https://doi.org/10.15388/vu.thesis.247
https://orcid.org/0000-0003-2188-0556

VILNIUS UNIVERSITY

Inga

SONGAILIENĖ

Bacterial defense islands: CRISPR-Cas and toxin-antitoxin systems

DOCTORAL DISSERTATION

Natural Sciences,
Biochemistry (N 004)

VILNIUS 2021

The work presented in this doctoral dissertation has been carried out at the Institute of Biotechnology, Life Sciences Centre of Vilnius University from 2013 to 2017.
The research was supported by Research Council of Lithuania.
The dissertation is defended on an external basis.

Scientific Advisor – **Prof. dr. Virginijus Šikšnys** (Vilnius University, Natural Sciences, Biochemistry – N 004).

Chairman – **Prof. dr. Edita Sužiedėlienė** (Vilnius University, Natural Sciences, Biochemistry – N 004).
Members:
Dr. Julija Armalytė (Vilnius University, Natural Sciences, Biochemistry – N 004);
Prof. habil. dr. Rimantas Daugelavičius (Vytautas Magnus University, Natural Sciences, Biochemistry – N 004);
Dr. Dukas Jurėnas (Centre National de la Recherche Scientifique (CNRS)/Aix-Marseille Université (AMU), Natural Sciences, Biochemistry – N 004);
Prof. dr. Rolandas Meškys (Vilnius University, Natural Sciences, Biochemistry – N 004).

The dissertation shall be defended at a public meeting of the Dissertation Defense Panel at 10:00 h on November 15, 2021 in a meeting room R401 of the Vilnius University Life Sciences Center.
Address: Saulėtekis ave. 7, R401, Vilnius, Lithuania
Tel. +370 68602086; e-mail: inga.songailiene@bti.vu.lt
The text of this dissertation can be accessed at the library of Vilnius university as well as on the website of Vilnius University:
www.vu.lt/lt/naujienos/ivykiu-kalendorius

https://doi.org/10.15388/vu.thesis.247
https://orcid.org/0000-0003-2188-0556

VILNIAUS UNIVERSITETAS

Inga

SONGAILIENĖ

Bakterijų gynybinės salos: CRISPR-Cas ir toksino-antitoksino sistemos

DAKTARO DISERTACIJA

Gamtos mokslai,
biochemija (N 004)

VILNIUS 2021

Disertacija rengta 2013 – 2017 metais Vilniaus universitete, Gyvybės mokslų centre, Biotechnologijos institute, Baltymų-nukleorūgščių sąveikos tyrimų skyriuje.
Mokslinius tyrimus rėmė Lietuvos mokslo taryba.
Disertacija ginama eksternu.

Mokslinis konsultantas:
Prof. dr. Virginijus Šikšnys (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Gynimo taryba:
Pirmininkė – **Prof. dr. Edita Sužiedėlienė** (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Nariai:
Dr. Julija Armalytė (Vilniaus universitetas, gamtos mokslai, biochemija – N 004);
Prof. habil. dr. Rimantas Daugelavičius (Vytauto Didžiojo universitetas, gamtos mokslai, biochemija – N 004);
Dr. Dukas Jurėnas (Prancūzijos nacionalinis mokslinių tyrimų centras, gamtos mokslai, biochemija – N 004);
Prof. dr. Rolandas Meškys (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Disertacija ginama viešame Gynimo tarybos posėdyje 2021 m. lapkričio mėn. 15 d. 10:00 val. Vilniaus universiteto Gyvybės mokslų centro R401 auditorijoje. Adresas: Saulėtekio al. 7, R401, Vilnius, Lietuva
tel. +37068602086; el. paštas *inga.songailiene@bti.vu.lt*

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

CONTENTS

LIST OF ABBREVIATIONS

(d)NTP	(deoxy)nucleoside triphosphate
Abi	abortive infection
AID	activation-induced cytidine deaminase
Ago	Argonaute
BREX	bacteriophage exclusion system
Cas	CRISPR-associated
Cascade	CRISPR-associated complex for antiviral defense
CBASS	cyclic oligonucleotide based defense systems
cOA	cyclic oligoadenylate
CRISPR	clustered regularly interspaced short palindromic repeats
CSR	class-switch recombination
crRNA	CRISPR ribonucleic acid
CTD	C-terminal domain
DISARM	defense island system associated with restriction-modification
DNase	deoxyribonuclease
Ec	*Escherichia coli*
EMSA	electrophoretic mobility shift assay
ER	endoplasmic reticulum
FRET	Forster resonance energy transfer
HEPN	Higher Eukaryotes and Prokaryotes Nucleotide Binding
MTase	methyltransferase
msDNA	multi-copy single-stranded DNA
NTD	N-terminal domain
pAgo	prokaryotic Argonautes
PAM	protospacer adjacent motif
PFS	protospacer flanking site
PDB	protein data bank
pre-crRNA	precursor crRNA
REase	restriction endonuclease
RF	replication fork
RM	restriction-modification
RNAP	RNA polymerase
RNase	ribonuclease
RNP	ribonucleoprotein
RNAi	RNA interference
ss	single-stranded
St	*Streptococcus thermophilus*

