in damaged ribosomes and halting protein production in the cell (Grela et al., 2019).

According to the current understanding of stalk proteins bL12 and P1/P2, their role involves increasing the local concentration of trGTPases, especially elongation factors, in the vicinity of ribosomes during translation (Liljas and Sanyal, 2018). This notion is supported by recent experiments. Single-molecule *in vivo* tracking of EF-Tu and ribosomes show that they colocalize in live *E. coli* with nearly four copies of EF-Tu per ribosome which matches copy

number of bL12 (Mustafi and Weisshaar, 2018, 2019). Elongation factor pooling by archaeal ribosome stalk was directly visualized *in vitro* using high-speed atomic force microscopy (Imai et al., 2020). Additionally, it is suggested that trGTPase interaction with the stalk could stabilize its position near SRL and promote GTP hydrolysis (Liljas and Sanyal, 2018), a property taken advantage of by RIPs to damage the SRL (Grela et al., 2019). Besides trGTPases and RIPs, no other proteins have been shown to interact with the ribosome stalk.

2. MATERIALS AND METHODS

2.1. Materials

2.1.1.Chemicals

All chemicals used in this study were purchased from Thermo Scientific, Invitrogen, Sigma-Aldrich, Carl Roth, Invitrogen and Fluka and of the highest purity grade available. Radiolabelled nucleotides were purchased from Perkin Elmer.

2.1.2.Proteins and kits

“GeneJET PCR Purification Kit”, “GeneJET Gel Extraction Kit”, “GeneJET Plasmid Miniprep Kit”, “T7 High Yield Transcription Kit”, “GeneJET RNA Cleanup and Concentration Micro Kit”, “FastDigest” restriction endonucleases, DreamTaq and Phusion DNA polymerases, T4 polynucleotide kinase (PNK), FastAP thermosensitive alkaline phosphatase, RNases A and T1, „PageRuler™ Unstained Protein Ladder“ (protein marker sizes in kDa: 10, 15, 20, 25, 30, 40, 50, 60, 70, 85, 100, 120, 150, 200), RiboLock RNase inhibitor and bovine serum albumin (BSA) were purchased from Thermo Scientific.

“NEBuilder HiFi DNA Assembly Master Mix”, micrococcal nuclease and *E. coli* 70S ribosomes were purchased from New England Biolabs.

All these products were used according to the recommendations by the manufacturer unless indicated otherwise. Reaction mixtures containing commercial enzymes were used in supplemented buffer solutions unless indicated otherwise.

Candidatus C. acidaminovorans Csm complex was obtained from Dalia Smalakytė (Smalakyte et al., 2024).

TEV protease was obtained from dr. Arūnas Šilanskas.

70SΔbL12 ribosomes were obtained from prof. Paul Clint Spiegel of Western Washington University (Carlson et al., 2017).

2.1.3.Plasmids

All plasmids used in this work with their descriptions, sources and links to sequences are listed in Appendix 1.

2.1.4. Oligonucleotides

Synthetic DNA and RNA oligonucleotides were purchased from Metabion and Eurofins Genomics. $tRNA^{fmet}$ was purchased from ChemBlock. cA_4 and cA_6 were purchased from Biolog Life Sciences institute. Synthetic oligonucleotides used as substrates and probes are listed in Appendix 2.

2.1.5. Bacterial strains and media

E. coli strain BL21(DE3) [F^- *omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm* (DE3)] was used for StCsm complex expression.

E. coli strain ER2566 [F^- λ^- *fhu*A2 *lon omp*T *lac*Z::T7.1 *gal sul*A11 Δ(*mcr*C-*mrr*)*114*::IS10 R(*mcr*-73::miniTn10)(*Tet*^{S})2 R(*zgb*-210::Tn10)(*Tet*^{S}) *end*A1 *dcm*] was used for Cami1 protein *in vivo* toxicity assays.

E. coli strain DH5α [F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1] was used for the cloning procedures.

E. coli strain DH10B [F- λ- *end*A1 *rec*A1 *gal*E15 *gal*K16 *nup*G *rps*L(Str^R) Δ*lac*X74 Φ80*lac*ZΔM15 *ara*D139 Δ(*ara-leu*)7697 *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC)] was used for Cami1 protein expression.

E. coli bacteria were cultivated in LB broth (1% (all w/v) peptone, 0.5% yeast extract, 0.5% NaCl in deionized water) or plated on LB agar, which also contained 1.5 % agar. Various supplements to these media are indicated where appropriate. Full list of supplements used:

- antibiotics: ampicillin (100 μg/mL; all final concentrations), carbenicillin, (100 μg/mL) streptomycin (50 μg/mL), chloramphenicol (30 μg/mL), kanamycin (50 μg/mL);
- inducers: IPTG, arabinose (concentrations are indicated where appropriate);
- glucose (1% w/v).

2.1.6. Buffers and other solutions

Protein purification

- StCsm complex purification buffer: 20 mM Tris–HCl (pH 8.5), 500 mM NaCl, 1 mM EDTA, 7 mM 2-mercaptoethanol.
- StCsm complex storage buffer: 10 mM Tris-HCl (pH 8.5), 300 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 50% (v/v) glycerol.

- Cami1 purification buffer: 20 mM Tris–HCl (pH 8.5), 500 mM NaCl, 7 mM 2-mercaptoethanol.
- Cami1 storage buffer: 10 mM Tris-HCl (pH 8.5), 300 mM KCl, 1 mM DTT, 0.1 mM EDTA, and 50% (v/v) glycerol.
- bL12 purification buffer: 50 mM Tris-HCl (pH 7.5), 60 mM NH_4Cl, 7 mM $MgCl_2$.
- bL12 storage buffer: 50 mM Tris-HCl (pH 7.5), 60 mM NH_4Cl, 7 mM $MgCl_2$, 25% glycerol.

Protein analysis

- 4X protein gel loading solution: 100 mM Tris-HCl (pH 6.8), 4% SDS (w/v), 20% (v/v) glycerol, 200 mM DTT, trace of bromophenol blue.
- 4% concentrating polyacrylamide gel: 4% acrylamide/N,N′-methylene-bisacrylamide (w/w; 37.5:1), 125 mM Tris-HCl (pH 6.8), 0,1 % SDS.
- 12% fractionating polyacrylamide gel: 12% acrylamide/N,N′-methylene-bisacrylamide (w/w; 37.5:1), 375 mM Tris-HCl (pH 8.8), 0,1 % SDS.
- Protein electrophoresis buffer: 25 mM Tris-HCl (pH 8.3), 190 mM glycine, 0.1 % SDS.
- "PageBlue Protein Staining Solution" (Thermo Fisher Scientific).
- SEC-MALS buffer: 25 mM Tris-HCl (pH 8.0), 250 mM NaCl and 0.01% (w/v) NaN_3.

Buffers and dyes for activity assays.

- StCsm binding buffer: 1X TAE, 10% (v/v) glycerol, 0.1 mg/ml BSA.
- StCsm reaction buffer: 33 mM Tris-acetate, 66 mM K-acetate, 0.1 mg/ml BSA, 0.25 U/µl Thermo Scientific RiboLock RNase Inhibitor and 0.5 mM $MnCl_2$.
- RN buffer: 33 mM Tris-acetate (pH 7.9), 66 mM K-acetate, 0.1 mg/ml BSA
- cOA synthesis buffer: 20 mM HEPES (pH 7.5), 50 mM KCl, 1 mM $Mg(CH_3COO)_2$, 0.1 mg/ml BSA
- RB buffer: 50 mM Tris-HCl (pH 7.5 at 37 °C), 70 mM NH_4Cl, 30 mM KCl, 7 mM Mg-acetate, 1 mM DTT
- cryoRB buffer: 20 mM HEPES-KOH (pH 7.5), 100 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM DTT, 2 mM spermidine, 0.05 mM spermine, 0.5 mM EDTA
- 1X PNK reaction buffer A: 50 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 5 mM DTT, 1 mM spermidine (Thermo Scientific).

Electrophoresis buffers and gels:

- 1X TAE buffer (Thermo Scientific): 40 mM Tris (pH 8.3), 20 mM acetic acid, 1 mM EDTA.
- 1% agarose gel: 1% agarose (w/v) in TAE buffer.
- 1X TBE buffer (Thermo Scientific): 89 mM Tris (pH 8.3), 89 mM boric acid, 2 mM EDTA.
- Denaturing 20% PAAG: 8 M urea, 20% acrylamide/N,N′-methylenebisacrylamide (w/w; 29:1) in 0.5X TBE buffer.
- Native 8% PAAG: 8% acrylamide/N,N′-methylenebisacrylamide (w/w; 29:1) in 1X TAE buffer.
- DNA Loading Dye: 25 mM EDTA, 0.01% (w/v) bromophenol blue, 95% (v/v) formamide (pH 9.0).
- 2X RNA Gel Loading Dye (Thermo Scientific): 98% formaldehyde, 25 mM EDTA, 0.025% bromophenol blue.

Buffers for northern blot:

- RNAsnap solution: 18 mM EDTA, 0.025% SDS, 1% β-mercaptoethanol and 95% formamide.
- Hybridization solution: 30 mM Na-citrate (pH 7.0), 300 mM NaCl, 1% SDS, 0.1 mg/mL salmon sperm DNA.
- Washing solution: 30 mM Na-citrate (pH 7.0), 300 mM NaCl, 0.1% SDS

pH was measured at 25 °C unless indicated otherwise.

2.2. Methods

2.2.1. Sequence analysis of CARF7 clade

Construction of an enriched set of CARF7 proteins. Initially, the recently classified CARF7 clade (Makarova et al., 2020a) containing 490 sequences was enriched with additional archaeal and prokaryotic homologs identified in non-redundant protein sequence database (“nr”) downloaded from the NCBI FTP site. To this end, for each sequence from the initial set, the CARF domain region was identified based on comparison with Sso1393, a structurally resolved CARF7 member (PDB ID: 3QYF) from *S. solfataricus* (Athukoralage et al., 2018). CARF domains were then extracted and corresponding HMM (hidden Markov model) profiles were constructed using HHpred (Steinegger et al., 2019) searches against the Uniclust30 database (Mirdita et al., 2017).

Domain boundaries were further inspected and adjusted if necessary. CARF domain sequences derived in this way were then used as queries to scan the "nr" database using BLAST (Altschul et al., 1997) and sequences with E-value = 0.01 or lower were retained. Every retained candidate protein was further checked for the presence of an intact CARF domain of the CARF7 type. This was done by constructing an HMM profile for every candidate sequence using HHsearch and comparing this profile against the initially constructed set of CARF domain HMMs. If an HMM profile of a candidate protein matched at least five CARF domains with probability above 90% and nearly complete sequence overlap, it was considered a significant hit, and the new domain was added to the set of CARF domains. CARF domain boundaries were adjusted if necessary. Every newly added CARF domain was in turn used as a query to scan the "nr" database until no new significant hits were found. In the next step, to remove identical sequences, the complete set of CARF7 sequences was filtered at 100% sequence identity using CD-HIT (Li and Godzik, 2006) and the resulting sequences were aligned using MAFFT (Katoh and Toh, 2008) in a high accuracy mode (L-INS-i). Multiple sequence alignment (MSA) was inspected manually and sequences lacking the CARF7 clade-specific motifs (Motif-I and Motif-IA) (Makarova et al., 2020a) were removed. The final non-redundant set consisted of 1521 CARF7 proteins.

Phylogenetic analysis. For phylogenetic analysis a smaller, non-redundant set of CARF7 proteins was prepared by clustering sequences with CD-HIT at 40% sequence identity. The resulting set of 385 CARF domain sequences was aligned using MUSCLE (-align mode) (Edgar, 2004). MSA was trimmed with trimAl (Capella-Gutiérrez et al., 2009) and a phylogenetic tree was constructed using IQTree (Minh et al., 2020).

Structural modelling and parsing into domains. To parse information about domains of CARF7 proteins, structural models were prepared for the entire set. Structural models for 755 out of 1521 CARF7 proteins were downloaded from UniProt repositories (UniProt Consortium, 2023). Remaining 766 proteins were modelled using AlphaFold 2 with default parameters (Jumper et al., 2021). Next, structural models of all 1521 CARF7 proteins were parsed into individual domains and annotated iteratively as follows. Initially, one of the structural models was manually split into domains and for each domain, except CARF domains identified previously, the structural-functional annotation was assigned based on sequence and structure comparison. Sequence HMM profiles of each domain were constructed and searched against PDB (Burley et al., 2017), ECOD (Cheng et al., 2014) and PFAM (Mistry et al., 2021) databases using HHsearch (Steinegger et al., 2019; Zimmermann et al.,

2018). Likewise, the 3D structure of each domain was compared with all annotated domains in ECOD using DaliLite (Holm, 2019). Based on the identified homologous structures, each domain was annotated accordingly, and its boundaries adjusted if necessary. Next, the structure of each of the annotated domains was compared with the remaining structural models from the CARF7 set using DaliLite, and annotations were transferred to the matched homologous domains. For the remaining unassigned regions in the models the procedure was repeated until all domains in the CARF7 protein set were identified and annotated.

Sequences and annotations of CARF7 clade proteins are listed in Supplementary Tables S1 to S3 published together with (Mogila et al., 2023).

2.2.2.Cloning and mutagenesis

pCsm2-TEV-tag. TEV protease recognition site was inserted into pET-Csm-tag plasmid (Appendix 1) between Csm2 and N-terminal StrepII affinity tag using Gibson assembly. For this a fragment corresponding to linearized pET-Csm-tag with a break between StrepII and *csm2* was amplified by PCR, mixed with a splint oligonucleotide containing TEV recognition site and recircularized using NEBuilder HiFi DNA Assembly Master Mix.

pStCsm_dRNase. pStCsm plasmid (Appendix 1) was double cut with FastDigest Mva1269I and the longer fragment was gel purified using GeneJET Gel Extraction Kit. This fragment was mixed with a PCR product of pCas/Csm_D33A containing *csm3* gene with D33A mutation and assembled using NEBuilder HiFi DNA Assembly Master Mix.

pUC18_S3/n and pUC18_NS. pUC18_S3/2 plasmid (Appendix 1) was double cut using FastDigest BamHI and FastDigest KpnI and purified using GeneJET PCR Purification Kit. Resulting fragment was mixed with annealed oligonucleotides designed to encode RNA either complementary to the full length of StCsm S3 crRNA (Tamulaitis et al., 2014) or non-complementary to it and assembled using NEBuilder HiFi DNA Assembly Master Mix ultimately yielding pUC18_S3/n and pUC18_NS, respectively.

Design of Cami1 expression plasmids. Four Cami1 proteins from *Allochromatium vinosum* DSM180 (NCBI protein reference sequence WP_012972290.1), *Caldilinea aerophila* DSM14535 (WP_044276641.1), *Candidatus* Cloacamonas acidaminovorans str. Evry (WP_015424585.1), *Caldicellulosiruptor hydrothermalis* 108 (WP_013404376.1), respectively, AvCami1, CaCami1, CCaCami1, and ChCami1, were selected for analysis. Gene sequences were optimized for expression in *E. coli*, fused to C-terminal

StrepII and His10 affinity purification tags and submitted to Twist Bioscience for DNA synthesis and cloning into pBAD/HisA vector (Invitrogen) resulting in pAvCami1, pCaCami1, pCCaCami1 and pChCami1 plasmids (Appendix 1).

Mutagenesis of pAvCami1. AvCami1 point mutants were constructed by separately amplifying two overlapping fragments of pAvCami1, pAvCami1-S11A or pAvCami1- S11A+Q210A+E212A (Appendix 1) plasmid using mutagenic primers in PCR. One overlap included mutation(s) to open reading frame of AvCami1 while the other overlap fully matched vector part of the plasmid. Fragments were used to reconstitute the mutant plasmid by Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix.

In all cases the Gibson assembly reaction mixture was prepared according to the manufacturer recommendations and was used to transform *E. coli* DH5α. Selected colonies were grown overnight in LB broth with appropriate antibiotics and plasmids were purified using GeneJET Plasmid Miniprep Kit. Sequences were confirmed by performing whole plasmid Nanopore sequencing (SeqVision).

2.2.3.Protein expression and purification

StCsm complex. E. coli BL21(DE3) was transformed with pCRISPR_S3, pCsm2N-TEV-Tag and pStCsm or pStCsm_dRNase plasmids for WT or RNase-dead StCsm complex production, respectively. Transformants were verified by colony PCR and used to inoculate overnight cultures in LB broth with Cm, Str and Ap antibiotics. After overgrowth the overnight cultures were 1:100 diluted in fresh LB broth with the same antibiotics and bacteria were grown at 37 °C with shaking at 200 rpm. The growth was monitored until bacterial optical density OD_{600} reached 0.6 when the expression of StCsm complex components was induced by adding IPTG to final 1 mM concentration and resuming the shaking for 3 h at 37 °C. Cells were harvested by centrifugation at 7000 rpm 4 °C for 10 min using F12-6x500 LEX rotor (Thermo Scientific) and disrupted by sonication in StCsm complex purification buffer supplemented with 2 mM phenylmethylsulfonyl fluoride (PMSF). The suspension was cleared by centrifugation at 18 000 rpm 4 °C for 1 h using JA-20 rotor (Beckman Coulter) and complexes from soluble fraction were captured by StrepII-affinity chromatography using StrepTrap HP (Cytiva) columns and eluted using StCsm complex purification buffer with 2.5 mM desthiobiotin. Complexes were subjected to size exclusion chromatography and fractionated according to complex size, which is dependent on crRNA maturation level

(Tamulaitis et al., 2014). crRNA samples from different fractions were extracted by phenol/chloroform and analysed by electrophoresis under denaturing conditions in 0.5X TBE buffer. Fractions in which fully matured 40 nt crRNA species was dominant were merged and incubated overnight with TEV protease (1 mg TEV per 50 A_{280} × mL units of StCsm complex) in StCsm complex purification buffer supplemented with 15 mM DTT at 4 °C. To discard His-tagged TEV protease and uncleaved StrepII-tagged complexes the mixture was subjected to HisTrap HP (Cytiva) and StrepTrap HP columns. The flow-through was collected, dialyzed against StCsm complex storage buffer and stored at -20 °C.

Cami1 proteins. *E. coli* DH10B was transformed with plasmids encoding WT and mutant Cami1 proteins (Appendix 1). Bacteria were incubated with shaking in LB broth supplemented with 25 μg/mL streptomycin and 50 μg/mL ampicillin at 37 °C until bacterial optical density OD_{600} reached 0.6. Cultures were induced with arabinose to final 0.2% concentration and incubated overnight at 16 °C. Biomass was collected by centrifugation at 7000 rpm 4 °C for 10 min using F12-6x500 LEX rotor (Thermo Scientific). Biomass was suspended (20% w/v) in Cami1 purification buffer supplemented with 2 mM PMSF and 40 mM imidazole and lysed by sonication. Lysates were cleared by centrifugation at 18 000 rpm 4 °C for 1 h using JA-20 rotor (Beckman Coulter). Supernatant was subjected to HisTrap 5 mL (Cytiva) column and proteins were eluted by linearly increasing imidazole concentration to 500 mM in Cami1 purification buffer. Fractions with eluted proteins were subjected to StrepTrap 1 mL (Cytiva) column and eluted with Cami1 purification buffer including 2.5 mM desthiobiotin. Eluted proteins were dialysed to Cami1 storage buffer and stored at -20 °C. Protein content was analysed by denaturing samples in protein gel loading solution and performing gel electrophoresis. Bands were depicted using PageBlue Protein Staining Solution (Thermo Scientific).

bL12. *E. coli* BL21(DE3) was transformed with pSV281-L12 (Appendix 1). Bacteria were incubated with shaking in LB broth supplemented with 50 μg/mL kanamycin at 37 °C until bacterial optical density OD_{600} reached 0.6. Cultures were induced with IPTG to final 0.5 mM concentration and incubated for 3 h at 37 °C. Biomass was collected by centrifugation at 7000 rpm 4 °C for 10 min using F12-6x500 LEX rotor (Thermo Scientific). Biomass was suspended (20% w/v) in bL12 purification buffer supplemented with 2 mM PMSF and 40 mM imidazole and lysed by sonication. Lysates were cleared by centrifugation at 18 000 rpm 4 °C for 1 h using JA-20 rotor (Beckman Coulter). Supernatant was subjected to HisTrap 5 mL (Cytiva) column and

proteins were eluted by linearly increasing imidazole concentration to 500 mM in bL12 purification buffer. Fractions with eluted proteins were subjected to TEV protease cleavage (mass ratio of TEV to bL12 equalling 1:100) overnight at 4 °C with cleavage reaction samples supplemented with 15 mM DTT. After cleavage the protein sample was subjected to desalting column (Cytiva) and proteins were eluted with bL12 purification buffer. Then the sample was subjected to HisTrap 1 mL (Cytiva) and the flow-through containing TEV-cleaved tag-less bL12 was collected. Eluted proteins were dialysed to bL12 storage buffer, snap frozen using liquid nitrogen and stored at -80 °C.

2.2.4. Preparation of RNA substrates

S3/a, S3/n and NS RNA. RNA substrates for StCsm binding and activity assays (Appendix 2) were produced by *in vitro* transcription. For this DNA matrices were amplified from pUC18_S3/2, pUC18_S3/n and pUC18_NS using appropriate DNA primers encoding a flanking T7 RNA polymerase promoter sequence at 5′-end of desired RNA coding region. Purified PCR products were used in the *in vitro* transcription reaction to obtain RNA substrates using TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific) according to manufacturer recommendations. RNAs were purified and concentrated using GeneJET RNA Cleanup and Concentration Micro Kit according to supplied protocol and stored at -80 °C.

For RNA binding experiments the substrates at 500 nM concentration were dephosphorylated using 0.05 u/μL FastAP in 0.5X PNK reaction buffer A for 10 min at 37 °C. Mixture was supplemented with EDTA to 5 mM and alkaline phosphatase was heat-inactivated for 10 min at 75 °C. PNK reaction buffer A was added to final 1.5X concentration and dephosphorylated RNA adjusted to 250 nM was 5′-radiolabelled using 0.25 μCi/μL [γ-^{32}P]-ATP (PerkinElmer) and 0.5 u/ μL T4 PNK (Thermo Scientific) for 30 min at 37 °C. Reaction mixture was again supplemented with EDTA to final 10 mM concentration, mixed with RNA Loading Dye (Thermo Scientific) and heated for 10 min and 75 °C. Electrophoresis of the reaction mixture in denaturing 20% PAAG was performed in 0.5X TBE to separate unincorporated label and gel pieces containing labelled substrates were excised. Gel was crushed and soaked with 1 mM EDTA solution overnight at 4 °C with shaking. Extracted RNA was purified using GeneJET RNA Cleanup and Concentration Micro Kit and concentration was estimated by comparing gel band intensities between gel-purified and unpurified substrates.

For fluorescence measurements each substrate was simultaneously hybridized to two fluorescently labelled DNA oligonucleotides “Green-probe” (ATTO532-labelled and complementary to 5′-end of S3/a, S3/n and NS RNA substrates) and “Red-probe” (ATTO647N-labelled and complementary to 3′-end of RNA substrates) (Appendix 2) by heating the RNA and DNA mixture to 95 °C and cooling to 4 °C at 1 K/min in 16 mM Tris-acetate, 33 mM K-acetate and 5 mM EDTA.

S3/iL RNA. Internally ATTO647N-labelled substrate S3/iL (Appendix 2) was ordered from Eurofins Genomics.

rib-UAG, rib-UAG-OMe, rib-NNN RNA. rib-UAG and rib-UAG-OMe RNA substrates were purchased from Metabion (Appendix 2). rib-UAG-OMe has the same sequence as rib-UAG with additional 2′-O-methyl (OMe) modifications in three consecutive nucleotides constituting “UAG” motif. rib-NNN was produced by *in vitro* transcription using TranscriptAid T7 High Yield Transcription Kit and purified as described above. For this a synthetic DNA duplex containing T7 promoter and a degenerate NNN motif in the template strand was used as transcription matrix.

These substrates were 5′-radiolabelled as described above except that for rib-UAG and rib-UAG-OMe the dephosphorylation step was skipped.

2.2.5. Electrophoretic mobility shift assays

50 pM of radiolabelled substrate (S3/a or NS) was incubated with 0.1–100 nM of WT StCsm complex in StCsm binding buffer for 10 min at room temperature. Reaction mixtures were analysed by electrophoresis on native 8% PAAG in 1X TAE and depicted using a phosphorimager. Each assay was performed in duplicate.

2.2.6. Fluorescence correlation spectroscopy

Time-resolved fluorescence correlation spectroscopy (FCS) measurements were carried out on a Microtime 200 setup (PicoQuant) in time-correlated single photon counting mode. Samples were excited in a confocal volume by two lasers at 532 nm and 638 nm with a power after all optical filters of 12.9 and 13.4 μW, respectively. The lasers were operated in pulsed interleaved excitation (PIE) mode with pulse widths of 50 ps at a repetition rate of 20 MHz. Reactions were mixed in a droplet on cleaned cover slides. Cover slides were sonicated for 10 min in acetone, 10 min in isopropanol and 20 min in 5 M KOH and afterwards thoroughly washed with milli-Q water and blown

dry with nitrogen. Reactions were prepared by mixing 10 nM (all final concentrations) of StCsm (WT or RNase-dead as indicated) with 2 nM of fluorescent DNA:RNA hybrid substrates or S3/iL (Appendix 2) in StCsm reaction buffer without $MnCl_2$ and incubated for 5 min. Reactions were started by adding $MnCl_2$ solution to final 0.5 mM concentration and fluorescence fluctuations were constantly registered at 25 °C.

All FCS measurements were analysed using PAM (PIE analysis with MATLAB), which is a software package that allows quantitative analysis of fluorescence microscopy data obtained using pulsed interleaved excitation (Schrimpf et al., 2018). Within the software, the cross-correlation (cross similarity) function $G_{ij}(\tau)$ is calculated according to:

$$G_{ij}(\tau) = \frac{\langle F_i(t)F_j(t+\tau)\rangle}{\langle F_i(t)\rangle\langle F_j(t)\rangle}, \tag{2.1}$$

where $\langle F_i(t)\rangle$ and $\langle F_j(t)\rangle$ are the time averaged fluorescence signals of the two fluorescent channels i and j, and $F_i(t)$ and $F_j(t+\tau)$ are the fluorescence intensities at a given real time-point (t) and at the time-point shifted by the lag time $(t+\tau)$ (Lakowicz, 2006). When data from a single fluorescence channel is analysed Equation 2.1 is reduced to autocorrelation (self-similarity) function by setting $i = j$ (Lakowicz, 2006):

$$G(\tau) = \frac{\langle F(t)F(t+\tau)\rangle}{\langle F\rangle^2}. \tag{2.2}$$

Exported correlation curves $G_{ij}(\tau)$ or $G(\tau)$ were then plotted and analysed using custom Python (version 3.8) scripts and fitted by three-dimensional molecule diffusion models (see text in Section 3.1.1) using a non-linear least squares method ("curve_fit"), which is part of the SciPy package (Virtanen et al., 2020). Cross- or autocorrelation analysis and subsequent diffusion model fitting was performed for each consecutive one-minute interval of fluorescence fluctuation signals allowing to track changes of labelled molecule integrity or diffusion during the real reaction time.

2.2.7. DNase activity assays

A fluorophore-quencher dual-labelled DNA oligonucleotide (FQ-DNA, Appendix 2) was used as a substrate in StCsm DNase activity assays. Reaction

mixtures were prepared in steps using StCsm reaction buffer. First 100 nM (all final concentrations) of FQ-DNA and 2 nM (unless indicated otherwise) of substrate RNA were preincubated for 5 min. Then WT or RNase-dead StCsm was mixed in to final 2 nM concentration, solution was transferred to 150 µl cuvette (Hellma Analytics). At 10 s post StCsm addition the cuvette was transferred to Cary Eclipse fluorescence spectrometer (Agilent) to record changes in fluorescence at 25 °C. The excitation wavelength was set to 495 nm and the emission wavelength was recorded at 518 nm, both with a slit width of 10 nm. The photomultiplier tube voltage was set to 550 V. The fluorescence signal was transformed to cleaved DNA using a linear scaling factor, which was obtained by measuring the saturated signal of known amounts of FQ-DNA being fully cleaved by micrococcal nuclease (New England Biolabs) (Appendix 4). For this FQ-DNA at concentrations of 10, 20, 30, 40, 50, 60, 80 and 100 nM were degraded by 10 U/µl micrococcal nuclease in supplied buffer supplemented with 0.5 mM $MnCl_2$. All fluorescence time traces were analysed using custom Python (version 3.8) scripts. All model fits (see text in Section 3.1.2) were performed using a nonlinear least-squares method ("curve_fit"), which is part of the SciPy package (Virtanen et al., 2020).

2.2.8.SEC-MALS

50-800 µg of each WT Cami1 protein sample was loaded on a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with SEC-MALS buffer. Light scattering signal in elutes was detected by miniDawn TREOS II detector, protein sample concentration was registered by Optilab T-rEX refractive index detector (Wyatt Technologies). Data were analysed in Astra software (Wyatt Technologies) using dn/dc value of 0.185 g/mL.

2.2.9.Protein crystallization and structure determination

For crystallization of WT AvCami1 an aliquot of protein was rebuffered using NAP-5 column with saline buffer (20 mM Tris-HCl (pH 8.5 at 25 °C), 150 mM NaCl, 3 mM DTT) and concentrated using 30 kDa cut-off centricons (Millipore) to the final concentration of 4-10 mg/mL. Crystals were prepared by mixing in a 1:1 ratio of the concentrated AvCami1 solution with the precipitating solution containing 0.1 M MOPS/HEPES (pH 7.5 at 25 °C), 0.2 M sodium formate, 0.2 M ammonium acetate, 0.2M sodium citrate, 0.2 M potassium sodium tartrate, 0.2 M sodium oxamate, 10% (v/v) PEG 20K, 40% (v/v)

PEG 400. Drops were equilibrated against the reservoir solution. Crystals grew at 19 °C in sitting-drop vapor-diffusion crystallization plates. The native apo-Cami1 dataset was collected to the nominal resolution of 1.7 Å at the EMBL/DESY Petra III P14 beamline (Hamburg, Germany) at 100 K. The SAD dataset of selenomethionine labelled Cami1 was collected at DESY Petra III P11 beamline (Hamburg, Germany). XDS (Kabsch, 2010), SCALA and TRUNCATE (Agirre et al., 2023) were used for data processing. The data collection and refinement statistics are presented in Appendix 5. The structure was solved using SAD protocol of Auto-Rickshaw, the EMBL-Hamburg automated crystal structure determination platform (Panjikar et al., 2005). The input diffraction data were prepared and converted for use in Auto-Rickshaw using programs of the CCP4 suite (Agirre et al., 2023). FA structure factors values were calculated using the program SHELXC (Sheldrick, 2010). Heavy atoms positions were found using the program SHELXD (Sheldrick, 2010). The correct hand for the substructure was determined and initial phases were calculated by SHELXE (Sheldrick, 2010). The initial phases were improved using the density modification program PARROT (v. 1.0.4) (Agirre et al., 2023). The initial model containing 739 residues was built by the program BUCANEER (v. 1.5) (Agirre et al., 2023). The model was improved by several cycles of refinement in Phenix (v. 1.19.2) (Afonine et al., 2012) and manual inspection in Coot (Emsley and Cowtan, 2004). Structural images and snapshots were prepared using PyMOL (Schrödinger, LLC, 2015).

2.2.10. Ring nuclease activity assays

Cyclic tetra-adenylate (cA_4) and cyclic hexa-adenylate (cA_6) were purchased from Biolog Life Science Institute. 10 μM of cA_4 or cA_6 was incubated with WT Cami1 proteins or AvCami1 CARF domain mutants at final 1 μM concentration in RN buffer for 2 h at 37 °C. Reactions were stopped by freezing with liquid nitrogen and subjected to HPLC-MS analysis (Methods 2.2.11). Reaction products were identified based on accurate mass and retention times. For time course experiments a larger reaction volume was set up, reactions were stopped at different time points (0, 4, 8, 16, 32, 64 min) and processed as above.

cA_4 cleavage kinetics. To directly visualize the hydrolysis of cA_4, α^{32}P-labeled cA_4 was prepared by mixing 0.2 μM *Candidatus* C. acidaminovorans Csm complex with 10 μM CCa target RNA, 0.5 mM ATP and 1 μM [α-^{32}P]ATP in cOA synthesis buffer and incubating at 37 °C for 4 h. Synthesis products were separated by denaturing PAGE (24% 19:1

acrylamide:bisacrylamide, 6 M urea in 0.5X TBE) and cA_4 was purified from the gel by phenol extraction and ethanol precipitation. α^{32}P-cA_4 hydrolysis reactions were conducted in RB buffer and contained 20 nM of α^{32}P-cA. Reactions were started by adding 200 nM of WT AvCami1 and were carried out at 37 °C. 2 µl aliquots were taken at indicated time points and reactions were quenched by mixing with equal amount of 2X RNA Gel Loading Dye (Thermo Fisher Scientific) preheated to 75 °C. Reaction products were analysed by denaturing PAGE (30% 19:1 acrylamide:bisacrylamide, 6 M urea in 0.5X TBE) and visualized by autoradiography. For single turnover reactions, k_{obs} were determined by fitting a single exponential to the substrate depletion data.

2.2.11. HPLC-MS analysis

To analyze the cA_4 and cA_6 hydrolysis products, electrospray ionization (ESI) mass spectrometry (MS) was performed in negative mode using an integrated reverse phase HPLC/ESI-MS system (1290 Infinity, Agilent Technologies/Q-TOF 6520, Agilent Technologies). For the cA_4 and cA_6 cleavage product analysis a Supelco Discovery®HS C18 column (7.5 cm × 2.1 mm, 3 µm) was used. Elution was performed with a linear gradient of solvents A (5 mM ammonium acetate in water, pH 7.0) and B (5 mM ammonium acetate in methanol, pH 7.0) at a flow rate of 0.3 ml/min at 30 °C as follows: 0–2 min, 0% B; 2–22 min, 20% B; 22–25 min, 50% B, 25–29 min 100% B. Ionization capillary voltage was set to 5000 V and fragmentor voltage was set to 150 V. The results were analyzed with Agilent MassHunter Qualitative Analysis software. The cA_4 and cA_6 hydrolysis products were annotated based on accurate mass, retention time from synthesized standards and deprotonated molecular ions (Appendix 6).

2.2.12. *In vitro* RNA cleavage assays

RNA cleavage assays were adapted from (Griffin et al., 2013) with modifications. 5′-radiolabelled rib-UAG, rib-UAG-OMe or rib-NNN substrates used in these assays were prepared as described in Methods 2.2.4. For ribosome complex formation 40 nM of labelled RNA substrate was mixed with 60 nM of unlabelled RNA substrate, 300 nM of *E. coli* 70S ribosomes and 400 nM of $tRNA^{fMet}$ in RB buffer and incubated for 30 min at 37 °C. Ribosome complexes were 2.5-fold diluted in RB buffer, supplemented with 10 µM cA_4 (final concentration) and reactions were initiated by adding an equal volume

of Cami1 WT or mutant proteins to 300 nM final concentration if not indicated otherwise. Single endpoint reactions were performed at 37 °C for 15 min. Reactions were quenched by adding an equal amount of 2X RNA Gel Loading Dye (Thermo Fisher Scientific) and heating at 70 °C for 10 min. Reaction products were separated on a denaturing 15% polyacrylamide gel and depicted by phosphorimaging.

RNA markers. To map the cleavage products, oligoribonucleotide markers were generated by RNase T1 (Thermo Fisher Scientific, 0.006 U/µL final concentration) treatment of rib-UAG for 15 min at 50 °C or by alkaline hydrolysis in 50 mM $NaHCO_3$ (pH 9.5), 1 mM EDTA at 95 °C for 15 min.

RelE mutant activity. For kinetic analysis 3 µM of WT AvCami1 protein or its RelE domain point mutants were used in reaction mixtures with ribosome complexes programmed with rib-UAG substrate as described above. Reactions were incubated at 15 °C for up to 2 h. Aliquots were removed at timed intervals and processed as described above. Decline of RNA substrate over reaction time was evaluated by densitometric analysis of gel autoradiographs. Reaction rate constant k_{obs} for WT and mutants were determined by fitting the data points to a single exponential:

$$S(t) = (100\% - S_\infty)e^{-k_{obs}t} + S_\infty, \qquad 2.3$$

where $S(t)$ is the proportion (in %) of RNA substrate at time t, k_{obs} is observed cleavage rate constant and S_∞ is the proportion of uncleaved RNA. Fitting was done using KYPLOT 2.0 software (Yoshioka, 2002). Mean and standard deviation of k_{obs} is reported from three technical replicates for each variant.

Different activators. Cleavage of rib-UAG RNA substrate by AvCami1 in the presence of different activators was done in activator concentration series. For this ribosome complexes were formed: 40 nM of labelled rib-UAG was mixed with 60 nM of unlabelled rib-UAG, 300 nM of *E. coli* 70S ribosomes and 400 nM of $tRNA^{fMet}$ in RB buffer and incubated for 30 min at 37 °C. Ribosome complexes were 2.5-fold diluted in RB buffer, supplemented with 0, 1, 10, 100, 1000 or 10000 nM (final concentrations) of either cA_4 or A_4>p and reactions were initiated by adding an equal volume of WT AvCami1 to 300 nM final concentration. Reactions were incubated for 15 min at 37 °C, quenched and analysed as described above.

Multiple turnover conditions. For reactions under multiple turnover conditions 50 nM of WT AvCami1 or S11A mutant was mixed with 1 µM

ribosomes programmed with 1.3 μM $tRNA^{fMet}$, 490 nM unlabelled rib-UAG and 10 nM labelled rib-UAG and 10 μM cA_4 (all final concentrations). Reactions were incubated at 37 °C. Aliquots were removed at timed intervals, processed and analysed as described above.

Substrate variants. Cleavage of rib-UAG-OMe RNA substrate which has 2′-OMe modified UAG codon nucleotides was tested in the same reaction setup as above using 1.5 μM of different WT Cami1 proteins. Similarly, the cleavage of rib-NNN substrate in which UAG codon is replaced by three random ribonucleotides was performed using 1.5 μM of Cami1 proteins. Reactions were incubated for 15 min at 37 °C. Reactions were quenched and analysed as described above. Proportion of cleaved substrate was determined by densitometric analysis of the autoradiographs.

2.2.13. Cryo-EM

Sample preparation. 500 nM (final concentrations are denoted) *E. coli* 70S ribosomes were mixed with 2.5 μM rib-UAG-OMe RNA substrate (Appendix 2) and 1 μM $tRNA^{fmet}$ in cryoRB buffer and incubated for 15 min at 37 °C. The programmed ribosomes were mixed with 10 μM cA_4 and 5 μM AvCami1-S11A-H343A, incubated for 5 min at 15 °C and chilled on ice for up to one hour before loading on grids. 1.2/1.3 Cu 300 mesh grids (Quantifoil) were glow discharged for one minute at 20 mA and coated with graphene oxide. For this graphene oxide dispersion (Sigma) was diluted in deionised water to 0.2 mg/mL and spun down at 300g for 10 s to remove aggregates. Supernatant was applied to a glow discharged grids, incubated for one minute and removed using blotting paper. Grids were washed three times with deionised water and dried for 5 min on blotting paper. Ribosome complex was applied to prepared grids and frozen in liquid ethane using Vitrobot Mark IV (Thermo Fisher) set at 4 °C and 95% humidity with 30 s wait time, 5 s blot time and +5 blot force. Two grids were prepared side-by-side using the same sample.

Data collection. Data collection from both grids was performed using EPU v3.0 software (Thermo Fisher Scientific) with the same settings. Grids were imaged on a Glacios Cryo-TEM (Thermo Fisher Scientific) operating at 200 kV equipped with a Falcon 3EC direct electron detector (Thermo Fisher Scientific) in counting mode at a nominal magnification of 92,000× corresponding to a calibrated pixel size of 1.12 Å. Exposures of 46.33 s were dose-fractionated into 30 frames with a total dose of 30.58 electrons/$Å^2$. The defocus targets ranged from −1.0 to −2.0 μm.

Image processing. Data processing was performed using cryoSPARC (v. 3.3.1) (Punjani et al., 2017). For both datasets patch-based motion correction and contrast transfer function (CTF) estimation was performed and micrographs with poor CTF fit or contamination were discarded. Particles of 300-350 Å in size were picked using an automatic blob-based picker and extracted using a box size of 400 px. Particles from both datasets were combined (totalling 642,401 particles) and processed as displayed in Appendix 7. After 2D classification and subsequent homo- and heterogeneous refinement steps, particles were split into five groups with only one (277,364 particles) showing faint density of AvCami1 in the A-site of 70S ribosome. This group was subjected to focused 3D classification without alignment using a mask covering A-site of the ribosome. Classes in which A-site was occupied by AvCami1 (158,387 particles) were selected for local motion correction and ribosome complex reconstruction. Using this set three local refinement jobs were performed in cryoSPARC by focusing on different parts of the ribosome complex: A-site, 30S subunit with EPA-sites, or 50S subunit. A composite map of the ribosome-AvCami1 complex was prepared using phenix.combine_focused_maps function in Phenix (v. 1.20.1-4487) and sharpened using phenix.auto_sharpen (Terwilliger et al., 2018). Map resolution was calculated using separately combined corresponding half-maps.

Model building and refinement. *E. coli* ribosome at a resolution of 2.00 Å (PDB ID: 7K00) (Watson et al., 2020) was rigid-body fitted into the composite map using UCSF-ChimeraX (v. 1.3) (Pettersen et al., 2021), then each protein and RNA chain was fitted separately. Additional ribosomal proteins bL10 and bL11 were fitted from another *E. coli* ribosome structure (PDB ID: 4YBB, chains DI and DJ, respectively) (Noeske et al., 2015). Starting model for $tRNA^{fMet}$ was taken from the ribosome structure (PDB ID: 5IQR) (Brown et al., 2016), however the map for the tRNAs is not well resolved, probably due to partial occupancy. Our crystal structure of the AvCami1 dimer was fitted in the remaining portion of the map using UCSF-ChimeraX. The AvCami1 structure was manually adjusted in the composite map using Coot (v. 0.9.7) (Emsley and Cowtan, 2004). The C-terminal domain of the 50S ribosomal protein L12 (PDB ID: 1RQS, residues 54-119) (Bocharov et al., 2004) was fitted near the wHTH domain of AvCami1 using UCSF-ChimeraX and manually adjusted in Coot. The map of mRNA was poor and allowed modelling only of 3 nt interacting with the P-site $tRNA^{fMet}$. The final model was obtained by several rounds of manual adjustment in Coot and real-space refinement in Phenix. Details of data collection, processing, and refinement statistics are summarized in Appendix 8.

2.2.14. bL12 interaction with AvCami1 assays

To assess Cami1 mRNA interferase activity dependence on bL12 we used WT *E. coli* 70S and 70SΔbL12 ribosomes and *E. coli* bL12 protein provided by (Carlson et al., 2017). WT, 70SΔbL12, reconstituted 70S ribosomes or a standalone bL12 protein were used to stimulate Cami1 activity. Reconstituted ribosomes were prepared before assays by incubating 400 nM 70SΔbL12 ribosomes with 4 μM bL12 in cryoRB buffer for 15 min at 37 °C. Single-point assays were performed for 15 min at 37 °C as described above using 40 nM of different ribosome complexes (or 400 nM for bL12 protein) with 20 nM of mRNA in cryoRB buffer. Reactions were initiated by adding 300 nM of Cami1 proteins (except for CCaCami1 used at 3000 nM). Reactions were supplemented with 10 μM cA_4.

For the analysis of AvCami1 wHTH domain mutants we incubated 10, 100 or 1000 nM of each mutant or WT AvCami1 protein with a preformed ribosome complex in RB buffer for 15 min at 37 °C in the presence of 10 μM cA_4. Reaction products were analysed using denaturing gel electrophoresis as described in 2.2.12. The mean and standard deviation of the percentage of cleaved substrate are reported from three technical replicates for each AvCami1 concentration.

2.2.15. Bio-Layer Interferometry (BLI)

BLI experiments were performed using an Octet K2 system (ForteBio). His-tagged S11A, S11A+T213R, S11A+Q210A+E212A or S11A+Q210A+E212A+E238A AvCami1 mutants was immobilized on Ni^{2+}-NTA sensor chips (ForteBio) as the ligand and the *E. coli* bL12 protein was used as the analyte. NTA biosensors were hydrated for 10 min at 25 °C in kinetics buffer prepared by supplementing RB buffer with 0.01% Triton X-100, 10 mM imidazole and 0.1% BSA. Experiments were set up in 96-well plate format using 200 μL reagent volumes. After a baseline step of 120 seconds, AvCami1 was immobilized on the NTA biosensors at a concentration of 200 nM for 120 seconds with the stage rotation rate set at 1000 rpm. A second baseline step of 120 seconds followed the immobilization to wash unbound AvCami1 and to allow for signal stabilization. The 60 seconds association, followed by 180 seconds dissociation protocol was used. The association was monitored by transferring the ligand biosensors to wells containing bL12 in a concentration of 500 nM, 1000 nM, 2000 nM and 4000 nM. A biosensor without AvCami1 ligand was used as a reference. Prior to each

measurement in series with increasing analyte concentration biosensors were regenerated by three cycles of 5 seconds incubation in 10 mM glycine pH 1.7 and 5 seconds in the kinetics buffer and recharged by incubation for 60 seconds in 10 mM $NiCl_2$. Assays were performed at 25 °C and mixtures were supplemented with 1 μM of cA_4 in appropriate measurement series. We report the amplitudes of bL12 binding response normalized by the amplitude of AvCami1 ligand loading response.

2.2.16. *In vivo* survival assay

Bacteria expressing *in vivo* stimulated StCsm were prepared as previously with modifications (Smalakyte et al., 2020). For the expression of StCsm proteins pStCsm_dNucleases plasmid was used. pCRISPR was used for expression of four crRNAs complementary to *Aequorea victoria* GFP mRNA encoded by pTarget. pTarget, pCRISPR and pStCsm_dNucleases were electroporated in concert into *E. coli* ER2566. Bacteria were recovered for 1 h in LB shaking at 37 °C and plated on LB agar with 25 μg/mL kanamycin, 30 μg/mL chloramphenicol, 25 μg/mL streptomycin and 1% w/v glucose. Transformants were confirmed by colony PCR. Single colonies of bacteria were picked to prepare chemically competent bacteria which were transformed with pAvCami1, pCaCami1, pCCaCami1, pChCami1, pAvCami1-S11A or pAvCami1-S11A-K317A mutant plasmids or empty pBAD/HisA vector. Overnight inoculates prepared from single colonies in LB with appropriate antibiotics and 1% glucose were 1:200 diluted in LB with antibiotics and incubated in a 96-well plate for 1 hour with shaking at 37 °C. IPTG and arabinose were added to each well to 0.1 mM and 0.2% final concentrations, respectively, and bacterial growth was monitored for 12 h by measuring OD_{600} every 10 min. Means with standard deviations of three biological replicates are reported for each experiment.

2.2.17. Northern blot analysis of mRNA levels in *E. coli*

Overnight *E. coli* ER2566 cultures carrying pTarget, pCRISPR, pStCsm_dNucleases and Cami1 variant encoding plasmid (pAvCami1 pCaCami1, pCCaCami1, pChCami1, pAvCami1-S11A, pAvCami1-S11A-K317A) or empty pBAD/HisA vector were 1:100 diluted with selective LB containing 1% w/v glucose and incubated at 37 °C with shaking for 3 h. 6 OD_{600} × mL units (12-15 mL) of each variant were collected by centrifugation at 3000 g for 5 min, LB with glucose was discarded. Bacteria were

resuspended in 10 mL of selective LB with 1 mM IPTG and 0.2% arabinose and additionally incubated at 37 °C for 1 h with shaking. RNA was extracted using RNAsnap method with adaptations (Stead et al., 2012). Briefly, cells were pelleted by centrifugation and resuspended in 500 μL of RNAsnap solution. Mixtures were incubated for 10 min in boiling hot water bath, centrifuged and the supernatant was collected. The RNA was separated using 6% denaturing polyacrylamide gel and blotted onto Hybond-N+ membranes (Amersham) using semi-dry blotting apparatus. Membranes were hybridized with 5′-radiolabeled DNA oligonucleotide probes complementary to endogenous *E. coli* genes (Appendix 2) in hybridization solution overnight at 40 °C with rotation. Membranes were washed three times with washing solution for 10 min at 40 °C. Probed RNA was visualized by phosphorimaging. DNA oligonucleotides were stripped from the membranes with boiling hot 0.1% SDS solution and the same membranes were reused for hybridization with different probes.

3. RESULTS AND DISCUSSION

Currently, type III CRISPR-Cas systems are divided into six subtypes from III-A to III-F (Makarova et al., 2020b). Of these, type III-A and III-B systems are well understood, and their effector complexes are characterized at structural and functional levels as summarized in Chapter 1.3.

The research group led by dr. Gintautas Tamulaitis is advancing the field of type III CRISPR-Cas by studying type III-A effector complex from *Streptococcus thermophilus* DGCC8004 (StCsm). Our studies showed that StCsm recognizes crRNA-complementary RNA targets which are cleaved by complex RNase Csm3 (Tamulaitis et al., 2014). Upon target RNA binding ssDNase (Kazlauskiene et al., 2016) and polymerase-cyclase (Kazlauskiene et al., 2017) activities are unleashed in Cas10 subunit. Functional analysis of StCsm deletional mutants revealed the roles of each subunit for these activities (Mogila et al., 2019). In addition to this we demonstrated that cA_6 synthesized by StCsm Cas10 polymerase-cyclase activates accompanying ancillary CARF-HEPN RNase Csm6 (Kazlauskiene et al., 2017; Smalakyte et al., 2020) exemplifying the cOA signalling pathway covered in Chapter 1.4.

This work is the continuation of type III CRISPR-Cas and StCsm research. I begin this part of the thesis by presenting the results of StCsm Cas10 activity regulation studies using fluorescence methods (Chapter 3.1). In the second part I present the structural and functional characterization of a novel ancillary CARF protein Cami1 which inhibits translation in response to cOA signalling (Chapter 3.2).