TA	toxin-antitoxin
TIRF	total internal reflection fluorescence microscopy
tracrRNA	trans-activating crRNA
WT	wild type

INTRODUCTION

Microorganisms living all around us and inside us make a huge impact to our lives. They can be major players in food production, our key commensals responsible for our immunity, digestion or vitamin production, as well as major cause of diseases. Due to rapidly growing antibiotic resistance in clinic, human health and well-being is severely threaten. Therefore, we need to explore the alternative ways to combat microbes. The most promising strategies as an alternative to antibiotics include phage therapy, in addition to antimicrobial peptides and antibodies. In phage therapy, species specific lytic phages are used to infect and kill certain bacteria. Therefore, natural mechanisms of bacteriophage resistance are crucial to investigate. The arms race between bacteria and phages gave rise to a number of nucleic acids targeting systems to eliminate phage genetic material (DNA or RNA). Fundamental research on phage resistance mechanisms led to the discovery of molecular tools for genetic engineering, such as restriction enzymes, transposases, bacterial argonautes, retrons or CRISPR-Cas systems. CRISPR-Cas defense systems are very abundant and are found in in more than 40 % of bacteria and nearly 90 % of archaea (Grissa et al., 2007). The CRISPR-Cas systems are divided into two major classes that are further divided into types and subtypes (Makarova et al., 2020). Class 2 systems encode a single protein effector (Cas9, Cas12, Cas13) and are widely used for genome engineering applications whereas the use of multiprotein class 1 systems is much less explored. Class 1 systems encode multiprotein effector complexes and crRNA guide. Differentially from class 2 where a single protein performs both the recognition of invading nucleic acids and the target cleavage, multiprotein class 1 systems recognize longer DNA or RNA targets and the complexes may destroy the targets (most of type III systems) or attract additional proteins to do that (type I). Class 1, type I CRISPR-Cas ribonucleoprotein effectors that are major interest of this thesis use a guide crRNA to recognise and unwind dsDNA forming DNA:RNA hybrids (R-loops). The R-loops further serve as docking sites for Cas3 nuclease-helicase that provides the target cleavage and complete elimination of certain phage DNA, so dsDNA recognition and targeting steps are separate for class 1, type I systems versus class 2. Type I systems characterized to date target dsDNA or ssDNA (I-D type systems, exclusively). Although, various aspects of type I CRISPR-Cas systems were studied earlier, their molecular mechanisms of action are still poorly understood.

The first part of this thesis focuses on fundamental aspects of class 1 type I-E and I-F CRISPR-Cas systems aiming to understand molecular mechanisms of target recognition and catalysis using a combination of biochemical and single molecule assays.

CRISPR-Cas systems are located in the genome part called "defense islands" where genes encoding different antiviral systems are clustered together. The second part of the thesis is dedicated to the biochemical and structural characterization of the toxin-antitoxin system that is present in the operon of the type I-D CRISPR-Cas system and may provide an extra layer of antiviral defense.

The goal of this study was to explore the mechanisms of action of type I CRISPR-Cas and CRISPR-Cas associated toxin-antitoxin systems. To achieve the goal, the following **objectives** were defined:

1. Establish the mechanism of R-loop formation and DNA cleavage of the type I-E Cascade complex;

2. Investigate the target recognition mechanism of type I-E Cascade and type I-F Csy complexes containing crRNA of different length;

3. Evaluate R-loop stability and target degradation by type I-E Cascade-Cas3 and type I-F Csy-Cas2/3 systems containing crRNA of different length in their effector complexes;

4. Determine the structural and molecular mechanisms of *Aphanizomenon flos-aquae* toxin-antitoxin system associated with I-D CRISPR-Cas system.

Scientific novelty and practical value:

Class 1 I type complexes recently have emerged as effective tools in genetic engineering of phages, prokaryotic and eukaryotic cells (Chapter 2.4.). The accurate DNA recognition is crucial for efficient gene silencing or target editing. The data obtained in this work provides the detailed mechanism of R-loop formation and cleavage of target DNA by type I CRISPR-Cas effector complexes. Single molecule experiments performed in this study identified different steps of the target recognition and degradation by I-E type Cascade complex that can explain the observed *in vivo* effects when CRISPR-Cas systems are applied for genome engineering or gene silencing. It was clearly demonstrated that the length increase of the crRNA in type I-E system does not change the mechanism of dsDNA targeting and Cas3-mediated cleavage, however, it ensures increased dissociation time from DNA and therefore it can be useful for gene silencing applications. The HEPN-MNT toxin-antitoxin (TA) system identified in cyanobacterium *Aphanizomeonon flos-aquae* is the first experimentally characterised TA system encoded in the same operon as I-D CRISPR-Cas effector complex. We show that HEPN toxin ribonuclease

activity is neutralized by covalent di-AMPylation that changes the conformation of the active site loop. Covalent di-AMPylation discovered in this work is a novel mechanism of toxin neutralisation. Furthermore, the active HEPN toxin cleaves 4 nt from 3′-stem of a range of tRNA, showing a new mechanism of toxin action on tRNA. We propose that HEPN-MNT toxin-antitoxin pair functions as a bacterial ATP-sensor able to detect cellular ATP levels. HEPN toxin can be applied for cleavge of a range of tRNA or inhibition of protein translation.

The major findings presented for defense in this thesis:

- I-E type Cascade complex recognize its target by unidirectional zipping mechanism of the protospacer starting from PAM sequence (**publication 1**)
- The proximal side of the R-loop causes the conformational changes termed "target locking" in the Cascade complex that triggers the target degradation by Cas3 nuclease (**publication 1**)
- I-E Cascade complex is able to accommodate crRNA of altered length and forms shorter or longer R-loops (ranging from 15 bp to 57 bp, respectively) depending on the spacer length (**publication 2 and 3**)
- Regardless of the crRNA length and the R-loop length formed, I-E Cascade or I-F Csy complexes recognize their targets and induce the target degradation by the mechanism used by WT complexes and use 3′-end target validation mechanism for dsDNA cleavage (**publication 2 and 4**)
- HEPN-MNT from *Aphanizomenon flos-aqaue* I-D CRISPR-Cas operon forms a toxin-antitoxin system where HEPN RNase is neutralized by di-AMPylation (**publication 5**)
- Unmodified active HEPN cleaves tRNA removing 4 nt from 3′-end: CCA and "discriminatory" nucleotide (**publication 5**)

1. Anti-phage defense strategies

Bacteriophages (phages) are obligatory parasites that use cellular resources to reproduce. The evolutionary fight between bacteria and phages gave rise to diverse strategies to block phage propagation (Doron et al., 2018; Gao et al., 2020). Numerous of anti-phage defense systems have been discovered to date, ranging from single enzyme activities to multiprotein action (Doron et al., 2018; Gao et al., 2020). Such strategies include blocking of the phage adsorbtion, restriction-modification (RM) systems, toxin-antitoxin (TA) modules, prokaryotic Argonautes (pAgo), BREX/Pgl systems, CRISPR-Cas, cyclic oligonucleotide-based systems (CBASS) and other (summarized in Table 1 and Figure 1) (Koonin et al., 2017; Lopatina et al., 2020). Major activities of the RM, pAgo or CRISPR-Cas defense systems rely on targeting of phage DNA or RNA, while others (TA, CBASS, RADAR) induce abortive infection by interfering with key cellular processes.

CRISPR system has been identified in 1987 in bacterial genomes when sequencing data emerged (Ishino et al., 1987), though the *in vivo* activity in *S. thermophilus* of CRISPR system and related Cas genes has been demonstrated more than two decades later (Barrangou et al., 2007)

The Table 1 summarizes current knowledge on anti-phage defense strategies:

Table 1. Intracellular anti-phage defense strategies.