Describing the performed experiments and obtained results I use the term "we" extensively. While it may seem odd to use the plural number in a doctoral thesis which is a personal work, the input from my colleagues was critical to achieve the research goals and could not be left unnoticed. Therefore, the results are presented as "ours" and individual contributions of my colleagues are mentioned where appropriate in the main text and emphasized in the Acknowledgements. This work is based on two publications with my personal contributions outlined in the List of publications.

3.1. StCsm complex regulation by target RNA

Previously, it was observed that StCsm, when activated by target RNA, slowly loses its ssDNase (Kazlauskiene et al., 2016) and polymerase-cyclase (Kazlauskiene et al., 2017) activities over time. This effect was linked to target

cleavage, however the timescales for these processes do not match – Cas10 activity diminishes in terms of hours, while RNA is cleaved within seconds (overviewed in Section 1.3.3). Since RNA binding, but not cleavage, is required for type III effector complex activation, we hypothesized that RNA cleavage products could remain bound to the complex which keeps it in the activated state. To test this, we decided to directly track RNA dissociation using fluorescence correlation spectroscopy and compare it with temporal decrease of Cas10 activity by measuring ssDNA cleavage rate dynamics.

3.1.1. Terminal RNA cleavage products are retained by StCsm

RNA is cleaved by StCsm in five positions every 6 nt by Csm3 which yields several short internal fragments and two terminal fragments located at both ends of the complex (Tamulaitis et al., 2014). To track the release of the different RNA fragments from StCsm, we utilized a dual-colour FCS setup in pulsed interleaved excitation mode, which allows monitoring the diffusion properties of multiple biomolecules simultaneously. Within the setup, double-fluorescently labelled RNA molecules are alternatingly illuminated with green and red laser pulses within the diffraction limited spot of the laser focus and the resulting fluorescence emission is measured in separate channels (see Methods 2.2.6). Fluctuations of fluorescence signal correspond to labelled RNA constantly entering and leaving the illuminated confocal volume either bound or released from the complex (Fig. 3.1A). Processing the resulting fluctuation trajectories using correlation analysis allows to monitor the proportions of bound and released RNA fragments after the cleavage.

For these experiments we produced an StCsm complex devoid of purification tags and enriched with homogenous crRNA containing S3 spacer from *S. thermophilus* DGCC8004 CRISPR2 region (Tamulaitis et al., 2014 and Methods 2.2.3). We designed 180 nt RNA substrates (Appendix 2) to be used for fluorescent labelling and assessed their binding to WT StCsm by electrophoretic mobility shift assays (EMSA). Substrates S3/a and S3/n are complementary to S3 crRNA and show high affinity to WT StCsm compared to non-complementary NS substrate (Fig. 3.1B). S3/a is not complementary to 5′-tag allowing it to activate StCsm Cas10 protein when bound in the complex while S3/n, being complementary to the whole crRNA including the 5′-tag, mimics CRISPR region antisense transcript imitating self-targeting scenario and Cas10 inhibition (see Section 1.3.3 and Appendix 3). Overall, EMSA shows that crRNA-complementary substrates used in this study bind tag-less StCsm

with K_d being in subnanomolar range as observed previously (Mogila et al., 2019; Tamulaitis et al., 2014).

For FCS we labelled the substrates by annealing 5′-ATTO647N-labelled ("Red-probe") and 3′-ATTO532-labelled ("Green-probe") DNA oligonucleotides to the 3′- and 5′-ends of the RNA, respectively (Appendices 2 and 3). Upon binding by StCsm, the red ATTO647N label to be located on the Cas10-proximal and green ATTO532 on the Cas10-distal side of the complex.

Before each FCS assay, StCsm complexes (at 10 nM) were pre-incubated with RNA substrates (at 2 nM) in the absence of divalent ions to ensure full crRNA-complementary substrate binding facilitated by the low StCsm K_d to target RNA. Then, the RNase reaction was initiated by the addition of Mn^{2+} (in form of $MnCl_2$) which enables rapid RNA cleavage (Tamulaitis et al., 2014). If the complex is bound to intact RNA or both terminal cleavage products, it often diffuses with both labels through the confocal volume resulting in mutually similar fluctuation patterns between red and green fluorophores displaying high cross-correlation. Upon terminal product release by StCsm, the fluorophores begin to travel through the laser focus separately in uncorrelated fashion (Fig. 3.1A). Most of the FCS measurements were performed by dr. Patrick Irmisch.

We first measured the RNA cleavage product release kinetics for the activating target RNA S3/a substrate. To estimate the quotients of joint (either not cleaved or bound together by StCsm after cleavage) and isolated diffusion of the RNA ends, we calculated cross-correlation functions between the two fluorescence signals (solid lines in Fig. 3.1C, see Methods 2.2.6). Cross-correlation functions were fitted by a three-dimensional diffusion model to describe the correlation signal G depending on the lag time τ:

$$G(\tau) = G_0 \frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)} \frac{1}{\sqrt{1 + \frac{\tau}{k^2 \tau_D}}}, \tag{3.1}$$

where G_0 is the cross-correlation amplitude, τ_D is the mean diffusion time and $k = 4.9$ (as obtained by the manufacturer) an instrument specific constant describing the vertical stretching of the confocal volume compared to a sphere (dashed lines in Fig. 3.1C, see Methods 2.2.6). The amplitude of the cross-correlation curve is proportional to the concentration of double-labelled species (Bacia and Schwille, 2007). In this case, G_0 was proportional to the

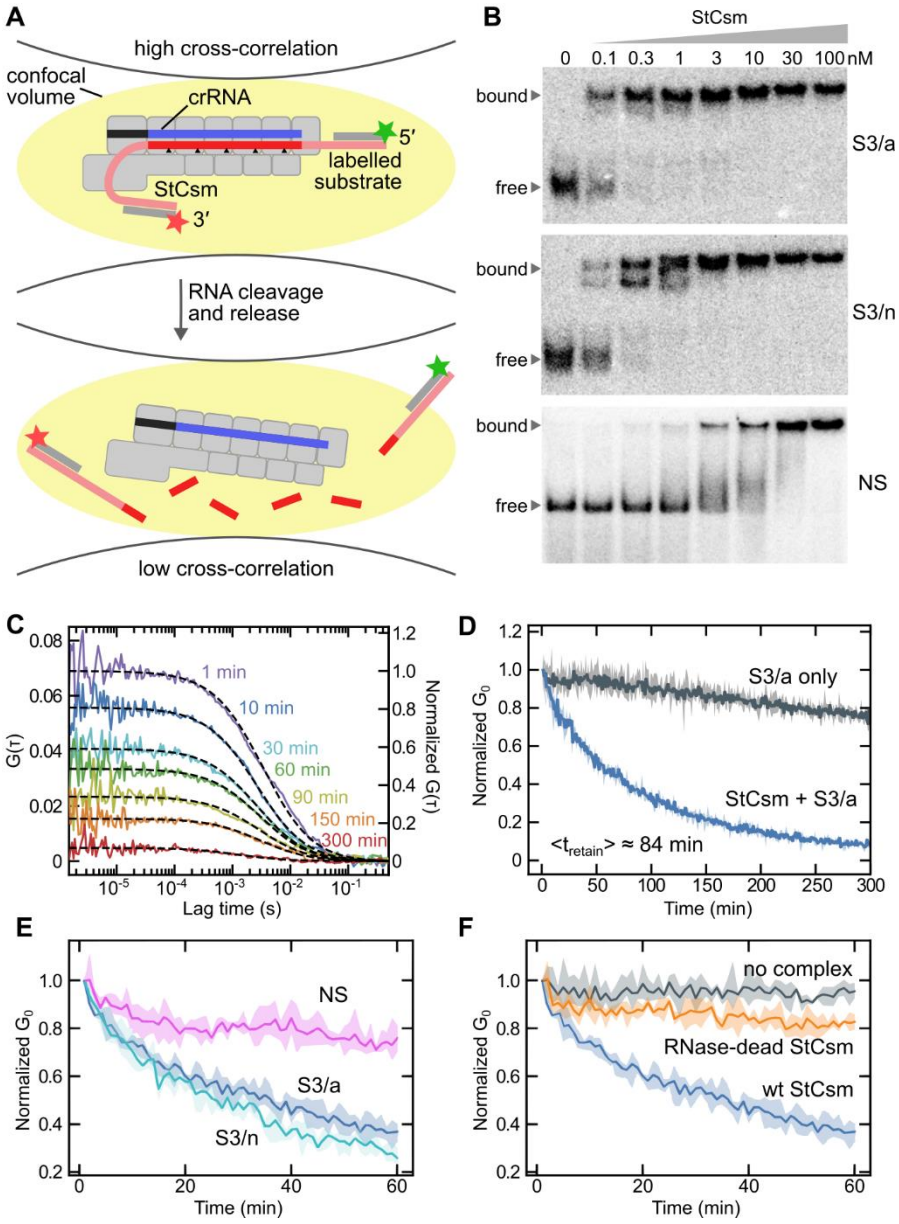


Fig. 3.1. Dual colour FCS analysis of RNA cleavage product retention by StCsm. (**A**) Scheme of StCsm and labelled RNA diffusing through the confocal excitation volume in the reaction mixture. 3′ and 5′ ends of RNA substrate are indicated. (**B**) StCsm affinity to RNA substrates assessed by EMSA. Each assay was performed twice. (**C**) Fluorescence cross-correlation curves as function of lag time measured for S3/a substrate at different time points of the reaction (coloured solid lines). Each curve was calculated from one-minute intervals of fluorescence signal fluctuations of a five-hour long reaction course. Fits of the data to three-dimensional diffusion model are

shown as dashed lines. Representative set of curves of three independent measurements are displayed. (**D**) Cross-correlation amplitude G_0 obtained from one-minute fluctuation traces as a function of wall-clock reaction time for the activating target RNA S3/a in absence (grey) and in presence of StCsm (blue). (**E**) Cross-correlation amplitudes of WT StCsm reactions on different labelled RNA substrates: S3/a (blue), S3/n (cyan), NS (magenta). (**F**) G_0 function of reaction time in reaction mixtures containing S3/a RNA with either WT StCsm (blue), RNase-dead StCsm (orange) or no complex (grey). In (D), (E) and (F) solid lines indicate the mean of three independent measurements, while coloured areas indicate the range between minimum and maximum measured values at each time point. All traces were normalized by the first data point. NS – non-specific RNA substrate.

concentration of joint RNA ends in the reaction mixture at selected one-minute time intervals (Fig. 3.1C). Plotting the G_0 as function of wall-clock reaction time over a five-hour long reaction course revealed an exponential decay (blue curve in Fig. 3.1D). As a control, we measured the same RNA substrate in the absence of StCsm displaying only a small decrease of G_0 during the incubation (grey curve in Fig. 3.1D). To extract quantitative parameters, we normalized the cross-correlation amplitude in the presence of StCsm by the amplitude measured in the absence of it. Normalized data was described by a single exponential decay:

$$R(t) = (1 - R_\infty)e^{-t/\langle t_{retain}\rangle} + R_\infty, \tag{3.2}$$

with $R(t)$ being amplitude at time t, R_∞ being remaining amplitude at infinite time reflecting uncleaved or unreleased RNA and $\langle t_{retain}\rangle$ being the mean RNA end retention duration. Fits to three independent measurements yielded $\langle t_{retain}\rangle$ = 84 ± 1 min.

We conducted a series of experiments with different RNA substrates showing that crRNA-complementary substrate (either S3/a or S3/n) cleavage by WT StCsm gives a substantial reduction of G_0 over time while StCsm incubation with non-specific NS substrate displays only modest drop of G_0 signifying no cleavage (Fig. 3.1E). While S3/n is complementary to both spacer and 5′-tag sequence of crRNA, this does not affect terminal cleavage product retention in cross-correlation mode compared to S3/a, which is not complementary to 5′-tag. Reaction mixtures of S3/a incubated with RNase-dead StCsm or without a complex show modest reduction of G_0 compared to those with WT StCsm (Fig. 3.1F). These results confirm that the substantial drop of cross-correlation amplitude is caused by specific RNA cleavage by Csm3 subunit of the complex and subsequent diffusion.

We explored the RNA fragment release further and investigated whether the drop of the cross-correlation amplitude could be attributed to the release of a single RNA end at one of the sides. To accomplish this, we reanalysed the fluorescence data of WT StCsm cleavage of double-labelled S3/a substrate by calculating separate autocorrelation functions (see Methods 2.2.6) for either the red or the green detection channel, corresponding to Cas10-proximal 3′-end terminal product or Cas10-distal 5′-end terminal product, respectively (Fig. 3.2A and Appendix 3). A shift of autocorrelation curves towards shorter timescales during the reaction course was observed. This is consistent with gradual accumulation of RNA ends released from the StCsm, since the free RNA ends diffuse faster through the confocal volume due to their smaller size (Ries and Schwille, 2012). For a quantitative description of the autocorrelation curves, we expanded the previous model by including two diffusing species to account for slowly diffusing RNA ends bound to StCsm, and rapidly diffusing free RNA ends.

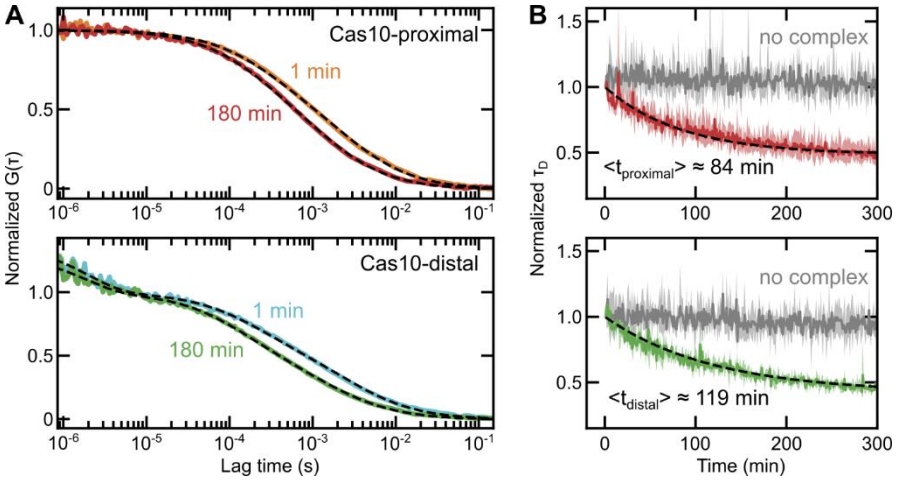


Fig. 3.2. Fluorescence autocorrelation spectroscopy discerns the retention of terminal RNA cleavage products. (**A**) Autocorrelation curves of the Cas10-proximal (red ATTO647N label) and Cas10-distal (green ATTO532 label) terminal products of the activating target RNA S3/a measured at 1 min and 180 min after reaction start (solid lines). Dashed lines show fits to models (see text) including slowly and rapidly diffusing species. Representative curves of three independent measurements are displayed. (**B**) Normalized (according to the first datapoint) mean diffusion time plotted as a function of reaction time for the Cas10-proximal and Cas10-distal end of the activating target RNA S3/a in the presence (red and green solid lines, respectively) and absence (grey solid lines) of StCsm. Fitted single exponential decay models (Equation 3.5) are shown as dashed lines.

Autocorrelation curves of the red fluorophore was fitted using updated model including two diffusing species with separated amplitudes G_1 and G_2 and diffusion times τ_1 and τ_2:

$$G(\tau) = G_1 \frac{1}{\left(1+\frac{\tau}{\tau_1}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_1}}} + G_2 \frac{1}{\left(1+\frac{\tau}{\tau_2}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_2}}}. \quad (3.3)$$

Similarly, for green autocorrelation a model accounting for two species was also utilized. Yet, a noticeable portion of green dyes were excited to the triplet state. The fluorophores do not exhibit fluorescence when in triplet state which lasts for a few microseconds in aqueous solutions (Ha and Tinnefeld, 2012). This blinking causes the increase of autocorrelation in microsecond range (Fig. 3.2A). Since it is not related to molecular diffusion occurring at the timescales of milliseconds, the model was expanded by an additional amplitude T and triplet-state lifetime τ_T (Lakowicz, 2006):

$$G(\tau) = \left(G_1 \frac{1}{\left(1+\frac{\tau}{\tau_1}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_1}}} + G_2 \frac{1}{\left(1+\frac{\tau}{\tau_2}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_2}}} \right) \left(1 + \frac{T}{1-T} e^{-\tau/\tau_T}\right). \quad (3.4)$$

Models agreed the experimentally determined autocorrelation curves over the full range of lag times (dashed lines in Fig. 3.2A). From the best-fit parameters, we calculated weighted mean diffusion time of both species τ_D for each consecutive one-minute interval of the reaction course:

$$\tau_D = \frac{G_1\tau_1 + G_2\tau_2}{G_1 + G_2}. \quad (3.5)$$

We plotted τ_D as a function of the reaction wall-clock time as a measure for the release of the respective RNA end (solid lines in Fig. 3.2B). The mean diffusion time again showed an exponential decay during the reaction in the presence of WT StCsm, while in its absence only an insignificant change could be observed. We fitted a single exponential decay model (analogous to Equation 3.2, marked by dashed lines in Fig. 3.2B) to the obtained time courses and obtained a mean retention duration $\langle t_{proximal} \rangle$ = 84 ± 2 min of Cas10-proximal product and $\langle t_{distal} \rangle$ = 119 ± 3 min of Cas10-distal product.

Assuming independent release, the mean retention duration for at least one of the ends could be obtained by:

$$\frac{1}{\langle t_{retain,ind} \rangle} = \frac{1}{\langle t_{proximal} \rangle} + \frac{1}{\langle t_{distal} \rangle}, \tag{3.6}$$

which would give $\langle t_{retain,ind} \rangle \approx 50$ min. However, the observed duration for both end release is $\langle t_{retain} \rangle = 84 \pm 1$ min (from Fig. 3.1D), which apparently is limited by Cas10-proximal product release. This disagreement suggests some degree of cooperation or interdependency between Cas10-proximal and distal RNA end product release requiring additional investigation. Nevertheless, autocorrelation analysis confirmed that both terminal ends are retained for more than an hour after cleavage.

Next, we tested on what timescales the internal fragments are released. To probe this, we used a 68 nt RNA substrate that carried an internal ATTO647N fluorophore (S3/iL, Appendix 2). The label is located within the binding region of Csm3.3 in StCsm complex (Appendix 3). We measured the fluorescence signal after the addition of Mn^{2+} ions to reaction mixture containing WT StCsm and S3/iL. Calculating autocorrelation functions revealed that the ternary complex with the S3/iL initially exhibits a comparable diffusion behaviour as the ternary complex with the end labelled target RNA (yellow curve in Fig. 3.3A). During the reaction, the autocorrelation curve shifted towards lower timescales significantly faster than that of the end labels (blue curve in Fig. 3.3A). To quantify our observations, we fitted a two-species diffusion model (Equation 3.3). The resulting fits reproduced the autocorrelation curves well (dashed lines in Fig. 3.3A). We calculated the mean diffusion time of both species (Equation 3.5) for the whole 60 min reaction course processed as previously in one-minute intervals. Plotting the resulting mean diffusion times versus time suggested a biphasic decay (blue trace in Fig. 3.3B). We fitted a biphasic decay model using a sum of two exponential decays to the diffusion time $\tau_D(t)$ data:

$$\tau_D(t) = A_1 e^{-t/\langle t_1 \rangle} + A_2 e^{-t/\langle t_2 \rangle} + \tau_\infty, \tag{3.7}$$

with τ_∞ being the diffusion time of the short fragments at the end the experiment and $A_1 = 1 - A_2 - \tau_\infty$ and A_2 being the amplitudes of the species exhibiting the mean retention times $\langle t_1 \rangle$ and $\langle t_2 \rangle$, respectively. The resulting fit described the data (dashed line in Fig. 3.3B) and provided the mean retention

times. Comparison of $\langle t_1 \rangle$ and $\langle t_2 \rangle$ suggested that internal products are released at different timescales. Thus, we assigned $\langle t_{internal}^{fast} \rangle$ = 1.0 ± 0.1 min and $\langle t_{internal}^{slow} \rangle$ = 25 ± 4 min. We attribute the fast phase to full target RNA cleavage and release of the final middle 6 nt product and the slow phase to intermediate ≥12 nt cleavage products, which could be retained slightly longer. The intermediate cleavage products are observed to accumulate when performing bulk *in vitro* RNA cleavage reactions (Mogila et al., 2019; Tamulaitis et al., 2014). We also performed a control measurement employing the RNase-dead mutant of StCsm. As expected, the mean diffusion time remained constant throughout the time course of this reaction (grey trace in Fig. 3.3B) since no cleavage products were produced.

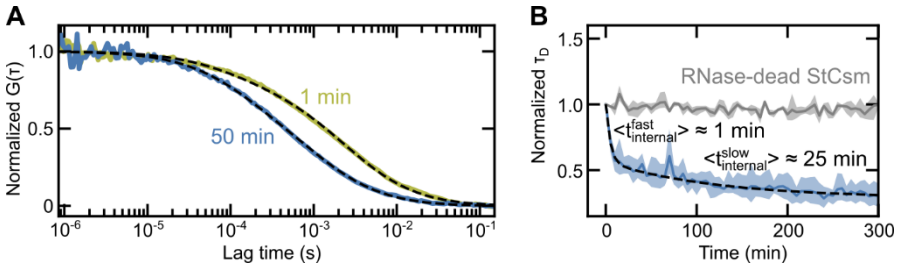


Fig. 3.3. Tracking of internally labelled RNA product release by fluorescence autocorrelation spectroscopy. (**A**) Autocorrelation curves of the internally libelled activating target RNA S3/iL, measured at 1 min and 50 min after reaction start (yellow and blue solid lines, respectively). Dashed lines are fits to a model including slowly and rapidly diffusing species. Representative curves of three independent measurements are displayed. (**B**) Normalized diffusion time as a function of reaction time for internally labelled activating target RNA S3/iL in the presence of StCsm (blue solid line) or the RNase-dead mutant (grey solid line). An exponential fit employing two decaying species is shown as a dashed line.

Overall, tracking of internal RNA cleavage products confirms that they are released by StCsm around two orders of magnitude faster than terminal products. This independently signifies the discrepancy between timescales of fast RNA cleavage and slow release of terminal RNA products.

3.1.2. Duration of Cas10 activity corresponds to timescale of RNA release

Having resolved the timescales of RNA cleavage product release, we next aimed to investigate the RNA-dependent temporal control of Cas10 subunit. For this we tracked the ssDNA cleavage activity by the HD domain of Cas10. We employed a 6 nt long single-stranded FQ-DNA oligonucleotide

(Appendix 2), labelled at the 5′-end with a fluorophore 6-FAM and at the 3′-end with a fluorescence quencher BHQ-1. Due to the proximity of the labels, the fluorescence emission of the substrate is effectively quenched. However, upon StCsm binding to RNA resulting in Cas10 activation and ssDNA cleavage, the quencher would become spatially separated from the fluorophore exhibiting observable fluorescence emission (Fig. 3.4A). To track the ssDNA cleavage activity of StCsm, we monitored the fluorescence signal after the addition of unlabelled target RNA. To relate the fluorescence signal to the amount of cleaved substrate, we degraded increasing amounts of FQ-DNA using micrococcal nuclease and mapped the used amount of ssDNA to fluorescence signal measured after completion of the reaction (Appendix 4). Fluorescence assays were performed by dr. Patrick Irmisch.

We first measured the upregulation of ssDNase activity using an RNA substrate S3/a (blue trace in Fig. 3.4B). The reactions were performed at 50-fold molar excess of ssDNA substrate over StCsm and RNA which were kept at equimolar amounts at 2 nM. Resulting time trajectory exhibited an increase of cleaved DNA which stabilized, indicating a decaying ssDNA cleavage activity. Overall, there was only a moderate ssDNA turnover per StCsm complex at the applied concentrations. Nevertheless, the acquired ssDNA cleavage kinetics were well described by a single exponential fit:

$$c(t) = A\left(1 - e^{-t/\langle t_{ssDNase} \rangle}\right), \tag{3.8}$$

with $c(t)$ being the concentration of cleaved (and thus fluorescent) DNA at time t, A being the total cleaved DNA and $\langle t_{ssDNase} \rangle$ being the mean duration of ssDNase reaction. It provided $\langle t_{ssDNase} \rangle$ = 68 ± 1 min, which agrees with the observed duration of terminal RNA product retention upon cleavage $\langle t_{retain} \rangle$ = 84 ± 1 min and is more similar to Cas10-proximal product retention duration $\langle t_{proximal} \rangle$ = 84 ± 2 min than distal product retention duration $\langle t_{distal} \rangle$ = 119 ± 3 min. The mean duration of ssDNase activity is substantially longer than timescales of internal product release (Fig. 3.3) and RNA cleavage (Tamulaitis et al., 2014). This data supports our initial hypothesis that ssDNase activity is stimulated by retained terminal RNA products.

To further validate our ssDNA cleavage assay, we performed several control assays. Measurements in the absence of target RNA (grey trace in Fig. 3.4B) or in the presence of the non-activating target RNA S3/n (cyan trace in Fig. 3.4B), which is fully complementary to crRNA 5′-tag, showed only a

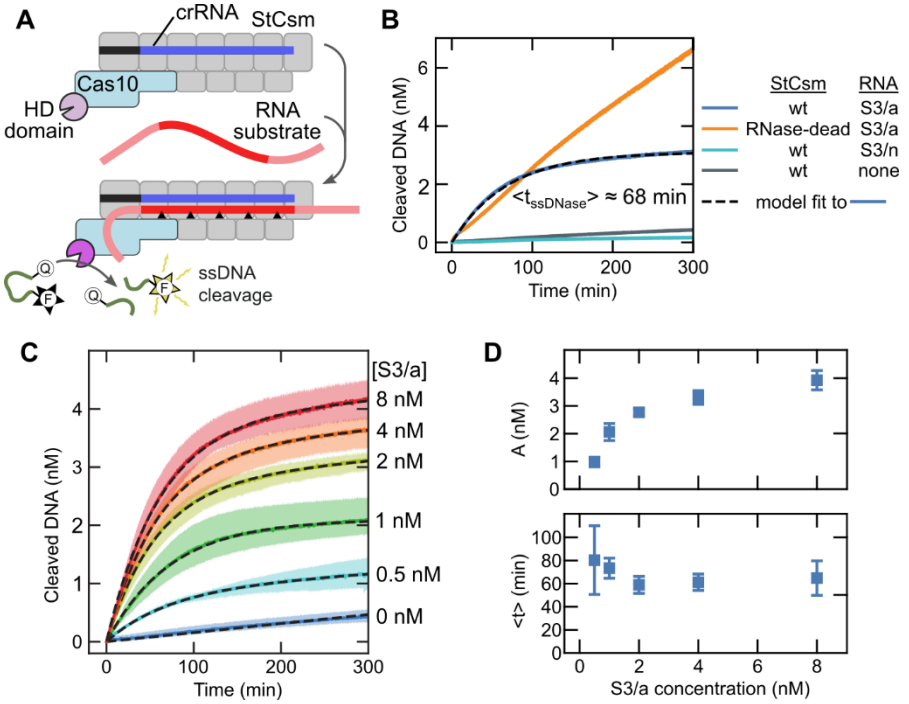


Fig. 3.4. Duration of StCsm Cas10 ssDNase activity measured by bulk fluorescence. (**A**) Scheme of StCsm complex activation by RNA and subsequent FQ-DNA substrate cleavage releasing active fluorophore. (**B**) Cleaved ssDNA as function of time measured for activating target RNA S3/a added to either WT StCsm complex (blue) or the RNase-dead mutant (orange). Additionally, control measurements in the presence of the non-activating target RNA S3/n (cyan) or in the absence of RNA (grey) are shown. Traces represent the mean of three repeated measurements. Reactions were initiated by adding 2 nM StCsm (WT or RNase-dead) to 2 nM target RNA, and 100 nM FQ-DNA (or 200 nM FQ-DNA in case of the RNase-dead mutant). Trace of ssDNA cleavage by S3/a RNA activated WT StCsm was fitted to exponential model (Equation 3.8, dashed line). (**C**) Fluorescence traces (solid lines) fitted to exponential model with background activity (Equation 3.9, dashed lines). Reactions were performed as in (B) with WT StCsm and increasing S3/a RNA concentrations. The range between the minimum and maximum values of three measurement repeats is shown as coloured areas. (**D**) The total amount of cleaved ssDNA during activation (amplitude, A) and mean duration $\langle t \rangle$ of the reactions obtained from fitting the traces plotted in (C). Error bars denote standard deviation of fitted values obtained from three independent measurements.

background ssDNA cleavage activity as observed previously for StCsm (Kazlauskiene et al., 2016). Furthermore, we performed ssDNA cleavage measurements using the RNase-dead StCsm and the activating target RNA S3/a (orange trace in Fig. 3.4B). Pronounced ssDNA cleavage activity was observed which continued mostly unchanged during the full measurement

period in contrast to WT StCsm. These controls validate our fluorescence assay and confirm previous observations that RNA cleavage is required to diminish the ssDNase (Kazlauskiene et al., 2016) activity in Cas10 subunit.

Previous studies of StCsm RNase activity under multiple turnover conditions yielded the RNA substrate turnover rate k_{cat} = 3.0 ± 0.6 min^{-1} which means that around three substrate RNA molecules are degraded by a single StCsm complex per minute (Kazlauskiene et al., 2016). Having in mind that under substoichiometric RNA concentrations the terminal RNA products are retained for $\langle t_{retain} \rangle$ = 84 ± 1 min (Fig. 3.1D), the RNA turnover data suggests that they are readily outcompeted by the excess of substrate. Assuming that our hypothesis linking RNA end retention and Cas10 activity stands, the RNA turnover should have a minimal effect on the amount of cleaved ssDNA, since the excess of RNA would be rapidly degraded and ssDNase would only be stimulated by the retained terminal products. Differently, under RNA deprivation the amount of activated StCsm complexes (and thus the amount of cleaved ssDNA) would be limited by the RNA substrate concentration.

Prompted by this, we tested the RNA concentration effect on ssDNase activity in the same reaction setup. To this end, we kept the StCsm and ssDNA concentrations constant and varied the amount of activating target RNA S3/a (Fig. 3.4C). To quantify the observations, we fitted the ssDNA cleavage kinetics to an exponential model with an additional constant-rate background:

$$c(t) = A\left(1 - e^{-t/\langle t \rangle}\right) + k_{bg}t, \tag{3.9}$$

with A being the total cleaved DNA during activation under exponential mode, $\langle t \rangle$ being the mean duration of the ssDNase reaction and k_{bg} being the background cleavage rate, pronounced in reaction mixtures without RNA as observed previously (Kazlauskiene et al., 2016). Fits described the cleavage kinetics for all tested RNA concentrations (dashed lines in Fig. 3.4C) providing the amount of ssDNA cleaved by activated complex and the mean duration of the ssDNase activity (Fig. 3.4D).

For substoichiometric amounts of RNA, the total amount of degraded ssDNA increased as we increased the RNA concentration. When the concentration of RNA substrate exceeded the concentration of StCsm, the total amount of cleaved ssDNA (A) increased only slightly. Importantly, the mean duration of ssDNase activity remained constant with $\langle t \rangle$ being in 60-80 min range regardless of the amount of S3/a substrate in reaction mixture.

These results agreed with our assumptions that with increasing RNA concentration the proportion of StCsm in Cas10-activated state also increases until it reaches saturation when RNA concentration exceeds than that of StCsm. The lifetime of Cas10-activated state (or the ssDNase activity duration) being independent of RNA concentration suggests that this property is intrinsic to the complex. Since it is similar to the timescale of terminal RNA cleavage product retention, these processes could be functionally coupled as discussed in the following section.

3.1.3. Outline of Csm regulation mechanism

Taken together, the FCS data on RNA cleavage product retention and bulk fluorescence measurements of ssDNase activity provide new insights into type III-A Csm temporal regulation mechanisms (Fig. 3.5).

Fluorescence cross-correlation analysis using double-labelled RNA substrate shows that upon target RNA cleavage by StCsm, both ends are retained in the complex for the mean duration $\langle t_{retain} \rangle = 84 \pm 1$ min (Fig. 3.1) which is in a sharp contrast to RNA cleavage undergoing in seconds in either single or multiple turnover conditions (Tamulaitis et al., 2014; Kazlauskiene et al., 2016) suggesting that terminal RNA cleavage products are retained unless an intact target RNA is available which outcompetes the products.

Autocorrelation analysis of individual product retention shows that Cas10-proximal product is retained for mean duration $\langle t_{proximal} \rangle = 84 \pm 2$ min and Cas10-distal product is retained for $\langle t_{distal} \rangle = 119 \pm 3$ min (Fig. 3.2) whereas internal products are released at faster timescales (Fig. 3.3). Results suggest that terminal product release could be coordinated: judged by individual retention times the Cas10-proximal product dissociation from the complex could trigger the release of Cas10-distal RNA, yet further studies are needed to test this speculation.

Due to 5′ to 3′ directionality of transcription machinery, Cas10-proximal part of the complex binds the nascent 3′-part of newly synthesized mRNA directing Cas10 subunit towards the transcription bubble. Prolonged retention of the RNA bound there could keep the complex tethered in the vicinity of transcribed DNA for effective cleavage of the DNA strand displaced by RNA polymerase (Fig. 1.6A). The relatively fast release of internal RNA cleavage products suggest that they are weakly bound in the disordered crRNA:target duplex after cleavage compared to protein-RNA interactions of the terminal products (Fig. 1.7B and C). Since Cmr and Csm complexes have minor differences in, respectively, Cmr3-Cas10 and Csm4-Cas10 dyad structures and

surface electrostatics (Section 1.3.3), it remains to be tested whether prolonged Cas10-proximal RNA cleavage product retention observed for StCsm extends to type III-B complexes.

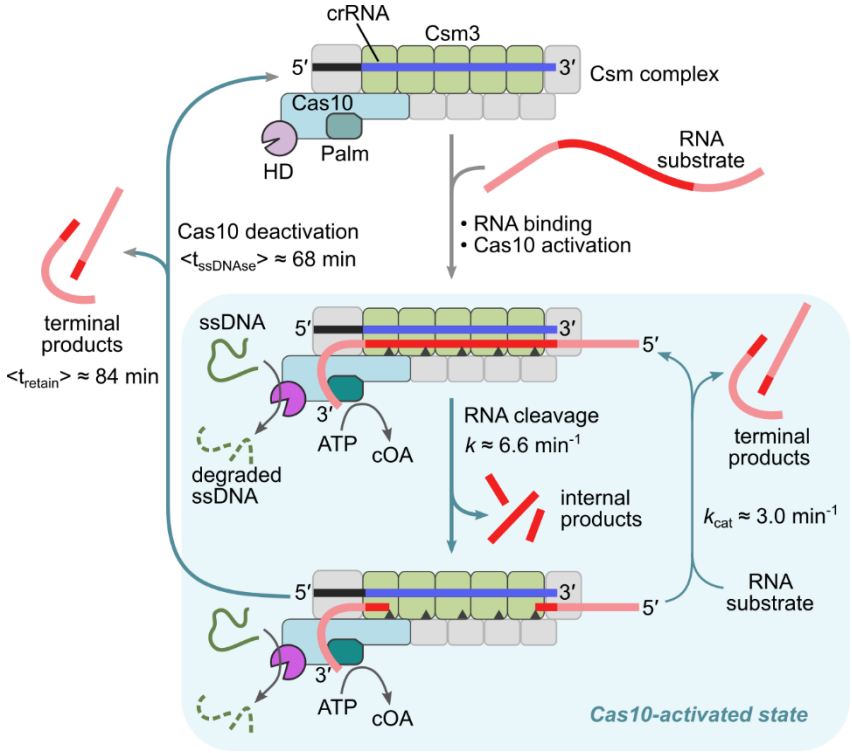


Fig. 3.5. Proposed Csm complex regulation mechanism. Scheme shows Csm complex activation by cognate RNA binding. Prolonged retention of RNA ends after cleavage by Csm3 subunits stimulates the activity of Cas10. Terminal RNA cleavage products are rapidly displaced under multiple turnover conditions without disrupting Cas10-activated state of the complex. See main text for a detailed description. RNA cleavage rate *k* under single turnover conditions (at 25 °C) and multiple turnover rate k_{cat} (at 37 °C) were obtained from (Tamulaitis et al., 2014) and (Kazlauskiene et al., 2016), respectively.

ssDNase activity experiments revealed that upon StCsm activation the complex remains in the activated state for $\langle t_{ssDNase} \rangle = 68 \pm 1$ min. The agreement between this value and the timescale of terminal RNA cleavage product retention suggests functional association. We propose that terminal products could keep type III CRISPR-Cas complexes in the activated state upon RNA cleavage (Fig. 3.5). In agreement to this, we see that disruption of RNA cleavage in RNase-dead StCsm mutant can prolong ssDNase activity (Fig. 3.4B). While both terminal products remain strongly bound, we

speculate that the PFS-containing Cas10-proximal product rather than Cas10-distal product could modulate Cas10 activity as the PFS complementarity to crRNA 5′-tag is checked for self vs. non-self discrimination in Csm/Cmr complexes (Section 1.3.3).

The amount of cleaved ssDNA scales with activator RNA concentration when it is lower than that of StCsm (Fig. 3.4D). This suggests that the proportion of activated complexes is limited by substoichiometric RNA concentration. The saturation effect observed at higher RNA concentrations can be explained by rapid RNA turnover by StCsm observed previously (Kazlauskiene et al., 2016). Retained terminal products could keep Cas10 in the active state upon RNA substrate degradation (Fig. 3.5, lower right). Such a property could ensure a robust defence in cases of high expression of foreign genes that tend to generate most of the spacers in type III CRISPR regions (Aviram et al., 2022; González-Delgado et al., 2019; Silas et al., 2016).

The first reports of ssDNase (Kazlauskiene et al., 2016) and polymerase-cyclase (Kazlauskiene et al., 2017) activities of, respectively, StCsm Cas10 HD and Palm domains featured elegant biochemical assays revealing that both activities diminish gradually over a timescale of tens of minutes upon stimulus by RNA binding. In the light of the results presented in this thesis, previous observations suggest that retained RNA cleavage products could keep not only ssDNase, but also polymerase-synthase activities of Cas10 protein turned on (Fig. 3.5). Under the scenario of an infection, the slow diminish of polymerase-synthase in Csm could lead to a prolonged synthesis of cOA and more pronounced response from cOA-binding CARF/SAVED effectors forming an additional line of defence in type III CRISPR-Cas systems presented in Chapter 1.4. However, a direct correlation between RNA retention and polymerase-cyclase activity remains to be established.

3.2. Characterization of CARF mRNA interferases

The auxiliary effector proteins often encoded in the vicinity of the CRISPR-Cas locus are typically composed of the cOA-sensing CARF or SAVED domains fused to a wide variety of effector domains. During infection the cOA signalling pathway reinforces type III CRISPR-Cas defence by inducing toxic activities of CARF or SAVED effectors. This either kills the infected cell or leads to its growth arrest thus preventing the spread of infection in the population (Chapter 1.4).

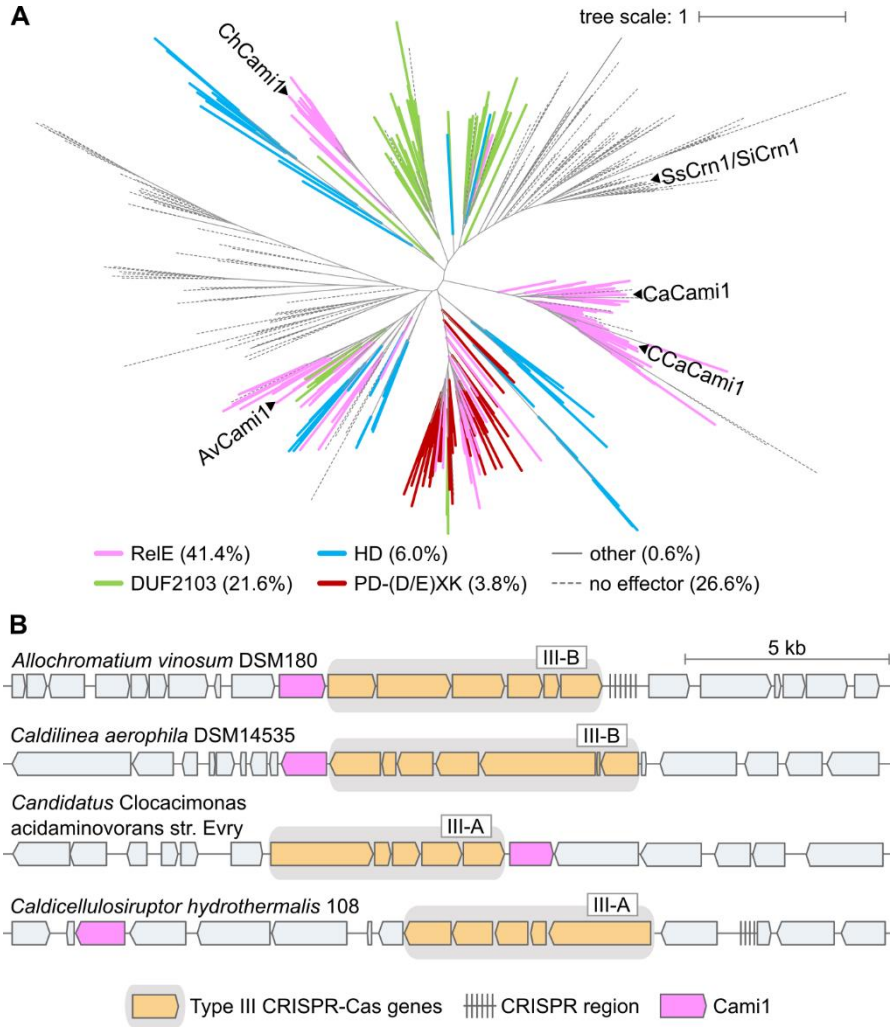


Fig. 3.6. Phylogenetic tree of revised CARF7 clade. (**A**) Phylogenetic tree of the CARF7 clade based on the subset of CARF domain sequences reduced to 40% identity (n = 385). Colouring of the tree leaves corresponds to the type of effector domain fused to CARF7 domain or CARF-wHTH domain core. Leaves corresponding to Cami1 proteins, selected for characterization in this study, and Crn1 proteins (SsCrn1 and SiCrn1) are indicated. Numbers in parentheses in the legend correspond to the proportion of given effector domain in the expanded CARF7 set. A tree scale corresponding to the average number of amino acid substitutions per site is indicated. Protein data used to construct this tree is presented in (Mogila et al., 2023). (**B**) Genomic regions of four bacteria encoding type III CRISPR-Cas (orange in grey background) and associated Cami1 proteins (magenta).

According to recent classification there are 10 major and several minor classes of CARF proteins (Section 1.4.1). These classes differ by the CRISPR-

Cas system they are associated with (if any) and by the range of effector domains they are fused with. Importantly, some CARF domains possess a ring nuclease activity to destroy the bound cOA to limit signalling response or self-deactivate (Section 1.4.2).

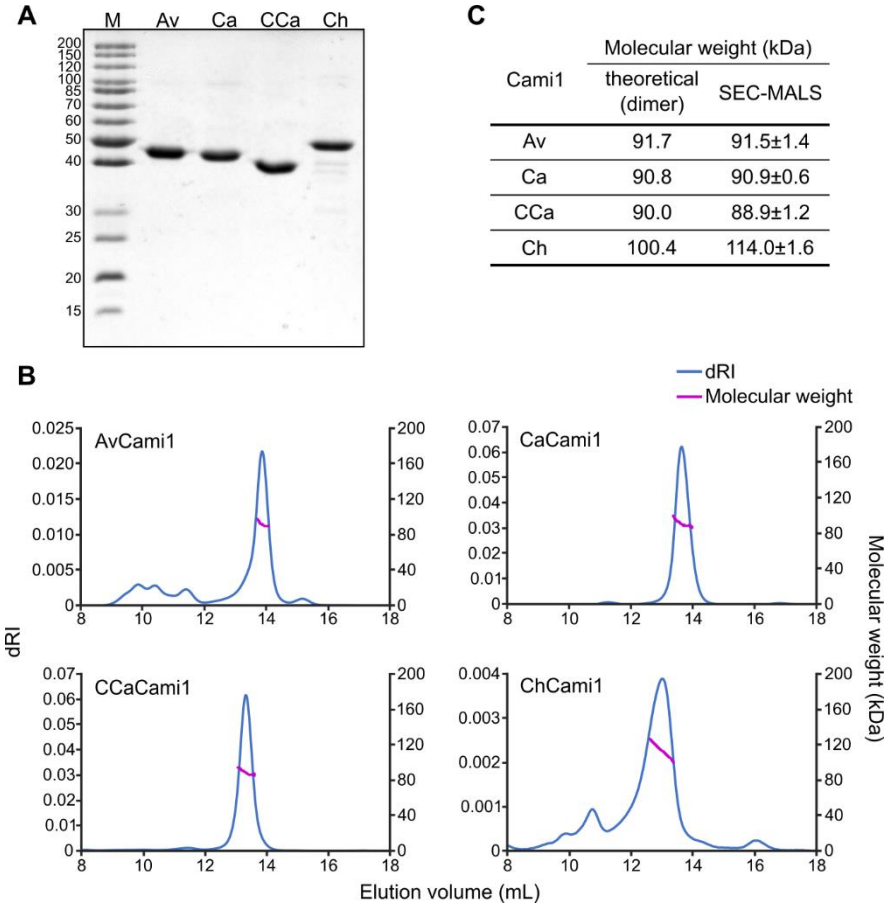
We focused on CARF7 group of CARF proteins formed of well-characterized ring nucleases or their fusions with various effector domains, mostly putative nucleases (Section 1.4.3). We enriched the previously defined CARF7 class with new sequence homologs (totalling 1521 sequences) and obtained structural models for the entire set. Next, using both sequence and predicted structural information, we identified domain architectures for each protein. Our analysis revealed that ~90% of CARF7 domains are fused to a wHTH domain, as exemplified by characterized Crn1 ring nucleases SsCrn1 from *Saccharolobus solfataricus* (Athukoralage et al., 2018) and SiCrn1 from *Sulfolobus islandicus* (Molina et al., 2021). This two-domain CARF7-wHTH core in most cases (~80%) is fused to additional effector domains, such as RelE (mRNA interferases), HD (phosphohydrolases), PD-(D/E)XK (endonucleases) or DUF2103 (function unknown) (Fig. 3.6A). Phylogenetic analysis of CARF domains revealed that the same type of associated effector domain (e.g., RelE) may be present in different subclades, while proteins in a single subclade, comprising closely related CARF7 domains, may feature different effector domains. This observation suggests that CARF7-wHTH and effector domain fusion events occurred multiple times independently during evolution. The analysis of CARF7 protein sequences was performed by dr. Albertas Timinskas.

The dominant architecture among CARF7 clade proteins is composed of CARF7-wHTH-RelE domains in which the C-terminal domain is similar to ribonuclease domain identified as RelE toxin in the type II toxin-antitoxin systems (Jurėnas and Van Melderen, 2020). Type II RelE family toxins are ribosome-dependent endoribonucleases (mRNA interferases), which bind ribosomal A-site to cleave the translated mRNA (see Section 1.5.1).

Owing to the architecture and genome location of CARF-wHTH-RelE proteins, we named them Cami1 (CRISPR-Cas-associated mRNA interferase 1). For further experimental analysis, we selected four Cami1 proteins from *Allochromatium vinosum* (AvCami1), *Caldilinea aerophila* (CaCami1), *Candidatus* Cloacimonas acidaminovorans (CCaCami1), and *Caldicellulosiruptor hydrothermalis* (ChCami1), which are all encoded in the vicinity of type III CRISPR-Cas loci (Fig. 3.6B) and belong to different sequence clusters in the CARF7 clade (Fig. 3.6A).

3.2.1.Purification of Cami1 proteins

We ordered synthetic genes of selected Cami1 proteins precloned to pBAD expression vector and fused with C-terminal His_{10} and StrepII tags. Proteins were expressed in *E. coli* and isolated using affinity chromatography (Fig. 3.7A). Molecular weight measurement using SEC-MALS showed that Cami1 proteins form dimers in solution (Fig. 3.7B and C) similarly to SsCrn1 and SiCrn1 (Athukoralage et al., 2018; Molina et al., 2021) ring nucleases from CARF7 group as well as other CARF proteins (Stella and Marraffini, 2023).



C

Cami1	Molecular weight (kDa)	
	theoretical (dimer)	SEC-MALS
Av	91.7	91.5±1.4
Ca	90.8	90.9±0.6
CCa	90.0	88.9±1.2
Ch	100.4	114.0±1.6

Fig. 3.7. Purification and stoichiometry analysis of Cami1 proteins selected for this study. (**A**) Purified Cami1 proteins analysed by SDS-PAGE. "M" – protein size ladder with indicated values in kDa. (**B**) SEC-MALS traces of Cami1 proteins. Blue lines – dRI response, magenta traces – molecular weight obtained from MALS response. (**C**) Absolute MW estimates of Cami1 proteins in solution obtained by SEC-

MALS. MW of a translation product corresponding to whole Cami1 ORF was used to calculate theoretical MW of dimers.

3.2.2. Crystal structure of AvCami1

To understand the structural arrangement of Cami1, we determined the crystal structure of AvCami1 at 1.7 Å resolution (Fig. 3.8A, PDB: 8PHB). AvCami1 protein was crystallized by Konstanty Keda and its structure was solved by dr. Giedrė Tamulaitienė.

The structure revealed a predicted three-domain architecture with the N-terminal CARF domain (spanning the first 186 amino acids), the middle wHTH domain (187 to 265) and the C-terminal RelE domain (266 to 381). The AvCami1 CARF domain is most similar to that of SsCrn1 ring nuclease (Athukoralage et al., 2018) (Fig. 3.8B). SsCrn1 active site residues S11 and K168 are spatially conserved in AvCami1 (S11 and K157, respectively), implying that CARF domain of AvCami1 is a ring nuclease.