Defense system	Activity	Year of discovery	Ref.
RM systems	Phage DNA cleavage	1952	(Luria and Human, 1952)
TA systems	Interferes with key cellular processes	1983	(Ogura and Hiraga, 1983)
CRISPR-Cas	crRNA-guided RNA or DNA cleavage	CRISPR repeats observed in 1987, 2007 detected activity *in vivo*	(Barrangou, 2007; Ishino et al., 1987)
Prokaryotic argonautes pAgo	ssDNA or ssRNA guides to target phage DNA or RNA	2009 bioinformatically, 2014 biochemical activity	(Makarova et al., 2009; Swarts et al., 2014)
BREX/Pgl systems	Methylates host DNA, the defense mechanism unknown	2014	(Goldfarb et al., 2014)
Small molecule defense	Organic molecules acting against phage replication	2018	(Kronheim et al., 2018)

DISARM	Methylates host DNA, further mechanism is unknown	2018	(Ofir et al., 2018)
New defense systems from pangenome: Thoeris, Gabija, Hachiman, Shedu, Septu, Lamassu, Zorya, Kiwa, Druantia	Role of most systems, except Gabija and Thoeris, is unknown	2018	(Doron et al., 2018; Ka et al., 2020)
Cyclic-oligonucleotide based anti-phage signaling system (CBASS)	Cyclases sense phage and produce signaling molecules, signaling molecules bind protein effectors and induce cell death	2019	(Cohen et al., 2019; Whiteley et al., 2019)
Thoeris	Cleaves NAD^+ and forms pores in cell membrane	2020	(Ka et al., 2020)
Gabija	Cleaves dsDNA as a type IIP restriction endonuclease when (d)NTP concentration is low	2020	(Silva, 2020)
Retrons	Reverse transcriptase forms a complex with msDNA and effector protein, retron needs RecBCD for function, biochemical details unknown	2020	(Bobonis et al., 2020a; Gao et al., 2020; Millman et al., 2020)
RADAR	Deaminates cellular RNA in response to phage infection	2020	(Gao et al., 2020)