The structure of AvCami1 RelE domain is similar to the RelE family toxins, the closest match being *Vibrio cholerae* toxin VcHigB2 (PDB: 5JA9) which itself adopts an EcRelE-like active centre (Hadži et al., 2017). Structural superposition revealed that VcHigB2 active site residues are conserved in the AvCami1 RelE domain (Fig. 3.8C). While most RelE family toxins are monomeric proteins (Jurėnas and Van Melderen, 2020), the RelE domains of AvCami1 form a dimer (Fig. 3.8A). Differently, dimeric AlphaFold 2 models of CaCami1 and CCami1 suggest that RelE domains of these proteins do not interact whereas model of ChCami1 shows a modest interaction surface between RelE domains (Fig. 3.8D). Taken together, structural data suggested that Cami1 could function as dimeric cOA-activated ribonucleases.

3.2.3. Ring nuclease activity

CARF effectors bind either cA_4 or cA_6 and some of them have an intrinsic ring nuclease activity allowing them to destroy their activators to self-inhibit and lower the immune response (Athukoralage et al., 2019; Athukoralage and White, 2021; Garcia-Doval et al., 2020; Smalakyte et al., 2020). We performed ring nuclease assays by incubating Cami1 proteins with an excess of cA_4 and cA_6 and analysed the reaction products using mass spectrometry (MS) which was performed by Audronė Rukšėnaitė. These experiments revealed

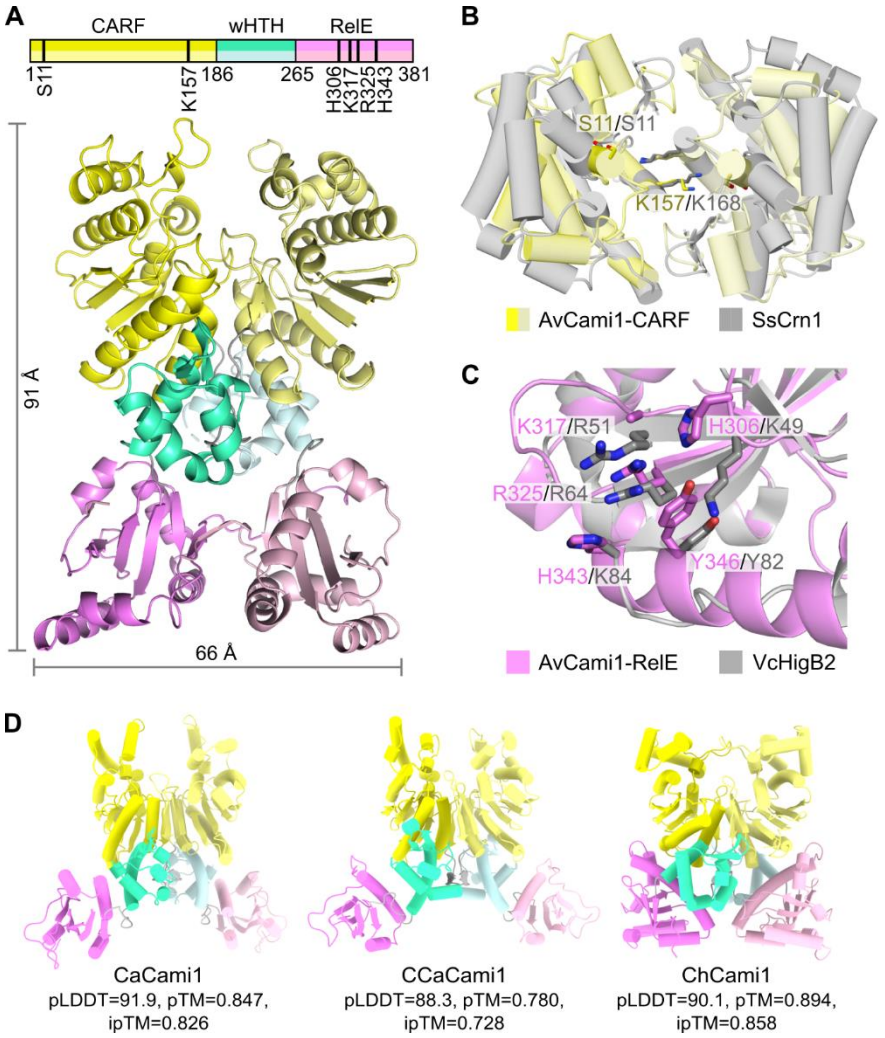


Fig. 3.8. Structural analysis of Cami1 proteins. (**A**) Overview of apo dimeric AvCami1 crystal structure comprised of CARF (yellow), wHTH (teal) and RelE-like (magenta) domains. Crystallographic data collection and refinement statistics are provided in Appendix 5. (**B**) Structural alignment of AvCami1 and SsCrn1 (PDB: 3QYF, RMSD 2.2 Å) CARF domains. Amino acids critical for SsCrn1 ring nuclease activity and corresponding AvCami1 residues are shown as sticks. (**C**) Structural alignment of AvCami1 RelE domain and *Vibrio cholerae* HigB2 toxin (PDB: 5JA9, RMSD 2.6 Å). Experimentally verified mRNA interferase active site residues of VcHigB2 and structurally analogous residues of AvCami1 are shown as sticks. (**D**) Full-length dimeric Ca, CCa and ChCami1 AlphaFold 2 models with indicated modelling scores. CARF, wHTH and RelE domains are coloured as in (A).

that cA_4 but not cA_6 is cleaved by Cami1 proteins (Fig. 3.9A, Appendix 6). Cami1 proteins converted cA_4 to A_4>p and A_2>p, similarly to other ring

nucleases and CARF effectors (Athukoralage et al., 2019, 2018; Molina et al., 2019, 2021) with an exception for CaCami1 protein which showed lower ring nuclease activity and produced no A_2>p. Taken together these results show that Cami1 proteins are ring nucleases and suggest that cA_4 is their cognate activator.

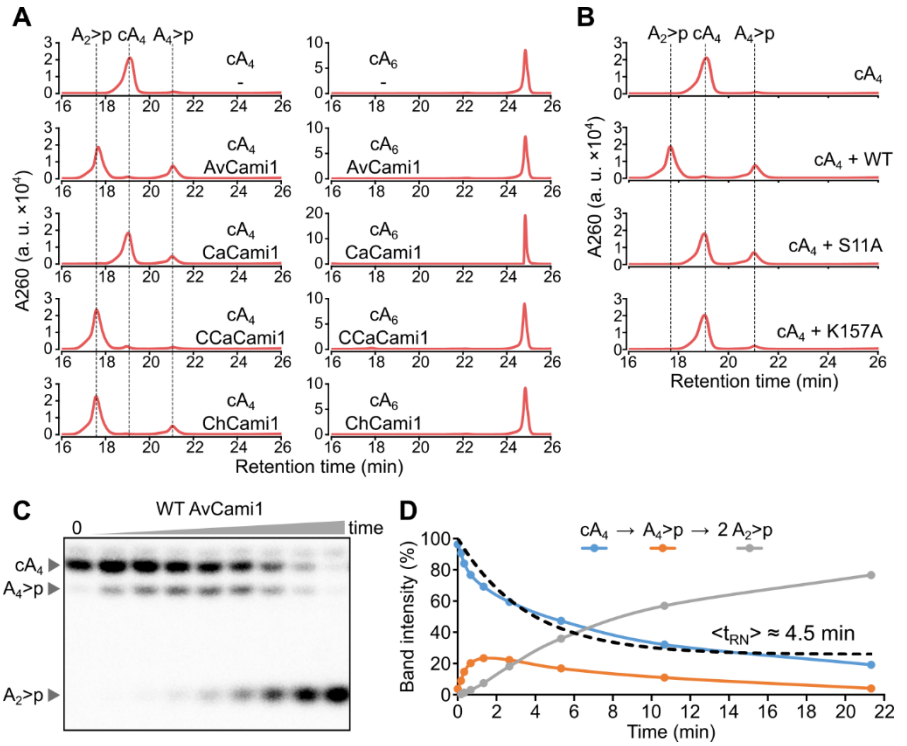


Fig. 3.9. Ring nuclease activity of Cami1 proteins. (**A**) MS analysis of reaction mixtures with cA_4 or cA_6 as substrates after incubation with different Cami1 proteins. Retention times of different reaction species with m/z ratios corresponding to cA_4 substrate or A_4>p or A_2>p linear products are indicated by dashed lines. (**B**) MS analysis of cA_4 cleavage by AvCami1 CARF domain point mutants compared to WT protein. Reaction products are indicated as in (A). (**C**) Products of α^{32}P-labeled cA_4 cleavage by AvCami1 under single turnover conditions at different time points. Representative radiogram of three independent experiments is presented. Time points: 0, 10, 20, 40, 80, 160, 320, 640, 1280 s. (**D**) Plotted densitometric analysis results of the radiogram displayed in (C). Fit to cA_4 depletion data is shown as dashed lines.

We additionally constructed and purified AvCami1 mutants of conservative amino acids S11A and K157A (Fig. 3.8B). By repeating ring nuclease assays we confirmed that aforementioned serine and lysine residues are critical for cA_4 cleavage (Fig. 3.9B) as reported for ring nucleases SsCrn1

(Athukoralage et al., 2018) and SiCrn1 (Molina et al., 2021) confirming the specific ring nuclease activity of the CARF domain.

We tracked the cleavage of radioactively labelled cA_4 at different time points under single-turnover conditions (Fig. 3.9C). Densitometric analysis of the radiogram revealed that the first cleavage event leads to initial accumulation of A_4>p which is then cleaved into two molecules of A_2>p. We fitted cA_4 depletion data to a single exponential (Fig. 3.9D):

$$S(t) = (100\% - S_\infty)e^{-t/\langle t_{RN} \rangle} + S_\infty, \qquad (3.10)$$

where $S(t)$ is the proportion (in %) of labelled cA_4 at reaction time t, $\langle t_{RN} \rangle$ is mean ring nuclease reaction duration and S_∞ is the proportion of uncleaved substrate at the end of experiment. Fitting yielded $\langle t_{RN} \rangle$ = 4.4 ± 0.7 min.

Ring nuclease assays confirmed that Cami1 proteins bind and cleave cA_4 over timescale of minutes akin to known CARF7 ring nucleases.

3.2.4.mRNA interferase activity

Sequence analysis (Fig. 3.6A) together with structural data (Fig. 3.8C) suggested that C-terminal domain of Cami1 proteins could function as a ribonuclease that cleaves mRNA in a ribosome-dependent manner similar to homologous RelE family toxins. To test this hypothesis, we adapted an assay that was previously used to monitor *E. coli* RelE toxin mediated RNA cleavage in programmed ribosomes (Griffin et al., 2013). Since EcRelE efficiently cleaves the mRNA at the UAG stop codon *in vitro* (Pedersen et al., 2003), we programmed *E. coli* 70S ribosomes with a short radiolabelled RNA substrate rib-UAG (Appendix 2) and $tRNA^{fMet}$ to lock the AUG codon of RNA in the ribosome P-site, thereby exposing the downstream UAG stop codon in the A-site (Fig. 3.10A, right). AvCami1 did not cleave free or ribosome-bound RNA without an activator, however the cleavage of ribosome-bound RNA was induced by cA_4 (Fig. 3.10A, left) as hinted by ring nuclease assays. Mapping of the cleavage site using alkaline and RNase T1 degradation of the substrate revealed that AvCami1 cleaves rib-UAG mainly after the 3rd nucleotide of the UAG codon ("product 1" in Fig. 3.10A) similarly to *Mycobacterium avium* RelE and *E. coli* YhaV (Choi et al., 2017; Goeders et al., 2013). A minor "product 2" is also formed due to the cleavage after the 2nd nucleotide of UAG codon or truncation of "product 1". A similar relaxed cleavage specificity was also observed for EcRelE *in vivo* (Goeders et al., 2013).

To test RelE domain active site predicted by structural comparison with VcHigB2 protein (Fig. 3.8C), we constructed and isolated AvCami1 RelE active site mutants H306A, K317A, R325A and H343A. We performed mRNA interference assays using these mutants in comparison with WT protein. We fitted the densitometric data of RNA degradation over time to single exponential decay model (see Method 2.2.12) to obtain cleavage rate constants. Selected mutations resulted in a 5- to 150-fold reduction in the mRNA cleavage rate with H343A substitution being the most critical for activity (Fig. 3.10B). Measured reaction rate for WT protein k_{mRNase} = 0.05 ± 0.01 s^{-1} corresponds to ribosome-bound RNA degradation in tens of seconds under tested reaction conditions. Mutant proteins were purified and assayed by Konstanty Keda.

Since the CARF domain performs a stepwise cleavage of the activator cA_4 with an intermediate linear A_4>p product, we were interested to check whether both molecules stimulate mRNA cleavage. At 10-fold step increases of cA_4 and A_4>p concentration we observed similar amounts of substrate cleaved in mRNA interferase assays (Fig. 3.10C) confirming that Cami1 does not self-inhibit during the initial step of cA_4 cleavage (Fig. 3.9C and D). In addition to this, a single AvCami1 protein can cleave multiple RNA molecules in programmed ribosomes as tested under substrate excess (Fig. 3.10D). Moreso, cA_4 cleavage cycle seemed unrelated to RNA substrate turnover since both WT and ring nuclease-dead S11A mutant degraded RNA at the same efficiency in programmed ribosomes under the excess of activator (Fig. 3.10D, right). These results suggest that once activated Cami1 proteins do not need to be reloaded with cA_4 after each mRNA cleavage event.

CaCami1, CCaCami1 and ChCami1 also cleaved RNA in a ribosome- and cA_4-dependent manner (Fig. 3.11A). CCaCami1 protein cleaved ribosome-bound rib-UAG substrate at lower activity which is observed only at higher protein concentrations. AvCami1 and CCaCami1 demonstrate cleavage both after the 2nd and the 3rd nucleotide whereas CaCami1 and ChCami1 homologs cleave mainly after the 2nd nucleotide in UAG codon. These results suggest subtle variation in the positioning of RelE active centre among Cami1 proteins in ribosome A-site.

In line with the proposed RelE mRNA cleavage mechanism, which involves a nucleophilic attack by a 2′-OH group on the adjacent phosphate (Jurėnas and Van Melderen, 2020; Neubauer et al., 2009), Cami1 did not cleave rib-UAG-OMe substrate in which 2′-O positions of UAG codon are methylated (Fig. 3.11B, Appendix 2). RNA substrates containing random nucleotides

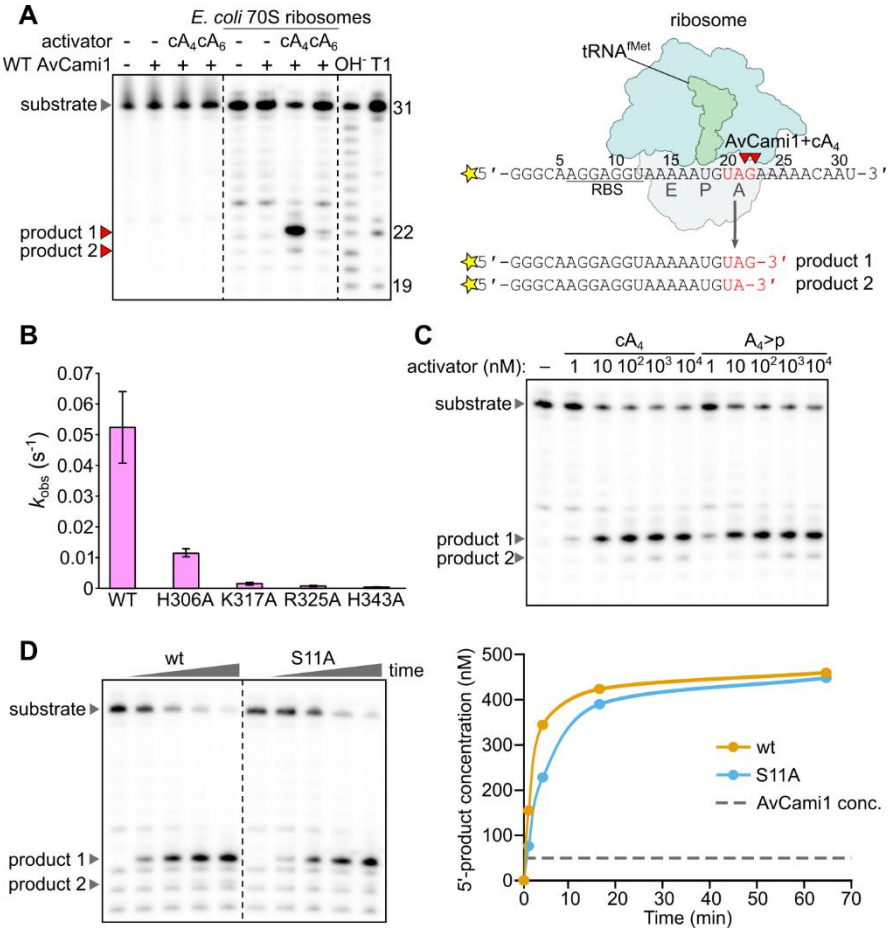


Fig. 3.10. AvCami1 protein cleaves RNA exposed in the A-site of programmed 70S ribosomes upon cA_4 stimulation. (**A**) Cleavage activity of AvCami1 in the presence of cOA and ribosomes. Reactions contained 20 nM (final concentrations are listed) of 5′-radiolabelled rib-UAG substrate and 300 nM of AvCami1 protein. In reactions with ribosomes the ribosome-RNA complexes were preformed additionally using 60 nM of *E. coli* 70S ribosomes and 80 nM of P-site $tRNA^{fMet}$ (see Methods 2.2.12). cOA activators were used at 10 μM. A schematic of the AvCami1 mRNA cleavage experiment is shown on the right. The 31 nt rib-UAG substrate consists of a ribosome-binding-site element (RBS) followed by a spacer and the P-site (AUG) and A-site (UAG) codons. OH^- and T1, alkaline hydrolysis and RNase T1 digestion markers, respectively. (**B**) mRNA interferase activity of AvCami1 RelE domain alanine point mutants compared to WT protein at 15 °C. k_{obs} – observed reaction rate constant. Mean and standard deviation of three technical replicates is reported. (**C**) Activation of WT AvCami1 mRNA interferase in *E. coli* 70S ribosomes programmed with rib-UAG substrate in the presence of cA_4 or A_4>p. Reactions were prepared as in (A) using different amounts of activators. (**D**) RNA hydrolysis activity of AvCami1 protein under multiple turnover conditions in programmed *E. coli* 70S ribosomes. Reactions contained 500 nM of labelled and unlabelled rib-UAG RNA, 1 μM of *E. coli*

70S ribosomes, 1.3 μM of $tRNA^{fMet}$, 10 μM cA_4 and 50 nM WT or S11A mutant AvCami1. Aliquots were removed at timed intervals and samples were analysed in PAGE under denaturing conditions.

(rib-NNN) instead of UAG at the A-site were also cleaved by Cami1 (Fig. 3.11C) in agreement with the relaxed sequence specificity of EcRelE (Goeders et al., 2013).

Collectively, we show here that Cami1 proteins are ribosome-dependent mRNA interferases activated by the cA_4 signalling.

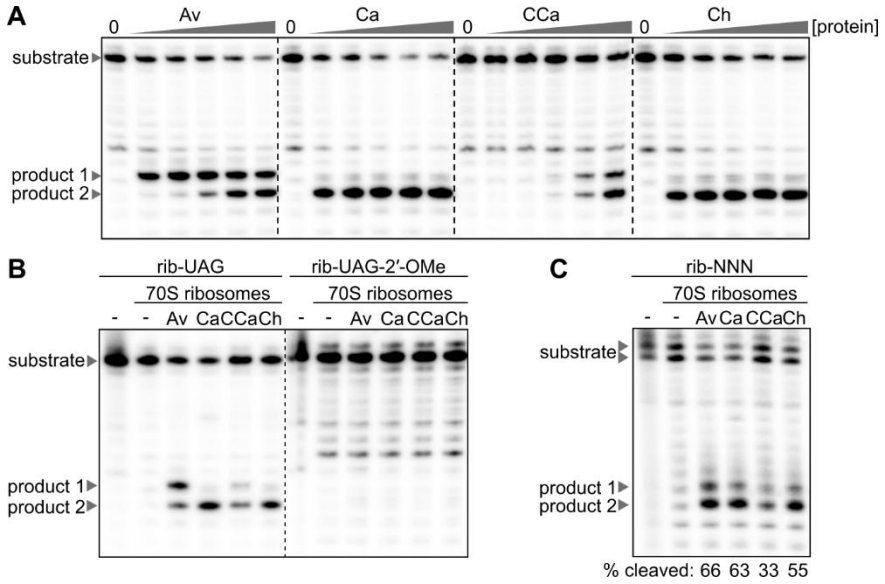


Fig. 3.11. mRNA interferase activity of Cami1 homologs. (**A**) RNA cleavage reactions of Cami1 proteins. Each protein was tested at 0, 30, 100, 300, 1000 or 3000 nM concentration using reaction components as in Fig. 3.10A. Activity was stimulated by adding 10 μM of cA_4. (**B**) Cami1 proteins activity assay using rib-UAG-OMe substrate in which UAG codon nucleotides exposed in the A-site are 2′-OMe modified. 1.5 μM Cami1 were activated by 10 μM cA_4 and reactions were prepared as in Fig. 3.10A. (**C**) Cami1 activity using substrate mix rib-NNN in which ribonucleotides exposed in the A-site are randomized. Reactions were prepared as in (B). Densitometric proportion of cleaved substrate is indicated under the radiogram.

3.2.5. Structure of AvCami1 bound to the ribosome

To gain insights into the structural mechanism of Cami1 activity, we determined a cryo-EM structure of cA_4-activated AvCami1-S11A-H343A

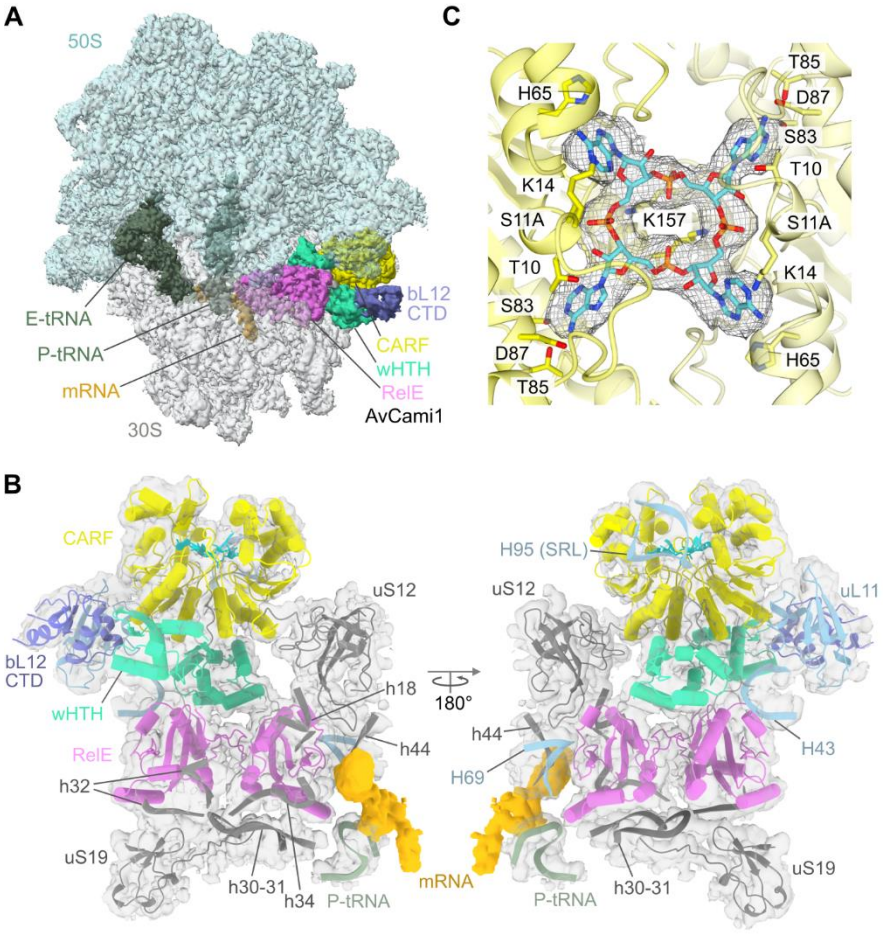


Fig. 3.12. Cryo-EM structure of cA_4 activated AvCami1 bound to 70S ribosome programmed with mRNA mimicking substrate and tRNA. (**A**) Overview of cA_4 activated AvCami1-S11A+H343A (yellow-turquoise-magenta) mutant bound to 70S ribosome (light blue and grey) programmed with rib-UAG-OMe mRNA substrate (orange) and $tRNA^{fMet}$ (green). AvCami1 interaction with ribosomal stalk protein bL12 (slate blue) is also observed. CTD, C-terminal domain. Cryo-EM image processing scheme is provided in Appendix 7 and data collection, refinement and validation statistics are provided in Appendix 8. (**B**) AvCami1 occupies the ribosomal A-site and interacts with the 16S (loops h18, h30-31, h32, h34 and h44) and the 23S (loops H43, H69 and H95) rRNA and with the P-site $tRNA^{fMet}$. AvCami1 dimer also interacts with the ribosomal proteins uS12 and uS19 of 30S subunit and uL11 (light blue) together with bL12 (violet) of 50S subunit. Cryo-EM map is displayed as a transparent grey surface. Structural elements are coloured as in (A). (**C**) cA_4 (cyan) recognition in CARF domain. Mesh depicts cryo-EM map of cA_4 molecule.

mutant bound to *E. coli* ribosome complex with rib-UAG-OMe mRNA mimic and $tRNA^{fMet}$ at nominal 3.7 Å resolution (Fig. 3.12A, PDB: 8PHJ). The

atomic structure of ribosome-AvCami1 complex was reconstructed by dr. Giedrė Tamulaitienė.

The complex contains two $tRNA^{fMet}$ molecules, one bound in the E-site and the other to the cognate AUG codon in the P-site. One AvCami1 subunit of the dimer occupies the ribosomal A-site and its RelE domain interacts with 16S and 23S rRNA as well as with the P-site $tRNA^{fMet}$, (Fig. 3.12B) similar to other RelE family toxins (Mansour et al., 2022; Neubauer et al., 2009). The RelE domain of another AvCami1 subunit makes additional contacts with 16S rRNA. The CARF domains of both subunits contact sarcin-ricin loop (SRL) H95 of 23S RNA. AvCami1 dimer also interacts with core ribosomal proteins uS12, uS19, uL11 and, unexpectedly, a ribosomal stalk protein bL12 (Fig. 3.12B). The diffused cryo-EM density of A-site RNA suggests that the substrate was flexible in the vicinity of the AvCami1 binding site (Fig. 3.12B).

cA_4 is bound in the central cavity of the CARF domain dimer. It adopts an approximately planar conformation supported by interactions between adenine nucleobases and sugar phosphate backbone of cA_4 and polar or charged residues of CARF domain (Fig. 3.12C).

As mentioned above, we found that AvCami1 wHTH domain which is exposed to the solution interacts with the C-terminal domain of bL12 ribosomal protein (Fig. 3.13A). bL12 is a key component of the prokaryotic ribosomal stalk. Stalk structure is conserved in all domains of life and plays an essential role in pooling of GTPase translation factors (Imai et al., 2020; also see Chapter 1.7). The observed AvCami1-bL12 interaction suggested that Cami1 proteins may use the ribosome stalk-mediated capture mechanism to enter the ribosomal A-site.

Following the hint posed by the cryo-EM structure regarding AvCami1 interaction with bL12, we performed point mutagenesis of CARF and wHTH domain bL12-binding surfaces and measured the mRNA interferase activityof these mutants. Mutations in wHTH domain compromised RNA cleavage to varying extents, requiring higher concentrations of mutants to achieve cleavage levels comparable to the WT protein (Fig. 3.13B). This result confirms the importance of bL12-CTD and wHTH interaction for the ribosome-dependent mRNA cleavage activity of AvCami1.

Further, we evaluated mRNA interferase activity of Cami1 homologs using bL12-deficient *E. coli* 70S ribosomes in the presence and absence of a standalone bL12 protein. Cami1 proteins showed only trace activity with bL12-depleted 70S ribosomes, however, the mRNA cleavage activity was partially reconstituted by supplementing bL12 protein to the bL12-depleted

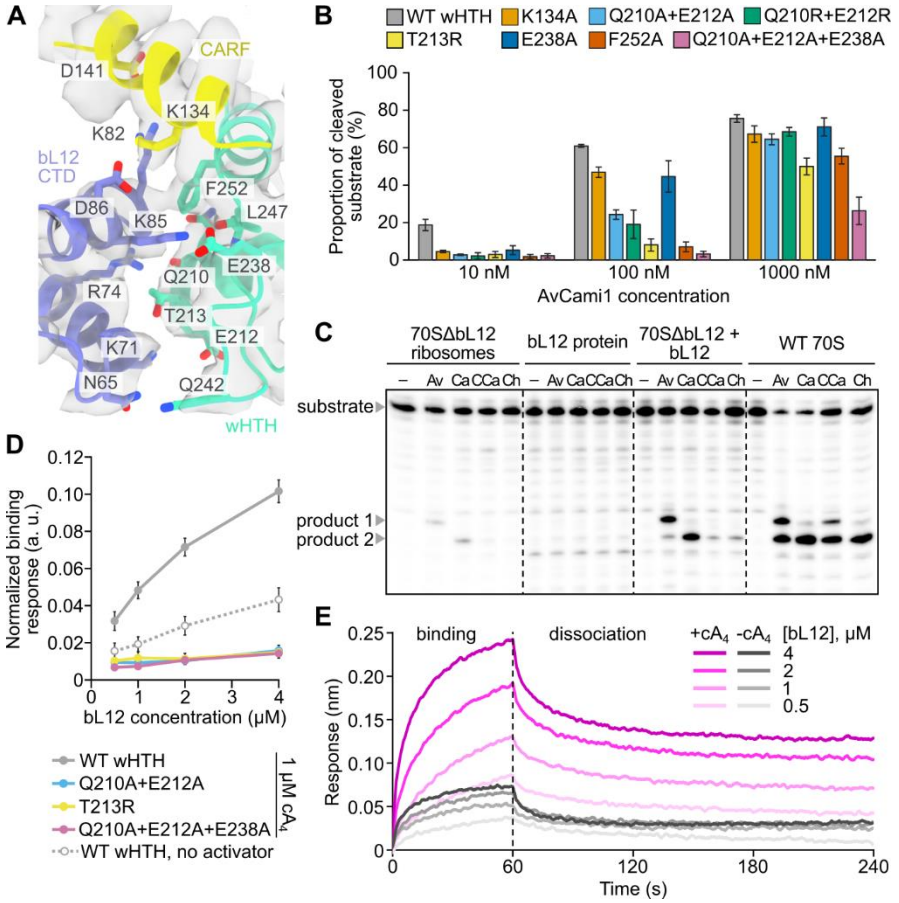


Fig. 3.13. Cami1 interaction with ribosomal bL12 is required for mRNA interferase activity. (**A**) Interaction surface between ribosomal bL12 protein and the distal wHTH domain of AvCami1. (**B**) wHTH and CARF surface alanine point mutation effect on mRNA interferase activity. Alanine substitutions were introduced in addition to S11A mutation resulting in reduced ring nuclease activity of surface mutants. Reactions were performed as in Fig. 3.10A using increasing concentrations of AvCami1 variants. Proportion of cleaved substrate was obtained from densitometric analysis as in Fig. 3.11C. (**C**) Ribosomal bL12 protein depletion effect on Cami1 mRNA interferase activity. bL12 depleted ribosomes (70SΔbL12), standalone bL12 proteins, reconstituted ribosomes (70SΔbL12+bL12) or WT 70S ribosomes were pre-incubated with rib-UAG and tRNA before assay. Reactions were initiated by adding Cami1 proteins at final concentration of 300 nM (3 μM in case of CCaCami1) under stimulation by 10 μM cA_4. (**D**) Amplitudes of bL12 binding BLI signal normalized by AvCami1 loading efficiency at various bL12 concentrations. bL12 binding by cA_4-activated AvCami1-S11A wHTH mutants were compared with AvCami1-S11A in the presence and the absence of cA_4. (**E**) BLI sensorgrams showing bL12 association and dissociation to AvCami1-S11A immobilized on the biosensor in the presence (magenta shades) or absence (grey shades) of 1 μM cA_4. bL12 concentration series are marked by different colour intensities. Representative sensorgrams of three

independent measurement series are displayed. The mean values with standard deviations from three technical replicates are reported in (B) and (D).

ribosomes (Fig. 3.13C). This result suggests that the ribosome stalk-mediated capture is common for Cami1 protein family.

We have also examined the interaction between isolated bL12 and AvCami1 using biolayer interferometry (BLI). For this AvCami1 was immobilized on a biosensor surface as the ligand and submerged into a buffer with bL12 protein which was used as the analyte. The amplitude of bL12 binding signal obtained in the presence of cA_4 was substantially higher compared to that in the absence of cA_4, confirming the importance of cA_4 in stabilizing the AvCami1 conformation favouring the interaction with bL12 (Fig. 3.13D and E). As expected, AvCami1 wHTH surface mutants displayed low binding to bL12 even in the presence of cA_4 (Fig. 3.13D). This result supports the mechanism of Cami1 being captured by the ribosome stalk upon activation by cA_4.

Comparison between apo and cA_4-activated ribosome-bound AvCami1 structures revealed relatively small conformational differences in the CARF dimer and a considerable 35° tilt of both wHTH domains (Fig. 3.14). A similar cork-screw motion of wHTH domains was also suggested in the post-catalytic structure of cA_4 ring nuclease SiCrn1 (Molina et al., 2021) (see Fig. 1.9), however the authors could not conclude whether this conformation change was induced due to cA_4 binding or cleavage. AvCami1-S11A-H343A and ribosome complex structure obtained using ring nuclease-deficient mutant suggests that the binding of cA_4 is sufficient for structural rearrangement of wHTH domains since S11A mutation disrupts the cleavage of the activator (Fig. 3.9B). wHTH movement exposes its surface for the capture by bL12 as supported by cA_4-dependant binding of isolated bL12 (Fig. 3.13D and E).

In addition to wHTH movement, the entire RelE dimer is shifted diagonally ~15 Å away from the CARF-wHTH core (Fig. 3.14). This breaks the symmetry of AvCami1 dimer and enables the proximal RelE domain to fit into the ribosome cleft where mRNA is exposed (Fig. 3.12A and B). As suggested by structural models, in other Cami1 dimers RelE domains do not interact (Fig. 3.8D) and thus may adopt more flexible conformations relative to CARF-wHTH core during activation and binding to the ribosomes. Also, it remains unknown whether the activation by cA_4 is sufficient to break the symmetry of AvCami1 or accommodation to the ribosome is additionally required to stabilize the asymmetrical structure of activated protein.

The structural data together with results of bL12 interaction assays allows us to conclude that upon cA_4-activation Cami1 proteins change the

conformation of wHTH domains to be captured by bL12-CTD and then are transferred to ribosomal A-site for mRNA cleavage.

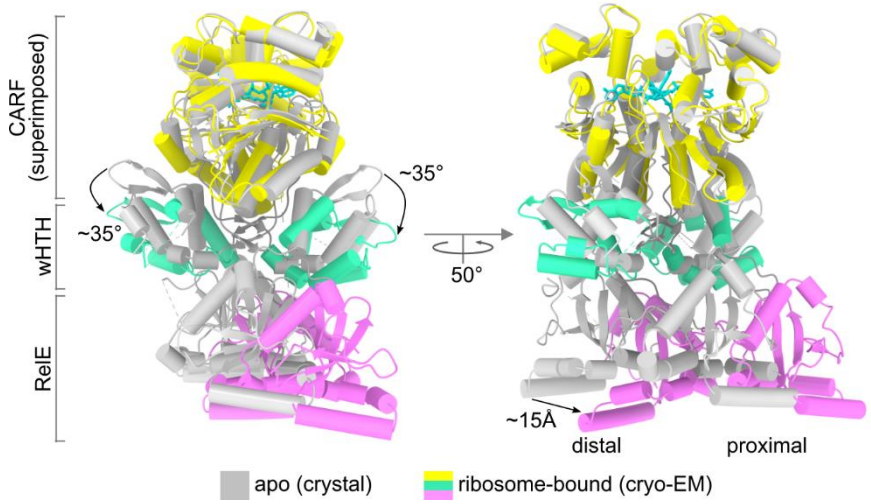


Fig. 3.14. Conformational change of AvCami1 upon cA$_4$-activation and ribosome binding. AvCami1 apo crystal structure (grey) superimposed over CARF domain dimer of cA$_4$-activated and ribosome-bound cryo-EM structure (coloured). "Distal" and "proximal" indicates the positioning of protein chains of AvCami1 dimer relative to ribosomal A-site.

3.2.6.*In vivo* toxicity of Cami1

To probe whether Cami1 could contribute to the type III CRISPR-Cas immunity we coupled Cami1 with the *S. thermophilus* type III-A CRISPR-Cas system that produces cOAs from cA$_3$ to cA$_6$ (including cA$_4$) in response to transcription from foreign DNA *in vivo* (Smalakyte et al., 2020).

The three-plasmid IPTG-inducible system in *E. coli* (Fig. 3.15A) ensured production of RNase and DNase-dead StCsm complex bound to crRNA targeted against custom transcript of a pTarget. Nuclease-deficient StCsm was used to avoid autoimmunity while allowing the perpetual production of cOAs when growing in media with inductor IPTG.

We transformed the cOA-producing *E. coli* with pCami1 plasmids encoding different Cami1 proteins to monitor their *in vivo* toxicity. We performed bacterial growth assays using either WT AvCami1, a ring nuclease-dead variant with S11A mutation or double S11A and K317A mutant which is additionally deficient in mRNA interferase activity (Fig. 3.10B). We

observed that +WT AvCami1 and +AvCami1-S11A cultures, but not +AvCami1-S11A-K317A, exhibited a growth defect upon the induction of target transcription and protein expression compared to bacteria with empty vector instead of pCami1 (Fig. 3.15B). This result shows that mRNA interferase activity of AvCami1 causes toxicity under constant stimulation by StCsm produced cA_4. Moreso, the growth of +AvCami1-S11A bacteria was repressed more than of +WT AvCami1 bacteria suggesting that CARF-mediated cA_4 cleavage could lower the intracellular concentration of the activator thus limiting the downstream activity of RelE effector domain.

To monitor mRNA cleavage activity of Cami1 *in vivo*, we used the same set of *E. coli* hosts and plasmids as in the bacterial growth assay (Fig. 3.15A). Cells were grown in a liquid medium with inducers and total RNA was extracted from them. Several abundant transcripts were depicted by northern blot using radioactively labelled hybridization probes (Appendix 2). No mRNA degradation was detectable in cells either lacking AvCami1 or encoding RelE domain mutant, but significant RNA degradation at multiple sites was observed for most transcripts after induction of WT and S11A AvCami1 (Fig. 3.15C). The translation of mRNAs cleaved in the coding sequence could cause stalling of the ribosomes since they do not reach the stop codon to disassemble (Müller et al., 2021). Stalled ribosomes in *E. coli* and other bacteria can be recovered by tmRNA, which releases ribosomes from damaged mRNAs and tags the nascent polypeptides from such ribosomes for further proteolysis (Müller et al., 2021, see also Chapter 1.6). Northern blot analysis of RNA extracted from cA_4-producing cells expressing AvCami1 revealed some traces of tmRNA cleavage products (Fig. 3.15C), indicating that cA_4-activated Cami1 could cleave tmRNAs together with mRNA, thereby interfering with ribosome recovery. Untranslated 5S ribosomal RNA was not targeted by AvCami1 (Fig. 3.15C).

Other Cami1 proteins also showed *in vivo* activity in heterologous *E. coli* host. Reduced cell growth was observed upon induction of WT CaCami1, CCaCami1 and ChCami1 similarly to AvCami1 (Fig. 3.15D). This result suggests that the conservation of translation machinery permits the spread of Cami1 protein genes during bacterial evolution (Fig. 3.6A).

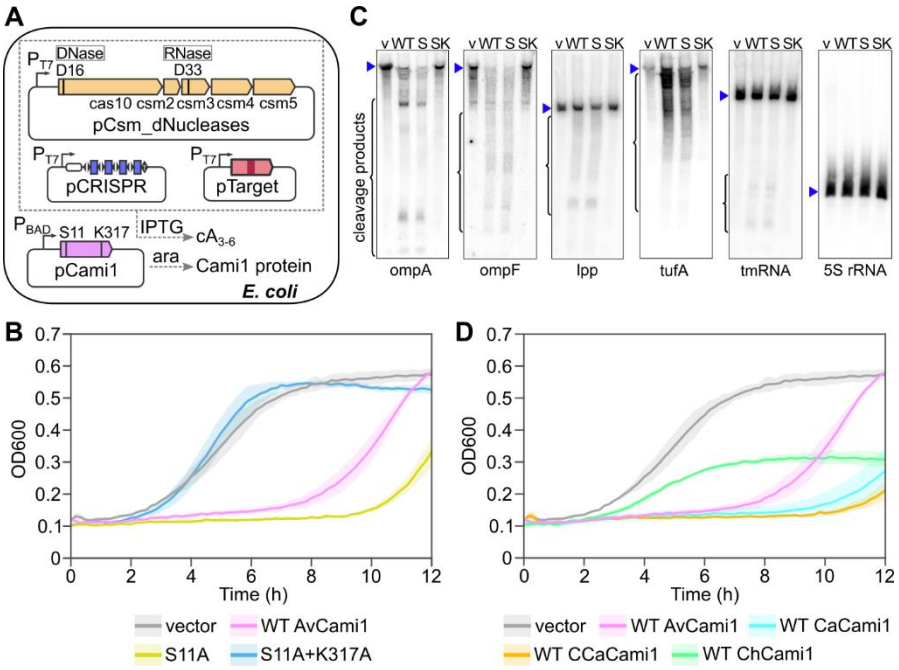


Fig 3.15. Cami1 proteins display cOA-regulated mRNA interferase activity in heterologous host. (**A**) The experimental setup (see Methods 2.2.16) to study Cami1 activity *in vivo*. *E. coli* bacteria coexpressing heterologous Cami1 protein and nuclease-deficient perpetually activated *S. thermophilus* type III-A CRISPR-Cas system were used for the bacterial growth assays (B and D) and RNA integrity analysis (C). (**B**) Bacterial growth curves. Growth was traced by measuring OD_{600} every 10 minutes. An empty plasmid vector without AvCami1 gene denoted as "vector". Lines and shades represent mean and SD, respectively (n = 3). (**C**) Northern blot analysis of endogenous lpp, ompA, ompF, tufA mRNA, tmRNA and 5S rRNA integrity in total RNA extracted from bacteria expressing activated StCsm and AvCami1 variants. "v", "vector" as in (B); "WT", WT AvCami1; "S" and "SK", S11A and S11A+K317A mutants of AvCami1, respectively. (**D**) Bacterial growth under activation of Cami1 homologs. Growth data is presented as in (B).

3.3. Mechanism of type III CRISPR-Cas defence employing Cami1

The results presented in this thesis provide details on how protection by type III CRISPR-Cas defence system is coordinated and how regulated mRNA interferase Cami1 is utilized in a downstream defence signalling pathway (Fig. 3.16). According to the established spatiotemporal regulation mechanism, an effector complex Csm/Cmr recognizes foreign transcripts and

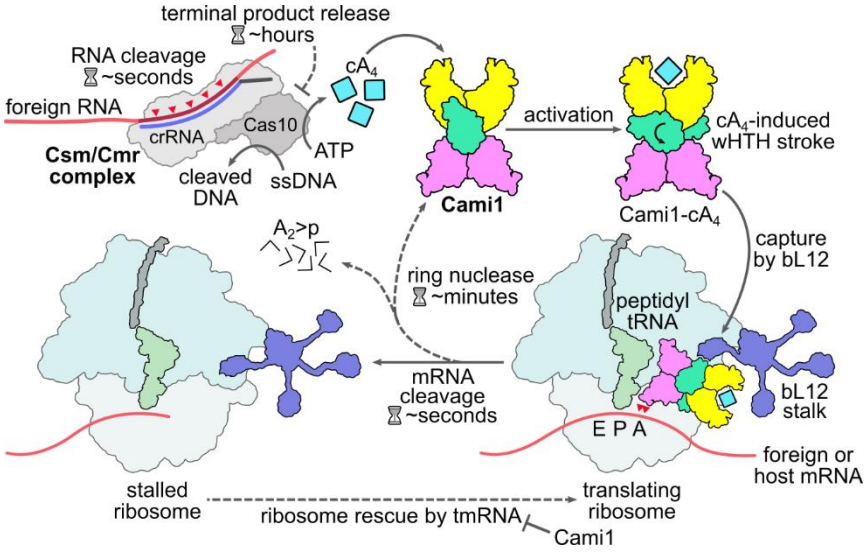


Fig. 3.16. Mechanism of Cami1 activation in type III-A CRISPR-Cas immunity. Infection is sensed by type Csm interference complex which binds crRNA-complementary foreign transcript. This triggers ssDNA degradation and synthesis of signalling molecules cA_4 in Cas10 subunit. Due to slow release of RNA cleavage products, on average Cas10 remains activated for an hour. cA_4 binding to Cami1 CARF domains (yellow) induces conformational changes of the protein that enable wHTH domain (turquoise) interaction with ribosomal stalk via bL12 protein (slate blue). Stalk-captured Cami1 is delivered into the A-site of a translating ribosome. RelE domain (magenta) of Cami1 starts cleaving mRNAs which inhibits translation and stops the production of both cellular and viral proteins. Cami1 also cleaves tmRNA, hindering one of the ribosome rescue pathways. Cleavage of cA_4 by ring nuclease activity of CARF domain deactivates Cami1 and dissociation of RNA cleavage products from Csm complexes deactivate Cas10 subunit.

activates Cas10 protein for co-transcriptional degradation of non-template DNA strand and production of cOAs for subsequent activation of ancillary CARF/SAVED proteins (reviewed in (Stella and Marraffini, 2023). However, the temporal control for Cas10 activity, especially its deactivation, remained puzzling since the cleavage of target RNA occurs in seconds whereas ssDNase and cOA synthase activities persist for tens of minutes (Kazlauskiene et al., 2017, 2016).

Our FCS-based RNA retention assays revealed the missing link explaining the regulation of type III-A effector complex (Chapter 3.1). We observed that the terminal RNA cleavage products dissociate slowly over period of hours compared to internal product released from the complex in minutes after rapid RNA cleavage. The prolonged Cas10 ssDNase activity which

diminishes over similar timescale indicates a functional association to RNA cleavage product retention.

Since the rate of cOA synthesis in Cas10 also exhibits a gradual diminish upon activation, it is tempting to speculate that both enzymatic activities of Cas10 protein are regulated in the same fashion, i.e. by RNA retention. A piece of evidence supporting this speculation comes from self vs. non-self discrimination in type III effectors (see Section 1.3.3). Once Csm/Cmr binds to an RNA target which is fully complementary to crRNA including its repeat-borne 5′-tag – a situation possible in case of antisense host CRISPR region transcription – both ssDNA degradation and cOA synthesis are inhibited. This shows that target RNA has a direct impact on both activities of Cas10 protein.

Due to prolonged Cas10 activity the intracellular concentration of cOAs builds up and ancillary effectors of CARF and SAVED families are activated. Here we structurally and functionally characterized Cami1 effector, a regulated mRNA interferase belonging to CARF protein group, and proposed its mechanism (Fig. 3.16). Cami1 effectors respond to cA_4 activators (Sections 3.2.3 and 3.2.4). According to our model, cA_4 binding induces conformational changes in dimeric Cami1 structure during which wHTH domain moves away from CARF dimer. This exposes a wHTH domain surface for interaction with ribosomal stalk protein bL12 (Section 3.2.5). Upon capture by ribosomal stalk AvCami1 accommodates into the ribosomal cleft to cleave the mRNA exposed in the A-site. Once activated Cami1 proteins cleave mRNAs in multiple ribosomes before shutting down due to cA_4 cleavage in CARF domain.

Truncated mRNAs cause staling of ribosomes resulting in overall translation inhibition. In case of phage infection, this could prevent the production of phage enzymes and coat proteins thus limiting their assembly. During translation shutdown the infectious nucleic acids can be destroyed by cellular nucleases, other defence systems, or ssDNase activity of Csm/Cmr complex.

Stalled ribosomes are alleviated by various ribosome rescue mechanisms, for example, trans-translation employing hybrid tmRNA (Chapter 1.6). However, we observed that endogenous *E. coli* tmRNA is prone to cleavage by Cami1 (Section 3.2.6). This exemplifies a strict Cami1 response resulting in a persistent translation inhibition lasting for as long as cA_4 is available.

According to the proposed model, if foreign nucleic acids are neutralized and the cell survives the infection, the dissociation of RNA cleavage products from Csm complex eventually causes the diminish of polymerase-cyclase activity and stops cOA production, while any available cA_4 molecules are cleaved by Cami1 or other ring nucleases. The emergence of these negative feedback loops suggests that such defence strategy enabled bacteria to survive

numerous infections on the course of fierce coevolution between them and their parasites.

3.4. Further directions

While the results presented in this thesis align to a dynamic model of type III CRISPR-Cas defence, it should be noted that most of the conclusions came from *in vitro* studies which do not provide the full picture of processes happening in live cell. Therefore, it remains to be tested whether the observed duration of Cas10 activities is similar during the infection under natural conditions. Also, it should be investigated whether the observed kinetic parameters can be generalized for diverse type III effectors as there are subtle structural differences between type III-A and III-B complexes in the vicinity of PFS (Section 1.3.3). In the same line, it remains to be tested how efficiently Cami1 effectors bind to the ribosome under natural settings and cleave the mRNA being actively translated. While we observed that Cami1 causes growth reduction when overexpressed (Fig. 3.15D) it should be investigated how naturally expressed Cami1 proteins compete against translation factors for ribosome binding as the latter are among the most abundant proteins in bacterial cells (Ishihama et al., 2008).

The pace of type III CRISPR-Cas ancillary effectors research is accelerating in past years providing new structures, functions, and mechanisms (Stella and Marraffini, 2023). Our comprehensive sequence analysis of CARF7 proteins allowed to enrich this group with CARF-[PD-(D/E)XK] proteins not reported in the original CARF domain classification (Makarova et al., 2020a). The PD-(D/E)XK domain is commonly found in various nucleases with restriction endonucleases being the best known example (Knizewski et al., 2007). Couple instances of CARF-[PD-(D/E)XK] fusions were recently reported, namely, Can1 which is a DNA nickase (McMahon et al., 2020) and Card1/Can2 (Rostøl et al., 2021; Zhu et al., 2021) which exhibits both DNase and RNase activities. These proteins belong to CARF4 group which shows no ring nuclease activity therefore their downregulation depends on other cOA degrading enzymes. Differently, CARF7, which known instances are ring nucleases, and PD-(D/E)XK fusion proteins could provide a more controlled nucleolytic response to the infection presumably avoiding infected cell death. Other CARF7 fusion variants to HD domains could exhibit nuclease or phosphatase activity (Aravind and Koonin, 1998) while CARF7-DUF2103 fusion proteins could function in the same manner as Cami1 since DUF2103 domain

shows distant structural homology to RelE domain (Mogila et al., 2023). cA_4 binding to known CARF7 proteins Crn1 and Cami1 induces a cork-screw twist of wHTH domains (Molina et al. 2021 and Fig. 3.14), therefore similar conformational changes could be expected in other effectors of this group.

Despite the diversity of cOA-responsive CARF and structurally related SAVED proteins (Makarova et al., 2020a), type III CRISPR-Cas systems also recruit effectors lacking these domains such as cA_3-activated DNase NucC in which the activator binds in the interface of a trimeric structure (Grüschow et al., 2021; Lau et al., 2020). Recently prof. Malcolm White′s group demonstrated that Cmr complex from *Bacteroides fragilis* produces SAM-AMP instead of cOA (Chi et al., 2023). The newly discovered activator binds to a CARF-unrelated transmembrane protein CorA leading to a loss of membrane integrity which causes cell dormancy or death of an infected cell. With ever growing sequencing capabilities more diverse ancillary effectors are being discovered with type III CRISPR-Cas systems at their core (Altae-Tran et al., 2023). These findings could help to reinforce biotechnologically important prokaryotes against infections or overcome disease causing bugs with novel medicine such as phage therapy circumventing antiviral defence systems.

Another potential line of research could stem from the observation that Cami1 interacts with the ribosome stalk protein bL12 (Section 3.2.5). Previously, it was demonstrated that ribosome stalk binds trGTPases to facilitate translation in both prokaryotes and eukaryotes (Chapter 1.7). Additionally, in eukaryotes the stalk could be hijacked by RIPs which poison ribosomes. To our knowledge, Cami1 is the first characterized prokaryotic stalk-binding protein besides trGTPases and, analogously to RIPs, it inhibits translation. Cami1 binds CTD of bL12 via wHTH domain. As different trGTPases bind bL12-CTD via their G (or G') domains, one could suggest that wHTH domains akin to those of Cami1 could be used by other unknown proteins to modulate ribosome functions. A comprehensive sequence-structure analysis and interactome studies of bL12 could provide hints about other stalk-interacting proteins. Since translation is one of the central processes in the cell, this data about its possible modulators could provide valuable insights into cell biology and lead to important applications.

CONCLUSIONS

1. After RNA cleavage by Csm3, *S. thermophilus* Csm complexes retain terminal products that maintain Cas10 DNase activity.
2. CARF-wHTH-RelE domain architecture constituting Cami1 proteins likely arose multiple times in CARF7 group phylogeny.
3. Cami1 proteins are activated by cA_4 which is cleaved in CARF domain.
4. RelE domain of activated Cami1 proteins cleaves mRNA exposed in the A site of the ribosome.
5. *In vivo* activated Cami1 cleaves translated mRNAs inhibiting bacterial growth.
6. wHTH domain of activated Cami1 proteins interact with ribosome stalk protein bL12 which facilitates mRNA cleavage in the ribosome.

SANTRAUKA

SANTRUMPOS

A_2>p/A_4>p	linijiniai di/tetraadenozino fosfatai su 2′,3′-cikliniu fosfatu
BLI	biosluoksnio interferometrija; pagal angl. *biolayer interferometry*
cA_3/cA_4/cA_6	ciklinis tri/tetra/heksaadenozino fosfatas
Cami	su CRISPR-Cas susijusi mRNR interferazė; pagal angl. *CRISPR-Cas associated mRNA interferase*
CARF	su CRSIPR susijusi Rossman sankloda; pagal angl. *CRISPR associated Rossmann fold*
Cas	su CRISPR susijęs; pagal angl. *CRISPR-associated*
cOA	cikliniai oligoadenilatai
CRISPR	pagal angl. *clustered regularly interspaced short palindromic repeats*
Crn1	su CRISPR susijusi žiedo nuklеazė; pagal angl. *CRISPR associated ring nuclease 1*
crRNR	CRISPR RNR
DNazė, RNazė	deosiribonukleazė, ribonukleazė
dRI	diferencialinis lūžio rodiklis (pagal angl. *differential refractive index*)
FCS	fluorescencijos koreliacijos spektroskopija; pagal angl. *fluorescence correlation spectroscopy*
IPTG	izopropil-β-D-1-tiogalaktopiranozidas
kb	1000 bazių porų
krio-EM	kriogeninio pavyzdžio elektroninė mikroskopija
MS	masių spektrometrija
nt	nukleotidas
OD_{600}	optinis tankis 600 nm bangos ilgio šviesai 1 cm kelyje
PDB	Protein Data Bank (baltymų struktūrų duomenų bazė)
RIP	ribosomą išveiklinantys baltymai; pagal angl. *ribosome inactivating proteins*
SAM	S-adenozilmetioninas
SEC-MALS	dydžio atskyrimo chromatografija – daugiakampė šviesos sklaida; pagal angl. *size exclusion chromatography – multiple angle light scattering*
TM	transmembraninis
tmRNR	transportinė-informacinė RNR (nuo angl. *tmRNA*)
vgDNR	viengrandinė DNR
wHTH	pagal angl. *winged helix-turn-helix*
WT	laukinis tipas; pagal angl. *wild type*

ĮVADAS

Nuožmi koevoliucija tarp bakterijų ir archėjų, kartu vadinamų prokariotais, bei jas infekuojančių virusų, vadinamų bakteriofagais arba fagais, lėmė prokariotinių apsaugos sistemų vystymąsi. Dėka šių sistemų įvairovės prokariotai geba apsisaugoti nuo fagų, kurie pasižymi didesniu skaitlingumu ir kintamumu (Mayo-Muñoz et al., 2023).

Prokariotuose yra paplitusios CRISPR-Cas sistemos, kurios suteikia adaptyvią apsaugą prieš svetimas nukleorūgštis (Makarova et al., 2020b). Defektyvios infekcijos metu šios sistemos įterpia trumpus svetimos nukleorūgšties fragmentus į šeimininko genome esantį CRISPR regioną (Barrangou et al., 2007). Į genomą įrašyta informacija apie buvusias infekcijas yra nurašoma į CRISPR regiono transkriptą, iš kurio subrandinamos individualias CRISPR RNR (crRNR) molekulės. Kartu su Cas (nuo angl. *CRISPR-associated*) baltymais crRNR sudaro kompleksus, kurie pakartotinių infekcijų metu pagal komplementarumą crRNR atpažįsta į ląstelę patekusias svetimas nukleorūgštis ir jas sunaikina (Wang et al., 2022). CRISPR-Cas sistemoms priskiriami įvairūs kompleksai, kurie gali atpažinti ir sunaikinti DNR, RNR arba abiejų tipų nukleorūgštis (Makarova et al., 2020b).

III tipo CRISPR-Cas sistemoms būdingi Csm ir Cmr kompleksai (priklausantys, atitinkamai, III-A ir III-B potipiams) infekcijos metu atpažįsta svetimus transkriptus ir sunaikina tiek juos, tiek tuo metu transkribuojamą DNR (Tamulaitis et al., 2017). Reguliuojamu DNaziniu aktyvumu pasižymi komplekso Cas10 subvienetas, kuris yra aktyvinamas kompleksui surišus svetimą RNR. Nors surišta RNR per kelias sekundes yra degraduojama komplekso RNazių Csm3 (Csm kompleksuose; Cmr kompleksuose atitinkamą funkciją atlieką Cmr4 subvienetai), Cas10 išlieka aktyvus bent kelias dešimtis minučių ir karpo vgDNR (Kazlauskiene et al., 2016). Siekiant paaiškinti šiuos rezultatus yra pasiūlytas modelis pagal kurį komplekse užlaikomi RNR karpymo reakcijos produktai galėtų stimuliuoti užsitęsiantį Cas10 aktyvumą (Kazlauskiene et al., 2016). Vis dėlto iki šiol šis modelis patikrintas nebuvo.

Neseniai nustatyta, kad Cas10 pasižymi dar vienu fermentiniu aktyvumu. Aktyvintas Cas10 iš ATP sintetina iš 3-6 AMP fragmentų sudarytus ciklinius oligoadenilatus (cOA), kurie aktyvina pagalbinius efektorius (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Šis polimerazinis-ciklazinis aktyvumas yra reguliuojamas panašiu būdu kaip ir DNazinis (Kazlauskiene et al., 2017). cOA aktyvina įvairius efektorius iš kurių labiausiai paplitę yra dimeriniai CARF domeną turintys baltymai, kurie yra skirstomi į devynias pagrindines grupes (CARF1-CARF9) (Makarova et al., 2020a). CARF domene surištas

cOA aktyvina su CARF sulietą efektorinį domeną. Efektoriniai domenai pasižymi fermentinių aktyvumų įvairove, kai kurie iš jų yra gerai funkciškai ištirti, nustatytos jų struktūros. Pvz., su CARF1, CARF2 arba CARF9 sulieti HEPN domenai pasižymi reguliuojamu RNaziniu aktyvumu (Garcia-Doval et al., 2020; Jia et al., 2019a; Samolygo et al., 2020; Smalakyte et al., 2020), su CARF4 sulieti į restrikcijos endonukleazes panašūs PD-(D/E)XK domenai pasižymi tiek DNaziniu, tiek RNaziniu aktyvumu (McMahon et al., 2020; Rostøl et al., 2021; Zhu et al., 2021), su membranomis yra susiję ir jas depoliarizuoti gali CARF8-TM efektoriai (Baca et al., 2024). Vis dėlto didžiosios dalies kitų efektorinių domenų funkcijos eksperimentiškai nėra nustatytos ir gali būti spėjamos pagal sekų panašumą į žinomus baltymus.

Dalis CARF7 grupės baltymų yra sulieti su efektoriais, kurie panašūs į RelE superšeimai priskiriamas mRNR interferazes (Makarova et al., 2020a). Pavieniai RelE baltymai yra gerai charakterizuoti toksinai būdingi II tipo toksino-antitoksino (TA) sistemoms, jie *Escherichia coli* yra neutralizuojami antitoksino RelB (Li et al., 2009). Esant mitybinių medžiagų nepritekliui, veikiant antibiotikams arba fago infekcijos metu RelB:RelE santykis ląstelėje gali sumažėti (Song ir Wood, 2020). Tokiu atveju laisvas RelE toksinas prisijungia ribosomos A srityje vykstant transliacijos elongacijai ir sukarpo transliuojamą mRNR. Tokiu būdu yra slopinama baltymų sintezė ir stabdomas ląstelės augimas (Christensen ir Gerdes, 2003; Neubauer et al., 2009).

RelE domenas, kaip dimerinių CARF baltymų efektorius, gali atrodyti neįprastas evoliucijos produktas, nes RelE baltymai veikia kaip pavieniai TA sistemų toksinai. Kita vertus, cOA surišantis CARF domenas galėtų reguliuoti CARF-RelE aktyvumą. Remiantis CARF-RelE ir RelE toksinų sekų tarpusavio panašumo duomenimis CARF-RelE baltymai šiame darbe vadinami Cami1 (nuo angl. *CRISPR-Cas-associated mRNA interferase 1*).

Šios disertacijos **objektai** yra *Streptococcus thermophilus* DGCC8004 III-A tipo CRISPR-Cas kompleksas Csm (toliau – StCsm) ir keturi Cami1 baltymai susiję su III tipo CRISPR-Cas sistemomis aptinkami skirtinguose šeimininkuose. Šio darbo **tikslai** yra detalizuoti StCsm aktyvumo reguliacijos mechanizmą ir nustatyti Cami1 funkcijas CRISPR-Cas gynyboje. Siekiant užsibrėžtų tikslų buvo iškelti šie **uždaviniai**:

1. Išmatuoti produktų užlaikymo trukmę po RNR karpymo reakcijos StCsm komplekse ir palyginti ją su Cas10 subvieneto DNazinio aktyvumo trukme.
2. Nustatyti filogenetinius ryšius tarp CARF7 grupės baltymų.
3. Nustatyti Cami1 aktyvatorius.
4. Nustatyti Cami1 efektorinio domeno fermentinį aktyvumą.

5. Parodyti Cami1 aktyvumą gyvose bakterijose.
6. Detalizuoti Cami1 reguliacijos ir aktyvumo mechanizmą pasitelkiant struktūrinius metodus.

Mokslinis naujumas ir praktinė vertė. Mokslinėje literatūroje yra išsamiai aprašyti CRISPR-Cas III-A ir III-B potipių Csm ir Cmr kompleksų veikimo principai, kurie atsikartoja tarp įvairių homologų, yra nustatytos šių kompleksų struktūros (Molina et al., 2020; Stella ir Marraffini, 2023; van Beljouw et al., 2022). Šiuo darbu buvo siekiama papildyti nusistovėjusį Csm/Cmr aktyvumo reguliacijos modelį. Nors šiuo metu žinoma, kad komplekso Cas10 subvienetas yra aktyvinamas kompleksui surišus taikinio RNR, jo aktyvumo nuslopinimo mechanizmas neaiškus. Šioje disertacijoje pristatomi fluorescencijos koreliacijos spektroskopijos (FCS) tyrimai parodė, kad nepaisant greitos RNR degradacijos, galiniai RNR karpymo produktai lieka surišti StCsm komplekse ilgiau negu valandą. Aktyvintame komplekse Cas10 subvienetas DNazinį aktyvumą išlaiko panašų laiko tarpą, kas žymi funkcinę sąsają tarp šių procesų. Papildytą modelį paremia anksčiau atliktų „savas-svetimas“ atpažinimo tyrimų rezultatai. Tuomet nustatyta, kad RNR substratų 3′ galinė sritis reguliuoja Cas10 aktyvumą priklausomai nuo to ar ji yra komplementari (t. y. atitinka „savą“, iš ląstelės CRISPR regiono kilusį taikinį) ar nekomplementari (atitinka „svetimą“ taikinį) atitinkamai crRNR 5′ sričiai (Li et al., 2024). Pirmuoju atveju Cas10 yra slopinamas, antruoju – aktyvinamas. Kartu paėmus šiame darbe ir mokslinėje literatūroje aprašytus rezultatus disertacijoje pateikiamas papildytas III tipo CRISPR-Cas reguliacijos modelis paaiškinantis kaip šie kompleksai yra išjungiami nuslopinus infekciją.

Disertacijoje pateikta su III tipo CRISPR-Cas susijusių CARF7 baltymų sekų analizė parodo kaip evoliucijos metu susiliejant baltymų domenams iš įvairių prokariotinių apsaugos sistemų susidaro nauji apsaugą suteikiantys baltymai. Filogenetinė analizė atskleidė, kad RelE superšeimos mRNR interferazės toksinas kelis kartus nepriklausomai susiliejo su CARF domenu susidarant polifiletinei Cami1 baltymų grupei. Biocheminiai aktyvumo ir *in vivo* toksiškumo tyrimai parodė, kad Cami1 jungiasi prie ribosomų ir karpo transliuojamą mRNR, kai yra aktyvinamas cA_4, kurį sintetina III tipo CRISPR-Cas kompleksas. CARF domenas, kuris suriša cA_4 aktyvatorių, jį ir perkerpa.

Struktūriniai Cami1 tyrimai pateikė naujų įžvalgų apie ribosomas. Cami1 į ribosomos A sritį patenka per ribosomos stiebo struktūrą susirišęs su stiebo baltymu bL12. Sąveika tarp bL12 C-galinio domeno ir cA_4 suaktyvinto Cami1 wHTH domeno yra būtina mRNR karpymui. Iki šiol buvo žinoma, kad bL12 sąveikauja tik su GTP surišančiais transliacijos faktoriais per jų G domeną (Liljas ir Sanyal, 2018). Tad šio darbo metu nustatyta sąveika tarp transliaciją

slopinančio Cami1 toksino ir prokariotinės ribosomos stiebo mokslinėje literatūroje yra aprašoma pirmą kartą.

Šioje disertacijoje pristatomi rezultatai pažymi reguliacijos mechanizmų svarbą prokariotinėse gynybos sistemose. Į šį aspektą būtina atsižvelgti daugėjant antibiotikams atsparių bakterijų, kurias medicinos, žemės ūkio bei maisto pramonės srityse siekiama kontroliuoti pasitelkiant fagus.

Ginamieji teiginiai:

1. *S. thermophilus* Csm kompleksas ilgam užlaiko galinius RNR karpymo produktus bei išlaiko Cas10 DNazinį aktyvumą po greitos RNR hidrolizės.
2. Evoliucijos eigoje Cami1 baltymai susiformavo susiliejant CARF ir į RelE panašiems domenams.
3. Cami1 CARF domenas suriša ir kerpa cA_4.
4. cA_4 suaktyvinti Cami1 baltymai kerpa ribosomoje surištą mRNR.
5. Aktyvintas Cami1 baltymas karpo transliuojamą mRNR gyvose *Escherichia coli* bakterijose.
6. Aktyvinto Cami1 wHTH domenas jungiasi prie ribosomos stiebo baltymo bL12 C-galinio domeno.

METODAI

Atliekant šį darbą pasitelkti įvairūs kompiuteriniai, fizikiniai, biocheminiai ir mikrobiologiniai metodai.

Laukinio tipo ir RNazinio aktyvumo mutanto (Csm3.D33A) StCsm kompleksų raiška buvo atlikta *E. coli* BL21(DE3) kamiene, kompleksai išgryninti giminingumo chromatografijos metodais. StCsm kompleksai su subrandinta crRNR atskirti atliekant gelfiltraciją. Išgrynintų kompleksų giminingumas RNR patikrintas elektroforetinio judrumo poslinkio metodu.

RNR karpymo produktų užlaikymo tyrimams naudoti hibridiniai RNR substratai su abiejuose galuose prilydytais dvejomis skirtingomis fluorescencinėmis žymėmis pažymėtais DNR oligonukleotidais (žr. schemą Appendix 3). RNR užlaikymo StCsm tyrimai atlikti FCS metodu, registruojant abiejų žymių fluorescencijos fluktuacijas vykstant RNR karpymo reakcijai. Duomenys analizuoti dvejopai. Tarpusavio koreliacijos režimu palyginant abiejų žymių fluktuaciją įvertintas abu galinius produktus surišusių StCsm kompleksų dalies pokytis reakcijos metu. Autokoreliacijos režimu abiejų žymių fluorescencijos fluktuacijos analizuotos atskirai įvertinant jomis pažymėtų biomolekulių difuzijos greičių pokyčius dėl RNR degradacijos ir disociacijos.

Eksperimentiniams duomenims pritaikyti trimatės difuzijos modeliai, žymių difuzijos pokyčiams įvertinti pritaikyti eksponentiniai modeliai.

Cas10 DNazinio aktyvumo trukmei ištirti panaudotas fluoroforu ir gesikliu pažymėtas vgDNR substratas. Reakcijos metu stebimiems fluorescencijos pokyčiams pritaikytas eksponentinis produkto susidarymo modelis ir apskaičiuota DNazės reakcijos trukmė.

Anksčiau išskirta CARF7 baltymų grupė sudaryta iš 490 sekų (Makarova et al., 2020a) buvo keliomis pakopomis praplėsta įtraukiant homologines sekas iš NCBI „nr“ baltymų sekų duomenų bazės naudojant BLAST (Altschul et al., 1997). Gautame galutiniame 1521 baltymų sekų sąraše CARF, wHTH ir efektoriniai domenai bei jų ribos buvo nustatytos remiantis jų struktūriniais modeliais paruoštais AlphaFold 2 metodu (Jumper et al., 2021). MUSCLE metodu (Edgar, 2004) išlygiavus išskirtų CARF7 domenų sekas paruoštas filogenetinis CARF7 baltymų grupės medis panaudojant IQTree (Minh et al., 2020). Tolimesniems tyrimams pasirinkti keturi su III tipo CRISPR-Cas sistemomis susiję Cami1 homologai iš *Allochromatium vinosum* DSM180, *Caldilinea aerophila* DSM14535, *Candidatus* Cloacamonas acidaminovorans str. Evry ir *Caldicellulosiruptor hydrothermalis* 108, atitinkamai, AvCami1, CaCami1, CCaCami1 ir ChCami1. Šiuos baltymus su prilietomis gimininingumo chromatografijos žymėmis koduojančios plazmidės (žr. Appendix 1) užsakytos iš Twist Bioscience.

Cami1 baltymų raiška vykdyta *E. coli* DH10B kamiene, baltymai išgryninti gimininingumo chromatografijos metodais. Cami1 baltymų oligomerinė būsena tirpale nustatyta SEC-MALS metodu. AvCami1 baltymo kristalai buvo paruošti garų difuzijos „sėdinčio lašo“ metodu, kristalinė AvCami1 struktūra nustatyta rentgeno spindulių kristalografijos metodu. Kitų Cami1 baltymų struktūros sumodeliuotos AlphaFold 2 (Jumper et al., 2021).

cA_4 arba cA_6 aktyvatorių karpymo Cami1 produktai analizuoti masių spektrometrijos metodu. AvCami1 cA_4 hidrolizės kinetikos tyrimas atliktas naudojant radioaktyviai žymėtą cA_4, reakcijos produktai analizuoti elektroforezės metodu ir detektuoti autoradiografiškai. Žiedo nukleazės aktyvaus centro mutantai buvo sukonstruoti Gibson surinkimo metodu, išgryninti kaip aprašyta prieš tai, jų aktyvumas analizuotas taip pat kaip ir laukinio tipo AvCami1.

Cami1 mRNR interferazinis aktyvumas analizuotas *in vitro* panaudojant prieš tai pasiruoštas radioaktyviu RNR substratu programuotas *E. coli* 70S ribosomas. AvCami1 RelE domeno mutantai paruošti ir išgryninti kaip aprašyta anksčiau, jų mRNR interferazinio aktyvumo greičio konstantos palygintos su laukinio tipo baltymo.

Panaudojant dvigubą žiedo nukleazės ir mRNR interferazės aktyvumo AvCami1-S11A-H343A mutantą krio-EM metodu nustatyta cA_4 surišusio AvCami1 struktūra programuotoje ribosomoje. Krio-EM paveikslų ir dalelių analizė atlikta naudojant cryoSPARC (Punjani et al., 2017), krio-EM tankio žemėlapis paruoštas naudojant cryoSPARC ir phenix (Terwilliger et al., 2018). Atominis modelis paruoštas naudojant Coot (Emsley ir Cowtan, 2004) ir phenix (Afonine et al., 2012) programinę įrangą. Struktūrinių duomenų atvaizdavimui ir palyginimui naudotos PyMOL (Schrödinger, LLC, 2015) ir ChimeraX (Pettersen et al., 2021) programos.

Sąveika su ribosomos stiebo baltymu bL12 įvertinta tikrinant Cami1 baltymų aktyvumą mRNR interferazės reakcijose naudojant ribosomas be bL12 arba į reakcijų mišinius pridedant atskirai išgryninto bL12. Analogiškai patikrintas ir AvCami1 wHTH taškinių mutantų aktyvumas. AvCami1 ir bL12 sąveika įvertinta naudojant biosluoksnio interferometrijos (angl. *biolayer interferometry*, BLI) metodą. Tam ant BLI sensorių imobilizuoti AvCami1 laukinio tipo baltymai arba wHTH domeno mutantai, baltymai aktyvinti cA_4 ir perkelti į bL12 baltymo tirpalą. Remiantis BLI sensorgramomis įvertintas sąveikos stiprumas.

Cami1 baltymų ir jų mutantų toksiškumas gyvoms *E. coli* bakterijoms įvertintas atliekant jų ir StCsm komplekso su jo taikiniu raišką. Bakterijų augimas įvertintas matuojant suspensijų optinį tankį. mRNR interferazinis aktyvumas gyvose bakterijose nustatytas iš bakterijų su aktyvintu StCsm kompleksu ir AvCami1 baltymais išgryninus visuminę RNR bei atlikus joje esančių mRNR detekciją „northern blot“ metodu.

REZULTATAI

Šiuo metu yra išskiriami šeši III tipo CRISPR-Cas sistemų potipiai nuo nuo III-A iki III-F (Makarova et al., 2020b). Dr. Gintauto Tamulaičio vadovaujamoje mokslinėje grupėje tiriame III-A potipio sistemos kompleksą Csm iš *S. thermophilus* DGC8004 (StCsm). Mūsų tyrimai atskleidė, kad StCsm atpažįsta crRNR komplementarius RNR taikinius, kurie yra perkerpami komplekso RNazės Csm3 (Tamulaitis et al., 2014), vieni pirmųjų parodėme, kad surišus taikinio RNR komplekse yra aktyvinamas Cas10 subvieneto DNazinis (Kazlauskiene et al., 2016) ir polimerazinis-ciklazinis (Kazlauskiene et al., 2017) aktyvumas. Taip pat nustatėme ir kitų komplekso subvienetų funkcijas (Mogila et al., 2019). Galiausiai parodėme, kad StCsm sintetina įvairaus

dydžio cOA, iš kurių cA_6 aktyvina susijusią CARF1-HEPN RNazę Csm6 (Kazlauskiene et al., 2017; Smalakyte et al., 2020).

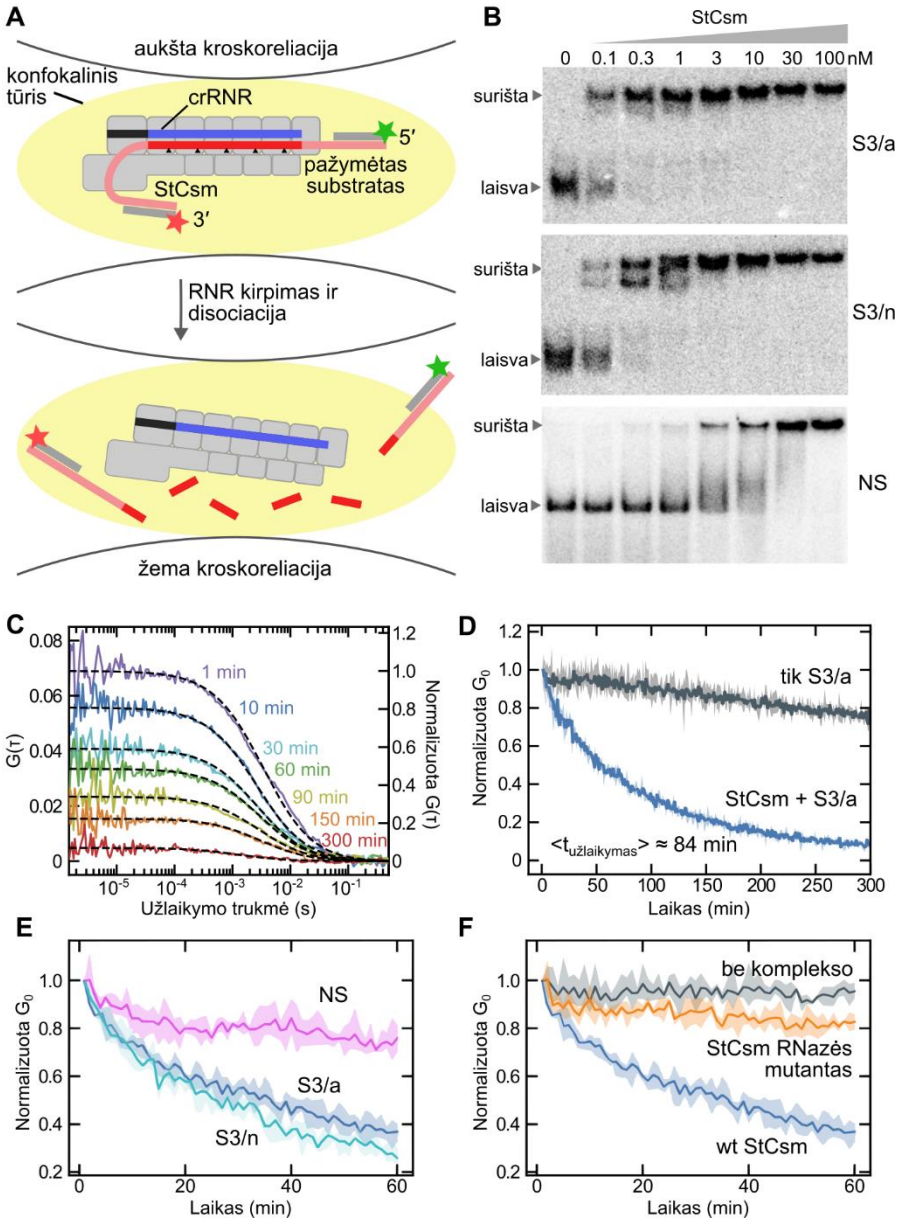
Šis darbas yra III tipo CRISPR-Cas sistemų ir StCsm komplekso tyrimų tąsa. Pirmojoje šio santraukos skyriaus dalyje aprašyti StCsm Cas10 subvieneto reguliacijos tyrimai pasitelkiant fluorescencinius metodus. Toliau pristatomi funkciniai ir struktūriniai pagalbinių baltymų Cami1 šeimos tyrimai, kurie atskleidė, kad Cami1 yra transliaciją slopinantys efektoriai aktyvinami cA_4.

❖ StCsm komplekso reguliacijos tyrimas

Ankstesnių tyrimų metu pastebėta, kad taikinio RNR aktyvintas StCsm DNazinį (Kazlauskiene et al., 2016) ir polimerazinį-ciklazinį (Kazlauskiene et al., 2017) išlaiko dešimtis minučių nors pati taikinio RNR yra degraduojama per greičiau negu minutę (Tamulaitis et al., 2014). Sukirpus RNR DNazinis ir polimerazinis-ciklazinis aktyvumas nuslopsta eksponentiškai (Kazlauskiene et al., 2017, 2016). Remiantis šiais rezultatais buvo iškelta hipotezė, jog RNR karpymo produktai lieka prisijungę prie StCsm komplekso ir aktyvina Cas10 subvienetą ilgiau negu vyksta RNR karpymas. Siekdami tai patikrinti, nusprendėme tiesiogiai sekti RNR karpymo produktų disociaciją iš komplekso pasitelkdami fluorescencinės koreliacinės spektroskopijos (FCS) metodą ir nustatytas produktų užlaikymo trukmes palyginti su Cas10 DNazinio aktyvumo trukme.

- Galiniai RNR karpymo produktai yra užlaikomi StCsm

StCsm kerpa taikinio RNR penkiose vietose, kas 6 nt susidarant keliems vidiniams ir dviems galiniams produktams (Tamulaitis et al., 2014). Mes paruošėme galuose skirtingomis fluorescencinėmis žymėmis pažymėtus RNR substratus ir vykdėme jų karpymo StCsm reakcijas. Reakcijos metu matavome mišinio fluorescenciją FCS metodu dviejų spalvų režimu. Tam dviejų žadinančių lazerių spinduliai buvo sufokusuoti į tą patį šviesos difrakcijos apribotą (konfokalinį) tūrį reakcijos mišinyje. StCsm surištoms taikinio RNR molekulėms difunduojant į arba iš konfokalinio tūrio (1A pav.) registruojamos fluorescencijos intensyvumo fluktuacijos dviejuose atskiruose kanaluose. Atlikus fluktuacijų duomenų koreliacinę analizę įvertinamas StCsm surištų (difunduojančių kartu) ir paleistų (difunduojančių atskirai) RNR fragmentų kiekio pokytis reakcijos metu.



1 pav. StCsm užlaikomų galinių RNR karpymo produktų analizė naudojant FCS metodą dviejų spalvų režimu. (A) Schema vaizduojanti StCsm ir fluorescenciškai pažymėtos RNR difuziją per konfokalinį tūrį. Nurodyti skirtingomis fluorescencinėmis žymėmis pažymėti 3′ ir 5′ RNR substrato galai. (B) StCsm gimininguma tyrime naudojamiems RNR substratams. Elektroforetinio judrumo poslinkio eksperimentai atlikti dukart. (C) Fluorescencijos kroskoreliacijos kreivės nustatytos skirtingais S3/a substrato karpymo reakcijos su WT StCsm etapais (ištisos spalvotos linijos). Kiekviena kreivė apskaičiuota pagal vienos minutės intervalų fluorescencinių signalų

duomenis (visa reakcijos trukmė yra penkios valandos). Duomenimis pritaikyti trimatės difuzijos modeliai pažymėti punktyrinėmis linijomis. (D) Kreivės vaizduojančios S3/a substrato kroskoreliacijos amplitudės G_0 priklausomybę nuo laiko reakcijos mišiniuose esant (mėlyna) ir nesant (pilka) StCsm. (E) Wt StCsm reakcijų su S3/a (mėlyna), S3/n (žydra) arba NS (purpurinė) RNR substratais kroskoreliacijos amplitudės. (F) Kroskoreliacijos amplitudės vykstant S3/a RNR karpymo reakcijoms su WT StCsm (mėlyna) arba StCsm RNazės mutantu (oranžinė). Kroskoreliacijos amplitudės mišinyje be StCsm pažymėtos pilka spalva. (D), (E) ir (F) paveiksluose ištisinės linijos atitinka trijų matavimų vidurkį, spalvoti plotai atitinka diapazoną tarp didžiausios ir mažiausios matavimų vertės. Visos kreivės normalizuotos pagal pirmąjį duomenų tašką. NS – nespecifinis RNR substratas.

Šiems eksperimentams paruoštas subrandinta homogeniška crRNR praturtintas StCsm kompleksas be gryninimo žymių. Ruošiant kompleksą naudotas S3 skirtuką iš *S. thermophilus* DGCC8004 CRISPR2 regiono turinčią crRNR koduojantis konstruktas (Tamulaitis et al., 2014). Mes paruošėme 180 nt dydžio RNR substratus (Appendix 2) ir nustatėme WT StCsm gimininguma šiems substratams elektroforetinio judrumo poslinkio metodu. Specifiniai S3/a ir S3/n substratai, kurie yra pilnai komplementarūs S3 skirtukui, pasižymėjo aukštu gimininguma lyginant su nekomplementariu NS substratu (1B pav.). S3/a nėra komplementarus 8 nt dydžio crRNR 5′-galinei žymai ir atitinka „svetimą“ RNR taikinį infekcijos metu, o S3/n, kuris yra komplementarus visai crRNR sekai, atitinka CRISPR regiono priešprasminį („savą“) transkriptą (Appendix 3); crRNR 5′-galinei žymai nekomplementarūs RNR substratai aktyvina Cas10, o komplementarūs šį subvienetą slopina (Kazlauskiene et al., 2016). StCsm gimininguma nukleorūgštims tyrimas parodė, kad šis kompleksas pasižymi aukštu gimininguma skirtukui komplementariems RNR substratams (K_d < 0.3 nM) kaip stebėta anksčiau (Mogila et al., 2019; Tamulaitis et al., 2014).

FCS tyrimams RNR substratus pažymėjome prilydydami 5′-ATTO647N („Red-probe“) ir 3′-ATTO532 („Green-probe“) pažymėtus DNR oligonukleotidus prie, atitinkamai, 3′ ir 5′ RNR galinių regionų (Appendix 2 ir Appendix 3). Šiems substratams prisijungus prie StCsm, raudonoji ATTO647N žymė atsiduria šalia Cas10 esančioje StCsm komplekso pusėje, o žalioji ATTO532 žymė – nuo Cas10 nutolusioje pusėje.

Prieš kiekvieną FCS matavimą StCsm kompleksai (10 nM) buvo inkubuojami kartu su pažymėtais RNR substratais (2 nM) reakcijos mišinyje be dvivalenčių jonų. Reakcija pradėta į reakcijos mišinį pridedant $MnCl_2$, kuris būtinas Csm3 RNaziniam aktyvumui (Tamulaitis et al., 2014). Jeigu kompleksas yra surišęs neperkirptą RNR arba jos galinius kirpimo produktus, jis per konfokalinį tūrį difunduoja su abiem fluoroforais, kas atitinka aukštą

kroskoreliaciją (angl. *cross-correlation*) tarp abiejų fluorescencijos signalų. StCsm paleidus bent vieną iš galinių produktų fluoroforai per konfokalinį tūrį difunduoja atskirai, signalai tokiu atveju nustoja koreliuoti (1A pav.). Didžiąją dalį FCS matavimų atliko dr. Patrick Irmisch.

Pirmiausiai nustatėme StCsm aktyvinančio S3/a RNR substrato karpymo produktų disociacijos dinamiką. Tuo tikslu apskaičiavome kroskoreliacijos funkcijas tarp abiejų fluoroforų signalų skirtingais RNR karpymo reakcijos eigos etapais (ištisinės linijos 1C pav.). Kroskoreliacinėms funkcijoms buvo pritaikytas trimatės difuzijos modelis:

$$G(\tau) = G_0 \frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)} \frac{1}{\sqrt{1 + \frac{\tau}{k^2 \tau_D}}}, \qquad [1]$$

kur G_0 yra kroskoreliacijos amplitude, τ_D vidutinis difuzijos laikas, $k = 4.9$ (gamintojo nurodyta) specifinė konstanta aprašanti vertikalų konfokalinio tūrio ištįsimą lyginant su sfera (punktyrinės linijos 1C pav.). Kroskoreliacijos amplitudė yra proporcinga abiem fluoroforais pažymėtų biomolekulių koncentracijai (Bacia ir Schwille, 2007). Šiuo atveju G_0 vertė proporcinga neperkirptos RNR arba StCsm komplekse surištų galinių karpymo produktų porų suminei koncentracijai. Atvaizdavus G_0 kaip funkciją nuo reakcijos laiko stebimas eksponentinis nykimas (mėlyna kreivė 1D pav.). Taip pat apskaičiavome G_0 kreivę ir mišinyje be StCsm komplekso (pilka kreivė 1D pav.), šiuo atveju amplitudės pokytis buvo nedidelis. Normalizavome kroskoreliacijos duomenis matuojant reakcijos mišinį su StCsm pagal duomenis iš mišinio be komplekso ir amplitudės pokyčiui pritaikėme eksponentinio nykimo modelį:

$$R(t) = (1 - R_\infty) e^{-t/\langle t_{užlaikymas} \rangle} + R_\infty, \qquad [2]$$

kur $R(t)$ yra amplitudė laiku t, R_∞ yra amplitudė atitinkanti neperkirptą arba nepaleistą RNR ir $\langle t_{užlaikymas} \rangle$ yra vidutinė RNR galų užlaikymo trukmė. Remiantis trijų nepriklausomų matavimų duomenimis $\langle t_{užlaikymas} \rangle$ = 84 ± 1 min, kas žymi, kad RNR karpymo reakcijos metu abu galiniai produktai vidutiniškai yra užlaikomi daugiau negu valandą.

Atlikome FCS matavimus ir su kitais RNR substratais. Reakcijos mišiniuose su WT StCsm naudojant crRNR komplementarius S3/a ir S3/n yra stebimas G_0 sumažėjimas laikui bėgant, tačiau naudojant NS substratą G_0 keičiasi nežymiai (1E pav.). Reakcijos mišiniuose, kuriuos S3/a buvo inkubuojamas su StCsm komplekso RNazės mutantu arba be komplekso, stebimas

nedidelis G_0 pokytis lyginant su matavimų, kuomet buvo naudotas WT StCsm, rezultatais (1F pav.). Šie rezultatai patvirtina, kad kroskoreliacijos amplitudė ilgainiui sumažėja dėl RNR karpymo Csm3 subvienetuose ir produktų difuzijos iš kompleksų.

FCS duomenis taip pat apdorojome vienos spalvos (autokoreliacijos) režimu siekdami patikrinti ar kroskoreliacijos amplitudės sumažėjimą StCsm karpant S3/a substratą nulėmė kurio nors vieno galinio produkto disociacija. Tam atskirai apskaičiavome autokoreliacijos kreives raudonajam ir žaliajam fluoroforui (2A pav.), kuriais, atitinkamai, buvo pažymėtos 3′-galo ir 5′-galo RNR substratų pusės. Pastebėjome, kad reakcijos metu autokoreliacijos kreivės pasislenka trumpesnių užlaikymo trukmių link, kas žymi, jog reakcijos metu padaugėja iš komplekso disociavusių galinių RNR produktų, kurie dėl mažesnės masės greičiau difunduoja per konfokalinį tūrį (Ries ir Schwille, 2012). Autokoreliacijos kreivėms pritaikėme praplėstą trimatės difuzijos modelį apimantį dviejų tipų (dviejų skirtingų masių) daleles – lėčiau difunduojančius StCsm komplekse užlaikytus RNR galus ir greičiau difunduojančius iš komplekso paleistus RNR galus.

Raudonojo fluoroforo autokoreliacijos kreivėms buvo pritaikytas praplėstas difuzijos modelis aprašantis dviejų skirtingų dalelių difuziją su atitinkamomis amplitudėmis G_1 ir G_2 bei difuzijos laikais τ_1 ir τ_2:

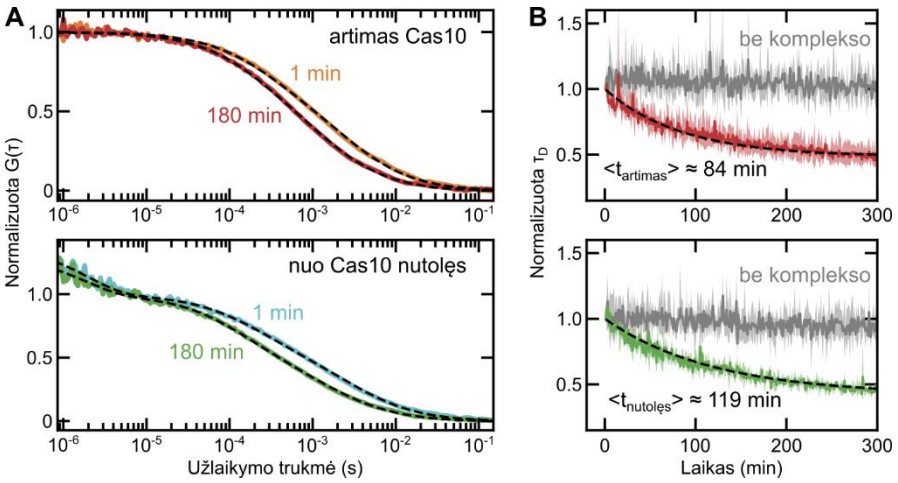
$$G(\tau) = G_1 \frac{1}{\left(1+\frac{\tau}{\tau_1}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_1}}} + G_2 \frac{1}{\left(1+\frac{\tau}{\tau_2}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_2}}}. \qquad [3]$$

Žaliojo fluoroforo autokoreliacijai taip pat taikytinas dviejų difunduojančių dalelių grupių modelis, tačiau dalyje žaliųjų fluoroforų buvo sužadinta tripletinė būsena, kurios metu fluorescencija nevyksta (Ha ir Tinnefeld, 2012). Dėl kelias mikrosekundes trunkančios tripletinės būsenos (blyksėjimo) yra stebimas autokoreliacijos padidėjimas atitinkamame užlaikymo trukmių intervale (2A pav.). Ši autokoreliacijos dedamoji nėra susijusi su milisekundžių trukmės molekulių difuzija, todėl modelis buvo praplėstas įtraukiant papildomą amplitudę T ir tripletinės būsenos trukmę τ_T (Lakowicz, 2006):

$$G(\tau) = \left(G_1 \frac{1}{\left(1+\frac{\tau}{\tau_1}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_1}}} + G_2 \frac{1}{\left(1+\frac{\tau}{\tau_2}\right)} \frac{1}{\sqrt{1+\frac{\tau}{k^2\tau_2}}} \right) \left(1 + \frac{T}{1-T} e^{-\tau/\tau_T}\right). \qquad [4]$$

Pritaikyti modeliai atitiko eksperimentiškai gautas autokoreliacijos kreives visame tirtame užlaikymo trukmių intervale (punktyrinės linijos 2A pav.). Pagal gautas greičiau ir lėčiau difunduojančių dalelių amplitudes ir difuzijos trukmes apskaičiavome abiejų grupių svertines vidutines difuzijos trukmes τ_D kiekviename minutės trukmės RNR karpymo reakcijos intervale:

$$\tau_D = \frac{G_1\tau_1 + G_2\tau_2}{G_1 + G_2}. \qquad [5]$$



2 pav. Fluorescencijos autokoreliacijos spektroskopijos būdu nustatytos galinių RNR karpymo produktų užlaikymo trukmės. (**A**) Šalia Cas10 esančio (pažymėto raudonuoju ATTO647N fluoroforu) ir nuo Cas10 nutolusio (pažymėto žaliuoju ATTO532 fluoroforu) galinių produktų autokoreliacijos kreivės nustatytos S3/a RNR karpymo reakcijos su WT StCsm metu. Pavaizduotos kreivės gautos apdorojus vienos minutės trukmės duomenis 1 ir 180 min po reakcijos pradžios (ištisinės linijos). Punktyrinėmis linijomis pažymėti pritaikyti modeliai (žr. tekste) aprašantys greičiau ir lėčiau difunduojančias daleles. Pavaizduotos reprezentatyvios kreivės iš trijų pakartojimų. (**B**) Galinių RNR karpymo produktų vidutinės difuzijos trukmės pokyčiai priklausomai nuo reakcijos laiko reakcijos mišinyje esant (raudona ir žalia ištisinės linijos) arba nesant (pilkos linijos) WT StCsm. Pritaikyti eksponentinio nykimo modeliai ([5] lygtis) pažymėti punktyrinėmis linijomis. Duomenys normalizuoti pagal pirmąjį kreivės tašką.

τ_D funkcija nuo reakcijos laiko atvaizduota ištisinėmis linijomis 2B pav. Vidutinės difuzijos trukmės pokyčiai reakcijos mišiniuose esant StCsm atitiko eksponentinį nykimą, tačiau mišiniuose be StCsm kito nežymiai. Pirmosioms kreivėms pritaikę eksponentinio nykimo modelį (analogišką [2] lygčiai, pažymėtą punktyrinėmis linijomis 2B pav.) nustatėme vidutinę užlaikymo trukmę

$\langle t_{artimas} \rangle$ = 84 ± 2 min šalia Cas10 esančiam 3′-galo produktui ir $\langle t_{nutolęs} \rangle$ = 119 ± 3 min nuo Cas10 nutolusiam 5′-galo produktui.

Darant prielaidą, kad abu galiniai produktai yra užlaikomi nepriklausomai, vidutinė bent vieno iš jų užlaikymo trukmė galėtų būti apskaičiuojama taip:

$$\frac{1}{\langle t_{užlakymas,nepr} \rangle} = \frac{1}{\langle t_{artimas} \rangle} + \frac{1}{\langle t_{nutolęs} \rangle}, \qquad [6]$$

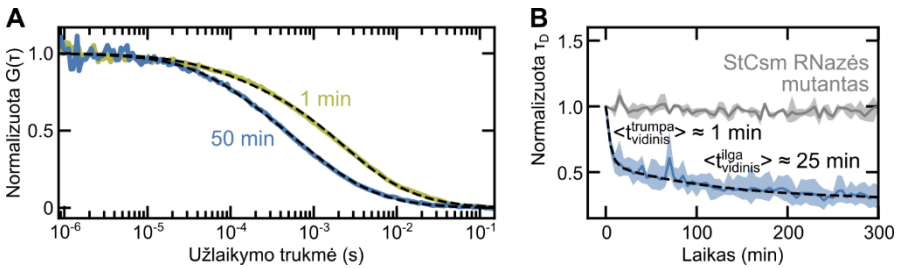
kas atitiktų $\langle t_{užlakymas,nepr} \rangle \approx 50$ min. Vis dėlto prieš tai kroskoreliacinės analizės metu nustatyta bendro galų užlaikymo trukmė yra $\langle t_{užlaikymas} \rangle$ = 84 ± 1 min (pagal 1D pav.). Šis neatitikimas galimai atsiranda dėl to, kad RNR galai po karpymo reakcijos yra paleidžiami ne atsitiktinai, o priklausomai vienas nuo kito, tačiau šiam spėjimui patikrinti reikėtų atskirų tyrimų. Nepaisant to autokoreliacijos kreivių analizė atskleidė, kad abu galiniai RNR karpymo produktai komplekse yra užlaikomi po ilgiau negu valandą.