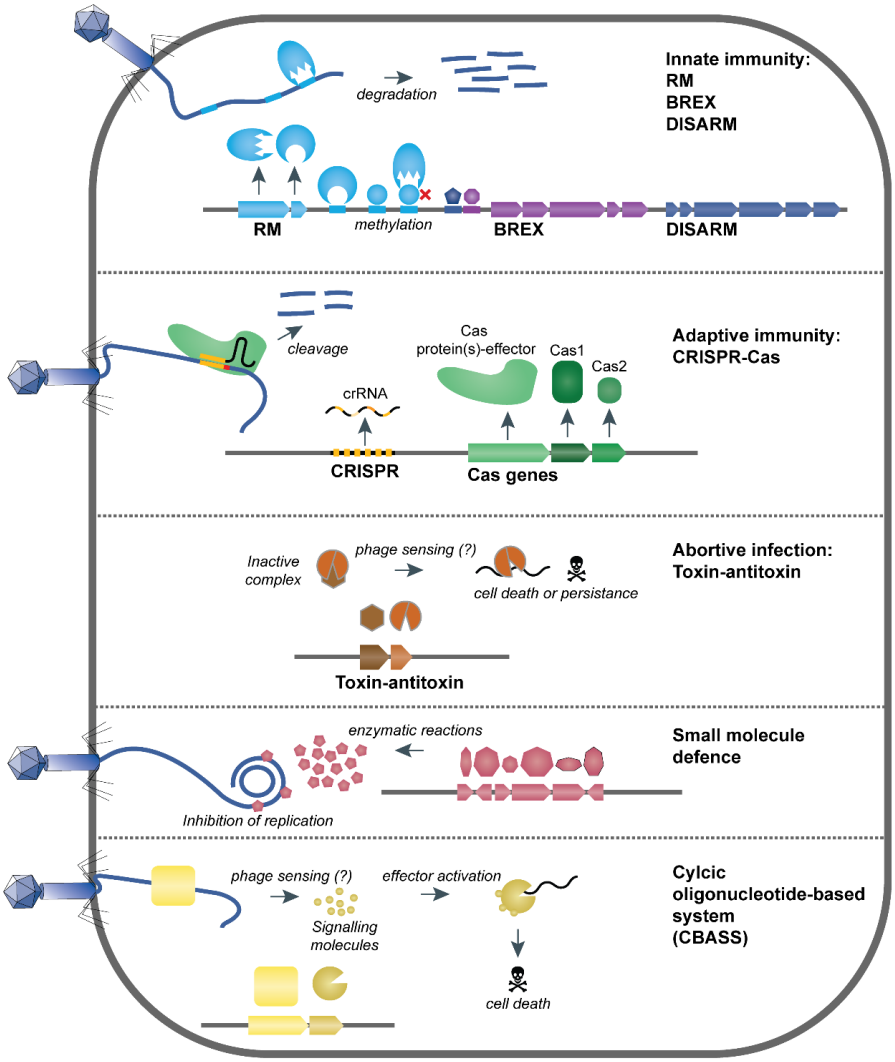


Figure 1. Intracellular bacterial phage defense strategies. Innate immunity strategies include restriction-modification (RM) systems, BREX and DISARM systems. Host DNA is protected by certain modification, e.g. methylation, while phage DNA remains unmodified and therefore is cleaved by specific nucleases. Adaptive immunity is ensured by CRISPR-Cas systems. Adaptation proteins Cas1-Cas2 (or also Csn2, Cas4) insert short DNA fragments into CRISPR region (adaptation). Later CRISPR region is transcribed to crRNA that serves as a guide for Cas proteins to destroy the phage. Toxin-antitoxin systems as a part of abortive infection are formed by antitoxin and stable toxin. Antitoxin is degraded upon phage infection and toxin (most often ribonuclease) ensures persister formation or cell death prior to phage burst. In small molecule defense strategy certain enzymes synthesize organic molecules (doxorubicin, daunorubicin) that serve as DNA intercalants and inhibit phage replication. Cyclic oligonucleotide-based systems CBASS contain an

enzymatic effector that senses phage components and synthesizes signaling molecules, e.g. cGMP-AMP, that activate enzymatic effectors and induce abortive infection. Figure prepared according to (Bernheim and Sorek, 2020).

2. CRISPR-Cas systems

2.1. Function of CRISPR-Cas systems

CRISPR-Cas systems are encoded in more than 40 % of bacteria and nearly 90 % of archaea (Grissa et al., 2007). CRISPR-Cas systems function as adaptive RNA-guided nucleic acid degradation systems (Figure 1). Recently, novel roles of CRISPR-Cas systems have been established: CRISPR in large phages are shown to combat other phage of the same species (Al-Shayeb et al., 2020) or mediate transposition rather than target cleavage (Klompe et al., 2019; Petassi et al., 2020).

The CRISPR-Cas immunity can be divided into three main stages: adaptation, ribonucleoptrotein (RNP) complex formation and interference. At the adaptation stage Cas1 and Cas2 (in some types, also Csn2 or Cas4) insert short foreign 30-35 nt sequences (spacers) from phages or plasmids into a CRISPR region (Sasnauskas and Siksnys, 2020). The CRISPR array is then transcribed into pre-crRNA precursor, that is further processed into mature crRNAs (Carte et al., 2010; Deltcheva et al., 2011). Lastly, Cas proteins form RNP complexes with crRNA and perform recognition and cleavage of invading nucleic acids (Figure 1). Different CRISPR-Cas systems can recognize either DNA or RNA targets. dsDNA recognizing RNP scan for a short protospacer adjacent motif (PAM) sequence, once it is found, dsDNA protospacer (sequence that is complementary to crRNA guide) hybridizes to crRNA starting from an initial ~8 bp sequence called „seed sequence“ and forms an R-loop structure.

2.2. Classification and action of CRISPR-Cas systems and their ancillary proteins

Current classification of CRISPR-Cas systems includes **2 classes**, 6 types and 33 subtypes. **Class 1** is the most abundant and contains multicomponenent RNP effector complexes (Cascade, Csy, Csm/Cmr complexes) whereas **class 2** systems encode a single-subunit effector RNP (Cas9, Cas12, Cas13) (Figure 2) (Makarova et al., 2018, 2019). The emerging metagenomic data may soon reveal more diversity of CRISPR-Cas systems.

Class 1 is further subdivided to three types (Figure 2a):

- **Type I** recognizes short nucleotide sequences termed PAM and uses crRNA guide to unwind dsDNA sequences called protospacer forming the R-loop structure on DNA target. The R-loop formation triggers Cas3 nuclease-helicase binding and target DNA cleavage (Brouns et al., 2008; Sinkunas et al., 2011, 2013; Westra et al., 2012); Some subtypes (I-F, I-B) of class 1 lack nuclease-helicase activity, but encode active Tn7-transposon (Peters et al., 2017). Recently biochemical activity of type I CRISPR-Cas systems that can employ crRNA-guided transposition have been demonstrated. The systems encode I-Fv type CRISPR-Cas system and TniQ transposition protein (Klompe et al., 2019).