Toliau tikrinome kaip ilgai yra užlaikomi vidiniai RNR karpymo produktai. Tam panaudojome skirtuko srityje ATTO647N fluoroforu pažymėtą 68 nt S3/iL substratą (Appendix 2). StCsm surišus šį substratą fluorescencinė žymė turėtų atsidurti priešais Csm3.3 (trečiąjį nuo Cas10) subvienetą (Appendix 3). S3/iL ir StCsm reakcijos mišinio fluorescencija pridėjus $MnCl_2$ buvo registruojama FCS metodu, duomenys analizuoti autokoreliacijos režimu. Reakcijos pradžioje autokoreliacijos kreivė (geltona linija 3A pav.) atitinka prieš tai analizuotas autokoreliacijos kreives (2A pav.), tačiau reakcijos metu dėl S3/iL karpymo ši autokoreliacijos kreivė į trumpesnių užlaikymo trukmių ruožą pasislenka greičiau (mėlyna linija 3A pav.). Autokoreliacijos kreivėms pritaikėme prieš tai naudotą dviejų difunduojančių dalelių grupių modelį (lygtis [3]), kuris atitiko eksperimentinius duomenis (punktyrinės linijos 3A pav.). Svertines vidutines difuzijos trukmes vienos minutės reakcijos intervaluose apskaičiavome taip pat, kaip ir prieš tai, pritaikę [5] lygtį. Atvaizdavę vidutinės difuzijos trukmės priklausomybę nuo reakcijos laiko pastebėjome iš dviejų fazių sudarytą pokytį (mėlyna linija 3B pav.). $\tau_D(t)$ duomenims pritaikėme nykimo modelį atitinkantį dviejų eksponentinio nykimo modelių sumą:

$$\tau_D(t) = A_1 e^{-t/\langle t_1 \rangle} + A_2 e^{-t/\langle t_2 \rangle} + \tau_\infty, \qquad [7]$$

kur τ_∞ yra santykinė trumpų fragmentų difuzijos trukmė eksperimento pabaigoje, $A_1 = 1 - A_2 - \tau_\infty$ ir A_2, atitinkamai, yra dviejų difunduojančių molekulių grupių, kurios vidutiniškai užlaikomos $\langle t_1 \rangle$ arba $\langle t_2 \rangle$, amplitudės.

Modelis atitiko duomenis (punktyrinė linija 3B pav.), pagal jį apskaičiuotos užlaikymo trukmės $\langle t_1 \rangle$ ir $\langle t_2 \rangle$. Jas palyginę priskyrėme tokias užlaikymo trukmių vertes: $\langle t_{vidinis}^{trumpa} \rangle$ = 1.0 ± 0.1 min ir $\left\langle t_{vidinis}^{ilga} \right\rangle$ = 25 ± 4 min. RNR karpymo StCsm metu susidaro ne tik galutiniai ir 6 nt ilgio vidiniai, bet ir tarpiniai ≥12 nt produktai (Mogila et al., 2019; Tamulaitis et al., 2014). Manome, kad trumpoji fazė atitinka galutinių vidinių karpymo produktų užlaikymą, o ilgoji fazė – tarpinių produktų užlaikymą. Panaudodami StCsm RNazinio aktyvumo mutantą atlikome kontrolinį matavimą tokiomis pačiomis sąlygomis. Kaip ir tikėtasi, vidutinė molekulių difuzijos trukmė šio matavimo metu išliko nepakitusi (pilka linija 3B pav.), kas žymi, jog RNR nebuvo karpoma.



3 pav. Fluorescencijos autokoreliacijos spektroskopijos būdu nustatytos vidinių RNR karpymo produktų užlaikymo trukmės. Vidiniu ATTO647N fluoroforu pažymėto S3/iL substrato karpymo su WT StCsm produktų autokoreliacijos kreivės. Pavaizduotos kreivės gautos apdorojus vienos minutės trukmės duomenis 1 ir 50 min po reakcijos pradžios (ištisinės linijos). Punktyrinėmis linijomis pažymėti pritaikyti modeliai aprašantys greičiau ir lėčiau difunduojančias daleles. Pavaizduotos reprezentatyvios kreivės iš trijų pakartojimų. (**B**) Vidinių RNR karpymo produktų vidutinės difuzijos trukmės pokyčiai priklausomai nuo reakcijos laiko naudojant WT StCsm (mėlyna ištisinė linija) arba StCsm RNazės mutantą (pilka linija). Pritaikytas dviejų fazių eksponentinio nykimo modelis ([7] lygtis) pažymėtas punktyrine linija. Duomenys normalizuoti pagal pirmąjį kreivės tašką.

Vidinių RNR karpymo produktų difuzijos matavimai parodė, kad šie produktai StCsm komplekse yra užlaikomi 10-100 kartų trumpiau negu galiniai RNR karpymo produktai. Tai leidžia teigti, kad greitą RNR karpymo kinetiką labiau atitinka vidinių, o ne galinių, produktų užlaikymo trukmė.

- Cas10 aktyvumo trukmė atitinka RNR galų užlaikymo trukmę

Tolimesniame etape tyrime nuo RNR priklausomo Cas10 subvieneto aktyvumo reguliaciją laike. Tam panaudojome 6 nt dydžio FQ-DNA oligonukleotidą, kuris pažymėtas 6-FAM fluoroforu 5′-gale ir BHQ-1 fluorescencijos

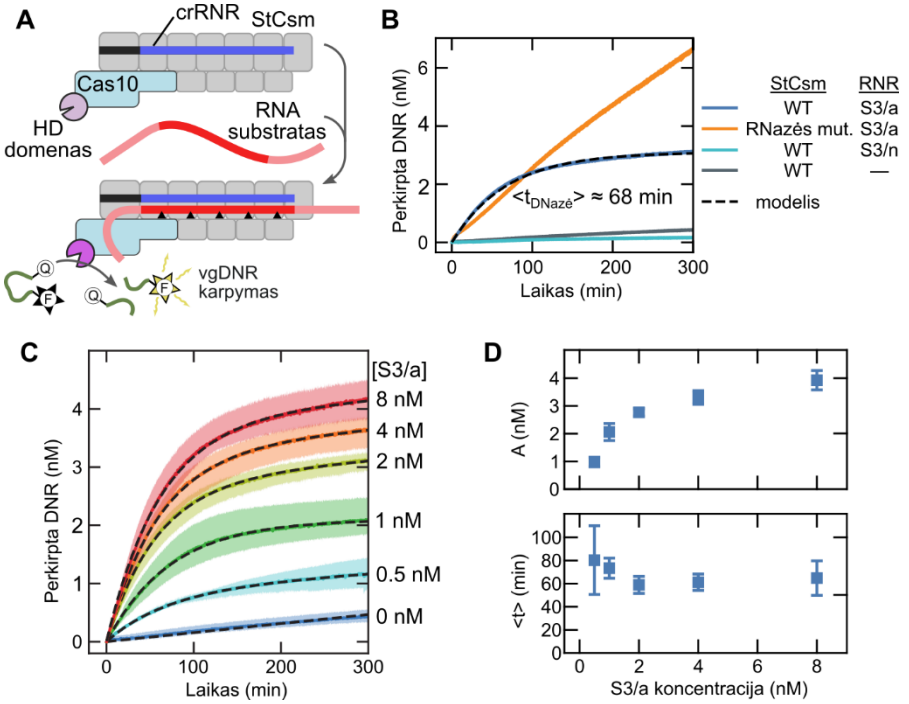
gesikliu 3′-gale (Appendix 2). Neperkirptame DNR substrate fluoroforas ir gesiklis yra šalia vienas kito, todėl fluorescencija yra gesinama. Reakcijos mišinyje StCsm kompleksą aktyvinus RNR substratu dėl Cas10 DNazinio aktyvumo DNR yra kerpama, fluoroforas atskiriamas nuo gesiklio ir stebimas fluorescencijos padidėjimas (4A pav.). Karpydami žinomos koncentracijos FQ-DNA substratą mikrokų nukleaze nustatėme fluorescencijos intensyvumo priklausomybę nuo perkirptos DNR kiekio (Appendix 4). Tolimesnius StCsm Cas10 DNazinio aktyvumo matavimus atliko dr. Patrick Irmisch.

Pirmiausiai ištyrėme StCsm Cas10 DNazę aktyvinimui naudodami S3/a RNR substratą (mėlyna linija 4B pav.). Reakcijos atliktos naudojant 50 kartų didesnį FQ-DNA perteklių lyginant su WT StCsm ir S3/a, kurių reakcijos mišinyje buvo po 2 nM. Reakcijos metu buvo stebimas didėjantis perkirptos DNR kiekis, kuris galiausiai stabilizavosi, kas žymi slopstantį DNazės aktyvumą. Perkirptos DNR kaupimuisi reakcijos metu buvo pritaikytas eksponentinio pokyčio modelis:

$$c(t) = A\left(1 - e^{-t/\langle t_{DNazė}\rangle}\right), \qquad [8]$$

kur $c(t)$ yra perkirpto DNR substrato koncentracija laiku t, A atitinka visos perkirptos DNR kiekį ir $\langle t_{DNazė}\rangle$ yra vidutinė DNazės aktyvumo trukmė reakcijos metu. Nustatyta $\langle t_{DNazė}\rangle$ = 68 ± 1 min vertė atitinka stebėtą galinių (bet ne vidinių) RNR karpymo produktų užlaikymo trukmę $\langle t_{užlaikymas}\rangle$ = 84 ± 1 min ir labiau panaši į šalia Cas10 esančio RNR galo užlaikymo trukmę $\langle t_{artimas}\rangle$ = 84 ± 2 min negu nuo Cas10 nutolusio galo užlaikymo trukmę $\langle t_{nutolęs}\rangle$ = 119 ± 3 min. Vidutinė DNazės aktyvumo trukmė yra ilgesnė negu vidinių substratų užlaikymo trukmė (3 pav.). Šie duomenys paremia mūsų hipotezę, jog komplekse užlaikyti galiniai RNR karpymo produktai stimuliuoja DNazinį aktyvumą.

Atlikome kontrolinius matavimus be RNR substrato (pilka linija 4B pav.) arba naudojant StCsm slopinantį S3/n substratą (žydra linija 4B pav.), kuris yra komplementarus crRNR 5′ sričiai. Abiem atvejais pastebėjome, kad DNazinis aktyvumas yra žemas. Reakcijos mišinyje be RNR stebimas foninis DNazės aktyvumas, kuris buvo stebėtas ir anksčiau (Kazlauskiene et al., 2016). Taip pat atlikome DNazinio aktyvumo matavimus naudodami StCsm RNazės mutantą ir aktyvinančią S3/a RNR (oranžinė linija 4B pav.). Šiuo atveju pastebėjome neslopstantį DNazinį aktyvumą, kas patvirtina ankstesnius pastebėjimus, kad RNR karpymas yra būtinas Cas10 DNazinio aktyvumo nuslopinimui po StCsm aktyvacijos (Kazlauskiene et al., 2016).



4 pav. StCsm Cas10 DNazės aktyvumo tyrimai. (**A**) StCsm komplekso DNazės aktyvacijos RNR schema. Perkirpus FQ-DNA substratą nuo gesiklio atskiriamas fluoroforas. (**B**) Perkirptos DNR kiekio priklausomybė nuo reakcijos laiko. Linijos atitinka trijų pakartojimų vidurkį. Reakcijos pradėtos sumaišant 2 nM StCsm (WT arba RNazės mutantą) kartu su 2 nM RNR substrato ir 100 nM FQ-DNA (RNazės mutanto atveju – 200 nM), reakcijos komponentai pateikti legendoje. DNR karpymo reakcijai naudojant WT StCsm ir S3/a pritaikytas eksponentinis modelis ([8] lygtis) pažymėtas punktyrine linija. (**C**) DNazės reakcijos keičiant aktyvinančios RNR koncentraciją reakcijos mišinyje. Spalvotos linijos atitinka trijų matavimų vidurkį, atitinkami spalvoti plotai žymi rėžius tarp mažiausios ir didžiausios matavimų vertės. Reakcijos atliktos kaip (B) naudojant WT StCsm ir nurodytas S3/a substrato koncentracijas. Punktyrinės linijos žymi pritaikytą eksponentinio kitimo su foniniu aktyvumu modelį ([9] lygtis). (**D**) Aktyvinto StCsm perkirptos DNR kiekis A ir vidutinė DNazės aktyvumo trukmė $\langle t \rangle$ priklausomai nuo naudotos RNR koncentracijos. Vertės gautos (C) pateiktiems duomenims pritaikius eksponentinio kitimo su foniniu aktyvumu modelį. Paklaidos nurodo standartinį nuokrypį.

Esant RNR substrato nepritekliui parodėme, kad jo karpymo produktai užlaikomi $\langle t_{užlaikymas} \rangle = 84 \pm 1$ min (1D pav.). Anksčiau nustatyta, kad esant substrato pertekliui StCsm apsisukimų skaičius yra $k_{\mathrm{cat}} = 3.0 \pm 0.6$ min^{-1}, kas atitinka tris perkirptas RNR substrato molekules viename komplekse per minutę (Kazlauskiene et al., 2016). Tai reiškia, kad padidinus RNR koncentraciją, substrato molekulės gali greitai pakeisti iš StCsm nepaleistus karpymo

produktus. Darant prielaidą, jog DNazinį aktyvumą iš tikrųjų stimuliuoja surišti RNR substrato galai, RNR substrato pakeitimas esant jo pertekliui turėtų tik nežymiai lemti suminį perkirptos DNR kiekį. Taip yra dėl to, kad RNR perteklius yra greitai (lyginant su užlaikymo trukme) degraduojamas dėl aukšto RNazinio aktyvumo, tad Cas10 santykinai ilgiau būtų stimuliuojamas ne surišto RNR substrato, o jo karpymo produktų. Esant RNR nepritekliui aktyvintų Cas10 subvienetų kiekis (ir, atitinkamai, perkirptos DNR kiekis) būtų apribotas substrato koncentracijos.

Siekdami tai patikrinti, išmatavome DNazinį aktyvumą naudodami įvairias S3/a RNR substrato koncentracijas išlaikydami WT StCsm ir FQ-DNA DNR substrato koncentracijas kaip prieš tai (4C pav.). Duomenims pritaikėme eksponentinį modelį atsižvelgiant į foninį aktyvumą:

$$c(t) = A\left(1 - e^{-t/\langle t \rangle}\right) + k_f t, \qquad [9]$$

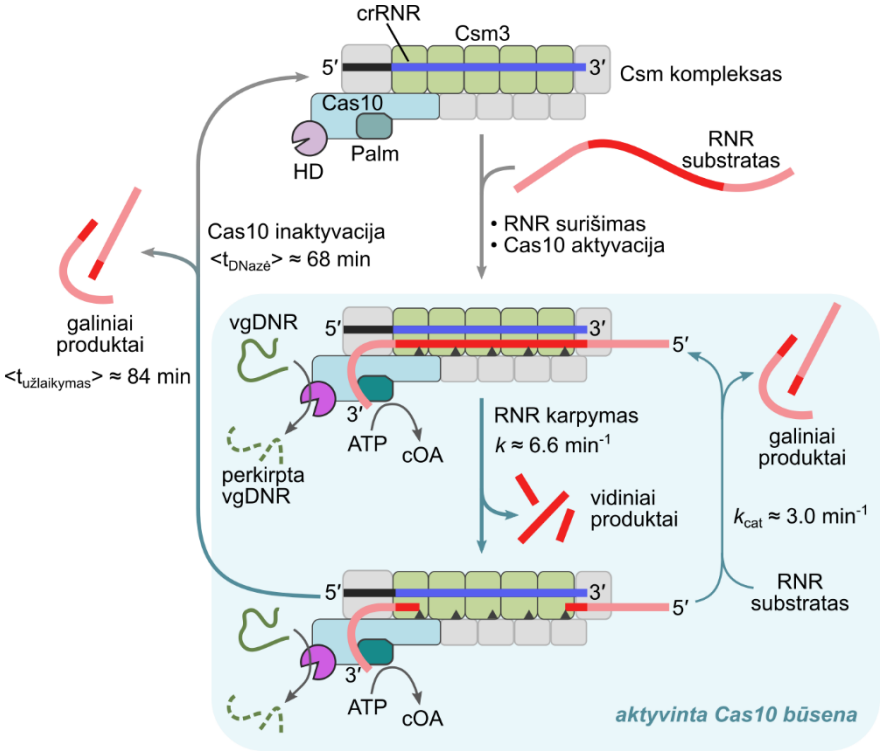
kur amplitudė A yra visas aktyvinto komplekso perkirptos DNR kiekis, $\langle t \rangle$ yra vidutinė DNazinio aktyvumo trukmė ir k_f yra foninio DNazinio aktyvumo, kuris reikšmingas mėginiuose be taikinio RNR (Kazlauskiene et al., 2016), greitis. Šį modelį pritaikę DNazinio aktyvumo duomenims (punktyrinės linijos 4C pav.) nustatėme perkirptos DNR kiekius ir DNazinio aktyvumo trukmes reakcijos mišiniuose su įvairiomis aktyvinančios RNR koncentracijomis (4D pav.).

Esant S3/a RNR nepritekliui aktyvinto komplekso perkirptos DNR kiekis A didėjo didinant RNR substrato koncentraciją. Esant RNR pertekliui A didėjo nežymiai didinant RNR koncentraciją. Keičiant RNR koncentraciją DNazinio aktyvumo trukmė $\langle t \rangle$ keitėsi nežymiai 60-80 min ribose.

Šie rezultatai atitinka prieš tai aprašytas prielaidas, jog didinant RNR substrato koncentraciją aktyvintų StCsm kompleksų dalis didėja iki tol, kol RNR koncentracija tampa didesnė negu StCsm. Cas10 aktyvumo trukmė nepriklauso nuo RNR koncentracijos ir yra artima galinių RNR kompleksų užlaikymo trukmei. Tai leidžia susieti šiuos du procesus, kaip aptariama kitame skyriuje.

- Pagrindiniai Csm komplekso reguliacijos mechanizmo bruožai

FCS eksperimentų RNR karpymo produktų užlaikymo duomenys bei DNazinio aktyvumo matavimai leidžia papildyti III-A tipo CRISPR-Cas efektorinio Csm komplekso aktyvumo reguliacijos mechanizmą (5 pav.).



5 pav. Siūlomas Csm komplekso reguliacijos mechanizmas. Schemoje pavaizduota Csm komplekso reguliacija prisijungus specifinei RNR. Csm3 subvienetų perkirptos RNR galiniai produktai stimuliuoja Cas10 aktyvumą. Esant RNR pertekliui galiniai produktai pakeičiami substrato molekulėmis nepertraukiant Cas10 aktyvumo. Žr. detalų mechanizmo aprašymą tekste. RNR karpymo greičio konstanta k esant substrato nepritekliui (esant 25 °C) ir apsisukimų skaičius k_{cat} (esant 37 °C) buvo nustatyti anksčiau, atitinkamai, (Tamulaitis et al., 2014) ir (Kazlauskiene et al., 2016).

Fluorescencijos kroskoreliacijos matavimai naudojant dviem fluoroforais pažymėtus RNR substratus parodė, kad StCsm sukarpius taikinio RNR abu galai yra užlaikomi komplekse ilgiau negu valandą, $\langle t_{užlaikymas} \rangle = 84 \pm 1$ min (1 pav.). Ši trukmė skiriasi nuo sekundes trunkančio RNR karpymo esant substrato nepritekliui arba pertekliui (Tamulaitis et al., 2014; Kazlauskiene et al., 2016). Šie duomenys leidžia teigti, kad RNR karpymo produktai yra užlaikomi nebent substrato molekulė komplekse juos pakeičia.

Tiriant abiejų galinių RNR produktų užlaikymą autokoreliacijos režimu nustatyta, kad šalia Cas10 esantis galas yra užlaikomas trumpiau negu nuo Cas10 nutolęs galas ($\langle t_{artimas} \rangle = 84 \pm 2$ min ir $\langle t_{nutolęs} \rangle = 119 \pm 3$ min) (2 pav.), o vidiniai produktai paleidžiami dar greičiau (3 pav.). Rezultatai neprieštarauja minčiai, jog galinių produktų užlaikymas tarpusavyje gali būti

koordinuotas – sprendžiant pagal užlaikymo trukmes pirma disocijuoja šalia Cas10 esantis produktas, o tuomet paleidžiamas kitoje komplekso pusėje surištas produktas. Norint detaliau patikrinti galimą kooperaciją reikėtų papildomų tyrimų.

Dėl 5′-3′ transkripcijos krypties šalia Cas10 esančioje pusėje surišama susintetintos RNR 3′-galinė dalis nukreipiant Cas10 į RNR polimerazės pusę. Šioje komplekso dalyje po karpymo surištas transkriptas išlaiko kompleksą šalia transkribuojamos DNR, kas galėtų padidinti RNR polimerazės išstumtos DNR grandinės karpymo efektyvumą (Tamulaitis et al., 2017; van Beljouw et al., 2022). Vis dėlto šioje komplekso dalyje tarp III-A tipo Csm ir III-B tipo Cmr, atitinkamai, Csm4-Cas10 ir Cmr3-Cas10 subvienetų porų, yra struktūrinių ir paviršiaus elektrostatikos skirtumų (Li et al., 2024), tad lieka nežinoma ar Cmr kompleksai yra užlaiko 3′-galo karpymo produktus.

Cas10 DNazės tyrimai parodė, kad po aktyvacijos StCsm išlieka aktyvintoje būsenoje apie valandą, $\langle t_{DNazė} \rangle = 68 \pm 1$ min. Ši vertė panaši į RNR užlaikymo trukmę, kas mums leidžia teigti, kad užlaikyti RNR produktai stimuliuoja Cas10 aktyvumą III-A tipo CRISPR-Cas kompleksuose (5 pav.). Šį mechanizmą taip pat paremia pastebėjimas, kad StCsm RNazės mutantas negali grįžti į neaktyvią būseną, nes RNR nėra kerpama ir paleidžiama (4B pav.). Nors abu galiniai RNR produktai yra ilgam surišami po RNR karpymo, manome, kad šalia Cas10 esantis, o ne nuo jo nutolęs produktas galėtų reguliuoti Cas10 subvienetą. Remiantis ankstesnių tyrimų rezultatais, būtent ši substrato dalis yra svarbi savas-svetimas atpažinimui ir, priklausomai nuo sekos, slopina arba aktyvina Cas10 (Elmore et al., 2016; Kazlauskiene et al., 2016; Li et al., 2024; T. Y. Liu et al., 2017; Wang et al., 2018).

Sukarpytos DNR kiekis priklauso nuo RNR koncentracijos esant RNR substrato nepritekliui (4D pav.), kas reiškia, kad tokiomis sąlygomis aktyvinto komplekso koncentracija yra apribota RNR substrato kiekio reakcijos mišinyje. Esant substrato pertekliui stebimas įsisotinimo efektas dėl greitos substrato degradacijos parodytos ankstesnių tyrimų metu (Kazlauskiene et al., 2016). Degradavus substrato perteklių užlaikyti karpymo produktai gali išlaikyti kompleksą aktyvintoje būsenoje kaip ir RNR nepritekliaus atveju (5 pav., apačioje dešinėje). Ši savybė leidžia užtikrinti apsaugą, jeigu užpuoliko genų raiška yra itin didelė. Pastebima, kad tokių genų fragmentai dažnai yra naujų skirtukų šaltinis III tipo CRISPR regionuose (Aviram et al., 2022; González-Delgado et al., 2019; Silas et al., 2016).

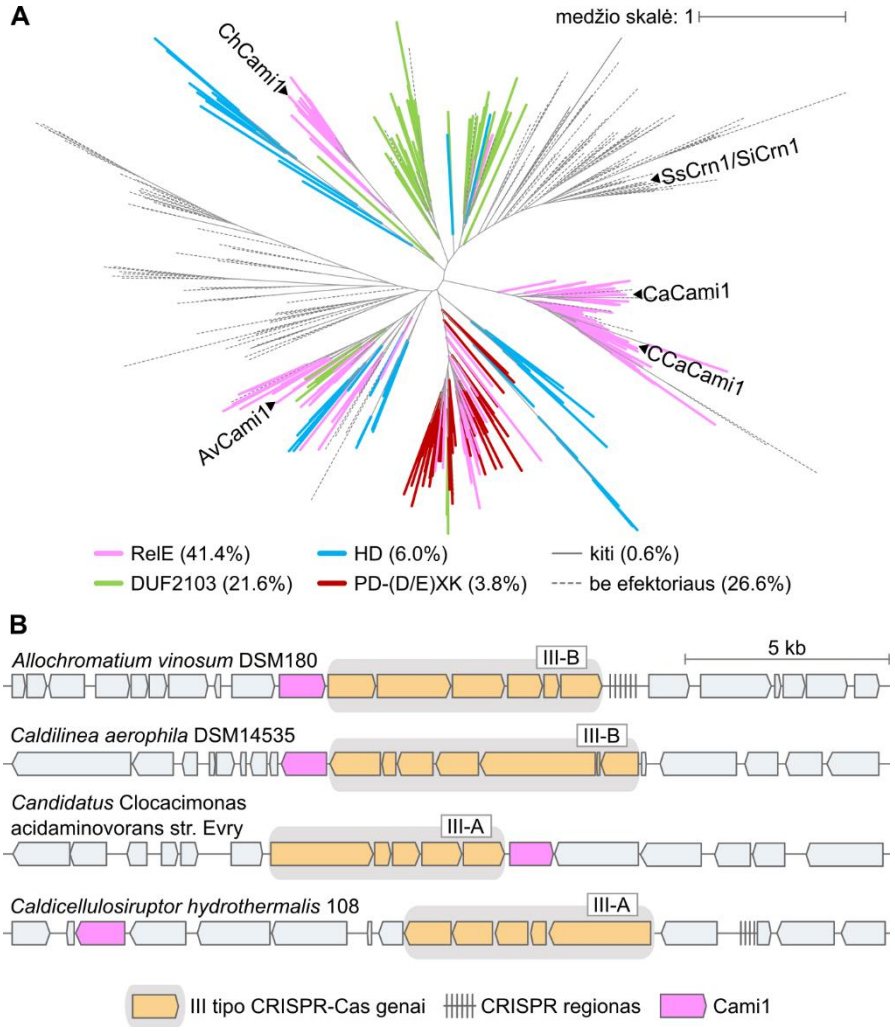
Pirmuosiuose darbuose, parodžiusiuose StCsm Cas10 HD domeno DNazinį (Kazlauskiene et al., 2016) ir Palm domeno polimerazinį-ciklazinį (Kazlauskiene et al., 2017) aktyvumą, pateikti biocheminiai eksperimentai

atskleidė, kad abu aktyvumai nuslopsta per kelias dešimtis minučių po aktyvinimo prisijungus RNR. Remiantis ir šioje disertacijoje pateiktais rezultatais, galima spėti, kad komplekse užlaikyti RNR karpymo produktai stimuliuoja ne tik DNazinį, bet polimerazinį-ciklazinį aktyvumą (5 pav.). Infekcijos metu tai reikštų, kad cOA yra sintetinami ilgesnį laiko tarpą, kas galėtų lemti labiau išreikštą pagalbinių CARF baltymų gynybinį atsaką. Galimą funkcinį ryšį tarp RNR produktų užlaikymo ir cOA sintezės trukmės lieka nustatyti ateityje.

❖ CARF mRNR interferazių charakterizavimas

Pagalbiniai baltymai sustiprinantys III tipo CRISPR-Cas sistemų atsaką dažniausiai yra sudaryti iš cOA surišančio CARF domeno ir efektorinio domeno (Makarova et al., 2020a). Šių baltymų genai yra išsidėstę šalia Csm ar Cmr kompleksą koduojančių genų (Makarova et al., 2020b). Efektoriniai domenai pasižymi didele įvairove, aktyvinti jie gali stabdyti infekuotos ląstelės augimą arba sukelti jos žūtį taip slopinant fago paplitimą bakterijų populiacijoje (Stella ir Marraffini, 2023).

Iš dešimties pagrindinių CARF baltymų grupių ėmėmės analizuoti CARF7 baltymus, kurie pasižymi žiedo nukleazės aktyvumu ir gali perkirpti surištą cA_4 aktyvatorių (Makarova et al., 2020a). Pasitelkdami kompiuterinius sekų paieškos metodus mes papildėme CARF7 grupę naujomis sekomis, paruošėme 1521 sekos dydžio CARF7 baltymų rinkinį bei nustatėme šių baltymų tretinių struktūrų modelius. Tarpusavyje lygindami CARF7 baltymų sekas arba struktūrinius modelius nustatėme šiuos baltymus sudarančius domenus ir jų ribas. Atlikta analizė atskleidė, kad 90% CARF domenų CARF7 baltymuose yra susilieję su wHTH domenais. Tokia CARF7-wHTH dviejų domenų architektūra būdinga gerai charakterizuotoms žiedo nukleazėms SsCrn1 (Athukoralage et al., 2018) ir SiCrn1 (Molina et al., 2021). 80% iš šių baltymų CARF7-wHTH domenų šerdis yra susiliejusi su papildomais efektoriniais



6 pav. CARF7 ir Cami1 sekų kompiuterinė analizė. (**A**) CARF7 baltymų grupės sekų filogenetinis medis sudarytas panaudojant iki 40% tarpusavio panašumo lygio nufiltruotas CARF7 domenų sekas (n = 385). Medžio atšakos nuspalvintos pagal prie CARF7 domeno arba CARF7-wHTH domenų šerdies prilietus efektorinius domenus. Tyrimams pasirinkti Cami1 baltymai ir anksčiau charakterizuoti SsCrn1 ir SiCrn1 pažymėti rodyklėmis. Legendoje skliausteliuose pateikti skaičiai atitinka nurodytus efektorinius domenus turinčių sekų dalį pilname CARF7 duomenų rinkinyje. Medžio skalė atitinka vieną aminorūgšties pakeitimą per poziciją. Duomenys šiam medžiui sudaryti yra publikuoti (Mogila et al., 2023). (**B**) Keturių bakterijų genomo regionai koduojantys III tipo CRISPR-Cas komplekso subvienetus (oranžinė) ir susijusius Cami1 baltymus.

domenais. Pagrindiniai yra RelE (mRNR interferazės), HD (fosfohidrolazės), PD-(D/E)XK (endonukleazės) arba nežinomos funkcijos DUF2103 (6A pav.).

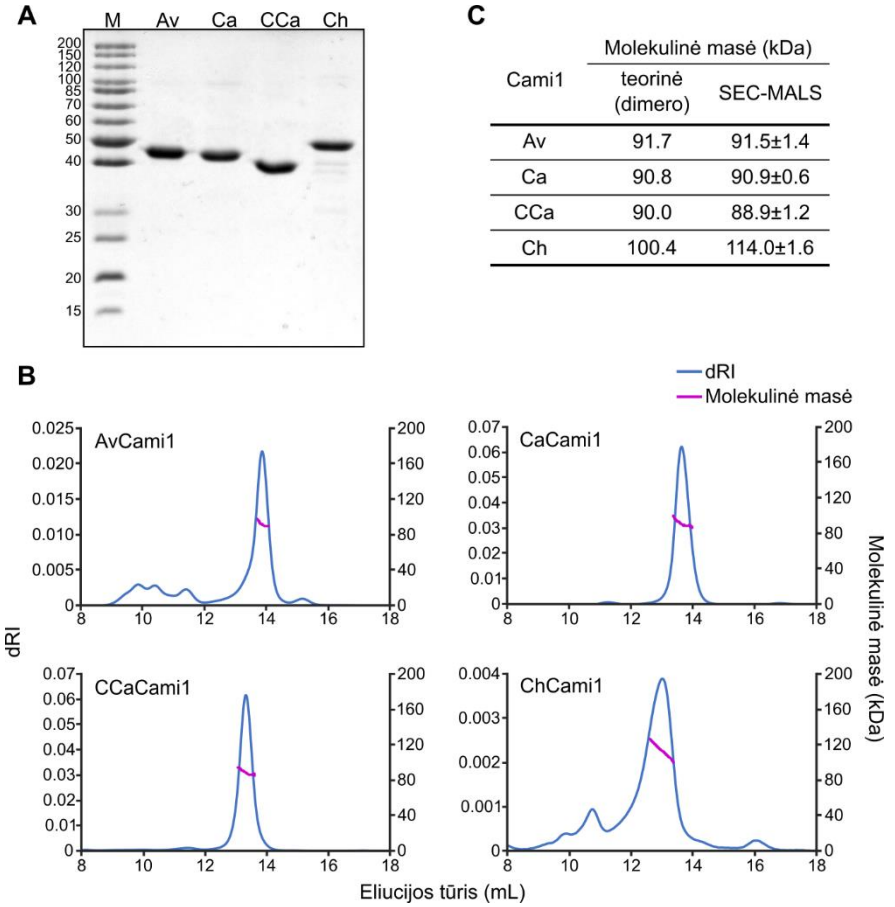
Filogenetinė CARF7 domenų analizė parodė, kad tos pačios šeimos efektoriniai domenai (pvz., RelE) gali būti aptinkami skirtingose filogenetinio medžio šakose ir atvirkščiai – tarpusavyje artimi CARF7 baltymai gali būti susilieję su skirtingais efektoriniais domenais. Tai leidžia teigti, kad CARF7-wHTH ir efektorinių domenų susiliejimai evoliucijos eigoje galėjo įvykti nepriklausomai daugiau nei kartą. Didžiąją dalį CARF7 grupės baltymų kompiuterinės analizės atliko dr. Albertas Timinskas.

Dažniausiai pasitaikanti domeninė architektūra CARF7 baltymuose yra CARF7-wHTH-RelE. Šių sekų C-galinis efektorinis domenas yra panašus į RelE superšeimos toksinus iš II tipo toksino-antitoksino sistemų (Jurėnas ir Van Melderen, 2020). II tipo RelE toksinai yra nuo ribosomų priklausomos endonukleazės, kurios dar vadinamoms mRNR interferazėmis. Jos prisijungia ribosomų A srityje ir karpo tuo metu transliuojamą mRNR (Christensen ir Gerdes, 2003).

CARF7-wHTH-RelE genai yra aptinkami šalia Csm/Cmr koduojančių genų, kas leidžia įtarti, jog šie baltymai yra funkciškai susiję su III tipo CRISPR-Cas apsauga. Remdamiesi domenine sudėtimi, šiuos baltymus pavadinome Cami1 (nuo angl. *CRISPR-Cas-associated mRNA interferase 1*). Tolimesniems eksperimentiniams tyrimams pasirinkome keturis Cami1 homologus iš *Allochromatium vinosum* (AvCami1), *Caldilinea aerophila* (CaCami1), *Candidatus* Cloacimonas acidaminovorans (CCaCami1) ir *Caldicellulosiruptor hydrothermalis* (ChCami1), kurie visi šeimininko genomuose yra koduojami netoli komplekso subvienetų genų (6B pav.) ir priklauso skirtingoms CARF7 filogenetinio medžio pagrindinėms šakoms (6A pav.).

- Cami1 baltymų gryninimas

Sintetiniai Cami1 baltymų genai su prilietomis His_{10} ir StrepII gryninimo žymėmis buvo perkelti į pBAD vektorius. Jų raiška atlikta *E. coli*, baltymai išgryninti giminingumo chromatografijos metodu (7A pav.). Molekulinės masės matavimai atlikti SEC-MALS metodu parodė, kad Cami1 formuoja dimerus tirpale (7B ir C pav.) kaip SsCrn1 (Athukoralage et al., 2018) ir SiCrn1 (Molina et al., 2021) CARF7 žiedo nukleazės bei kiti CARF baltymai (Stella ir Marraffini, 2023).



C

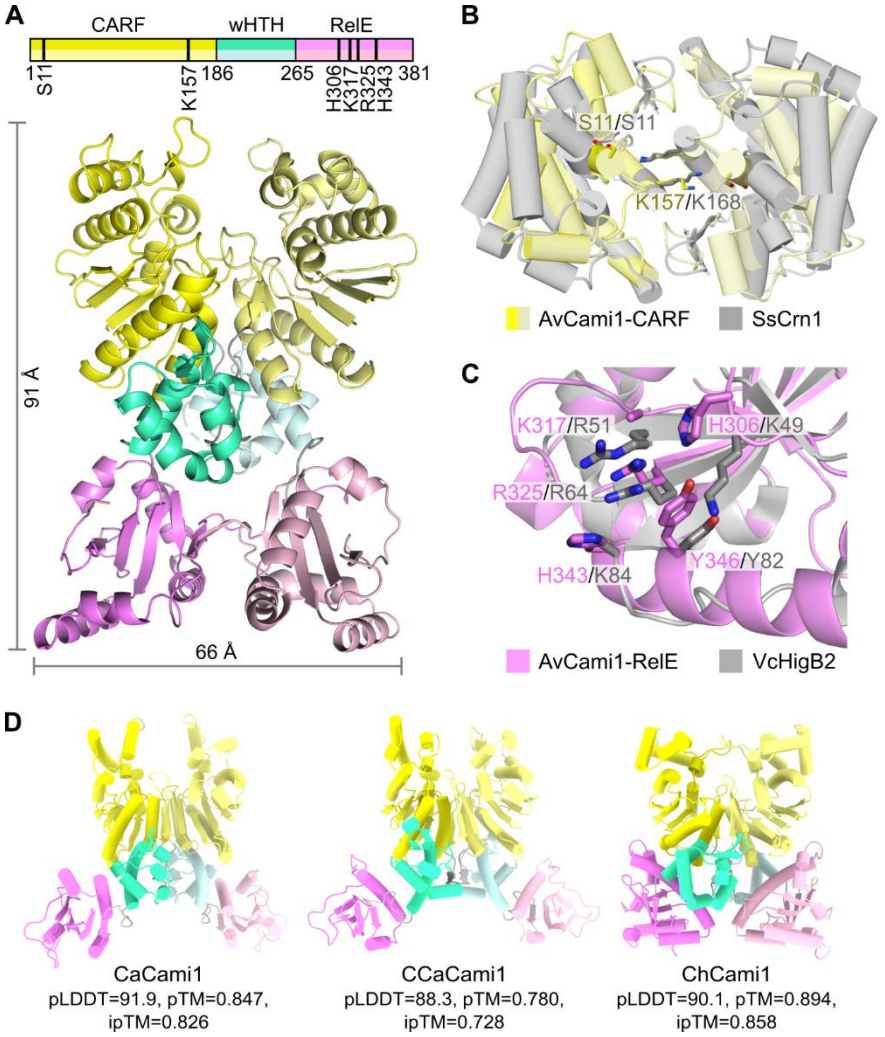
Cami1	Molekulinė masė (kDa)	
	teorinė (dimero)	SEC-MALS
Av	91.7	91.5±1.4
Ca	90.8	90.9±0.6
CCa	90.0	88.9±1.2
Ch	100.4	114.0±1.6

7 pav. Cami1 baltymų gryninimo ir stechiometrinės analizės rezultatai. (**A**) Išgryninti Cami1 baltymai analizuoti gelyje denatūruojančiomis sąlygomis. „M“ – baltymų masės markeris, pateiktos vertės atitinka molekulinę masę kDa. (**B**) Cami1 baltymų SEC-MALS analizės duomenys. Mėlynos linijos – dRI atsakas, proporcingas baltymo koncentracijai; rožinės linijos – molekulinė masė apskaičiuota pagal šviesos sklaidos (MALS) duomenis. (**C**) Molekulinės masės matavimo tirpale rezultatai naudojant SEC-MALS. Transliacijos produkto, kuris atitinka Cami1 genų atviro skaitymo rėmelį, molekulinė masė buvo panaudota apskaičiuojant teorinę dimerų molekulinę masę.

- Kristalinė AvCami1 struktūra

Mes nustatėme AvCami1 baltymo kristalinę struktūrą 1.7 Å skiriamąja geba (8A pav., PDB: 8PHB). AvCami1 kristalus struktūriniams tyrimams paruošė Konstanty Keda, baltymo struktūrą nustatė dr. Giedrė Tamulaitienė.

Struktūroje stebimi prieš tai nuspėti N-galinis CARF7 (1-186 aminorūgš-



8 pav. Cami1 baltymų struktūrinė analizė. (**A**) AvCami1 kristalinė struktūra. CARF, wHTH ir RelE domenai pažymėti, atitinkamai, geltona, melsvai žalia ir rožine spalvomis. Kristalografinių duomenų surinkimo ir apdorojimo statistiniai rodikliai pateikti Appendix 5. (**B**) CARF domenų iš AvCami1 ir SsCrn1 (PDB: 3QYF, RMSD 2.2 Å) struktūrų palyginimas. Pažymėtos SsCrn1 žiedo nukleazės aktyvumui svarbios aminorūgštys ir jas atitinkančios AvCami1 aminorūgštys. (**C**) AvCami1 RelE domeno ir VcHigB2 toksino (PDB: 5JA9, RMSD 2.6 Å) struktūrų palyginimas. Pavaizduotos eksperimentiškai nustatytų mRNR interferazės aktyviojo centro ir jas atitinkančių AvCami1 aminorūgščių šoninės grandinės. (**D**) Pilno ilgio Ca, CCa ir ChCami1 AlphaFold 2 struktūriniai modeliai su modeliavimo patikimumo įverčiais. CARF, wHTH ir RelE domenai nuspalvinti kaip (A).

tys), vidurinis wHTH (187-265 aminorūgštys) ir C-galinis RelE (266-381 aminorūgštys) domenai. Iš kitų eksperimentiniais metodais nustatytų baltymų

struktūrų CARF7 domenas panašiausias į SsCrn1 žiedo nukleazę (8B pav.) (Athukoralage et al., 2018). AvCami1 S11 ir K157 aminorūgštys užima panašias pozicijas erdvėje kaip ir SsCrn1 aktyviojo centro konservatyvios aminorūgštys S11 ir K168, kas leidžia teigti, kad AvCami1 irgi galėtų pasižymėti žiedo nukleazės aktyvumu.

AvCami1 RelE domeno struktūra yra panašiausia į RelE superšeimos VcHigB2 baltymą (PDB: 5JA9) iš *Vibrio cholerae*, kuriam būdingas į *E. coli* RelE panašus aktyvusis centras (Hadži et al., 2017). Perklojus struktūras matoma, kad kai kurios teigiamai įkrautos AvCami1 RelE domeno aminorūgštys erdvėje atitinka eksperimentiškai nustatytas VcHigB2 aktyviojo centro aminorūgštis (8C pav.). Nors didžioji dauguma RelE superšeimos toksinų yra monomeriniai baltymai (Jurėnas ir Van Melderen, 2020), AvCami1 RelE domenai sąveikauja tarpusavyje formuodami dimerą (8A pav.). CaCami1 ir CCaCami1 AlphaFold 2 struktūriniuose modeliuose sąveikos tarp RelE domenų nėra, ChCami1 sąveikos paviršius mažesnis negu AvCami1 kristalinėje struktūroje (8D pav.). Struktūriniai duomenys rodo, kad Cami1 galėtų veikti kaip cOA aktyvinamos ribonukleazės.

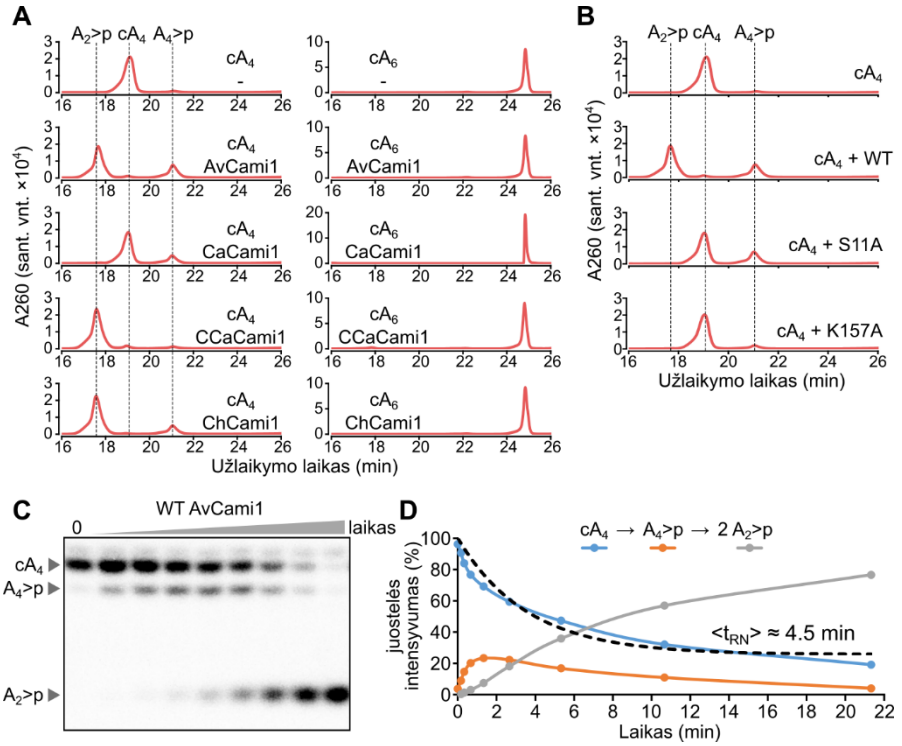
- Žiedo nukleazės aktyvumas

CARF baltymus dažniausiai aktyvina cA_4 arba cA_6, o kai kurie iš šių baltymų pasižymi žiedo nukleazės aktyvumu ir sunaikina savo aktyvatorių, kas apriboja jų imuninį atsaką (Athukoralage et al., 2019; Athukoralage ir White, 2021; Garcia-Doval et al., 2020; Smalakyte et al., 2020). Siekdami patikrinti ar Cami1 pasižymi žiedo nukleazės aktyvumu, paruošėme Cami1 baltymų ir cA_4 arba cA_6 mišinius, kuriuos analizavome masių spektrometrijos (MS) metodu. MS matavimus atliko Audronė Rukšėnaitė.

Eksperimentai atskleidė, kad Cami1 karpo cA_4, bet ne cA_6 (9A pav., Appendix 6). Reakcijos metu Cami1 baltymai cA_4 verčia į A_4>p ir A_2>p, panašiai kaip ir kitos žiedo nukleazės (Athukoralage et al., 2018, 2019; Molina et al., 2019, 2021). CaCami1 pasižymėjo žemesniu aktyvumu, reakcijos susidarė tik nedidelis kiekis A_4>p, bet ne A_2>p. Šie rezultatai patvirtino, kad Cami1 yra žiedo nukleazės. Jų karpomas cA_4 yra galimas aktyvatorius.

Mes paruošėme du AvCami1 CARF domeno mutantus S11A ir K157A koduojančias plazmides, mutantinius baltymus išgryninome ir patikrinome jų žiedo nukleazės aktyvumą. Šie konservatyvių aminorūgščių mutantai pasižymėjo žemesniu žiedo nukleazės aktyvumu (9B pav.) kaip ir atitinkami SsCrn1

(Athukoralage et al., 2018) ir SiCrn1 (Molina et al., 2021) baltymų mutantai, kas patvirtina specifinį AvCami1 žiedo nukleazės aktyvumą.



9 pav. Cami1 žiedo nukleazės aktyvumo tyrimai. (**A**) Cami1 žiedo nukleazės reakcijos mišinių MS analizė naudojant cA_4 ir cA_6 substratus. Reakcijos produktai, kurių m/z vertės atitinka cA_4, A_4>p arba A_2>p yra pažymėti punktyrinėmis linijomis. (**B**) AvCami1 CARF domenų mutantų žiedo nukleazės aktyvumo palyginimas su WT baltymu naudojant MS. Reakcijos produktai pažymėti kaip (A). (**C**) α-^{32}P pažymėto cA_4 karpymo AvCami1 pertekliumi reakcijos produktai. Pateikta reprezentatyvi radiogrma iš trijų eksperimentų. Analizuoti mėginiai, kurie paimti praėjus 0, 10, 20, 40, 80, 160, 320, 640 ir 1280 s po reakcijos pradžios. (**D**) Radiogramos, kuri pavaizduota (C), densitometrinės analizės rezultatai. cA_4 substrato kiekio nykimo duomenims pritaikytas [10] modelis pavaizduotas punktyrine linija.

Mes taip pat nustatėme santykinius cA_4 karpymo reakcijos metu susidarančių produktų kiekius bėgant laikui (9C pav.). Tam naudojome radioaktyviai pažymėtą substratą, reakcijas atlikome esant AvCami1 baltymo pertekliui. Densitometrinė radiogramų analizė atskleidė, kad reakcija vyksta pakopomis: pirmiausiai susidaro tarpinis A_4>p junginys, kuris yra perkerpamas į dvi galutinio A_2>p produkto molekules. cA_4 karpymo duomenims pritaikėme eksponentinio nykimo modelį:

$$S(t) = (100\% - S_\infty)e^{-t/\langle t_{RN} \rangle} + S_\infty, \quad [10]$$

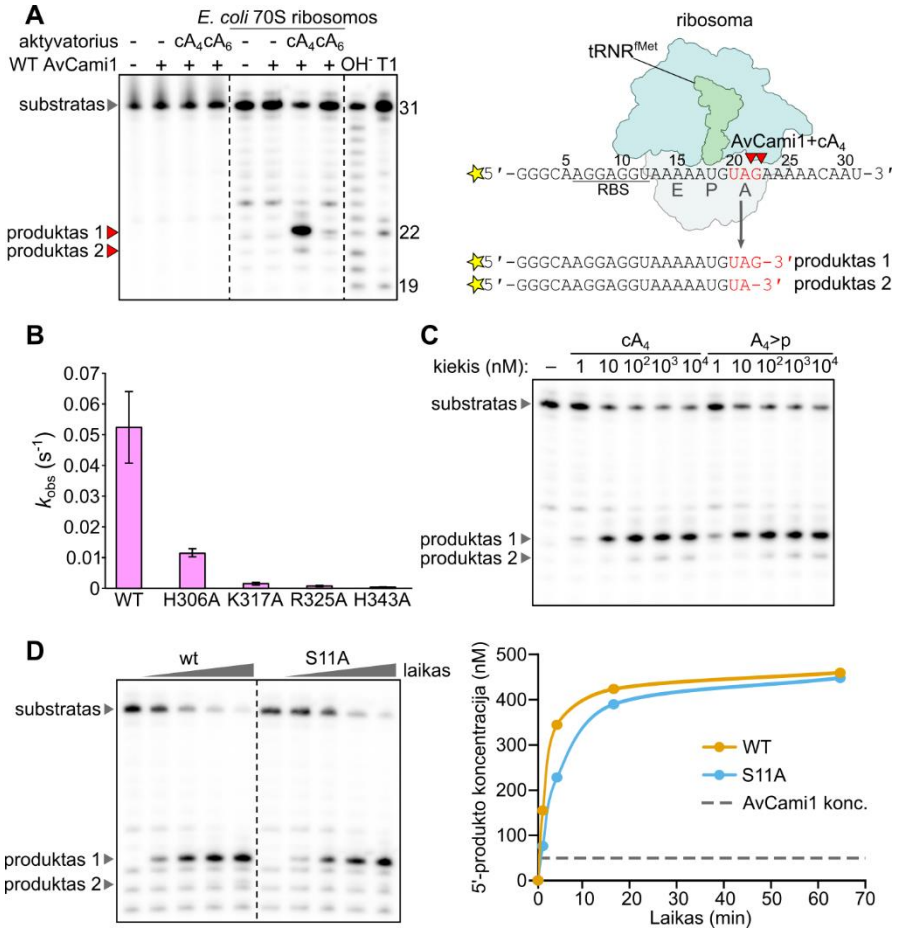
kur $S(t)$ yra pažymėto cA_4 dalis (%) reakcijos mišinyje laiku t, $\langle t_{RN} \rangle$ yra vidutinė žiedo nukleazės reakcijos trukmė ir S_∞ yra nekerpamo substrato kiekis. Pritaikius modelį nustatėme $\langle t_{RN} \rangle$ = 4.4 ± 0.7 min.

Žiedo nukleazės aktyvumo tyrimai parodė, kad AvCami1 baltymas suriša ir karpo cA_4 kelių minučių trukmės reakcijos metu.

- mRNR interferazės aktyvumas

Sekų (6A pav.) ir struktūros (8C pav.) analizės duomenys rodo, kad C-galinis Cami1 baltymų domenas galėtų veikti kaip ribonukleazė, kuri karpo ribosomoje surištą mRNR, panašiai kaip pavieniai RelE toksinai. Siekiant tai patikrinti, mes atlikome ribosomoje surišto radioaktyvaus RNR substrato karpymo reakcijas su AvCami1, kaip tai anksčiau buvo atlikta tiriant *E. coli* RelE toksino aktyvumą (Griffin et al., 2013). Kadangi EcRelE aktyviai kerpa mRNR UAG kodono vietoje *in vitro*, mes paruošėme atitinkamą rib-UAG substratą (Appendix 2), kurį sumaišius kartu su *E. coli* 70S ribosomomis ir $tRNR^{fMet}$ jo AUG kodonas yra užrakinamas P srityje, o UAG kodonas išsidėsto A srityje (10A pav., dešinėje). Nustatėme, kad nesant aktyvatorių AvCami1 nekerpa laisvos arba ribosomų surištos RNR, tačiau cA_4 aktyvintas AvCami1 kerpa ribosomoje surištą substratą (10A pav., kairėje). Pagal šarminės substrato degradacijos ir RNazės T1 karpymo produktus nustatėme, kad AvCami1 aktyviausiai kerpa rib-UAG substratą už trečiojo UAG kodono nukleotido („produktas 1“ 10A pav.). Reakcijos metu taip pat susidaro nedaug „produkto 2“ po substrato kirpimo už antrojo UAG nukleotido arba „produkto 1“ pritrumpinimo. Panašiu atpalaiduotu specifiškumu pasižymi ir EcRelE *in vivo* (Goeders et al., 2013).

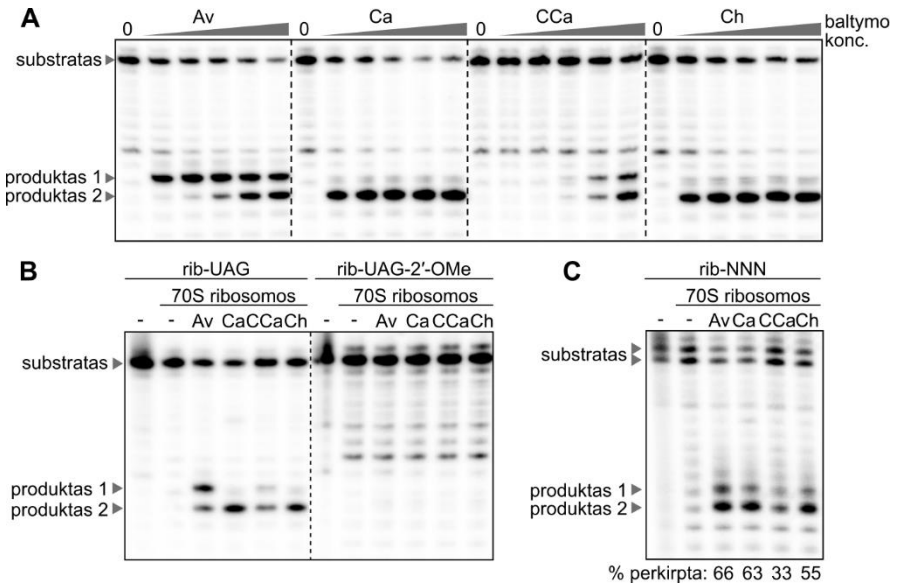
Siekiant patikrinti ar RelE domeno aktyviojo centro aminorūgštys atitinka VcHigB2 aktyvųjį centrą (8C pav.), buvo paruošti šie AvCami1 mutantai: H306A, K317A, R325A ir H343A. Atlikę mRNR interferazės reakcijas ir bei densitometrinę reakcijos produktų analizę palyginome mutantų ir WT AvCami1 karpymo greičio konstantas duomenims pritaikę eksponentinio nykimo modelį. Mutacijos sumažino aktyvumą nuo 5 iki 150 kartų, didžiausiu poveikiu pasižymėjo H343A mutacija (10B pav.). Mutantinius baltymus išgrynino ir jų aktyvumą nustatė Konstanty Keda.



10 pav. cA_4 aktyvintas AvCami1 baltymas kerpa ribosomoje surištą ir A srityje prieinamą RNR. (**A**) AvCami1 RNR karpymo aktyvumas esant cOA ir ribosomų. Reakcijose buvo 20 nM (nurodytos galutinės koncentracijos) 5′-gale žymėto rib-UAG substrato ir 300 nM AvCami1 baltymo. Reakcijose su ribosomomis buvo suformuoti jų kompleksai su RNR substratu naudojant 60 nM *E. coli* 70S ribosomų ir 80 nM $tRNR^{fMet}$. AvCami1 aktyvinti naudojant 10 μM cA_4 arba cA_6. Reakcijos schema pavaizduota dešinėje. 31 nt rib-UAG substratas susideda iš ribosomos prisijungimo srities (RBS), P ir A srityse atitinkamai atsiduriančių AUG ir UAG kodonų. OH^- ir T1 – atitinkamai, šarminės hidrolizės ir RNazės T1 karpymo standartai. (**B**) AvCami1 WT ir RelE domeno mutantų mRNR interferazės aktyvumas 15 °C temperatūroje. k_{obs} – stebima reakcijos greičio konstanta, nustatyta rib-UAG substrato karpymo duomenims pritaikius eksponentinio nykimo modelį. Kiekvienu atveju pateiktas vidurkis ir standartinis nuokrypis atlikus tris techninius pakartojimus. (**C**) AvCami1 mRNR interferazės aktyvinimas cA_4 arba A_4>p. Reakcijos paruoštos kaip (A) naudojant įvairius aktyvatoriaus kiekius. (**D**) AvCami1 aktyvumas esant substrato pertekliui. Reakcijos paruoštos naudojant 500 nM žymėtos rib-UAG RNR, 1 μM *E. coli* 70S ribosomų, 1.3 μM $tRNR^{fMet}$, 10 μM cA_4 ir 50 nM WT AvCami1 arba S11A mutanto.

Pavyzdžiai iš reakcijos mišinio paimti nurodytais laiko intervalais ir analizuoti kaip (A).

CARF domenas cA_4 aktyvatorių karpo susidarant tarpiniam A_4>p junginiui. Nustatėme, kad abu junginiai aktyvina AvCami1 (10C pav.). Tirdami aktyvumą esant substrato pertekliui nustatėme, kad vienas aktyvinto AvCami1 dimeras perkerpa RNR keliuose kompleksuose su ribosomomis (10D pav.). RNR karpymas nesusijęs su cA_4 karpymo ciklu, nes tiek WT, tiek S11A žiedo nukleazės mutantas substrato perteklių karpė panašiu efektyvumu esant cA_4 pertekliui (10D pav., dešinėje). Remiantis šiais rezultatais galima teigti, kad Cami1 baltymai nėra pakartotinai aktyvinami naujos cA_4 molekulės po kiekvienos mRNR kirpimo reakcijos.



11 pav. Cami1 homologų mRNR interferazės aktyvumas. (**A**) Cami1 baltymų ribosomose surištų RNR karpymo reakcijos. Kiekvienas baltymas tikrintas esant 0, 30, 100, 300, 1000 arba 3000 nM jo koncentracijai naudojant reakcijos sudėtį kaip 10A pav. aprašyme. Baltymai aktyvinti 10 μM cA_4. (**B**) Cami1 baltymų aktyvumas tikrintas naudojant rib-UAG-OMe substratą, kurio A srities nukleotidai yra 2′-OMe modifikuoti. 1.5 μM Cami1 baltymų buvo aktyvinti 10 μM, reakcijos paruoštos ir analizuotos kaip 10A pav. aprašyme. (**C**) Cami1 aktyvumas tikrintas naudojant rib-NNN substratų mišinį susidedantį iš RNR turinčių įvairius A srities kodonus. Reakcijos paruoštos kaip (B). Densitometriškai nustatytas perkirpto substrato kiekis nurodytas po radiograma.

CaCami1, CCaCami1 ir ChCami1 taip pat karpo ribosomoje surištą RNR priklausomai nuo cA_4 (11A pav.). CCaCami1 pasižymėjo žemesniu aktyvumu, kuris stebimas reakcijose naudojant aukštesnes baltymo koncentracijas. AvCami1 ir CCaCami1 kerpa RNR už antrojo ir trečiojo A srityje esančio kodono nukleotidų, CaCami1 ir ChCami1 kerpa už antrojo nukleotido. Šie rezultatai rodo, kad skirtinguose Cami1 baltymų RelE domeno aktyvusis centras užima įvairias pozicijas ribosomos A srityje.

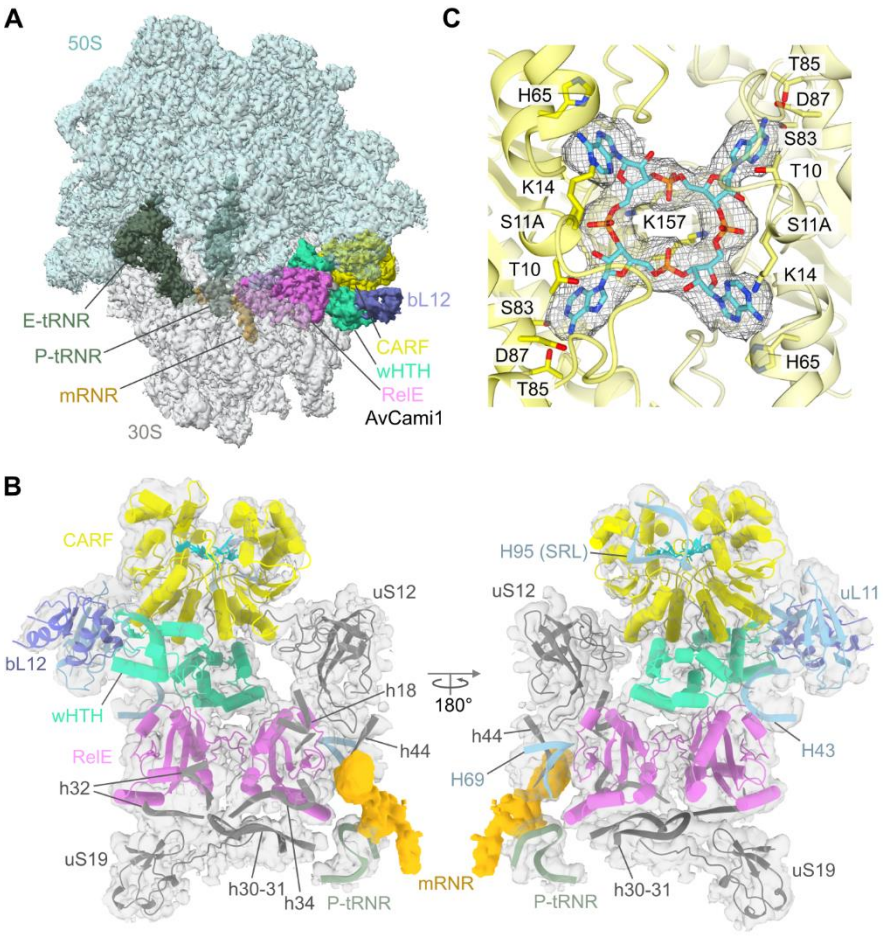
Pagal pasiūlytą RelE toksinų mRNR karpymo mechanizmą, baltymo aktyvusis centras pakeičia substrato konformaciją taip, kad 2′-OH atlieka nukleofilinę ataką į šalia esantį fosfatą (Jurėnas ir Van Melderen, 2020; Neubauer et al., 2009). Naudojant rib-UAG-OMe substratą, kurio trys A srityje atsiduriantys nukleotidai yra 2′-OMe modifikuoti (Appendix 2), pastebėjome, kad nei vienas iš tirtų Cami1 baltymų tokio substrato nekirpo (11B pav.). Cami1 baltymai karpė rib-NNN substratų mišinį, sudaryta iš įvairius A srities kodonus turinčių RNR (11C pav.), toks atpalaiduotas specifiškumas RNR sekai būdingas ir *E. coli* RelE toksinui (Goeders et al., 2013).

Kartu šie rezultatai patvirtina kompiuterinės analizės rezultatus ir parodo, kad Cami1 baltymai yra nuo ribosomų priklausomos mRNR interferazės, kurias aktyvina cA_4.

- Prie ribosomos prisijungusio AvCami1 struktūra

Siekdami detaliau suprasti Cami1 aktyvumo mechanizmą nustatėme cA_4 aktyvinto dvigubo AvCami1-S11A-H343A mutanto, prisijungusio prie *E. coli* ribosomos surišusios rib-UAG-OMe ir $tRNR^{fMet}$, krio-EM struktūrą (12A pav., PDB: 8PHJ, struktūros vidutinė skiriamoji geba 3.7 Å). Atominę ribosomos-AvCami1 komplekso struktūrą rekonstravo dr. Giedrė Tamulaitienė.

Ribosomos komplekse stebimos dvi $tRNR^{fMet}$ molekulės: viena yra prisijungusi P srityje ir sąveikauja su surišta RNR, kita išsidėsčiusi E srityje. AvCami1 baltymo dimeras yra prisijungęs A srityje taip, kad vieno AvCami1 subvieneto RelE domenas sąveikauja su 16S ir 23S rRNR, bei P srities tRNR (12B pav.). Panašios sąveikos būdingos ir kitiems RelE šeimos toksinams rišantis su ribosoma (Mansour et al., 2022; Neubauer et al., 2009). Nutolusio AvCami1 subvieneto RelE domenas sąveikauja su 16S rRNR. CARF domenų dimeras išsidėsto šalia sarcino-ricino kilpos (SRL) H95 iš 23S rRNR. AvCami1 taip pat sąveikauja su šerdiniais ribosomos baltymais uS12, uS19, uL11 ir su ribosomos stiebo baltymu bL12 (12B pav.). A srityje surištos RNR krio-EM tankis yra išsklidęs, kas galėtų reikšti, kad substratas RelE domeno prisijungimo vietoje yra judrus (12B pav.)



12 pav. cA_4 aktyvinto AvCami1 prisijungusio ribosomos komplekse su mRNR imituojančiu substratu ir tRNR krio-EM struktūra. (**A**) cA_4 aktyvinto dvigubo AvCami1-S11A+H343A mutanto (geltona, melsvai žalia, rožinė) prisijungusio prie ribosomos (šviesiai mėlyna ir pilka) komplekso su rib-UAG-OMe RNR substratu (oranžinė) ir $tRNR^{fMet}$ (žalia) bendra struktūra. Struktūroje stebima AvCami1 sąveika su ribosomos stiebo baltymo bL12 C-galiniu domenu (mėlyna). Krio-EM paveikslėlių apdorojimo schema yra pateikta Appendix 7, duomenų surinkimo, apdorojimo ir patikrinimo statistiniai rodikliai pateikti Appendix 8. (**B**) AvCami1 prisijungia A srityje ir sąveikauja su 16S (h18, h30-31, h32, h34, h44 kilpos) ir 23S (H43, H69, H95 kilpos) rRNR ir P srities $tRNR^{fMet}$. AvCami1 dimeras taip pat sąveikauja su ribosomos baltymais uS12 ir uS19 iš mažojo subvieneto (pilka) ir uL11 (šviesiai mėlyna), bei bL12 (mėlyna) iš didžiojo subvieneto. Krio-EM žemėlapis atvaizduotas kaip permatomas pilkas paviršius. Struktūriniai AvCami1 elementai nuspalvinti kaip (A). (**C**) CARF domenų dimere prisijungęs cA_4 (žydra). cA_4 molekulės krio-EM tankis atvaizduotas kaip pilkas tinklelis.

cA_4 yra prisijungęs centrinėje AvCami1 baltymo CARF domenų dimero ertmėje. Aktyvatorius yra plokščios konformacijos, kurią palaiko cA_4 adenino

nukleobazių bei fosfodiesterinio karkaso sąveikos su įkrautomis arba polinėmis CARF domenų aminorūgštimis (12C pav.)

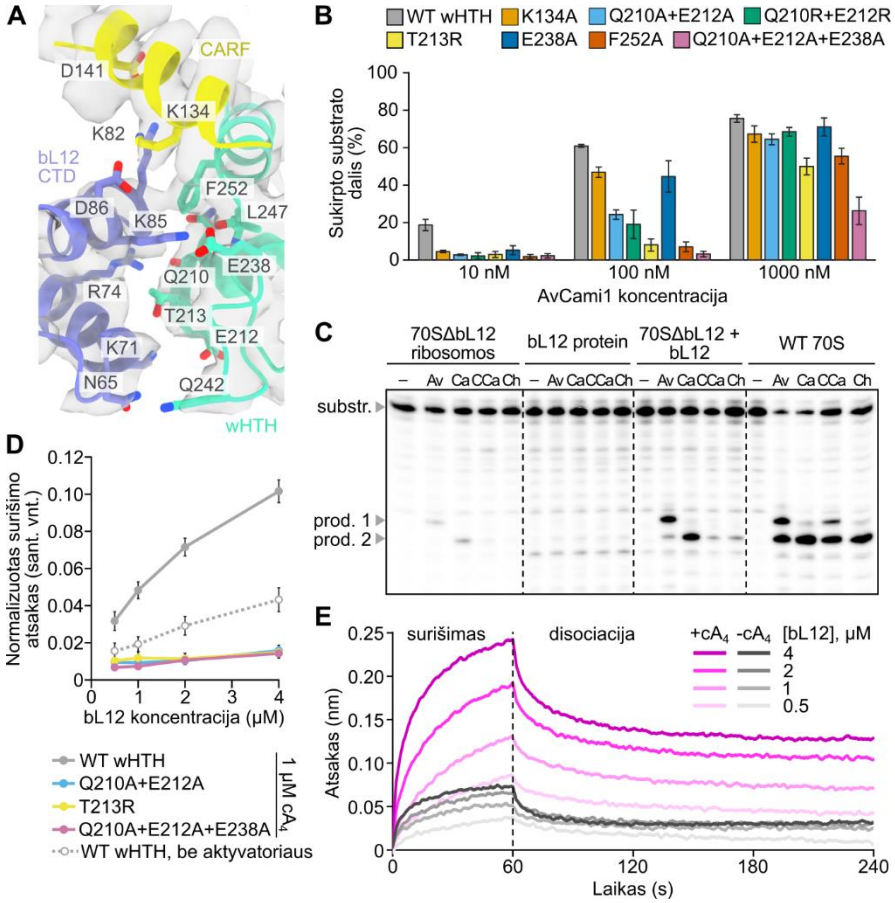
Analizuodami ribosomos-AvCami1 struktūros duomenis pastebėjome, kad į tirpalą nukreiptas AvCami1 wHTH domenas sąveikauja su bL12 ribosomos baltymo C-galiniu domenu (13A pav.). bL12 yra pagrindinis prokariotinių ribosomų stiebo struktūros elementas. Ribosomos stiebo, kurio struktūra yra konservatyvi visuose organizmuose, funkcija yra padidinti vietinę GTPazinių transliacijos faktorių koncentraciją (Imai et al., 2020; Liljas ir Sanyal, 2018). Mūsų pastebėta sąveika tarp AvCami1 ir bL12 leidžia spėti, kad Cami1 baltymai panaudoja sąveiką su ribosomos stiebu, kad patektų į ribosomos A sritį.

Siekdami patikrinti šią idėją sukonstravome CARF ir wHTH domenų paviršiaus aminorūgščių mutantus, kuriuose sąveika su bL12 būtų panaikinta. Tirti wHTH domeno mutantai pasižymėjo žemesniu mRNR interferaziniu aktyvumu (13B pav.). Šis rezultatas patvirtina spėjimą, kad AvCami1 wHTH ir bL12 sąveika svarbi mRNR interferaziniam aktyvumui.

Tęsdami šios sąveikos tyrimus patikrinome Cami1 baltymų aktyvumą reakcijose panaudodami bL12 neturinčias mutantines ribosomas, esant arba nesant atskirai pridėto išgryninto bL12 baltymo. Cami1 baltymų aktyvumas naudojant mutantines ribosomas buvo silpnas, tačiau išaugo į reakcijos mišinį pridėjus bL12 (13C pav.). Šis rezultatas rodo, kad sąveika su stiebo struktūra yra svarbi Cami1 homologų aktyvumui.

Galiausiai sąveiką tarp bL12 ir AvCami1 wHTH domeno patikrinome tiesiogiai BLI metodu. Tuo tikslu AvCami1 baltymas buvo prikabintas ant BLI sensoriaus paviršiaus kaip ligandas. Sensorius su AvCami1 panardintas į tirpalą su bL12 baltymu, kuris veikė kaip analitė. Pastebėjome, kad tirpale esant cA_4, bL12 prisijungimo signalo amplitudė yra aukštesnė negu atliekant bandymą be cA_4. Tai reiškia, kad cA_4 stabilizuoja AvCami1 konformaciją, reikalingą wHTH domeno sąveikai su bL12 (13D ir E pav.). Kaip ir tikėtasi, AvCami1 wHTH paviršiaus mutantai surišo nedidelį bL12 kiekį net ir esant cA_4 (13D pav.). BLI eksperimentai leido patvirtinti, kad po aktyvacijos cA_4 Cami1 yra „pagaunamas“ bL12 baltymo per wHTH domeną.

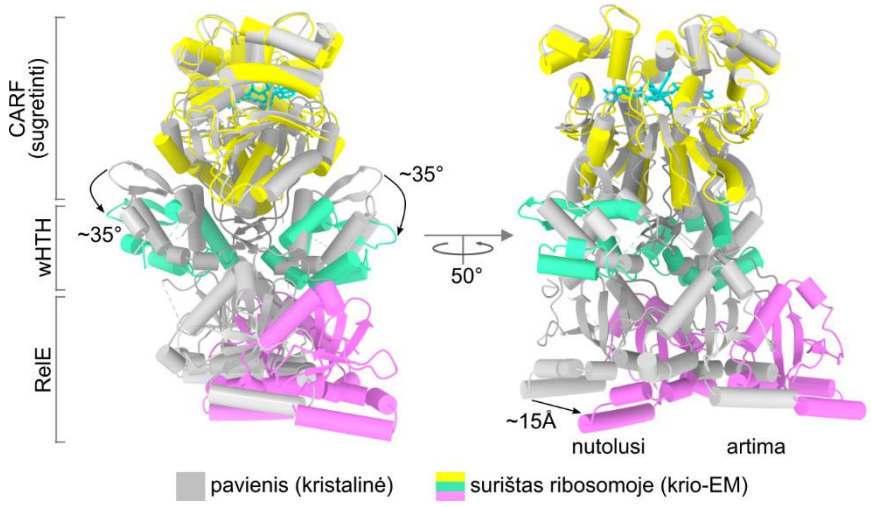
Per CARF domenų dimerą užkloję apo ir cA_4 aktyvinto bei prie ribosomos prisijungusio AvCami1 baltymo struktūras pastebėjome, kad abu wHTH domenai nulinksta tolyn nuo CARF domenų 35° laipsnių kampu (14 pav.). Panašus wHTH ir CARF domenų tarpusavio konformacijų pokytis yra stebimas ir cA_4 žiedo nukleazės SiCrn1 struktūrose (Molina et al., 2021), tačiau šis tyrimas neatskleidė, ar SiCrn1 konformaciniai pokyčiai įvyksta dėl cA_4 suri-



13 pav. Cami1 sąveika su bL12 yra reikalinga mRNR interferaziniam aktyvumui pasireikšti. (**A**) Sąveikos paviršius tarp AvCami1 wHTH domeno ir bL12 baltymo. (**B**) wHTH ir CARF paviršiaus mutacijų įtaka mRNR interferaziniam aktyvumui. Paviršiaus mutacijos buvo įvestos S11A mutacijos, sumažinančios žiedo nukleazės aktyvumą, fone. Reakcijos atliktos kaip 10A pav. aprašyme naudojant didėjančias AvCami1 variantų koncentracijas, perkirpto substrato kiekis įvertintas densitometriškai kaip 11C pav. (**C**) Cami1 homologų aktyvumo ribosomose be bL12 tyrimas. Ribosomos be bL12 (70SΔbL12), pavieniai bL12 baltymai, surinktos ribosomos (70SΔbL12+bL12) ir WT 70S ribosomos buvo sumaišytos su rib-UAG ir $tRNR^{fMet}$ prieš pridedant Cami1 baltymų iki 300 nM (CCaCami1 atveju 3 μM) aktyvinamų 10 μM cA_4. (**D**) bL12 surišimo ant AvCami1 prisijungusio sensoriaus BLI signalai normalizuoti pagal AvCami1 užnešimo efektyvumą. bL12 surišimas cA_4 aktyvintuose AvCami1-S11A ir wHTH paviršiaus mutantuose buvo lyginamas su AvCami1-S11A esant arba nesant cA_4. (**E**) BLI sensorgramos vaizduojančios bL12 prisijungimą prie imobilizuoto AvCami1-S11A ir disociaciją esant (rožinės kreivės) arba nesant (pilkos kreivės) 1 μM cA_4. Kreivės iš eksperimentų su skirtinga bL12 koncentracija yra skirtingo spalvos intensyvumo. Pateiktos reprezentatyvios sensorgramos iš trijų eksperimentų serijų. (B) ir (D) pateikti vidurkiai ir standartiniai nuokrypiai iš trijų techninių pakartojimų.

šimo, ar dėl kirpimo. AvCami1-S11A-H343A ir ribosomos komplekso struktūroje konformacinis wHTH domenų pokytis stebimas surišus cA_4, kuris nėra kerpamas dėl mutacijos žiedo nukleazės aktyviajame centre (9B pav.). cA_4 aktyvintame AvCami1 nulinkus wHTH domenams yra atveriamas sąveikos su bL12 paviršius, tai patvirtina BLI eksperimentuose stebima nuo cA_4 priklausoma bL12 sąveika su AvCami1 (13D ir E pav.).

RelE dimeras taip pat keičia savo poziciją kitų AvCami1 domenų atžvilgiu ir pasislenka per ~15Å įstrižai nuo CARF-wHTH šerdies (14 pav.). Šis judesys panaikina dimero simetriją, dėl ko išsikišęs RelE domenas gali tilpti į ribosomos A sritį ir pasiekti ten esančią mRNR (12A ir B pav.). Cami1 homologų struktūriniuose modeliuose RelE domenai tarpusavyje nesąveikauja (8D pav.), tad po aktyvinimo ir prisijungimo prie ribosomų jie gali įgyti lankstesnes konformacijas. Lieka nežinoma ar cA_4 prisijungimo pakanka panaikinti AvCami1 simetriją ar vis dėlto ši konformacija yra įgaunama jungiantis prie ribosomos.



14 pav. AvCami1 konformacijos pokyčiai surišus cA_4 ir prisijungus prie ribosomos. Pavienio AvCami1 baltymo kristalinė struktūra (pilka) per CARF domenus yra sugretinta su cA_4 surišusio ir prie ribosomos prisijungusio baltymo (spalvota) struktūra. „Nutolusi“ ir „artima“ atitinka AvCami1 dimero subvienetus ribosomos A srities atžvilgiu.

Apibendrinant, struktūriniai ir sąveikos su bL12 tyrimai atskleidė, kad surišus aktyvatorių cA_4 pasikeičia Cami1 baltymų konformacija ir atveriamas wHTH domeno sąveikos su bL12 C-galiniu domenu paviršius. Sąveikos su

bL12 dėka Cami1 RelE domenai efektyviau prisijungia ribosomos A srityje ir karpo ten esančią mRNR.

- Cami1 toksiškumas *in vivo*

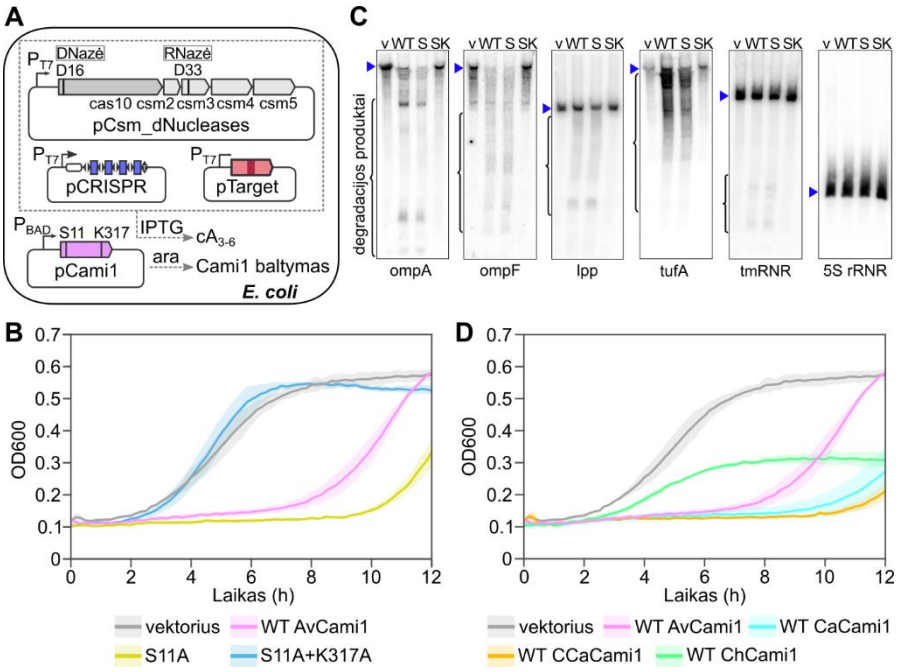
Siekdami patikrinti ar Cami1 išties gali dalyvauti III tipo CRISPR-Cas gynybiniame atsake pabandėme aktyvinti Cami1 *E. coli* bakterijose panaudodami *S. thermophilus* III-A tipo CRISPR-Cas sistemos StCsm kompleksą, kuris surišęs taikinį Svetimos DNR transkripte *in vivo* sintetina cOA nuo cA_3 iki cA_6 (įskaitant cA_4) (Smalakyte et al., 2020).

Tyrimui panaudotos trys IPTG indukuojamos plazmidės, koduojančios StCsm RNazinio ir DNazinio aktyvumo mutantą (pCsm_dNucleases) ir crRNR (pCRISPR), kuri yra nutaikyta prieš taikinį transkribuojamą nuo pTarget (15A pav.). Nukleaziniu aktyvumu pasižymintis StCsm mutantas buvo pasirinktas siekiant išvengti papildomo toksiškumo, tačiau jo laukinio tipo Palm domenas gali sintetinti cOA auginant šias plazmides turinčias bakterijas terpėje su IPTG.

Siekdami patikrinti Cami1 baltymų toksiškumą *in vivo,* cOA gaminančias *E. coli* transformavome pCami1 plazmidėmis, koduojančiomis skirtingus Cami1 baltymus arba jų mutantus, kurių raiška aktyvinama arabinoze. Bakterijų augimo bandymus atlikome panaudodami WT AvCami1, žiedo nukleazės S11A mutantą arba dvigubą S11A ir K317A mutantą, kuris papildomai pasižymėjo žemu mRNR interferazės aktyvumu (10B pav.). Pastebėjome, kad +WT AvCami1 ir +AvCami1-S11A bakterijos, bet ne +AvCami1-S11A-K317A, terpėje su induktoriais IPTG ir arabinoze augo daug lėčiau negu bakterijos su tuščiu vektoriumi vietoje pCami1 (15B pav.). Šis rezultatas rodo, kad AvCami1 mRNR interferazės aktyvumas sukelia toksiškumą vykstant nuolatinei cA_4 sintezei StCsm komplekse. +AvCami1-S11A bakterijos augo lėčiau negu +WT AvCami1 bakterijos. Spėjame, kad šis efektas nulemtas AvCami1 žiedo nukleazės aktyvumo dėl kurio +WT AvCami1 bakterijų vidinė cA_4 koncentracija buvo žemesnė, ir tai galėjo lemti žemesnį mRNR interferazės aktyvumą.

Norėdami tiesiogiai patikrinti mRNR karpymo aktyvumą gyvose bakterijose, panaudojome tas pačias *E. coli* ląsteles ir plazmides kaip ir bakterijų augimo bandymuose (15A pav.). Iš šių ląstelių, augintų terpėje su induktoriais, buvo surinktos išskirta visuminė RNR. Panaudojant radioaktyviai pažymėtus hibridizacijos zondus (Appendix 2) „Northern blot“ metodu buvo vizualizuoti keli gausiausi šeimininko transkriptai. mRNR nebuvo degraduojama ląstelėse be AvCami1 arba koduojančiose jo mRNR interferazės aktyviojo

centro mutantą, tačiau mRNR išskirta iš WT AvCami1 arba S11A mutantą koduojančių bakterijų buvo sudegraduota (15C pav.).



15 pav. Cami1 baltymai pasižymi cOA reguliuojamu mRNR interferazės aktyvumu heterologiniame šeimininke. (**A**) Keturiomis plazmidėmis transformuota *E. coli* koduojanti heterologinį StCsm kompleksą ir Cami1 baltymus. Šiose bakterijose IPTG sukelia StCsm komponentų ir taikinio raišką, kas lemia cOA sintezę, arabinozė sukelia Cami1 raišką. Šios bakterijos naudotos tiriant jų augimą (B ir D dalys) bei RNR integralumą (C). (**B**) Bakterijų vykdančių aktyvinto AvCami1 variantų raišką augimo kreivės. Bakterijų augimas įvertintas matuojant OD_{600} 10 min laiko intervalais. (**C**) Endogeninių *E. coli* baltymus koduojančių lpp, ompA, ompF, tufA mRNR, tmRNR ir 5S rRNR integralumo tyrimas „Northern blot“ metodu panaudojant iš AvCami1 paveiktų bakterijų išgrynintą visuminę RNR. „v“ – tuščio vektoriaus kontrolė, „WT“ – WT AvCami1, „S“ ir „SK“, atitinkamai AvCami1 S11A ir S11A+K317A mutantai. (**D**) Bakterijų vykdančių aktyvintų Cami1 homologų raišką augimo kreivės. Bandymai atlikti kaip (B). Tuščio vektoriaus kontrolė (B) ir (D) pažymėta kaip „vektorius“. Linijos ir spalvoti plotai (B) ir (D) atitinka trijų biologinių pakartojimų vidurkį ir standartinį nuokrypį.

Ribosomos, kurios transliuoja koduojančios sekos dalyje perkirptą mRNR, gali užstrigti, nes nepasiekia transliacijos terminaciją sukeliančio stop kodono (Müller et al., 2021). Užstrigusias ribosomas *E. coli* ir kitose bakterijose gali išlaisvinti hibridinė tmRNR, turinti tRNR ir mRNR savybių, kuri ribosomoje pakeičia mRNR matricą ir koduoja degradacijos žymę, kuria

pažymimas nebaigtas sintetinti peptidas (Müller et al., 2021). „Northern blot“ analizė parodė, kad tmRNR taip pat gali būti karpoma AvCami1, veikiausiai mRNR regione (15C pav.). Šis aktyvumas galėtų apsunkinti užstrigusių ribosomų išlaisvinimą kol nesibaigia cA_4 sintezė. 5S rRNR, kuri nėra transliuojama, nebuvo ir karpoma AvCami1 *in vivo* (15C pav.).

Kiti Cami1 baltymai taip pat apsunkina *E. coli* augimą, kai yra *in vivo* aktyvinami StCsm sintetinamo cA_4 (15D pav.). Šis rezultatas leidžia teigti, kad transliacijos aparato konservatyvumas leido Cami1 baltymų genams paplisti skirtingose bakterijose (6A pav.)

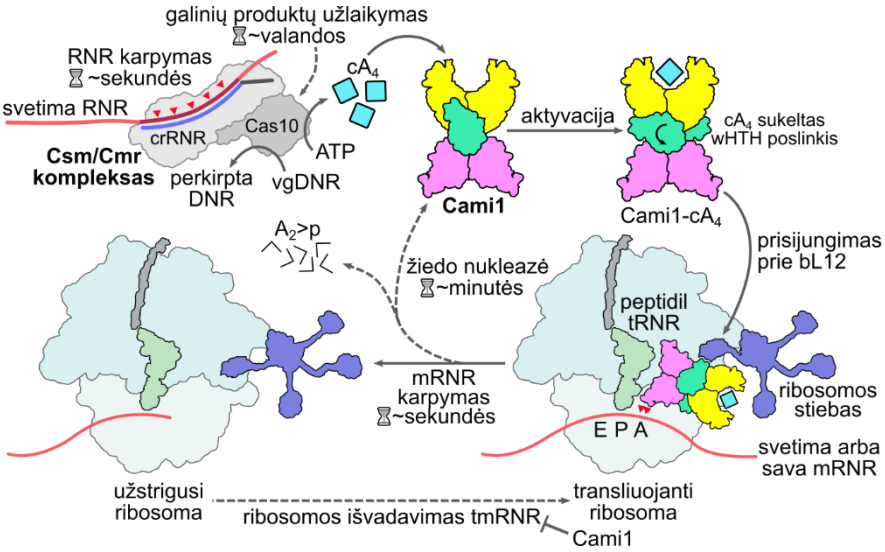
❖ III tipo CRISPR-Cas ir Cami1 gynybos mechanizmas

Remdamiesi atliktų tyrimų rezultatais, pasiūlėme apibendrintą III-A tipo CRISPR-Cas apsaugos sistemos mechanizmą, kuris paaiškina, kaip yra reguliuojama ši sistema ir kaip Cami1 padeda ginantis nuo bakteriją infekuojančių nukleorūgščių (16 pav.). Pagal prieš tai pasiūlytą reguliacijos mechanizmą, efektoriniam Csm kompleksui atpažinus svetimus transkriptus, yra aktyvinamas komplekso Cas10 subvienetas, kurio HD domenas degraduoja transkribuojamą DNR, o Palm domenas iš ATP sintetina cOA. Šios molekulės aktyvina su apsaugos sistema susijusius CARF/SAVED baltymus (Stella ir Marraffini, 2023). Nors šiame modelyje detalizuoti sistemos aktyvacijos etapai, nebuvo paaiškinta kaip sistemos aktyvumas yra nuslopinamas pasibaigus infekcijai. Ankstesnių tyrimų metu pastebėta, kad Cas10 subvieneto aktyvumas nuslopsta kompleksui sukarpius taikinio RNR, tačiau RNR hidrolizė įvyksta per kelias sekundes, o DNazinis ir polimerazinis-ciklazinis aktyvumas išlieka dešimtis minučių (Kazlauskiene et al., 2016, 2017).

Mūsų atlikti FCS eksperimentai leido susieti Csm komplekso RNR karpymą ir jo Cas10 subvieneto aktyvumo nuslopimą. Tyrimų metu nustatėme, kad kompleksui sukarpius taikinio RNR, jos galiniai produktai disocijuoja iš lėto: šalia Cas10 esantis 3′-galo produktas disocijuoja per $\langle t_{artimas} \rangle$ = 84 ± 2 min, o nuo Cas10 nutolęs 5′-galo produktas kompleksą palieka per $\langle t_{nutolęs} \rangle$ = 119 ± 3 min. Cas10 DNazinis aktyvumas nuslopsta per panašų laiką, nustatyta vidutinė DNazinio aktyvumo trukmė po aktyvacijos yra $\langle t_{DNazė} \rangle$ = 64 ± 1 min. Panašumas tarp šių verčių leidžia funkciškai susieti RNR užlaikymą komplekse su Cas10 DNaziniu aktyvumu.

Kadangi anksčiau buvo pastebėta, kad cOA sintezė po RNR karpymo irgi slopsta iš lėto (Kazlauskiene et al., 2017), galima spėti, kad abu Cas10 fermentiniai aktyvumai yra reguliuojami priklausomai nuo RNR užlaikymo. Šį

spėjimą netiesiogiai galima paremti pasigilinus į savos-svetimos RNR atpažinimą III tipo CRISPR-Cas kompleksuose. Kai Csm/Cmr kompleksai suriša RNR taikinį, kurio seka atitinka šeimininko CRISPR regiono priešprasminį transkriptą, ir DNazinis, ir polimerazinis-ciklazinis Cas10 aktyvumai yra slopinami (Li et al., 2024). Tai rodo, kad taikinio RNR turi įtakos abiejų Cas10 subvieneto fermentinių aktyvumų reguliacijai.



16 pav. Cami1 aktyvacija III-A tipo CRISPR-Cas apsaugos sistemoje. Svetimi transkriptai yra specifiškai surišami Csm komplekse. Tai aktyvina Csm komplekso Cas10 subvieneto DNazinį ir cA_4 sintazinį aktyvumą. Dėl lėtos RNR karpymo produktų disociacijos iš Csm komplekso Cas10 gali išlikti aktyvus vidutiniškai bent valandą. cA_4 prisirišus prie apsaugos sistemos papildomo Cami1 baltymo CARF domenų dimero (geltonas), wHTH domene (melsvai žalias) yra sukeliami konformaciniai pokyčiai, kurių dėka wHTH gali sąveikauti su ribosomos stiebo baltymu bL12 (mėlynas). Stiebo pagautas Cami1 yra pernešamas į ribosomos A sritį. Cami1 baltymo RelE domenas (rožinė) kerpa A srityje esančią mRNR, kas sukelia transliacijos slopinimą ląstelėje, sustabdo užpuoliko ir šeimininko baltymų gamybą. Cami1 taip pat kerpa ir tmRNR, taip slopindamas pagrindinį užstrigusių ribosomų išvadavimo mechanizmą. Pasibaigus infekcijai cA_4 karpymas CARF domene nuslopina Cami1 mRNR interferazinį aktyvumą. Spėjame, kad RNR produktų disociacija iš efektorinio komplekso nuslopina cA_4 sintezę.

Dėl ištęsto Cas10 aktyvumo ląstelėje ilgiau gali būti sintetinami cOA, kurie aktyvina įvairius pagalbinius CARF/SAVED efektorius, tarp jų ir šiame darbe aprašytą Cami1. Mūsų atlikti biocheminiai ir struktūriniai tyrimai atskleidė, kad Cami1 yra cA_4 reguliuojama mRNR interferazė, kuri surišusi aktyvatorių patenka į ribosomos A sritį ir karpo tuo metu ribosomoje transliuojamą

mRNR (16 pav.). Struktūriniai aktyvinto ir ribosomoje prisijungusio Cami1 baltymo tyrimai atskleidė, kad Cami1 surišus cA_4 aktyvatorių pasikeičia jo konformacija, wHTH domenai nutolsta nuo CARF domenų. Taip atveriamas sąveikos paviršius per kurį Cami1 prisijungia prie ribosomos stiebo baltymo bL12. Stiebo pagautas Cami1 baltymas patenka į A sritį, kur perkerpa mRNR. Šis prisijungimo-karpymo ciklas gali kartotis, kol CARF domene yra sukarpomas aktyvatorius cA_4.

Transliacijos metu koduojančioje srityje perkirptos mRNR sukelia ribosomų užstrigimą ant mRNR 3′-galo. Fago infekcijos metu tai gali sustabdyti greitą svetimų baltymų sintezę, kas trukdytų fago pasidauginimui. Dėl Cami1 aktyvumo sustojus transliacijai, infekuota ląstelė išlošia laiko sunaikinti svetimas nukleorūgštis, ką atlieka Csm kompleksas, taip pat kitos gynybos sistemos ar ląstelės nukleazės.

Užstrigusios ribosomos gali būti išlaisvinamos pasitelkiant įvairius ribosomų išvadavimo mechanizmus, iš kurių labiausiai yra paplitusi *trans*-transliacija panaudojant hibridinę tmRNR (Müller et al., 2021). Mūsų tyrimai *in vivo* parodė, kad Cami1 gali kirpti tmRNR ir taip riboti priešlaikinį transliacijos atstatymą dar nepasibaigus infekcijai.

Remiantis mūsų pasiūlytu modeliu, jeigu svetimos nukleorūgštys yra sunaikinamos ir ląstelė išgyvena infekciją, RNR disociacija iš Csm kompleksų galiausiai nuslopina cOA sintezę, o užsilikę cA_4 yra sukarpomi Cami1 arba kitų žiedo nukleazių. Tokių neigiamo atsako grandinių išsivystymas leidžia spėti, kad šios gynybos sistemos bakterijoms padėjo įveikti daugybę infekcijų aršios jų ir jas infekuojančių virusų koevoliucijos metu.

❖ Tolimesni darbai

Disertacijoje pateikti rezultatai leido papildyti III-A tipo CRISPR-Cas apsaugos mechanizmą įžvalgomis apie šios gynybos sistemos reguliaciją. Vis dėlto pagrindinės darbo išvados yra suformuluotos remiantis *in vitro* tyrimais, kurių sąlygos skiriasi nuo ląstelėje vykstančių procesų. Taigi lieka nežinoma, ar stebimos RNR užlaikymo ir Cas10 aktyvumo trukmės gyvoje ląstelėje atitinka nustatytąsias *in vitro* bandymų metu. Taip pat būtų galima patikrinti ar Csm kompleksui būdingos šių procesų trukmės yra panašios ir III-B CRISPR-Cas kompleksų Cmr atveju, mat šie du kompleksai pasižymi struktūriniais skirtumais Cas10 subvieneto aplinkoje (Li et al., 2024). Detalesni Cami1 tyrimai gyvose bakterijose leistų patikrinti, kaip efektyviai šie baltymai jungiasi prie veikiančių ribosomų. Nors tyrimų metu pastebėjome, kad vykdant Cami1

raišką nuo daugiakopijinių raiškos vektorių šis baltymas slopina augimą, lieka nustatyti, kaip efektyviai šis baltymas, esant natūraliai jo raiškai, konkuruoja su transliacijos faktoriais, kurie yra vieni gausiausių ląstelės baltymų (Ishihama et al., 2008).

Paskutiniais metais pagalbinių III tipo CRISPR-Cas efektorių tyrimų publikuojama vis daugiau atskleidžiant jų struktūras, funkcijas ir mechanizmus (Stella ir Marraffini, 2023). Mūsų atlikta CARF7 grupės baltymų sekų paieška ir analizė leido papildyti šią grupę CARF7-[PD-(D/E)XK] baltymais, kurie nebuvo pristatyti pirminėje CARF domenų klasifikacijoje (Makarova et al., 2020a). PD-(D/E)XK domenas yra dažnai aptinkamas įvairių nukleaziniu aktyvumu pasižyminčių baltymų struktūrose, geriausiai žinomas pavyzdys yra restrikcijos endonukleazės (Knizewski et al., 2007). Neseniai buvo aprašyti kiti CARF-[PD-(D/E)XK] baltymai: Can1, kuris yra DNR nikazė (McMahon et al., 2020) ir Card1/Can2 baltymai, kurie pasižymi tiek RNaziniu, tiek DNaziniu aktyvumais (Rostøl et al., 2021; Zhu et al., 2021). Šie baltymai priklauso CARF4 grupei, kuri nepasižymi žiedo nukleaziniu aktyvumu. Kiti mūsų analizuoti CARF7 baltymai, turintys sulietą HD domeną, galėtų pasižymėti reguliuojamu nukleaziniu arba fosfataziniu aktyvumu (Aravind ir Koonin, 1998). CARF7-DUF2103 baltymų efektorinio DUF2103 domeno struktūra yra panaši į RelE domeno struktūrą, kas sufleruoja panašią šio baltymo funkciją (Mogila et al., 2023). CARF7 baltymams Crn1 ir Cami1 surišant cA_4 stebimas wHTH domenų konformacinis pokytis galėtų būti būdingas ir kitiems CARF7 grupės baltymams.

Nepaisant didelės CARF ir struktūriškai giminingų SAVED baltymų įvairovės (Makarova et al., 2020a), III tipo CRISPR-Cas sistemose taip pat aptinkami pagalbiniai baltymai neturintys šių domenų, kaip, pavyzdžiui, trimerinė DNazė NucC, kuri aktyvuojasi subvienetų sąveikos paviršiuje surišusi cA_3 (Grüschow et al., 2021; Lau et al., 2020). Neseniai prof. Malcolm White grupė parodė, kad Cmr kompleksas iš *Bateroides fragilis* vietoje cOA gamina SAM-AMP signalines molekules (Chi et al., 2023). Naujai atrastas aktyvatorius aktyvina membraninį baltymą CorA pažeidžiantį plazmidės membranos integralumą infekcijos metu, kas galėtų sukelti žūtį arba stabdyti ląstelės augimą. Didėjant sekoskaitos apimtims atrandama vis naujų pagalbinių efektorių, susijusių su III tipo CRISPR-Cas sistemomis (Altae-Tran et al., 2023). Charakterizavus šiuos efektorius juos būtų galima panaudoti išvedant naujus biotechnologijoje svarbių bakterijų kamienus atsparesnius fagų infekcijoms. Išsiaiškinus šių ir kitų apsaugos sistemų veikimo mechanizmus galėtume riboti šias sistemas turinčias bakterijas naujoviška fagų terapija medicinoje.