  This thesis is focused on class 1 type I-E, I-F and I-D CRISPR-Cas systems. The genetic composition of the systems is shown in Figure 2b. The type I-E, I-F and I-D RNP complexes differ by their constituent proteins. I-E type complex contains 5 subunits (Cas5, Cas7, Cas6, Cse2 (small subunit) and Cse1 (large subunit)) (Mulepati et al., 2014; Zhao et al., 2014). I-F is composed of four subunits and lack a separate small subunit (Cas5, Cas7, Cas6, Csy8f (large subunit)) (Rollins et al., 2019). Composition of I-D type complex is still under debate, but it is proposed to contain four essential subunits (Cas7, Cas5 and Cas10d (large subunit), Cas11 (small subunit)) (McBride et al., 2020).
- **Type III** systems show different catalytic activities. Initially, Csm or Cmr complexes use crRNA guide to bind and cleave ssRNA what triggers non-specific ssDNA degradation and synthesis of cyclic signalling molecules – cyclic oligoadenylates cOA_{4-6} by the effector complexes (Estrella et al., 2016; Kazlauskiene et al., 2016, 2017; Niewoehner et al., 2017; Samai et al., 2015; Tamulaitis et al., 2014). cOA_{4-6} functions as a signalling molecule that activates CRISPR-Cas ancillary proteins with cOA-binding CARF domains. In Type III-A CRISPR-Cas systems Csm6 contains regulatory CARF domain and HEPN domain that acts as cOA activated RNase to cleave ssRNA (mRNA) (Kazlauskiene et al., 2017; Niewoehner et al., 2017; Rouillon et al., 2018);
- **Type IV** systems typically lack nucleases (except IV-C subtype that encodes active HD domain) and contain DinG family helicase. These systems perform *in vivo* plasmid interference and are exclusively located in plasmids, mobile genetic elements or prophages, the biochemical mechanisms of type IV systems are yet unknown (Pinilla-Redondo et al., 2020).

Class 2 is further subdivided into three types (Figure 2a):

- **Type II** effector complexes (as exemplified by Cas9) use crRNA guide together with tracrRNR to recognize dsDNA containing PAM sequence

at 3′-protospacer end and cleave dsDNA target by HNH and RuvC-domains (Gasiunas et al., 2012; Jinek et al., 2012).

- **Type V** (Cas12) contains RuvC-domain and different variants are reported to cleave dsDNA recognizing PAM sequence at 5′-protospacer end (Karvelis et al., 2020; Zetsche et al., 2015). For some dsDNA recognition further triggers degradation of ssDNA (Li et al., 2018, Chen et al., 2018), while some are shown to target ssDNA only (Harrington et al., 2018). Subtype V-K Cas12k encodes Tn7-like transposons that are active upon crRNA-guided transposition together with transposition proteins TnsB, TnsC and TniQ (Petassi et al., 2020).
- **Type VI** systems (Cas13, C2c2) contain HEPN domains and possess induced non-specific ssRNA cleavage activity upon target RNA binding what leads to cleavage of transcript RNA (Abudayyeh et al., 2016; Gootenberg et al., 2017). ssRNA binding via regulatory α-helix from Cas13 helical domain regulates conformation of HEPN domain that is responsible for non-specific ssRNA cleavage (Liu et al., 2017b; Zhang et al., 2018). The feature of the type V and type VI systems to non-specifically cleave ssDNA or ssRNA upon target recognition has been employed for sensitive one-stage virus detection assays (Gootenberg et al., 2017; Harrington et al., 2018; Kellner et al., 2019). For some Cas13 proteins (*Leptotrichia shahii* LshCas13a and *Bergeyella zoohelcum* Cas13b) specific sequence on ssRNA termed PFS (Protospacer Flanking Site) is required for target cleavage (Abudayyeh et al., 2016; Smargon et al., 2017), while this PFS sequence is not essential for some other Cas13 (*Leptotrichia buccalis* LbuCas13a) (Konermann et al., 2018).

  Mechanisms of action of different Cas complexes and CRISPR-Cas associated proteins are reviewed in Table 2.

CRISPR-Cas associated (ancillary) proteins are additional genes that are not necessary for CRISPR-Cas activity, but overall help to combat phage infection. These proteins are best characterized in the context of type III systems. They are activated by signals produced by Csm/Cmr complexes (cOA) or help to deactivate these signalling molecules after their action. cOA may activate Csm6, Csx1 with CARF-HEPN domains that cleave mRNA. Other cOA activated proteins include PD-D/EXK nucleases Can1 that nicks dsDNA plasmids or Can2 that nicks dsDNA plasmids and cleaves ssRNA (Kazlauskiene et al., 2017; McMahon et al., 2020; Niewoehner et al., 2017; Zhu et al. 2021). Furthermore, ancillary proteins Csx3, Crn1 with active CARF domain or RING-nucleases that deactivate cOA also play vital role in cellular signalling (Athukoralage et al., 2018, 2020; Brown et al., 2020; Smalakyte et al., 2020).