Taip pat būtų galima tęsti Cami1 ir ribosomos stiebo sąveikos tyrimus. Anksčiau buvo parodyta, kad ribosomų stiebas suriša GTPazinius transliacijos faktorius, kas padidina transliacijos efektyvumą prokariotuose ir eukariotuose (Liljas ir Sanyal, 2018). Eukariotuose prie stiebo gali jungtis ir ribosomas sugadinantys RIP baltymai. Mūsų žiniomis Cami1 yra pirmasis charakterizuotas baltymas, kuris jungiasi prie prokariotinės ribosomos stiebo neskaitant GTPazinių transliacijos faktorių. Kaip ir RIP baltymai, Cami1 slopina transliaciją. Cami1 jungiasi prie stiebo bL12 baltymo C-galinio domeno per wHTH domeną. Kadangi skirtingos transliacinės GTPazės sąveikauja su bL12 C-galiniu domenu per tarpusavyje panašius G (arba G') domenus, galima spėti, kad ir kiti dar nežinomi baltymai, turintys į Cami1 panašų wHTH domeną, taip pat galėtų per jį jungtis prie ribosomos stiebo. Atlikus nuodugnią sekų-struktūrų bei tarpusavio sąveikos analizę, būtų galima rasti potencialių naujų su ribosomomis sąveikaujančių baltymų. Kadangi transliacija yra vienas kertinių procesų ląstelėje, duomenys apie naujus ribosomų veiklą moduliuojančius efektorius galėtų pateikti naujų įžvalgų ląstelės biologijoje ar vesti prie pritaikymų biotechnologijoje.

IŠVADOS

1. Csm3 perkirpus RNR *S. thermophilus* Csm kompleksas užlaiko galinius karpymo produktus, kurie stimuliuoja Cas10 DNazės aktyvumą.
2. CARF7 baltymų grupėje CARF-wHTH-RelE domeninę sudėtį turintys Cami1 baltymai tikėtinai susidarė per kelis atskirus kartus susiliejant domenams.
3. Cami1 baltymus aktyvina cA_4, kuris yra sukarpomas CARF domene.
4. Aktyvinto Cami1 baltymo RelE domenas kerpa mRNR prieinamą ribosomos A srityje
5. *In vivo* aktyvintas Cami1 kerpa transliuojamas mRNR, kas slopina bakterijų augimą.
6. Aktyvinto Cami1 baltymo wHTH domenas sąveikauja su ribosomos stiebo baltymu bL12, kas palengvina patekimą į ribosomą ir mRNR kirpimą.

REFERENCES

1. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr 68, 352–367. https://doi.org/10.1107/S0907444912001308
2. Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C.B., Baslé, A., Beilsten-Edmands, J., Borges, R.J., Brown, D.G., Burgos-Mármol, J.J., Berrisford, J.M., Bond, P.S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A.G., Cowtan, K.D., Croll, T.I., Debreczeni, J.É., Devenish, N.E., Dodson, E.J., Drevon, T.R., Emsley, P., Evans, G., Evans, P.R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E.F., Gerstel, M., Gildea, R.J., Hatti, K., Hekkelman, M.L., Heuser, P., Hoh, S.W., Hough, M.A., Jenkins, H.T., Jiménez, E., Joosten, R.P., Keegan, R.M., Keep, N., Krissinel, E.B., Kolenko, P., Kovalevskiy, O., Lamzin, V.S., Lawson, D.M., Lebedev, A.A., Leslie, A.G.W., Lohkamp, B., Long, F., Malý, M., McCoy, A.J., McNicholas, S.J., Medina, A., Millán, C., Murray, J.W., Murshudov, G.N., Nicholls, R.A., Noble, M.E.M., Oeffner, R., Pannu, N.S., Parkhurst, J.M., Pearce, N., Pereira, J., Perrakis, A., Powell, H.R., Read, R.J., Rigden, D.J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G.M., Shelley, K.L., Simkovic, F., Simpkin, A.J., Skubak, P., Sobolev, E., Steiner, R.A., Stevenson, K., Tews, I., Thomas, J.M.H., Thorn, A., Valls, J.T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K.S., Winn, M.D., Winter, G., Wojdyr, M., Yamashita, K., 2023. The CCP4 suite: integrative software for macromolecular crystallography. Acta Crystallogr D Struct Biol 79, 449–461. https://doi.org/10.1107/S2059798323003595
3. Akabane, S., Ueda, T., Nierhaus, K.H., Takeuchi, N., 2014. Ribosome rescue and translation termination at non-standard stop codons by ICT1 in mammalian mitochondria. PLoS Genet 10, e1004616. https://doi.org/10.1371/journal.pgen.1004616
4. Akarsu, H., Bordes, P., Mansour, M., Bigot, D.-J., Genevaux, P., Falquet, L., 2019. TASmania: A bacterial Toxin-Antitoxin Systems database. PLoS Comput. Biol. 15, e1006946. https://doi.org/10.1371/journal.pcbi.1006946
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. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new

generation of protein database search programs. Nucleic Acids Res 25, 3389–3402. https://doi.org/10.1093/nar/25.17.3389
7. Amitai, G., Sorek, R., 2016. CRISPR–Cas adaptation: insights into the mechanism of action. Nature Reviews Microbiology 14, 67–76. https://doi.org/10.1038/nrmicro.2015.14
8. Amunts, A., Brown, A., Toots, J., Scheres, S.H.W., Ramakrishnan, V., 2015. Ribosome. The structure of the human mitochondrial ribosome. Science 348, 95–98. https://doi.org/10.1126/science.aaa1193
9. Anantharaman, V., Aravind, L., 2003. New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. Genome Biol 4, R81. https://doi.org/10.1186/gb-2003-4-12-r81
10. Andreev, D., Hauryliuk, V., Terenin, I., Dmitriev, S., Ehrenberg, M., Shatsky, I., 2008. The bacterial toxin RelE induces specific mRNA cleavage in the A site of the eukaryote ribosome. RNA 14, 233–239. https://doi.org/10.1261/rna.693208
11. Aravind, L., Koonin, E.V., 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci 23, 469–472. https://doi.org/10.1016/s0968-0004(98)01293-6
12. Arenz, S., Nguyen, F., Beckmann, R., Wilson, D.N., 2015. Cryo-EM structure of the tetracycline resistance protein TetM in complex with a translating ribosome at 3.9-Å resolution. Proc Natl Acad Sci U S A 112, 5401–5406. https://doi.org/10.1073/pnas.1501775112
13. 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
14. Athukoralage, J.S., McMahon, S.A., Zhang, C., Grüschow, S., Graham, S., Krupovic, M., Whitaker, R.J., Gloster, T.M., White, M.F., 2020. An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. Nature 577, 572–575. https://doi.org/10.1038/s41586-019-1909-5
15. 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
16. Athukoralage, J.S., White, M.F., 2021. Cyclic oligoadenylate signalling and regulation by ring nucleases during type III CRISPR defence. RNA. https://doi.org/10.1261/rna.078739.121
17. Aviram, N., Thornal, A.N., Zeevi, D., Marraffini, L.A., 2022. Different modes of spacer acquisition by the Staphylococcus epidermidis type III-A CRISPR-Cas system. Nucleic Acids Res 50, 1661–1672. https://doi.org/10.1093/nar/gkab1299
18. Baca, C.F., Yu, Y., Rostøl, J.T., Majumder, P., Patel, D.J., Marraffini, L.A., 2024. The CRISPR effector Cam1 mediates membrane

depolarization for phage defence. Nature 625, 797–804. https://doi.org/10.1038/s41586-023-06902-y
19. Bacia, K., Schwille, P., 2007. Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. Nat Protoc 2, 2842–2856. https://doi.org/10.1038/nprot.2007.410
20. Ban, N., Beckmann, R., Cate, J.H.D., Dinman, J.D., Dragon, F., Ellis, S.R., Lafontaine, D.L.J., Lindahl, L., Liljas, A., Lipton, J.M., McAlear, M.A., Moore, P.B., Noller, H.F., Ortega, J., Panse, V.G., Ramakrishnan, V., Spahn, C.M.T., Steitz, T.A., Tchorzewski, M., Tollervey, D., Warren, A.J., Williamson, J.R., Wilson, D., Yonath, A., Yusupov, M., 2014. A new system for naming ribosomal proteins. Curr Opin Struct Biol 24, 165–169. https://doi.org/10.1016/j.sbi.2014.01.002
21. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712. https://doi.org/10.1126/science.1138140
22. Bernard, P., Couturier, M., 1992. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J Mol Biol 226, 735–745. https://doi.org/10.1016/0022-2836(92)90629-x
23. Bhoobalan-Chitty, Y., Johansen, T.B., Di Cianni, N., Peng, X., 2019. Inhibition of Type III CRISPR-Cas Immunity by an Archaeal Virus-Encoded Anti-CRISPR Protein. Cell 179, 448-458.e11. https://doi.org/10.1016/j.cell.2019.09.003
24. Bocharov, E.V., Sobol, A.G., Pavlov, K.V., Korzhnev, D.M., Jaravine, V.A., Gudkov, A.T., Arseniev, A.S., 2004. From structure and dynamics of protein L7/L12 to molecular switching in ribosome. J Biol Chem 279, 17697–17706. https://doi.org/10.1074/jbc.M313384200
25. Bøggild, A., Sofos, N., Andersen, K.R., Feddersen, A., Easter, A.D., Passmore, L.A., Brodersen, D.E., 2012. The crystal structure of the intact E. coli RelBE toxin-antitoxin complex provides the structural basis for conditional cooperativity. Structure 20, 1641–1648. https://doi.org/10.1016/j.str.2012.08.017
26. Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading) 151, 2551–2561. https://doi.org/10.1099/mic.0.28048-0
27. Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., van der Oost, J., 2008. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science 321, 960–964. https://doi.org/10.1126/science.1159689
28. Brown, A., Fernández, I.S., Gordiyenko, Y., Ramakrishnan, V., 2016. Ribosome-dependent activation of stringent control. Nature 534, 277–280. https://doi.org/10.1038/nature17675

29. Burley, S.K., Berman, H.M., Kleywegt, G.J., Markley, J.L., Nakamura, H., Velankar, S., 2017. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. Methods Mol Biol 1607, 627–641. https://doi.org/10.1007/978-1-4939-7000-1_26
30. Burman, N., Belukhina, S., Depardieu, F., Wilkinson, R.A., Skutel, M., Santiago-Frangos, A., Graham, A.B., Livenskyi, A., Chechenina, A., Morozova, N., Zahl, T., Henriques, W.S., Buyukyoruk, M., Rouillon, C., Saudemont, B., Shyrokova, L., Kurata, T., Hauryliuk, V., Severinov, K., Groseille, J., Thierry, A., Koszul, R., Tesson, F., Bernheim, A., Bikard, D., Wiedenheft, B., Isaev, A., 2024. A virally-encoded tRNA neutralizes the PARIS antiviral defence system. Nature. https://doi.org/10.1038/s41586-024-07874-3
31. 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
32. Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348
33. Carlson, M.A., Haddad, B.G., Weis, A.J., Blackwood, C.S., Shelton, C.D., Wuerth, M.E., Walter, J.D., Spiegel, P.C., 2017. Ribosomal protein L7/L12 is required for GTPase translation factors EF-G, RF3, and IF2 to bind in their GTP state to 70S ribosomes. FEBS J 284, 1631–1643. https://doi.org/10.1111/febs.14067
34. Carte, J., Wang, R., Li, H., Terns, R.M., Terns, M.P., 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes & Development 22, 3489–3496. https://doi.org/10.1101/gad.1742908
35. Chadani, Y., Ito, K., Kutsukake, K., Abo, T., 2012. ArfA recruits release factor 2 to rescue stalled ribosomes by peptidyl-tRNA hydrolysis in Escherichia coli. Mol Microbiol 86, 37–50. https://doi.org/10.1111/j.1365-2958.2012.08190.x
36. Chadani, Y., Matsumoto, E., Aso, H., Wada, T., Kutsukake, K., Sutou, S., Abo, T., 2011a. trans-translation-mediated tight regulation of the expression of the alternative ribosome-rescue factor ArfA in Escherichia coli. Genes Genet Syst 86, 151–163. https://doi.org/10.1266/ggs.86.151
37. Chadani, Y., Ono, K., Kutsukake, K., Abo, T., 2011b. Escherichia coli YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. Mol Microbiol 80, 772–785. https://doi.org/10.1111/j.1365-2958.2011.07607.x
38. 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 Microbiology Reviews 39, 428–441. https://doi.org/10.1093/femsre/fuv023
39. Chen, X., Wolin, S.L., 2023. Transfer RNA halves are found as nicked tRNAs in cells: evidence that nicked tRNAs regulate expression of an RNA repair operon. RNA 29, 620–629. https://doi.org/10.1261/rna.079575.122
40. Cheng, H., Schaeffer, R.D., Liao, Y., Kinch, L.N., Pei, J., Shi, S., Kim, B.-H., Grishin, N.V., 2014. ECOD: an evolutionary classification of protein domains. PLoS Comput Biol 10, e1003926. https://doi.org/10.1371/journal.pcbi.1003926
41. 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. https://doi.org/10.1038/s41586-023-06620-5
42. Chi, H., White, M.F., 2024. RNA processing by the CRISPR-associated NYN ribonuclease. https://doi.org/10.1101/2024.04.08.588522
43. Chicherin, I.V., Dukhalin, S.V., Khannanov, R.A., Baleva, M.V., Levitskii, S.A., Patrushev, M.V., Sergiev, P.V., Kamenski, P., 2021. Functional Diversity of Mitochondrial Peptidyl-tRNA Hydrolase ICT1 in Human Cells. Front Mol Biosci 8, 716885. https://doi.org/10.3389/fmolb.2021.716885
44. Choi, J., Chen, W., Minkina, A., Chardon, F.M., Suiter, C.C., Regalado, S.G., Domcke, S., Hamazaki, N., Lee, C., Martin, B., Daza, R.M., Shendure, J., 2022. A time-resolved, multi-symbol molecular recorder via sequential genome editing. Nature 608, 98–107. https://doi.org/10.1038/s41586-022-04922-8
45. Choi, W., Yamaguchi, Y., Lee, J.-W., Jang, K.-M., Inouye, M., Kim, S.-G., Yoon, M.-H., Park, J.-H., 2017. Translation-dependent mRNA cleavage by YhaV in Escherichia coli. FEBS Lett 591, 1853–1861. https://doi.org/10.1002/1873-3468.12705
46. 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
47. Chou-Zheng, L., Hatoum-Aslan, A., 2019. A type III-A CRISPR-Cas system employs degradosome nucleases to ensure robust immunity. Elife 8. https://doi.org/10.7554/eLife.45393
48. Christensen, S.K., Gerdes, K., 2003. RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. Mol. Microbiol. 48, 1389–1400. https://doi.org/10.1046/j.1365-2958.2003.03512.x
49. Clementi, N., Polacek, N., 2010. Ribosome-associated GTPases: the role of RNA for GTPase activation. RNA Biol 7, 521–527. https://doi.org/10.4161/rna.7.5.12467
50. Cocozaki, A.I., Ramia, N.F., Shao, Y., Hale, C.R., Terns, R.M., Terns, M.P., Li, H., 2012. Structure of the Cmr2 Subunit of the CRISPR-Cas

RNA Silencing Complex. Structure 20, 545–553. https://doi.org/10.1016/j.str.2012.01.018
51. Davidson, A.R., Lu, W.-T., Stanley, S.Y., Wang, J., Mejdani, M., Trost, C.N., Hicks, B.T., Lee, J., Sontheimer, E.J., 2020. Anti-CRISPRs: Protein Inhibitors of CRISPR-Cas Systems. Annu. Rev. Biochem. 89, 309–332. https://doi.org/10.1146/annurev-biochem-011420-111224
52. Davydov, I.I., Wohlgemuth, I., Artamonova, I.I., Urlaub, H., Tonevitsky, A.G., Rodnina, M.V., 2013. Evolution of the protein stoichiometry in the L12 stalk of bacterial and organellar ribosomes. Nat Commun 4, 1387. https://doi.org/10.1038/ncomms2373
53. De Bruyn, P., Girardin, Y., Loris, R., 2021. Prokaryote toxin-antitoxin modules: Complex regulation of an unclear function. Protein Sci 30, 1103–1113. https://doi.org/10.1002/pro.4071
54. Dedrick, R.M., Jacobs-Sera, D., Bustamante, C.A.G., Garlena, R.A., Mavrich, T.N., Pope, W.H., Reyes, J.C.C., Russell, D.A., Adair, T., Alvey, R., Bonilla, J.A., Bricker, J.S., Brown, B.R., Byrnes, D., Cresawn, S.G., Davis, W.B., Dickson, L.A., Edgington, N.P., Findley, A.M., Golebiewska, U., Grose, J.H., Hayes, C.F., Hughes, L.E., Hutchison, K.W., Isern, S., Johnson, A.A., Kenna, M.A., Klyczek, K.K., Mageeney, C.M., Michael, S.F., Molloy, S.D., Montgomery, M.T., Neitzel, J., Page, S.T., Pizzorno, M.C., Poxleitner, M.K., Rinehart, C.A., Robinson, C.J., Rubin, M.R., Teyim, J.N., Vazquez, E., Ware, V.C., Washington, J., Hatfull, G.F., 2017. Prophage-mediated defence against viral attack and viral counter-defence. Nat Microbiol 2, 16251. https://doi.org/10.1038/nmicrobiol.2016.251
55. Deep, A., Liang, Q., Enustun, E., Pogliano, J., Corbett, K.D., 2024. Architecture and activation mechanism of the bacterial PARIS defence system. Nature. https://doi.org/10.1038/s41586-024-07772-8
56. Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602–607. https://doi.org/10.1038/nature09886
57. Diaconu, M., Kothe, U., Schlünzen, F., Fischer, N., Harms, J.M., Tonevitsky, A.G., Stark, H., Rodnina, M.V., Wahl, M.C., 2005. Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. Cell 121, 991–1004. https://doi.org/10.1016/j.cell.2005.04.015
58. 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. https://doi.org/10.1126/science.aar4120
59. Duarte, I., Nabuurs, S.B., Magno, R., Huynen, M., 2012. Evolution and diversification of the organellar release factor family. Mol Biol Evol 29, 3497–3512. https://doi.org/10.1093/molbev/mss157

60. Dy, R.L., Richter, C., Salmond, G.P.C., Fineran, P.C., 2014. Remarkable Mechanisms in Microbes to Resist Phage Infections. Annual Review of Virology 1, 307–331. https://doi.org/10.1146/annurev-virology-031413-085500
61. 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
62. Elmore, J.R., Sheppard, N.F., Ramia, N., Deighan, T., Li, H., Terns, R.M., Terns, M.P., 2016. Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. Genes Dev 30, 447–459. https://doi.org/10.1101/gad.272153.115
63. Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126–2132. https://doi.org/10.1107/S0907444904019158
64. Ernits, K., Saha, C.K., Brodiazhenko, T., Chouhan, B., Shenoy, A., Buttress, J.A., Duque-Pedraza, J.J., Bojar, V., Nakamoto, J.A., Kurata, T., Egorov, A.A., Shyrokova, L., Johansson, M.J.O., Mets, T., Rustamova, A., Džigurski, J., Tenson, T., Garcia-Pino, A., Strahl, H., Elofsson, A., Hauryliuk, V., Atkinson, G.C., 2023. The structural basis of hyperpromiscuity in a core combinatorial network of type II toxin-antitoxin and related phage defense systems. Proc Natl Acad Sci U S A 120, e2305393120. https://doi.org/10.1073/pnas.2305393120
65. Estrella, M.A., Kuo, F.-T., Bailey, S., 2016. RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. Genes Dev 30, 460–470. https://doi.org/10.1101/gad.273722.115
66. Feaga, H.A., Quickel, M.D., Hankey-Giblin, P.A., Keiler, K.C., 2016. Human Cells Require Non-stop Ribosome Rescue Activity in Mitochondria. PLoS Genet 12, e1005964. https://doi.org/10.1371/journal.pgen.1005964
67. Feaga, H.A., Viollier, P.H., Keiler, K.C., 2014. Release of nonstop ribosomes is essential. mBio 5, e01916. https://doi.org/10.1128/mBio.01916-14
68. Feng, S., Chen, Y., Kamada, K., Wang, H., Tang, K., Wang, M., Gao, Y.-G., 2013. YoeB-ribosome structure: a canonical RNase that requires the ribosome for its specific activity. Nucleic Acids Res 41, 9549–9556. https://doi.org/10.1093/nar/gkt742
69. Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., Salmond, G.P.C., 2009. The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proc Natl Acad Sci U S A 106, 894–899. https://doi.org/10.1073/pnas.0808832106
70. Franco, M.K., Koutmou, K.S., 2022. Chemical modifications to mRNA nucleobases impact translation elongation and termination. Biophys Chem 285, 106780. https://doi.org/10.1016/j.bpc.2022.106780

71. Fricke, T., Smalakyte, D., Lapinski, M., Pateria, A., Weige, C., Pastor, M., Kolano, A., Winata, C., Siksnys, V., Tamulaitis, G., Bochtler, M., 2020. Targeted RNA Knockdown by a Type III CRISPR-Cas Complex in Zebrafish. CRISPR J 3, 299–313. https://doi.org/10.1089/crispr.2020.0032
72. Gajiwala, K.S., Burley, S.K., 2000. Winged helix proteins. Curr Opin Struct Biol 10, 110–116. https://doi.org/10.1016/s0959-440x(99)00057-3
73. 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
74. 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
75. Garza-Sánchez, F., Schaub, R.E., Janssen, B.D., Hayes, C.S., 2011. tmRNA regulates synthesis of the ArfA ribosome rescue factor. Mol Microbiol 80, 1204–1219. https://doi.org/10.1111/j.1365-2958.2011.07638.x
76. Ge, X., Mandava, C.S., Lind, C., Åqvist, J., Sanyal, S., 2018. Complementary charge-based interaction between the ribosomal-stalk protein L7/12 and IF2 is the key to rapid subunit association. Proc Natl Acad Sci U S A 115, 4649–4654. https://doi.org/10.1073/pnas.1802001115
77. Ghisotti, D., Finkel, S., Halling, C., Dehò, G., Sironi, G., Calendar, R., 1990. Nonessential region of bacteriophage P4: DNA sequence, transcription, gene products, and functions. J Virol 64, 24–36. https://doi.org/10.1128/JVI.64.1.24-36.1990
78. Giudice, E., Macé, K., Gillet, R., 2014. Trans-translation exposed: understanding the structures and functions of tmRNA-SmpB. Front Microbiol 5, 113. https://doi.org/10.3389/fmicb.2014.00113
79. Goeders, N., Drèze, P.-L., Van Melderen, L., 2013. Relaxed cleavage specificity within the RelE toxin family. J. Bacteriol. 195, 2541–2549. https://doi.org/10.1128/JB.02266-12
80. González-Delgado, A., Mestre, M.R., Martínez-Abarca, F., Toro, N., 2019. Spacer acquisition from RNA mediated by a natural reverse transcriptase-Cas1 fusion protein associated with a type III-D CRISPR-Cas system in Vibrio vulnificus. Nucleic Acids Res 47, 10202–10211. https://doi.org/10.1093/nar/gkz746
81. 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 gkae603. https://doi.org/10.1093/nar/gkae603
82. Grabe, G.J., Giorgio, R.T., Wieczór, M., Gollan, B., Sargen, M., Orozco, M., Hare, S.A., Helaine, S., 2024. Molecular stripping underpins derepression of a toxin-antitoxin system. Nat Struct Mol Biol. https://doi.org/10.1038/s41594-024-01253-2
83. Grela, P., Szajwaj, M., Horbowicz-Drożdżal, P., Tchórzewski, M., 2019. How Ricin Damages the Ribosome. Toxins (Basel) 11, E241. https://doi.org/10.3390/toxins11050241
84. Griffin, M.A., Davis, J.H., Strobel, S.A., 2013. Bacterial toxin RelE: a highly efficient ribonuclease with exquisite substrate specificity using atypical catalytic residues. Biochemistry 52, 8633–8642. https://doi.org/10.1021/bi401325c
85. 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 gkab1190. https://doi.org/10.1093/nar/gkab1190
86. Guegler, C.K., Laub, M.T., 2021. Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. Mol Cell 81, 2361-2373.e9. https://doi.org/10.1016/j.molcel.2021.03.027
87. 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
88. Guyomar, C., D'Urso, G., Chat, S., Giudice, E., Gillet, R., 2021. Structures of tmRNA and SmpB as they transit through the ribosome. Nat Commun 12, 4909. https://doi.org/10.1038/s41467-021-24881-4
89. Ha, T., Tinnefeld, P., 2012. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annu Rev Phys Chem 63, 595–617. https://doi.org/10.1146/annurev-physchem-032210-103340
90. Hadži, S., Garcia-Pino, A., Haesaerts, S., Jurėnas, D., Gerdes, K., Lah, J., Loris, R., 2017. Ribosome-dependent Vibrio cholerae mRNAse HigB2 is regulated by a β-strand sliding mechanism. Nucleic Acids Research 45, 4972–4983. https://doi.org/10.1093/nar/gkx138
91. Hale, C., Kleppe, K., Terns, R.M., Terns, M.P., 2008. Prokaryotic silencing (psi)RNAs in Pyrococcus furiosus. RNA 14, 2572–2579. https://doi.org/10.1261/rna.1246808
92. Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., Terns, M.P., 2009. RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. Cell 139, 945–956. https://doi.org/10.1016/j.cell.2009.07.040

93. Handa, Y., Inaho, N., Nameki, N., 2011. YaeJ is a novel ribosome-associated protein in Escherichia coli that can hydrolyze peptidyl-tRNA on stalled ribosomes. Nucleic Acids Res 39, 1739–1748. https://doi.org/10.1093/nar/gkq1097
94. Harms, A., Brodersen, D.E., Mitarai, N., Gerdes, K., 2018. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. Mol. Cell 70, 768–784. https://doi.org/10.1016/j.molcel.2018.01.003
95. Hatoum-Aslan, A., Maniv, I., Marraffini, L.A., 2011. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proceedings of the National Academy of Sciences 108, 21218–21222. https://doi.org/10.1073/pnas.1112832108
96. Hatoum-Aslan, A., Maniv, I., Samai, P., Marraffini, L.A., 2014. Genetic Characterization of Antiplasmid Immunity through a Type III-A CRISPR-Cas System. Journal of Bacteriology 196, 310–317. https://doi.org/10.1128/JB.01130-13
97. Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W., Marraffini, L.A., 2013. A Ruler Protein in a Complex for Antiviral Defense Determines the Length of Small Interfering CRISPR RNAs. Journal of Biological Chemistry 288, 27888–27897. https://doi.org/10.1074/jbc.M113.499244
98. Helena-Bueno, K., Rybak, M.Y., Ekemezie, C.L., Sullivan, R., Brown, C.R., Dingwall, C., Baslé, A., Schneider, C., Connolly, J.P.R., Blaza, J.N., Csörgő, B., Moynihan, P.J., Gagnon, M.G., Hill, C.H., Melnikov, S.V., 2024. A new family of bacterial ribosome hibernation factors. Nature 626, 1125–1132. https://doi.org/10.1038/s41586-024-07041-8
99. Helgstrand, M., Mandava, C.S., Mulder, F.A.A., Liljas, A., Sanyal, S., Akke, M., 2007. The ribosomal stalk binds to translation factors IF2, EF-Tu, EF-G and RF3 via a conserved region of the L12 C-terminal domain. J Mol Biol 365, 468–479. https://doi.org/10.1016/j.jmb.2006.10.025
100. Holm, L., 2019. Benchmarking fold detection by DaliLite v.5. Bioinformatics 35, 5326–5327. https://doi.org/10.1093/bioinformatics/btz536
101. Høyland-Kroghsbo, N.M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-Denomy, J., Bassler, B.L., 2017. Quorum sensing controls the Pseudomonas aeruginosa CRISPR-Cas adaptive immune system. Proc Natl Acad Sci U S A 114, 131–135. https://doi.org/10.1073/pnas.1617415113
102. Hudson, C.M., Lau, B.Y., Williams, K.P., 2014. Ends of the line for tmRNA-SmpB. Front Microbiol 5, 421. https://doi.org/10.3389/fmicb.2014.00421
103. 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
104. Hurley, J.M., Cruz, J.W., Ouyang, M., Woychik, N.A., 2011. Bacterial toxin RelE mediates frequent codon-independent mRNA cleavage from

the 5’ end of coding regions in vivo. J. Biol. Chem. 286, 14770–14778. https://doi.org/10.1074/jbc.M110.108969
105. Hurley, J.M., Woychik, N.A., 2009. Bacterial Toxin HigB Associates with Ribosomes and Mediates Translation-dependent mRNA Cleavage at A-rich Sites. J. Biol. Chem. 284, 18605–18613. https://doi.org/10.1074/jbc.M109.008763
106. Imai, H., Uchiumi, T., Kodera, N., 2020. Direct visualization of translational GTPase factor pool formed around the archaeal ribosomal P-stalk by high-speed AFM. Proc Natl Acad Sci U S A 117, 32386–32394. https://doi.org/10.1073/pnas.2018975117
107. Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F.U., Kerner, M.J., Frishman, D., 2008. Protein abundance profiling of the Escherichia coli cytosol. BMC Genomics 9, 102. https://doi.org/10.1186/1471-2164-9-102
108. Ito, K., Chadani, Y., Nakamori, K., Chiba, S., Akiyama, Y., Abo, T., 2011. Nascentome analysis uncovers futile protein synthesis in Escherichia coli. PLoS One 6, e28413. https://doi.org/10.1371/journal.pone.0028413
109. Ito, K., Honda, T., Suzuki, T., Miyoshi, T., Murakami, R., Yao, M., Uchiumi, T., 2014. Molecular insights into the interaction of the ribosomal stalk protein with elongation factor 1α. Nucleic Acids Res 42, 14042–14052. https://doi.org/10.1093/nar/gku1248
110. Ivanova, N., Pavlov, M.Y., Felden, B., Ehrenberg, M., 2004. Ribosome rescue by tmRNA requires truncated mRNAs. J Mol Biol 338, 33–41. https://doi.org/10.1016/j.jmb.2004.02.043
111. Iyer, L.M., Aravind, L., 2002. The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates. BMC Genomics 3, 33. https://doi.org/10.1186/1471-2164-3-33
112. Jackson, R.N., Wiedenheft, B., 2015. A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. Mol Cell 58, 722–728. https://doi.org/10.1016/j.molcel.2015.05.023
113. Jansen, R., Embden, J.D.A. van, Gaastra, W., Schouls, L.M., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 43, 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x
114. Jia, N., Jones, R., Yang, G., Ouerfelli, O., Patel, D.J., 2019a. CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. Mol. Cell. https://doi.org/10.1016/j.molcel.2019.06.014
115. Jia, N., Mo, C.Y., Wang, C., Eng, E.T., Marraffini, L.A., Patel, D.J., 2019b. 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

116. Jiang, W., Samai, P., Marraffini, L.A., 2016. Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. Cell 164, 710–721. https://doi.org/10.1016/j.cell.2015.12.053
117. 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
118. Jung, T.-Y., An, Y., Park, K.-H., Lee, M.-H., Oh, B.-H., Woo, E., 2015. Crystal Structure of the Csm1 Subunit of the Csm Complex and Its Single-Stranded DNA-Specific Nuclease Activity. Structure 23, 782–790. https://doi.org/10.1016/j.str.2015.01.021
119. Jurėnas, D., Fraikin, N., Goormaghtigh, F., Van Melderen, L., 2022. Biology and evolution of bacterial toxin-antitoxin systems. Nat Rev Microbiol. https://doi.org/10.1038/s41579-021-00661-1
120. Jurėnas, D., Van Melderen, L., 2020. The Variety in the Common Theme of Translation Inhibition by Type II Toxin-Antitoxin Systems. Front Genet 11, 262. https://doi.org/10.3389/fgene.2020.00262
121. Kabsch, W., 2010. XDS. Acta Crystallogr D Biol Crystallogr 66, 125–132. https://doi.org/10.1107/S0907444909047337
122. Karzai, A.W., Susskind, M.M., Sauer, R.T., 1999. SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA). EMBO J 18, 3793–3799. https://doi.org/10.1093/emboj/18.13.3793
123. Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform 9, 286–298. https://doi.org/10.1093/bib/bbn013
124. 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
125. 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
126. Keiler, K.C., Waller, P.R., 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

127. Kelly, A., Arrowsmith, T.J., Went, S.C., Blower, T.R., 2023. Toxin-antitoxin systems as mediators of phage defence and the implications for abortive infection. Curr Opin Microbiol 73, 102293. https://doi.org/10.1016/j.mib.2023.102293
128. Knizewski, L., Kinch, L.N., Grishin, N.V., Rychlewski, L., Ginalski, K., 2007. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. BMC Struct Biol 7, 40. https://doi.org/10.1186/1472-6807-7-40
129. Koga, M., Otsuka, Y., Lemire, S., Yonesaki, T., 2011. Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system. Genetics 187, 123–130. https://doi.org/10.1534/genetics.110.121798
130. Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., Inokuchi, H., 1994. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli. Proc Natl Acad Sci U S A 91, 9223–9227. https://doi.org/10.1073/pnas.91.20.9223
131. Koonin, E.V., 2016. Viruses and mobile elements as drivers of evolutionary transitions. Philos Trans R Soc Lond B Biol Sci 371, 20150442. https://doi.org/10.1098/rstb.2015.0442
132. Koripella, R.K., Sharma, M.R., Bhargava, K., Datta, P.P., Kaushal, P.S., Keshavan, P., Spremulli, L.L., Banavali, N.K., Agrawal, R.K., 2020. Structures of the human mitochondrial ribosome bound to EF-G1 reveal distinct features of mitochondrial translation elongation. Nat Commun 11, 3830. https://doi.org/10.1038/s41467-020-17715-2
133. Kotta-Loizou, I., Giuliano, M.G., Jovanovic, M., Schaefer, J., Ye, F., Zhang, N., Irakleidi, D.A., Liu, X., Zhang, X., Buck, M., Engl, C., 2022. The RNA repair proteins RtcAB regulate transcription activator RtcR via its CRISPR-associated Rossmann fold domain. iScience 25, 105425. https://doi.org/10.1016/j.isci.2022.105425
134. Kumar, V., Chen, Y., Ero, R., Ahmed, T., Tan, J., Li, Z., Wong, A.S.W., Bhushan, S., Gao, Y.-G., 2015. Structure of BipA in GTP form bound to the ratcheted ribosome. Proc Natl Acad Sci U S A 112, 10944–10949. https://doi.org/10.1073/pnas.1513216112
135. Kurata, T., Brodiazhenko, T., Alves Oliveira, S.R., Roghanian, M., Sakaguchi, Y., Turnbull, K.J., Bulvas, O., Takada, H., Tamman, H., Ainelo, A., Pohl, R., Rejman, D., Tenson, T., Suzuki, T., Garcia-Pino, A., Atkinson, G.C., Hauryliuk, V., 2021. RelA-SpoT Homolog toxins pyrophosphorylate the CCA end of tRNA to inhibit protein synthesis. Molecular Cell S1097276521004524. https://doi.org/10.1016/j.molcel.2021.06.005
136. Labrie, S.J., Samson, J.E., Moineau, S., 2010. Bacteriophage resistance mechanisms. Nature Reviews Microbiology 8, 317–327. https://doi.org/10.1038/nrmicro2315
137. Lakowicz, J.R., 2006. Principles of fluorescence spectroscopy, 3rd ed. ed. Springer, New York.

138. 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. https://doi.org/10.1016/j.molcel.2019.12.010
139. 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
140. Lee, K.-Y., Lee, B.-J., 2023. Dynamics-Based Regulatory Switches of Type II Antitoxins: Insights into New Antimicrobial Discovery. Antibiotics (Basel) 12, 637. https://doi.org/10.3390/antibiotics12040637
141. LeRoux, M., Culviner, P.H., Liu, Y.J., Littlehale, M.L., Laub, M.T., 2020. Stress Can Induce Transcription of Toxin-Antitoxin Systems without Activating Toxin. Mol Cell 79, 280-292.e8. https://doi.org/10.1016/j.molcel.2020.05.028
142. LeRoux, M., Laub, M.T., 2022. Toxin-Antitoxin Systems as Phage Defense Elements. Annu Rev Microbiol 76, 21–43. https://doi.org/10.1146/annurev-micro-020722-013730
143. LeRoux, M., Srikant, S., Teodoro, G.I.C., Zhang, T., Littlehale, M.L., Doron, S., Badiee, M., Leung, A.K.L., Sorek, R., Laub, M.T., 2022. The DarTG toxin-antitoxin system provides phage defence by ADP-ribosylating viral DNA. Nat Microbiol 7, 1028–1040. https://doi.org/10.1038/s41564-022-01153-5
144. Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., Sorek, R., 2015. CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature 520, 505–510. https://doi.org/10.1038/nature14302
145. Li, G.-W., Burkhardt, D., Gross, C., Weissman, J.S., 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 157, 624–635. https://doi.org/10.1016/j.cell.2014.02.033
146. Li, G.-Y., Zhang, Y., Inouye, M., Ikura, M., 2009. Inhibitory mechanism of Escherichia coli RelE-RelB toxin-antitoxin module involves a helix displacement near an mRNA interferase active site. J. Biol. Chem. 284, 14628–14636. https://doi.org/10.1074/jbc.M809656200
147. Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158
148. 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
149. Liljas, A., Sanyal, S., 2018. The enigmatic ribosomal stalk. Q Rev Biophys 51, e12. https://doi.org/10.1017/S0033583518000100

150. Lillestøl, R., Redder, P., Garrett, R.A., Brügger, K., 2006. A putative viral defence mechanism in archaeal cells. Archaea 2, 59–72. https://doi.org/10.1155/2006/542818
151. Lin, J., Alfastsen, L., Bhoobalan-Chitty, Y., Peng, X., 2023. Molecular basis for inhibition of type III-B CRISPR-Cas by an archaeal viral anti-CRISPR protein. Cell Host Microbe 31, 1837-1849.e5. https://doi.org/10.1016/j.chom.2023.10.003
152. Lin, J., Fuglsang, A., Kjeldsen, A.L., Sun, K., Bhoobalan-Chitty, Y., Peng, X., 2020. DNA targeting by subtype I-D CRISPR–Cas shows type I and type III features. Nucleic Acids Res. gkaa749. https://doi.org/10.1093/nar/gkaa749
153. Liu, J., Li, Q., Wang, X., Liu, Z., Ye, Q., Liu, T., Pan, S., Peng, N., 2023. An archaeal virus-encoded anti-CRISPR protein inhibits type III-B immunity by inhibiting Cas RNP complex turnover. Nucleic Acids Res 51, 11783–11796. https://doi.org/10.1093/nar/gkad804
154. Liu, T., Liu, Z., Ye, Q., Pan, S., Wang, X., Li, Y., Peng, W., Liang, Y., She, Q., Peng, N., 2017. Coupling transcriptional activation of CRISPR-Cas system and DNA repair genes by Csa3a in Sulfolobus islandicus. Nucleic Acids Res 45, 8978–8992. https://doi.org/10.1093/nar/gkx612
155. Liu, T.Y., Iavarone, A.T., Doudna, J.A., 2017. RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. PLoS One 12, e0170552. https://doi.org/10.1371/journal.pone.0170552
156. Liu, T.Y., Liu, J.-J., Aditham, A.J., Nogales, E., Doudna, J.A., 2019. Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble. Nat Commun 10, 3001. https://doi.org/10.1038/s41467-019-10780-2
157. 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. https://doi.org/10.1016/j.cell.2020.05.019
158. Maehigashi, T., Ruangprasert, A., Miles, S.J., Dunham, C.M., 2015. Molecular basis of ribosome recognition and mRNA hydrolysis by the E. coli YafQ toxin. Nucleic Acids Res. 43, 8002–8012. https://doi.org/10.1093/nar/gkv791
159. Majumdar, S., Zhao, P., Pfister, N.T., Compton, M., Olson, S., Glover, C.V.C., Wells, L., Graveley, B.R., Terns, R.M., Terns, M.P., 2015. Three CRISPR-Cas immune effector complexes coexist in Pyrococcus furiosus. RNA 21, 1147–1158. https://doi.org/10.1261/rna.049130.114
160. Makarova, K.S., Anantharaman, V., Grishin, N.V., Koonin, E.V., Aravind, L., 2014. CARF and WYL domains: ligand-binding regulators of

prokaryotic defense systems. Front Genet 5, 102. https://doi.org/10.3389/fgene.2014.00102
161. 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. https://doi.org/10.1093/nar/gkaa635
162. 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
163. Maniv, I., Jiang, W., Bikard, D., Marraffini, L.A., 2016. Impact of Different Target Sequences on Type III CRISPR-Cas Immunity. Journal of Bacteriology 198, 941–950. https://doi.org/10.1128/JB.00897-15
164. Mansour, M., Giudice, E., Xu, X., Akarsu, H., Bordes, P., Guillet, V., Bigot, D.-J., Slama, N., D'urso, G., Chat, S., Redder, P., Falquet, L., Mourey, L., Gillet, R., Genevaux, P., 2022. Substrate recognition and cryo-EM structure of the ribosome-bound TAC toxin of Mycobacterium tuberculosis. Nat Commun 13, 2641. https://doi.org/10.1038/s41467-022-30373-w
165. Maracci, C., Rodnina, M.V., 2016. Review: Translational GTPases. Biopolymers 105, 463–475. https://doi.org/10.1002/bip.22832
166. Marraffini, L.A., Sontheimer, E.J., 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 463, 568–571. https://doi.org/10.1038/nature08703
167. 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
168. McBride, T.M., Schwartz, E.A., Kumar, A., Taylor, D.W., Fineran, P.C., Fagerlund, R.D., 2020. Diverse CRISPR-Cas Complexes Require Independent Translation of Small and Large Subunits from a Single Gene. Mol Cell 80, 971-979.e7. https://doi.org/10.1016/j.molcel.2020.11.003
169. 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
170. Mets, T., Kurata, T., Ernits, K., Johansson, M.J.O., Craig, S.Z., Evora, G.M., Buttress, J.A., Odai, R., Wallant, K.C., Nakamoto, J.A., Shyrokova, L., Egorov, A.A., Doering, C.R., Brodiazhenko, T., Laub, M.T., Tenson, T., Strahl, H., Martens, C., Harms, A., Garcia-Pino, A.,

Atkinson, G.C., Hauryliuk, V., 2024. Mechanism of phage sensing and restriction by toxin-antitoxin-chaperone systems. Cell Host Microbe S1931-3128(24)00174–4. https://doi.org/10.1016/j.chom.2024.05.003
171. Millman, A., Melamed, S., Amitai, G., Sorek, R., 2020. Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. Nat Microbiol. https://doi.org/10.1038/s41564-020-0777-y
172. Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., 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 expanding arsenal of immune systems that protect bacteria from phages. https://doi.org/10.1101/2022.05.11.491447
173. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015
174. Mirdita, M., von den Driesch, L., Galiez, C., Martin, M.J., Söding, J., Steinegger, M., 2017. Uniclust databases of clustered and deeply annotated protein sequences and alignments. Nucleic Acids Res 45, D170–D176. https://doi.org/10.1093/nar/gkw1081
175. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., Bateman, A., 2021. Pfam: The protein families database in 2021. Nucleic Acids Res 49, D412–D419. https://doi.org/10.1093/nar/gkaa913
176. 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
177. 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
178. Mohr, G., Silas, S., Stamos, J.L., Makarova, K.S., Markham, L.M., Yao, J., Lucas-Elío, P., Sanchez-Amat, A., Fire, A.Z., Koonin, E.V., Lambowitz, A.M., 2018. A Reverse Transcriptase-Cas1 Fusion Protein Contains a Cas6 Domain Required for Both CRISPR RNA Biogenesis and RNA Spacer Acquisition. Mol Cell 72, 700-714.e8. https://doi.org/10.1016/j.molcel.2018.09.013
179. Mojica, F.J., Díez-Villaseñor, C., Soria, E., Juez, G., 2000. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol 36, 244–246. https://doi.org/10.1046/j.1365-2958.2000.01838.x

180. 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
181. 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 gkac923. https://doi.org/10.1093/nar/gkac923
182. 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 gkab1130. https://doi.org/10.1093/nar/gkab1130
183. 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
184. 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
185. Moreno, S., Muriel-Millán, L.F., Rodríguez-Martínez, K., Ortíz-Vasco, C., Bedoya-Pérez, L.P., Espín, G., 2022. The ribosome rescue pathways SsrA-SmpB, ArfA, and ArfB mediate tolerance to heat and antibiotic stresses in Azotobacter vinelandii. FEMS Microbiol Lett 369, fnac104. https://doi.org/10.1093/femsle/fnac104
186. Müller, C., Crowe-McAuliffe, C., Wilson, D.N., 2021. Ribosome Rescue Pathways in Bacteria. Front. Microbiol. 12, 652980. https://doi.org/10.3389/fmicb.2021.652980
187. Mustafi, M., Weisshaar, J.C., 2019. Near Saturation of Ribosomal L7/L12 Binding Sites with Ternary Complexes in Slowly Growing E. coli. J Mol Biol 431, 2343–2353. https://doi.org/10.1016/j.jmb.2019.04.037
188. Mustafi, M., Weisshaar, J.C., 2018. Simultaneous Binding of Multiple EF-Tu Copies to Translating Ribosomes in Live Escherichia coli. mBio 9, e02143-17. https://doi.org/10.1128/mBio.02143-17
189. Naganuma, T., Nomura, N., Yao, M., Mochizuki, M., Uchiumi, T., Tanaka, I., 2010. Structural basis for translation factor recruitment to the eukaryotic/archaeal ribosomes. J Biol Chem 285, 4747–4756. https://doi.org/10.1074/jbc.M109.068098
190. Neubauer, C., Gao, Y.-G., Andersen, K.R., Dunham, C.M., Kelley, A.C., Hentschel, J., Gerdes, K., Ramakrishnan, V., Brodersen, D.E., 2009. The Structural Basis for mRNA Recognition and Cleavage by the Ribosome-

Dependent Endonuclease RelE. Cell 139, 1084–1095. https://doi.org/10.1016/j.cell.2009.11.015
191. 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
192. Niewoehner, O., Jinek, M., 2016. Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA 22, 318–329. https://doi.org/10.1261/rna.054098.115
193. Noeske, J., Wasserman, M.R., Terry, D.S., Altman, R.B., Blanchard, S.C., Cate, J.H.D., 2015. High-resolution structure of the Escherichia coli ribosome. Nat Struct Mol Biol 22, 336–341. https://doi.org/10.1038/nsmb.2994
194. Nuñez, J.K., Lee, A.S.Y., Engelman, A., Doudna, J.A., 2015. Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. Nature 519, 193–198. https://doi.org/10.1038/nature14237
195. Ofir, G., Sorek, R., 2018. Contemporary Phage Biology: From Classic Models to New Insights. Cell 172, 1260–1270. https://doi.org/10.1016/j.cell.2017.10.045
196. Ogura, T., Hiraga, S., 1983. Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc Natl Acad Sci U S A 80, 4784–4788. https://doi.org/10.1073/pnas.80.15.4784
197. Oh, B.K., Apirion, D., 1991. 10Sa RNA, a small stable RNA of Escherichia coli, is functional. Mol Gen Genet 229, 52–56. https://doi.org/10.1007/BF00264212
198. Osawa, T., Inanaga, H., Sato, C., Numata, T., 2015. Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. Molecular Cell 58, 418–430. https://doi.org/10.1016/j.molcel.2015.03.018
199. Otsuka, Y., Yonesaki, T., 2012. Dmd of bacteriophage T4 functions as an antitoxin against Escherichia coli LsoA and RnlA toxins. Mol Microbiol 83, 669–681. https://doi.org/10.1111/j.1365-2958.2012.07975.x
200. Overgaard, M., Borch, J., Jørgensen, M.G., Gerdes, K., 2008. Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. Mol Microbiol 69, 841–857. https://doi.org/10.1111/j.1365-2958.2008.06313.x
201. 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. https://doi.org/10.1038/s41586-021-03886-5
202. Pallesen, J., Hashem, Y., Korkmaz, G., Koripella, R.K., Huang, C., Ehrenberg, M., Sanyal, S., Frank, J., 2013. Cryo-EM visualization of the ribosome in termination complex with apo-RF3 and RF1. Elife 2, e00411. https://doi.org/10.7554/eLife.00411

203. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S., Tucker, P.A., 2005. Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. Acta Crystallogr D Biol Crystallogr 61, 449–457. https://doi.org/10.1107/S0907444905001307
204. Pape, T., Wintermeyer, W., Rodnina, M.V., 1998. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. EMBO J 17, 7490–7497. https://doi.org/10.1093/emboj/17.24.7490
205. Parikka, K.J., Le Romancer, M., Wauters, N., Jacquet, S., 2017. Deciphering the virus-to-prokaryote ratio (VPR): insights into virus-host relationships in a variety of ecosystems. Biol Rev Camb Philos Soc 92, 1081–1100. https://doi.org/10.1111/brv.12271
206. Patterson, A.G., Jackson, S.A., Taylor, C., Evans, G.B., Salmond, G.P.C., Przybilski, R., Staals, R.H.J., Fineran, P.C., 2016. Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. Mol Cell 64, 1102–1108. https://doi.org/10.1016/j.molcel.2016.11.012
207. Payne, L.J., Hughes, T.C.D., Fineran, P.C., Jackson, S.A., 2024. New antiviral defences are genetically embedded within prokaryotic immune systems. https://doi.org/10.1101/2024.01.29.577857
208. Payne, L.J., Meaden, S., Mestre, M.R., Palmer, C., Toro, N., Fineran, P.C., Jackson, S.A., 2022. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. Nucleic Acids Res 50, W541–W550. https://doi.org/10.1093/nar/gkac400
209. Pecota, D.C., Wood, T.K., 1996. Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. J Bacteriol 178, 2044–2050. https://doi.org/10.1128/jb.178.7.2044-2050.1996
210. Pedersen, K., Zavialov, A.V., Pavlov, M.Y., Elf, J., Gerdes, K., Ehrenberg, M., 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell 112, 131–140.
211. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci 30, 70–82. https://doi.org/10.1002/pro.3943
212. Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R.A., Randau, L., Sørensen, S.J., Shah, S.A., 2020. Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. Nucleic Acids Res 48, 2000–2012. https://doi.org/10.1093/nar/gkz1197
213. Pizzolato-Cezar, L.R., Spira, B., Machini, M.T., 2023. Bacterial toxin-antitoxin systems: Novel insights on toxin activation across populations and experimental shortcomings. Curr Res Microb Sci 5, 100204. https://doi.org/10.1016/j.crmicr.2023.100204

214. Pourcel, C., Salvignol, G., Vergnaud, G., 2005. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology (Reading) 151, 653–663. https://doi.org/10.1099/mic.0.27437-0
215. Prysak, M.H., Mozdzierz, C.J., Cook, A.M., Zhu, L., Zhang, Y., Inouye, M., Woychik, N.A., 2009. Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. Mol. Microbiol. 71, 1071–1087. https://doi.org/10.1111/j.1365-2958.2008.06572.x
216. Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., Wagner, R., 2010. Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. Mol Microbiol 75, 1495–1512. https://doi.org/10.1111/j.1365-2958.2010.07073.x
217. 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
218. Pyenson, N.C., Gayvert, K., Varble, A., Elemento, O., Marraffini, L.A., 2017. Broad Targeting Specificity during Bacterial Type III CRISPR-Cas Immunity Constrains Viral Escape. Cell Host & Microbe 22, 343-353.e3. https://doi.org/10.1016/j.chom.2017.07.016
219. Rae, C.D., Gordiyenko, Y., Ramakrishnan, V., 2019. How a circularized tmRNA moves through the ribosome. Science 363, 740–744. https://doi.org/10.1126/science.aav9370
220. Richter, R., Rorbach, J., Pajak, A., Smith, P.M., Wessels, H.J., Huynen, M.A., Smeitink, J.A., Lightowlers, R.N., Chrzanowska-Lightowlers, Z.M., 2010. A functional peptidyl-tRNA hydrolase, ICT1, has been recruited into the human mitochondrial ribosome. EMBO J 29, 1116–1125. https://doi.org/10.1038/emboj.2010.14
221. Ries, J., Schwille, P., 2012. Fluorescence correlation spectroscopy. Bioessays 34, 361–368. https://doi.org/10.1002/bies.201100111
222. Rodnina, M.V., 2018. Translation in Prokaryotes. Cold Spring Harb Perspect Biol 10. https://doi.org/10.1101/cshperspect.a032664
223. Rodnina, M.V., Savelsbergh, A., Katunin, V.I., Wintermeyer, W., 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. Nature 385, 37–41. https://doi.org/10.1038/385037a0
224. Rostøl, J.T., Marraffini, L.A., 2019. Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. Nat Microbiol. https://doi.org/10.1038/s41564-018-0353-x
225. 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. https://doi.org/10.1038/s41586-021-03206-x

226. 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
227. Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.-L., Lieberman, E., Garry, D., Rocha, E.P.C., Bernheim, A., Bikard, D., 2022. Phages and their satellites encode hotspots of antiviral systems. Cell Host Microbe 30, 740-753.e5. https://doi.org/10.1016/j.chom.2022.02.018
228. Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., Marraffini, L.A., 2015. Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. Cell 161, 1164–1174. https://doi.org/10.1016/j.cell.2015.04.027
229. Samatova, E., Daberger, J., Liutkute, M., Rodnina, M.V., 2020. Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding. Front Microbiol 11, 619430. https://doi.org/10.3389/fmicb.2020.619430
230. 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
231. Schrimpf, W., Barth, A., Hendrix, J., Lamb, D.C., 2018. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophys J 114, 1518–1528. https://doi.org/10.1016/j.bpj.2018.02.035
232. Schrödinger, LLC, 2015. The PyMOL molecular graphics system, version 1.8.
233. Schureck, M.A., Repack, A., Miles, S.J., Marquez, J., Dunham, C.M., 2016. Mechanism of endonuclease cleavage by the HigB toxin. Nucleic Acids Res. 44, 7944–7953. https://doi.org/10.1093/nar/gkw598
234. Sheldrick, G.M., 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr D Biol Crystallogr 66, 479–485. https://doi.org/10.1107/S0907444909038360
235. Shimizu, Y., 2012. ArfA recruits RF2 into stalled ribosomes. J Mol Biol 423, 624–631. https://doi.org/10.1016/j.jmb.2012.08.007
236. Shmakov, S.A., Makarova, K.S., Wolf, Y.I., Severinov, K.V., Koonin, E.V., 2018. Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. Proc. Natl. Acad. Sci. U.S.A. 115, E5307–E5316. https://doi.org/10.1073/pnas.1803440115
237. Silas, S., Lucas-Elio, P., Jackson, S.A., Aroca-Crevillén, A., Hansen, L.L., Fineran, P.C., Fire, A.Z., Sánchez-Amat, A., 2017a. Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. Elife 6, e27601. https://doi.org/10.7554/eLife.27601
238. Silas, S., Makarova, K.S., Shmakov, S., Páez-Espino, D., Mohr, G., Liu, Y., Davison, M., Roux, S., Krishnamurthy, S.R., Fu, B.X.H., Hansen,

L.L., Wang, D., Sullivan, M.B., Millard, A., Clokie, M.R., Bhaya, D., Lambowitz, A.M., Kyrpides, N.C., Koonin, E.V., Fire, A.Z., 2017b. On the Origin of Reverse Transcriptase-Using CRISPR-Cas Systems and Their Hyperdiverse, Enigmatic Spacer Repertoires. mBio 8, e00897-17. https://doi.org/10.1128/mBio.00897-17

239. Silas, S., Mohr, G., Sidote, D.J., Markham, L.M., Sanchez-Amat, A., Bhaya, D., Lambowitz, A.M., Fire, A.Z., 2016. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. Science 351, aad4234. https://doi.org/10.1126/science.aad4234
240. Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., Zaher, H.S., 2014. An active role for the ribosome in determining the fate of oxidized mRNA. Cell Rep 9, 1256–1264. https://doi.org/10.1016/j.celrep.2014.10.042
241. Smalakyte, D., Kazlauskiene, M., F Havelund, J., Rukšėnaitė, A., Rimaite, A., Tamulaitiene, G., Færgeman, N.J., Tamulaitis, G., Siksnys, V., 2020. Type III-A CRISPR-associated protein Csm6 degrades cyclic hexa-adenylate activator using both CARF and HEPN domains. Nucleic Acids Res. https://doi.org/10.1093/nar/gkaa634
242. Smalakyte, D., Ruksenaite, A., Sasnauskas, G., Tamulaitiene, G., Tamulaitis, G., 2024. Filament formation activates protease and ring nuclease activities of CRISPR Lon-SAVED. Mol Cell S1097-2765(24)00735–4. https://doi.org/10.1016/j.molcel.2024.09.002
243. 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. https://doi.org/10.1016/j.molcel.2020.07.008
244. Song, S., Wood, T.K., 2020. A Primary Physiological Role of Toxin/Antitoxin Systems Is Phage Inhibition. Front Microbiol 11, 1895. https://doi.org/10.3389/fmicb.2020.01895
245. Sridhara, S., Rai, J., Whyms, C., Goswami, H., He, H., Woodside, W., Terns, M.P., Li, H., 2022. Structural and biochemical characterization of in vivo assembled Lactococcus lactis CRISPR-Csm complex. Commun Biol 5, 279. https://doi.org/10.1038/s42003-022-03187-1
246. Staals, R.H.J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D.W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J.J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P.J., Doudna, J.A., Heck, A.J.R., Yonekura, K., van der Oost, J., Shinkai, A., 2013. Structure and Activity of the RNA-Targeting Type III-B CRISPR-Cas Complex of Thermus thermophilus. Molecular Cell 52, 135–145. https://doi.org/10.1016/j.molcel.2013.09.013
247. 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
248. Stead, M.B., Agrawal, A., Bowden, K.E., Nasir, R., Mohanty, B.K., Meagher, R.B., Kushner, S.R., 2012. RNAsnap™: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res. 40, e156. https://doi.org/10.1093/nar/gks680
249. Steens, J.A., Zhu, Y., Taylor, D.W., Bravo, J.P.K., Prinsen, S.H.P., Schoen, C.D., Keijser, B.J.F., Ossendrijver, M., Hofstra, L.M., Brouns, S.J.J., Shinkai, A., van der Oost, J., Staals, R.H.J., 2021. SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation. Nat Commun 12, 5033. https://doi.org/10.1038/s41467-021-25337-5
250. Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S.J., Söding, J., 2019. HH-suite3 for fast remote homology detection and deep protein annotation. BMC Bioinformatics 20, 473. https://doi.org/10.1186/s12859-019-3019-7
251. Stella, G., Marraffini, L., 2023. Type III CRISPR-Cas: beyond the Cas10 effector complex. Trends Biochem Sci S0968-0004(23)00261-X. https://doi.org/10.1016/j.tibs.2023.10.006
252. Sternberg, S.H., Richter, H., Charpentier, E., Qimron, U., 2016. Adaptation in CRISPR-Cas Systems. Molecular Cell 61, 797–808. https://doi.org/10.1016/j.molcel.2016.01.030
253. Tak, U., Walth, P., Whiteley, A.T., 2023. Bacterial cGAS-like enzymes produce 2',3'-cGAMP to activate an ion channel that restricts phage replication. bioRxiv 2023.07.24.550367. https://doi.org/10.1101/2023.07.24.550367
254. 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. Molecular Cell 56, 506–517. https://doi.org/10.1016/j.molcel.2014.09.027
255. Tamulaitis, G., Venclovas, Č., Siksnys, V., 2017. Type III CRISPR-Cas Immunity: Major Differences Brushed Aside. Trends Microbiol. 25, 49–61. https://doi.org/10.1016/j.tim.2016.09.012
256. Tanzawa, T., Kato, K., Girodat, D., Ose, T., Kumakura, Y., Wieden, H.-J., Uchiumi, T., Tanaka, I., Yao, M., 2018. The C-terminal helix of ribosomal P stalk recognizes a hydrophobic groove of elongation factor 2 in a novel fashion. Nucleic Acids Res 46, 3232–3244. https://doi.org/10.1093/nar/gky115
257. Taylor, D.W., Zhu, Y., Staals, R.H.J., Kornfeld, J.E., Shinkai, A., van der Oost, J., Nogales, E., Doudna, J.A., 2015. Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. Science 348, 581–585. https://doi.org/10.1126/science.aaa4535

258. Tejada-Arranz, A., de Crécy-Lagard, V., de Reuse, H., 2020. Bacterial RNA Degradosomes: Molecular Machines under Tight Control. Trends Biochem Sci 45, 42–57. https://doi.org/10.1016/j.tibs.2019.10.002
259. Terwilliger, T.C., Sobolev, O.V., Afonine, P.V., Adams, P.D., 2018. Automated map sharpening by maximization of detail and connectivity. Acta Crystallogr D Struct Biol 74, 545–559. https://doi.org/10.1107/S2059798318004655
260. Tesson, F., Planel, R., Egorov, A., Georjon, H., Vaysset, H., Brancotte, B., Néron, B., Mordret, E., Atkinson, G., Bernheim, A., Cury, J., 2024. A Comprehensive Resource for Exploring Antiphage Defense: DefenseFinder Webservice, Wiki and Databases. https://doi.org/10.1101/2024.01.25.577194
261. Thomas, E.N., Kim, K.Q., McHugh, E.P., Marcinkiewicz, T., Zaher, H.S., 2020. Alkylative damage of mRNA leads to ribosome stalling and rescue by trans translation in bacteria. Elife 9, e61984. https://doi.org/10.7554/eLife.61984
262. UniProt Consortium, 2023. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res 51, D523–D531. https://doi.org/10.1093/nar/gkac1052
263. Urban, J.H., Vogel, J., 2007. Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res 35, 1018–1037. https://doi.org/10.1093/nar/gkl1040
264. 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 eabk2718. https://doi.org/10.1126/science.abk2718
265. van Beljouw, S.P.B., Sanders, J., Rodríguez-Molina, A., Brouns, S.J.J., 2022. RNA-targeting CRISPR-Cas systems. Nat Rev Microbiol. https://doi.org/10.1038/s41579-022-00793-y
266. Van Melderen, L., Bernard, P., Couturier, M., 1994. Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. Mol Microbiol 11, 1151–1157. https://doi.org/10.1111/j.1365-2958.1994.tb00391.x
267. Vassallo, C.N., Doering, C.R., Littlehale, M.L., Teodoro, G.I.C., Laub, M.T., 2022. A functional selection reveals previously undetected antiphage defence systems in the E. coli pangenome. Nat Microbiol. https://doi.org/10.1038/s41564-022-01219-4
268. Venclovas, Č., 2016. Structure of Csm2 elucidates the relationship between small subunits of CRISPR-Cas effector complexes. FEBS Letters 590, 1521–1529. https://doi.org/10.1002/1873-3468.12179
269. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ.,

Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors, 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

270. 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
271. Wang, J.Y., Pausch, P., Doudna, J.A., 2022. Structural biology of CRISPR-Cas immunity and genome editing enzymes. Nat Rev Microbiol. https://doi.org/10.1038/s41579-022-00739-4
272. Wang, L., Mo, C.Y., Wasserman, M.R., Rostol, J.T., Marraffini, L.A., Liu, S., 2018. Dynamics of Cas10 Govern Discrimination between Self and Nonself in Type III CRISPR-Cas Immunity. https://doi.org/10.1101/369744
273. Watson, B.N.J., Steens, J.A., Staals, R.H.J., Westra, E.R., van Houte, S., 2021. Coevolution between bacterial CRISPR-Cas systems and their bacteriophages. Cell Host Microbe 29, 715–725. https://doi.org/10.1016/j.chom.2021.03.018
274. Watson, Z.L., Ward, F.R., Méheust, R., Ad, O., Schepartz, A., Banfield, J.F., Cate, J.H., 2020. Structure of the bacterial ribosome at 2 Å resolution. Elife 9, e60482. https://doi.org/10.7554/eLife.60482
275. 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
276. Wiegand, T., Wilkinson, R., Santiago-Frangos, A., Lynes, M., Hatzenpichler, R., Wiedenheft, B., 2023. Functional and Phylogenetic Diversity of Cas10 Proteins. CRISPR J 6, 152–162. https://doi.org/10.1089/crispr.2022.0085
277. Wimmer, F., Beisel, C.L., 2019. CRISPR-Cas Systems and the Paradox of Self-Targeting Spacers. Front Microbiol 10, 3078. https://doi.org/10.3389/fmicb.2019.03078
278. Xia, P., Dutta, A., Gupta, K., Batish, M., Parashar, V., 2022. Structural basis of cyclic oligoadenylate binding to the transcription factor Csa3 outlines crosstalk between Type-III & Type-I CRISPR systems. J Biol Chem 101591. https://doi.org/10.1016/j.jbc.2022.101591
279. 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. https://doi.org/10.1038/s41586-024-07815-0

280. Yosef, I., Goren, M.G., Qimron, U., 2012. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Research 40, 5569–5576. https://doi.org/10.1093/nar/gks216
281. Yoshioka, K., 2002. KyPlot — A User-oriented Tool for Statistical Data Analysis and Visualization. Computational Statistics 17, 425–437. https://doi.org/10.1007/s001800200117
282. 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
283. Zhang, Hongtai, Shi, M., Ma, X., Liu, M., Wang, N., Lu, Q., Li, Z., Zhao, Y., Zhao, H., Chen, H., Zhang, Huizhi, Jiang, T., Ouyang, S., Huo, Y., Bi, L., 2024. Type-III-A structure of mycobacteria CRISPR-Csm complexes involving atypical crRNAs. Int J Biol Macromol 129331. https://doi.org/10.1016/j.ijbiomac.2024.129331
284. Zhang, T., Tamman, H., Coppieters 't Wallant, K., Kurata, T., LeRoux, M., Srikant, S., Brodiazhenko, T., Cepauskas, A., Talavera, A., Martens, C., Atkinson, G.C., Hauryliuk, V., Garcia-Pino, A., Laub, M.T., 2022. Direct activation of a bacterial innate immune system by a viral capsid protein. Nature 1–9. https://doi.org/10.1038/s41586-022-05444-z
285. Zhang, Y., Inouye, M., 2009. The inhibitory mechanism of protein synthesis by YoeB, an Escherichia coli toxin. J Biol Chem 284, 6627–6638. https://doi.org/10.1074/jbc.M808779200
286. 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 Research. https://doi.org/10.1093/nar/gkab073
287. Zhu, X., Ye, K., 2012a. Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. FEBS Lett 586, 939–945. https://doi.org/10.1016/j.febslet.2012.02.036
288. Zhu, X., Ye, K., 2012b. Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. FEBS Letters 586, 939–945. https://doi.org/10.1016/j.febslet.2012.02.036
289. 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

APPENDICES

Appendix 1. Plasmids used in this work.

Name	Description	Source	Link to sequence
pET-Csm2-tag	Encodes *S. thermophilus* DGCC8004 Csm2 protein with fused StrepII affinity tag to N-terminus downstream of P_{T7}. Based on pETDuet-1 vector (Novagen).	(Tamulaitis et al., 2014)	https://benchling.com/s/seq-bh4qYLl21L0RpYWuBqTK
pCas/Csm_D33A	Encodes *S. thermophilus* DGCC8004 type III-A CRISPR-Cas protein locus (*cas6-cas10-csm2-csm3-csm4-csm5-csm6-csm6'*) downstream of P_{T7}. Csm3 encoding gene carries a D33A mutation. Based on pCDFDuet-1 vector (Novagen).	(Tamulaitis et al., 2014)	https://benchling.com/s/seq-TbuNIr6ws04hJkHrKht0
pBAD/HisA	Empty pBAD/HisA vector.	Thermo Fisher Scientific	-
StCsm complex components for FCS and activity assays			
pCsm2-TEV-tag	Encodes *S. thermophilus* DGCC8004 Csm2 protein with fused StrepII affinity tag and TEV protease recognition sequence to N-terminus. Prepared from pET-Csm2-tag as described in Methods 2.2.2.	This work	https://benchling.com/s/seq-t6LB9L8Qx1XUstqUX862
pCRISPR_S3	Encodes a synthetic CRISPR locus consisting of five *S. thermophilus* DGCC8004 CRISPR2 region repeats interspaced by four identical copies of S3 spacer of the same region downstream of P_{T7}. Based on pACYCDuet-1 vector (Novagen).	(Tamulaitis et al., 2014)	https://benchling.com/s/seq-ImgWSnyAdXKnwUROnIpj

pStCsm	Encodes *S. thermophilus* DGCC8004 Cas6, Cas10, Csm2, Csm3, Csm4, and Csm5 proteins downstream of P_{T7}. Based on pCDFDuet-1 vector (Novagen). Named as pCas/Csm_ΔCsm6'ΔCsm6 in source publication.	(Mogila et al., 2019)	https://benchling.com/s/seq-NEdoGMvJTB3n6ut9Al1L
pStCsm_dRNase	Similar to pStCsm with D33A substitution in Csm3 gene. Prepared from pCas/Csm_D33A and pStCsm using Gibson assembly as described in Methods 2.2.2.	This work	-
	Matrices of RNA substrates for StCsm		
pUC18_S3/2	Encodes S3/a RNA substrate in pUC18 vector.	(Tamulaitis et al., 2014)	https://benchling.com/s/seq-TLXrrcseI7xdcmAcNWHv
pUC18_S3/n	Encodes S3/n RNA substrate in pUC18 vector. Prepared from pUC18_S3/2 using Gibson assembly as described in Methods 2.2.2.	This work	https://benchling.com/s/seq-hg2m4sqIVtcSDceXIYZX
pUC18_NS	Encodes NS RNA substrate in pUC18 vector. Prepared from pUC18_S3/2 using Gibson assembly as described in Methods 2.2.2.	This work	https://benchling.com/s/seq-msWW4nBIeoN9SCxYVBS7
	WT Cami1 proteins		
pAvCami1	*A. vinosum* DSM180 Cami1 protein (NCBI protein reference sequence WP_012972290.1) with fused C-terminal StrepII and His_{10} affinity tags encoded downstream of P_{BAD}. Synthesized by Twist Bioscience.	This work	https://benchling.com/s/seq-oHpybdGJ3r9CRBwFndkB
pCaCami1	*C. aerophila* DSM14535 Cami1 protein (WP_044276641.1) with fused C-terminal StrepII and His_{10} affinity tags encoded downstream of P_{BAD}. Synthesized by Twist Bioscience.	This work	https://benchling.com/s/seq-9d9cQhNiAUBYyw5UEIAX

pCCaCami1	*Candidatus* C. acidaminovorans str. Evry Cami1 protein (WP_015424585.1) with fused C-terminal StrepII and His_{10} affinity tags encoded downstream of P_{BAD}. Synthesized by Twist Bioscience.	This work	https://benchling.com/s/seq-3aPoiKSOAGW0gnMrEVPf
pChCami1	*Caldicellulosiruptor hydrothermalis* 108 Cami1 protein (WP_013404376.1) with fused C-terminal StrepII and His_{10} affinity tags encoded downstream of P_{BAD}. Synthesized by Twist Bioscience.	This work	https://benchling.com/s/seq-0B9syIKHfLP9v6D39PuV
	AvCami1 CARF domain mutants		
pAvCami1-S11A	Similar to pAvCami1 with introduced S11A mutation. Prepared from pAvCami1. Preparation of this and subsequent mutants is described in Methods 2.2.2.	This work	-
pAvCami1-K157A	Similar to pAvCami1 with introduced K157A mutation. Prepared from pAvCami1.	This work	-
	AvCami1 RelE domain mutants		
pAvCami1-H306A	Similar to pAvCami1 with introduced H306A mutation. Prepared from pAvCami1.	This work	-
pAvCami1-K317A	Similar to pAvCami1 with introduced K317A mutation. Prepared from pAvCami1.	This work	-
pAvCami1-R325A	Similar to pAvCami1 with introduced R325A mutation. Prepared from pAvCami1.	This work	-
pAvCami1-H343A	Similar to pAvCami1 with introduced H343A mutation. Prepared from pAvCami1.	This work	-
	AvCami1 CARF and RelE domain mutants		

pAvCami1-S11A+K317A	Similar to pAvCami1 with introduced S11A and K317A mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+H343A	Similar to pAvCami1 with introduced S11A and H343A mutations. Prepared from pAvCami1-S11A.	This work	-
AvCami1 mutants and bL12 for wHTH-bL12 interaction studies			
pAvCami1-S11A+K134A	Similar to pAvCami1 with introduced S11A and K134A mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+Q210A+E212A	Similar to pAvCami1 with introduced S11A, Q210A and E212A mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+Q210R+E212R	Similar to pAvCami1 with introduced S11A, Q210R and E212R mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+T213R	Similar to pAvCami1 with introduced S11A, T213R mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+E238A	Similar to pAvCami1 with introduced S11A, E238A mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+F252A	Similar to pAvCami1 with introduced S11A, F252A mutations. Prepared from pAvCami1-S11A.	This work	-
pAvCami1-S11A+Q210A+E212A+E238A	Similar to pAvCami1 with introduced S11A, Q210A, E212A and E238A mutations. Prepared from pAvCami1-S11A+Q210A+E212A.	This work	-
pSV281-L12	Encodes *E. coli* bL12 protein with N-terminal His_6 affinity tag and TEV protease cleavage site downstream of P_{T7}.	(Carlson et al., 2017)	https://benchling.com/s/seq-9hLtN6IElhqRp4rxIQsy
StCsm components and target for Cami1 in vivo toxicity assays			
pStCsm_dNucleases	Similar to pStCsm with two single-residue substitutions: D33A in *csm3* gene and D16A in *cas10* gene.	(Smalakyte et al., 2020)	https://benchling.com/s/seq-rp2i46L3UoH8QsKcZLTv

	Named as pCas/Csm_dCsm3_dHDCas10_ΔCsm6′ΔCsm6 in source publication.		
pCRISPR	Similar to pCRISPR_S3, encodes four identical spacers complementary to mRNA of EGFP gene downstream of T7 promoter. Based on pACYCDuet-1 vector (Novagen). Named as pCRISPR_EGFP in source publication.	(Fricke et al., 2020)	https://benchling.com/s/seq-2Hj9qTTlJU6iLuBg8zUJ
pTarget	Encodes EGFP protein downstream of T7 promoter. Based on pCOLADuet-1 vector (Novagen). Named as pTarget_GFP in source publication.	(Smalakyte et al., 2020)	https://benchling.com/s/seq-kBJLZeHkXQmsBF218Uyi

Appendix 2. Substrates and probes used in this study.

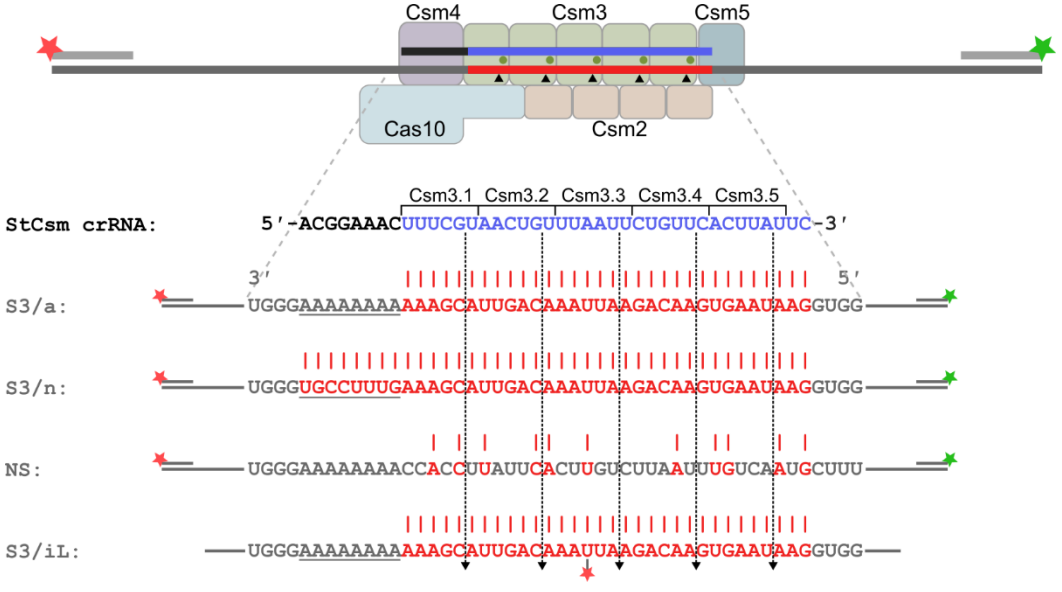
Name	Type	Sequence (5′ to 3′) with modifications	Description
		Substrates for StCsm binding and activity assays[1]	
S3/a	RNA	GGGAAACGACGGCCAGUGCCAAGCUUGCAUGCCUGCAGGUCGACUCUAGA GGAUCCCCAAAUAUAAGGUGGAAUAAGUGAACAGAAUUAAACAGUUACGA AAAAAAAAAAGGGUACCGAGCUCGAAUUCGUAAUCAUGGUCAUAGCUGUU UCCUGUGUGAAAUUGUUAUCCGCUCACAAU	Activating RNA substrate, complementary to spacer and non-complementary to 5′-tag of S3 crRNA bound by StCsm.
S3/n	RNA	GGGAAACGACGGCCAGUGCCAAGCUUGCAUGCCUGCAGGUCGACUCUAGA GGAUCCCCAAAUAUAAGGUGGAAUAAGUGAACAGAAUUAAACAGUUACGA AAGUUUCCGUGGGUACCGAGCUCGAAUUCGUAAUCAUGGUCAUAGCUGUU UCCUGUGUGAAAUUGUUAUCCGCUCACAAU	Non-activating RNA substrate, complementary to both spacer and 5′-tag of S3 crRNA.
NS	RNA	GGGAAACGACGGCCAGUGCCAAGCUUGCAUGCCUGCAGGUCGACUCUAGA GGAUCCCCAAAUAUAAUUUCGUAACUGUUUAAUUCUGUUCACUUAUUCCA CCAAAAAAAAGGGUACCGAGCUCGAAUUCGUAAUCAUGGUCAUAGCUGUU UCCUGUGUGAAAUUGUUAUCCGCUCACAAU	RNA substrate which is non-complementary to S3 crRNA.
S3/iL	RNA	GGGAUCCCCAAAUAUAAGGUGGAAUAAGUGAACAGAAU**[U-ATTO647N]** AAACAGUUACGAAAAAAAAAAGGGUACC	Internally labelled activating RNA substrate, complementary to spacer and non-complementary to 5′-tag of S3 crRNA
FQ-DNA	DNA	**[6FAM]**-TTAGCT-**[BHQ1]**	DNA substrate for StCsm DNase assays labelled with 6FAM fluorophore and BHQ1 quencher.
		Probes for FCS	
Green-probe	DNA	ATGCAAGCTTGGCACTGGCCGTCGTTTCCC-**[ATTO532]**	ATTO532-labelled probe complementary to 5′-end of S3/a, S3/n and NS substrates.

Red-probe	DNA	**[ATTO647N]-**ATTGTGAGCGGATAACAATTTCACACAGGA	ATTO647N-labelled probe complementary to 3′-end of S3/a, S3/n and NS substrates.
Substrate for* Candidatus *C. acidaminovorans Csm complex activation			
CCa target RNA	RNA	GGGUGAAGAGCAAUGAGCUCUCGAGGUGCGAUAUCGCUCUUCCCAGUGUA	Used to induce cA4 synthesis from ATP in CCaCsm. Substrate was obtained from D. Smalakytė (Smalakyte et al., 2024)
***Substrates for Cami1 activity assays*[2]**			
rib-UAG	RNA	GGGCAAGGAGGUAAAAAUGUAGAAAAACAAU	RNA substrate for *in vitro* mRNA interferase assays. Sequence was obtained from (Griffin et al., 2013).
rib-UAG-OMe	RNA	GGGCAAGGAGGUAAAAAUG**[OMe-U][OMe-A][OMe-G]**AAAAACAAU	RNA substrate corresponding to rib-UAG with 2′-O-methylated (OMe) UAG motif used as a control for mRNA interference assays and cryoEM.
rib-NNN	RNA	GGGCAAGGAGGUAAAAAUGNNNAAAAACAAU	RNA substrate corresponding to rib-UAG with NNN motif instead of UAG used in mRNA interference assays.
Northern blot probes for E. coli endogenous RNA			
DNA-ompA	DNA	ACACCCAGGCTCAGCATGCCGTTGTCC	Probe for ompA mRNA. Sequence was obtained from (Hurley et al., 2011).
DNA-ompF	DNA	AAACCAAGACGGGCATAGGTC	Probe for ompF mRNA (Hurley et al., 2011).
DNA-lpp	DNA	TTACTTGCGGTATTTAGTAGCC	Probe for lpp mRNA (Hurley et al., 2011).

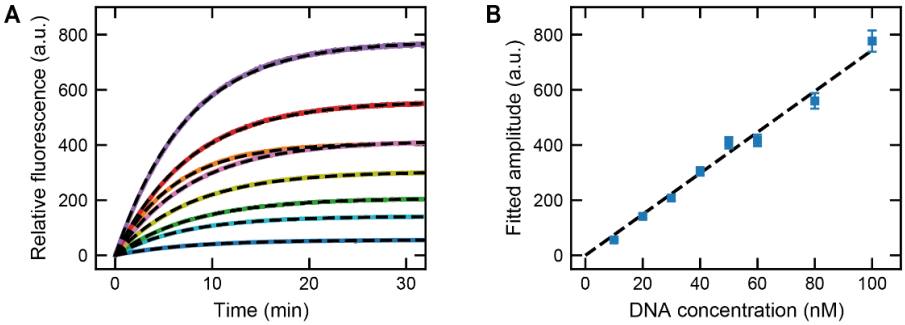
DNA-tufA	DNA	TGAGAAGTGTTGATGGTGATACC	Probe for tufA mRNA (Hurley et al., 2011).
DNA-tm	DNA	GCCCCTCGGCATGCACC	Probe for tmRNA (Christensen and Gerdes, 2003).
DNA-5S	DNA	CTACGGCGTTTCACTTCTGAGTTC	Probe for 5S rRNA (Urban and Vogel, 2007).

[1] Blue denotes sequence complementary to StCsm S3 crRNA, green and red indicate hybridization sites for "Green-probe" and "Red-probe" DNA oligonucleotides, respectively. Underlined sequence corresponds to the portion of substrate aligning with 5′-tag of crRNA upon substrate binding to StCsm complex. Modifications are **bolded**. See also Appendix 3.
[2] Blue denotes codons exposed in the A-site of *E. coli* 70S ribosomes programmed by $tRNA^{fMet}$ binding to AUG codon (underlined) of the substrate. Modifications are **bolded**.

Appendix 3. Scheme of substrates used in FCS assays. Full RNA and probe DNA sequences are listed in Appendix 2. StCsm cleavage positions are indicated by dashed arrows.



Appendix 4. FQ-DNA cleavage by micrococcal nuclease used to relate fluorescence intensities to cleaved single-stranded DNA. (**A**) Measured fluorescence traces (solid lines) together with an exponential fit (dashed lines) for increasing concentrations of FQ-DNA. (**B**) Fitted amplitudes (solid scatter points) obtained in (A) in dependence of the FQ-DNA concentration together with a linear fit (dashed line). The linear fit provided a linear scaling factor of 7.44 fluorescence units = 1 nM cleaved DNA (linked to Fig. 3.4).



Appendix 5. Crystallographic data collection and refinement statistics.

Data Collection	**Native**	**Se-Met**
Wavelength	0.9686	0.9739
Resolution range, Å	64.71-1.70 (1.79-1.70)	45.75-1.80 (1.90-1.80)
Space group	P 2_1	P 2_1
Unit cell	65.238 90.428 65.616 90.000 99.555 90.000	65.650 91.497 65.878 90.000 99.719 90.000
Total / unique reflections	578268 (83153)/ 81871 (11877)	942970 (134944)/ 70781 (10053)
Multiplicity	7.1 (7.0)	13.3 (13.4)
Completeness (%)	99.3 (99.0)	99.6 (97.5)
Mean I / sigma(I)	17.7 (2.5)	18.4 (3.6)
Wilson B-factor	30.1	31.1
R_{merge}[a]	0.042 (0.669)	0.070 (0.766)
$CC_{1/2}$[b]	0.999 (0.875)	0.999 (0.918)
Refinement		
Reflections used in refinement (total / R-free)	161196/16039	
R_{work}[c]/ R_{free}[d], %	18.58/21.47	
Number of non-hydrogen atoms (macromolecules / ligands / solvent)	5989/0/367	
RMS (bonds / angles)	0.007/0.820	
Ramachandran: favoured / allowed / outliers (%)	98.23/1.50/0.27	
Rotamer outliers (%)	1.50	
Clashscore	2.71	
Average B-factor (protein / ligand / solvent)	35.48/-/37.86	

Data collection statistics for the highest-resolution shell are shown in parentheses.

[a] $R_{merge} = \frac{\sum_{hkl}\sum_{i=1}^{n}|I_i(hkl)-\langle I(hkl)\rangle|}{\sum_{hkl}\sum_{i=1}^{n} I_i(hkl)}$, where $I_i(hkl)$ is the intensity of the measured reflection (hkl) and n denotes multiplicity.

[b] $CC_{1/2}$ is the correlation coefficient of the half datasets.

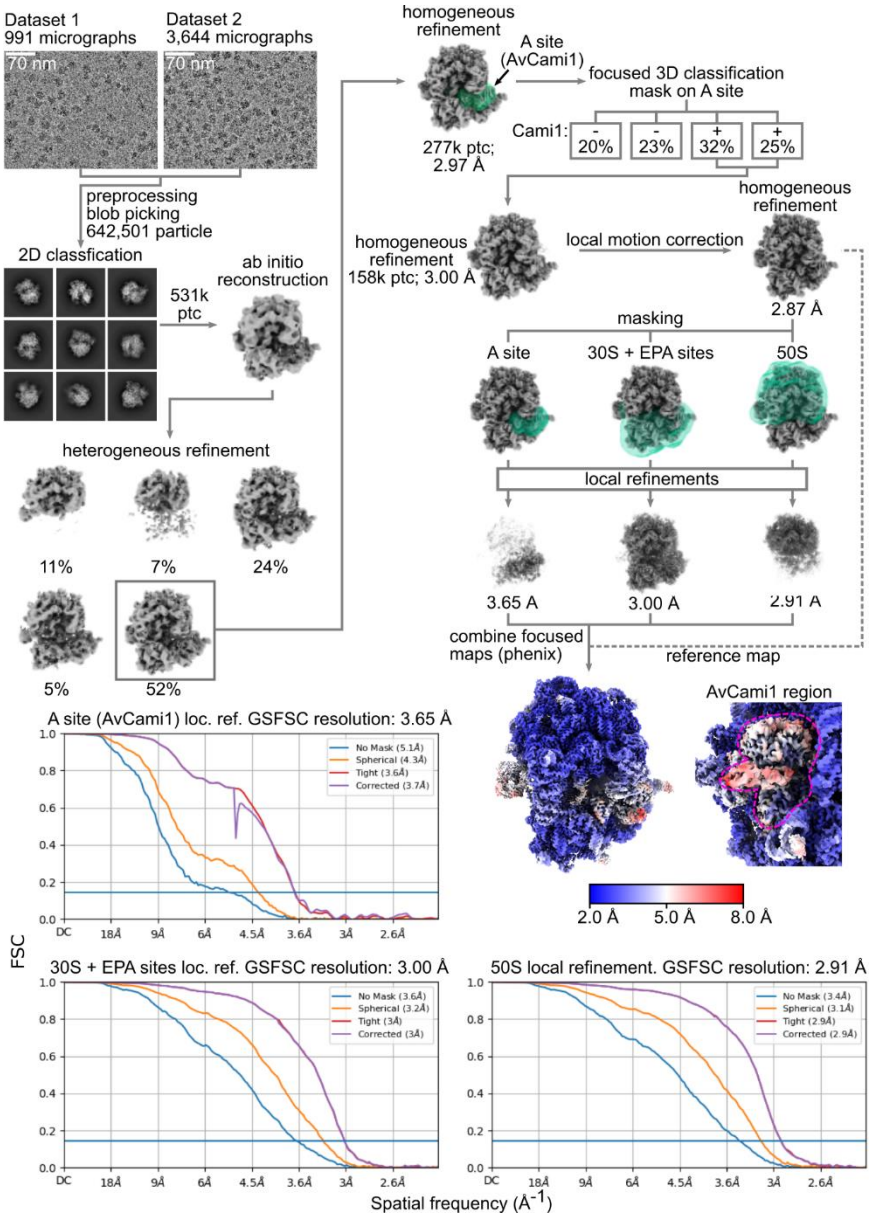
[c] $R_{work} = \frac{\sum||F_{obs}|-|F_{calc}||}{\sum|F_{obs}|}$, where F_{obs} and F_{calc} are observed and calculated structure factors.

[d] $R_{free} = \frac{\sum T||F_{obs}|-|F_{calc}||}{\sum T|F_{obs}|}$, where T is a test dataset of ~10% of the total reflections randomly chosen and set aside prior to refinement.

Appendix 6. Table of cA_4 hydrolysis (by AvCami1) product m/z ratios obtained by MS.

retention time of the peak		17.6'		19.0'		21.0'	
compound		A_2>p		cA_4		A_4>p	
ion		$(M-H)^-$	$(M-2H)^{-2}$	$(M-H)^-$	$(M-2H)^{-2}$	$(M-H)^-$	$(M-2H)^{-2}$
m/z	theoretical	657.0978	328.0452	1315.2028	657.0978	1315.2028	657.0978
	cA_4 only	n/d	n/d	1315.1956	657.0959	n/d	657.0956
	cA_4 + AvCami1	657.0956	328.0453	n/d	657.0976	1315.1942	657.0956
	cA_4 + CaCami1	n/d	n/d	1315.1975	657.0971	1315.1952	657.0969
	cA_4 + CCaCami1	657.0971	328.0460	n/d	657.0969	n/d	657.0966
	cA_4 + ChCami1	657.0971	328.0461	n/d	657.0903	1315.1980	657.0973

Appendix 7. Scheme of cryo-EM image processing.



Appendix 8. Cryo-EM data collection, refinement and validation statistics.

cA$_4$-bound AvCami1 in complex with 70S ribosome (PDB 8PHJ, EMD-1767)

Data collection and processing			
Magnification	92000	Axial symmetry	C1
Voltage (kV)	200	Initial particle images (no.)	642,401
Electron exposure (e–/Å^2)	30.58	Final particle images (no.)	158,387
Defocus range (μm)	-1.0/-2.0	Sphericity	0.982
Pixel size (Å)	1.12	Composite map resolution (Å) FSC threshold	3.4 0.143
Refinement			
Initial model used (PDB code)	7K00, 8PHB	R.m.s. deviations Bond lengths (Å) Bond angles (°)	 0.005 0.496
Model resolution (d_{FSC_model}) (Å) FSC threshold	3.0 0.143	Validation MolProbity score Clashscore Poor rotamers (%)	 1.62 3.58 2.16
Map sharpening *B* factor (Å^2)	-113.8	CaBLAM outliers (%) Cβ outliers (%) Peptide plane (%) Cis proline/general Twisted proline/general	0 0 3.0/0 0/0
Model composition Chains Non-hydrogen atoms Protein residues Nucleotides Ligands (ZN, MG) Water	 61 152472 6749 4642 243 0	Ramachandran plot Favoured (%) Allowed (%) Disallowed (%)	 96.62 3.37 0.02
B factors (Å^2) Protein Nucleotide Ligands	min/max/mean 3.48/82.89/35.10 3.15/138.82/38.06 1.84/102.25/20.65		

LIST OF PUBLICATIONS

This thesis is based on two original publications listed below with my contributions outlined.

1. Patrick Irmisch, Irmantas Mogila, Brighton Samatanga, Gintautas Tamulaitis, and Ralf Seidel. Retention of the RNA Ends Provides the Molecular Memory for Maintaining the Activation of the Csm Complex. *Nucleic Acids Research* 52, no. 7 (April 24, 2024): 3896–3910. doi: 10.1093/nar/gkae080.

My contributions: purified StCsm complexes together with Konstanty Keda and designed crRNA length-based selection protocol, prepared the RNA substrates and tested their affinity to StCsm, performed the unpublished groundwork testing StCsm activity towards hybrid RNA:DNA substrates in bulk and upon immobilization, performed initial FCS measurements together with B. S., contributed to the writing of the manuscript and addressing reviewer comments.

2. Irmantas Mogila*, Giedre Tamulaitiene*, Konstanty Keda, Albertas Timinskas, Audrone Ruksenaite, Giedrius Sasnauskas, Česlovas Venclovas, Virginijus Siksnys, Gintautas Tamulaitis. Ribosomal Stalk-Captured CARF-RelE Ribonuclease Inhibits Translation Following CRISPR Signaling. *Science* 382, no. 6674 (December 2023): 1036–41. doi: 10.1126/science.adj2107.

My contributions: participated in study design and conceptualisation, performed mutagenesis (with K. K.), purified Cami1 and bL12 proteins (with K. K.), designed and performed most of the biochemical assays (in vitro cOA and RNA cleavage, BLI) and all of the in vivo experiments, prepared cryoEM samples, collected and analysed cryoEM data, contributed to the writing of the manuscript and addressing reviewer comments.

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

CONFERENCE PRESENTATIONS

1. Irmantas Mogila, Konstanty Keda, Giedrė Tamulaitienė, Albertas Timinskas, Česlovas Venclovas, Gintautas Tamulaitis, Virginijus Šikšnys. Characterization of CARF-RelE proteins associated with Type III CRISPR-Cas systems. EMBO|EMBL Symposium: The Complex Life of RNA 2020. Heidelberg, Germany (virtual event). 2020 10 07-09. Poster presentation.

2. Irmantas Mogila, Giedre Tamulaitiene, Konstanty Keda, Albertas Timinskas, Česlovas Venclovas, Gintautas Tamulaitis, Virginijus Siksnys. Structural studies of CRISPR-Cas related toxin. CRISPR 2022. Boston, Massachusetts, USA. 2022 06 12-16. Poster presentation.

3. Irmantas Mogila, Giedre Tamulaitiene, Konstanty Keda, Albertas Timinskas, Audrone Ruksenaite, Giedrius Sasnauskas, Česlovas Venclovas, Virginijus Siksnys, Gintautas Tamulaitis. Translation inhibiting effector in type III CIRSPR-Cas defense. CRISPR 2023. Würzburg, Germany. 2023 06 27 – 07 01. Poster presentation.

4. Irmantas Mogila, Giedrė Tamulaitienė, Konstanty Keda, Albertas Timinskas, Audronė Rukšėnaitė, Giedrius Sasnauskas, Česlovas Venclovas, Virginijus Šikšnys, Gintautas Tamulaitis. Transliacijos slopinimas priešvirusinėje CRISPR-Cas gynyboje (Translation inhibition in antiviral CRISPR-Cas defence). Bioateitis 2023. Vilnius, Lithuania. 2023 11 23. Oral presentation. Award for the best presentation (2nd place).

5. Irmantas Mogila. Molecular memory prolongs CRISPR-Cas complex activity. Baltic Photonics 2024. Vilnius, Lithuania. 2024 10 14 – 18. Invited oral presentation.

All listed presentations were delivered by the author.

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisor dr. Gintautas Tamulaitis for introducing me with the field of academic research, helping me to grow as a scientist and giving independence to tackle scientific questions my way. Your openness when discussing arising matters, study design and future projects gave me the sense of mutual companionship allowing to feel both supported and acknowledged.

A special thanks goes to prof. dr. Virginijus Šikšnys for accepting me as an undergraduate freshman to the scientific group back in 2012. Your contributions to bacterial immunity research field, novel ideas during discussions and vast array of interests kept me impressed during these years.

I was very fortunate to be mentored by dr. Miglė Kazlauskienė during my first years in a scientific laboratory. Your attentive supervision, effective teaching and critical evaluation put me on the right track of improvement as a researcher. Rigorous and reflective yet creative and warm-hearted you inspire me both professionally and personally. I thank you deeply.

I am grateful to students Konstanty Keda and Germantas Pašilis for the motivation and passion for our research. Seeing your willingness to learn and the resulting improvement, while fruitfully challenging, gives me optimism and helps to realize that academic impact is not only about the science being published, but also the knowledge being passed on directly.

As for publishing, this thesis would not be defensible without the two required associated publications (thank you, Lithuanian doctoral system!), which, in turn, would not have been possible without the input from co-authors. I thank our collaborators prof. dr. Ralf Seidel and dr. Brighton Samatanga for accepting and supervising me during a visit to their laboratory in Leipzig University and helping to lay grounds for StCsm regulation studies. I am grateful to dr. Patrick Irmisch for resiliently pushing this project to the happy end. I thank my academic consultant dr. Giedrė Tamulaitienė for her work solving the structures of Cami1 and, in general, for expressing the beauty of nature in angstrom resolution. I thank dr. Giedrius Sasnauskas for countless advice regarding laboratory equipment including the invaluable help with cryoEM. I thank dr. Albertas Timinskas and prof. dr. Česlovas Venclovas for the computational analysis of CARF proteins and discussions regarding sequence search and protein modelling from which I learned a lot. Again, I thank Konstanty Keda for the help with protein purification and crystallization and Cami1 mutant activity assays. I am grateful to Audronė Rukšėnaitė for performing the mass spectrometry analysis of our never-ending samples. I thank

prof. dr. P. Clint Spiegel of Western Washington University for a kind gift of bL12 expression plasmid and bL12-deficient *E. coli* ribosomes.

My gratitude extends to everyone in the Department of Protein-DNA Interactions. Your advice and insights helped me day-to-day to solve the scientific problems I faced. I am especially thankful to dr. Arūnas Šilanskas for guiding through protein purification techniques and taming the chromatography systems. The whole group has proven their synergy not only at the lab bench, but also during lab retreats which made many unforgettable experiences. Thank you!

I was lucky to overlap with brilliant (former or soon-to-be former) doctoral students: Jonas Juozapaitis, dr. Greta Bigelytė, dr. Edvardas Golovinas, dr. Andrius Merkys, Antanas Vaitkus, Algirdas Grybauskas and others. Thank you for sharing laughter or sorrow over lunch, for the serious discussions and geeky parties. I express exclusive gratitude to my academic sibling Dalia Smalakytė for the mutual understanding, respect and support throughout all these years.

I am thankful to Monika Šimoliūnienė, Juta Rainytė, dr. Daina Bujanauskienė, Dalia Smalakytė and Jonas Juozapaitis for our endeavour enabling doctoral student representation at Life Sciences Center. Seeing that the representation has gained momentum, I can conclude that the time spent preparing to and attending various faculty meetings was meaningful and worth the effort. I thank the head of the Doctoral school dr. Daiva Baltriukienė for welcoming us as representatives.

I thank to dr. Patrick Pausch and dr. Saulius Serva for the thorough review of this thesis and suggestions improving the text and enhancing presentation of the findings.

My deepest gratitude goes to my family for the immeasurable support, encouragement, and most importantly, love. I thank my dearest Gintarė for being there when I needed, for the tolerance of me working into late evenings or weekends and for the patience listening to my lengthy explanations of the contents of this thesis. I thank our children for helping me realize what is truly important in life.

ABOUT THE AUTHOR

Irmantas Mogila (b. 1993) began his academic research career as an undergraduate freshman. By studying CRISPR-Cas defence systems he accumulated knowledge of various biochemical, biophysical, structural and computational methods enabling him to decipher their functions and mechanisms. His research was published in well regarded scientific journals and presented in conferences attended by leading experts in the bacterial defence field.

Irmantas earned his bachelor's degree with *cum laude* honours (2012-2016) and master's degree with *magna cum laude* honours (2016-2018) in biochemistry, both from Vilnius University. During doctoral studies he received a scholarship for academic accomplishments by Research Council of Lithuania. Irmantas' doctoral research has been recognized with the International Birnstiel Award, awarded by the Research Institute of Molecular Pathology (IMP) and Max Birnstiel Foundation in Vienna, Austria.

NOTES

NOTES

NOTES

NOTES

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.