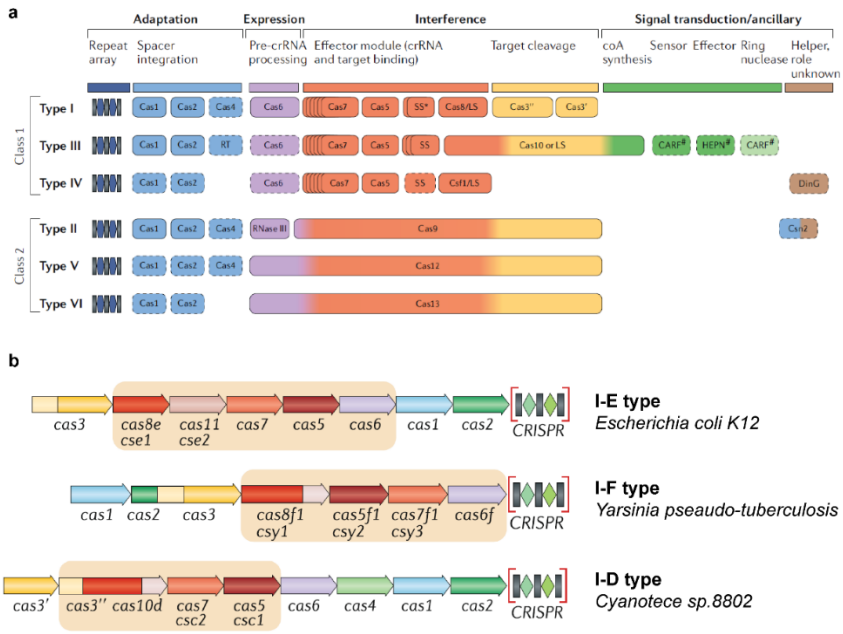


Figure 2. Classification of CRISPR-Cas systems. a. Functional modules of class 1 and class 2 CRISPR-Cas systems are divided to adaptation genes (Cas1 and Cas2, Cas4, Csn2, RT), pre-crRNA processing module (Cas6 for Class 1). The genes that are essential for the function of certain types are encircled in a straight line, while the genes that are not essential and might be missing are shown in a dashed line. Adapted from (Makarova et al., 2020). b. The genes that are characteristic of type I-E, I-F and I-D type of CRISPR-Cas systems. The proteins forming a multiprotein effector complex are shaded in brown.

Table 2. Mechanisms of action of class 1 and class 2 CRISPR-Cas systems and ancillary CRISPR proteins.

Class 1				
Type	**Effector**	**crRNA-guided targeting**	**Additional activity**	**Ref.**
I-E	Cascade + Cas3	5′PAM dependant dsDNA → ssDNA fragments	Not reported	(Sinkunas et al., 2013; Westra et al., 2012)
I-F	Csy + Cas3			(Rollins et al., 2019)
I-C	Cascade + Cas			(Hochstrasser et al., 2016)
I-D	Cascade + Cas3		ssDNA → cleavage to 6 nt fragments	(Lin et al., 2020)

I-F variant	Cascade + TniQ	5′PAM dependant dsDNA → transposition of dsDNA cargo	Not reported	(Klompe et al., 2019)
IV	Cascade	Unknown	Unknown	(Pinilla-Redondo et al., 2020)
III-A	Csm complex	ssRNA cleavage to 6 nt fragments	ssDNA cleavage; Synthesis of signaling molecule cOA	(Estrella et al., 2016; Hatoum-Aslan et al., 2014; Kazlauskiene et al., 2016; Samai et al., 2015; Tamulaitis et al., 2014)
III-V	Cmr complex			(Hale et al., 2014; Han et al., 2018; Zhu and Ye, 2014)
	Ancillary protein	***Activation mechanism***	***Activity***	***Ref.***
III-A	Csm6	Cyclic oligoadenylate cOA_{4-6}	ssRNA cleavage	(Kazlauskiene et al., 2017; Niewoehner et al., 2017)
III-B	Csx1			(Grüschow et al., 2019)
III-B	Csx3	Bind and cleave cOA_4	Deactivation of CARF-HEPN RNases	(Brown et al., 2020)
III-A	Ring nucleases			(Athukoralage et al., 2018)
III-D	Crn1			(Athukoralage et al., 2020)
III-A	Card1	Cyclic oligoadenylate cOA_4	ssDNA cleavage *in vivo*, ssDNA and ssRNA cleavage *in vitro*	(Rostøl et al., 2021)
III-A	Can1		Nicks dsDNA plasmid	(McMahon et al., 2020)
III-A	Can2		Nicks dsDNA plasmids,	(Zhu et al., 2021)

			cleaves ssRNA	
Class 2				
Type	**Effctor**	**crRNA-guided targeting**	**Additional activity**	**Ref.**
II	Cas9	3′PAM dependant dsDNA → dsDNA cleavage, single cleavage site	Not reported	(Gasiunas et al., 2012; Jinek et al., 2012)
V	Cas12	5′PAM dependant dsDNA → dsDNA cleavage, single cleavage site	Collateral ssDNA cleavage, activated by dsDNA target binding	(Karvelis et al., 2020; Li et al., 2018; Zetsche et al., 2015)
VI	Cas13	3′ PFS dependant ssRNA → ssRNA cleavage	Collateral non-specific ssRNA trans-cleavage, activated by ssRNA target binding	(Abudayyeh et al., 2016; Gootenberg et al., 2017)

2.3. Class 1 systems – structures and functions

Class 1 systems are the most widespread and diverse (Makarova et al., 2020). They contain multisubunit ribonucleoprotein (RNP) complexes.

Type I effector complexes form multisubunit RNP from crRNA that is bound by 3′-end hairpin after the pre-crRNA is cleaved by Cas6 (Charpentier et al., 2015). Type I RNP targets dsDNA and recognizes PAM sequence at 5′-protospacer end that is present in the invader DNA, but not host DNA (Brouns et al., 2008). PAM recognition is followed by R-loop formation and target degradation by Cas3 (Mulepati and Bailey, 2013; Sinkunas et al., 2011, 2013; Westra et al., 2012). In the currently characterized type I CRISPR-Cas systems, PAM recognition is mediated by Cse1 (large subunit) and occurs from minor groove side following by DNA bending and a wedge-mediated R-loop formation (Figure 3c) (Hayes et al., 2016; Rollins et al., 2019; Xiao et al., 2017). In general, type I-E and I-F CRISPR-Cas systems form multisubunit surveillance complexes resembling ‘seahorse-shaped’ structures (Figure 3, Figure 4).

In **type I-E** Cascade (CRISPR-associated complex for antiviral defense) crRNA is in a pseudo helical A-conformation, anchored by Cas5, Cse1 and Cas7 at 5′-end and Cas6 at 3′-hairpin crRNA end. Cas7 forms a backbone with 6-nt kinks on the guide crRNA. When Cascade forms the

R-loop, DNA:RNA duplex is disrupted at every 6-th nt by Cas7 structural finger forming "broken-ladder" DNA:RNA duplex (Jackson, 2014; Mulepati et al., 2014; Van Erp et al., 2015; Zhao et al., 2014). The dsDNA target binding induces ~10 Å Cse2 sliding downwards to Cse1 and subsequent tilting of Cse1 by 14 Å closer to ssDNA, thus suggesting that small subunit Cse2 may validate full R-loop recognition (Figure 3d) (Xiao et al., 2017). Once full-length R-loop is formed, Cascade complex undergoes conformational rearrangements mediated by Cas7 and Cse2 termed "R-loop locking". In *Escherichia coli* (Ec) Ec-Cascade interactions between Asp22 from Cas7 and Arg107 or Arg101 from Cse2 ensure this target validation step (Figure 3e) ("R-loop locking" in *Streptococcus thermophilus* (St) St-Cascade complex is discussed in **publication 1**) (Rutkauskas et al., 2015; Szczelkun et al., 2014a; Van Erp et al., 2015). The complex contains a ssDNA part that is required for recruitment of Cas3 nuclease-helicase that further performs initial R-loop nicking in the non-target strand (NTS) (Sinkunas et al., 2013; Westra et al., 2012; Xiao et al., 2018). Nicked R-loop with 10-nt non-target strand (NTS) overhang forms the most stable complex between Cascade and Cas3 (Figure 3f). Cas3 is a HD-nuclease (which is absent in type I-D) and SF2-helicase responsible for DNA degradation (Hochstrasser et al., 2014; Sinkunas et al., 2011; Westra et al., 2012). Cas3 binds Cascade after R-loop formation by interacting with Cse1 (large subunit) N-terminal domain NTD by its helicase domain and with Cse1 C-terminal domain by its HD-domain containing three main interaction surfaces (Figure 3g) (Huo et al., 2014; Xiao et al., 2018). SF2-helicase uses ATP and displaces ssDNA in 3′-5′-direction by repetitive reeling by 3 nt steps to feed it to HD-nuclease domain (Loeff et al., 2018).

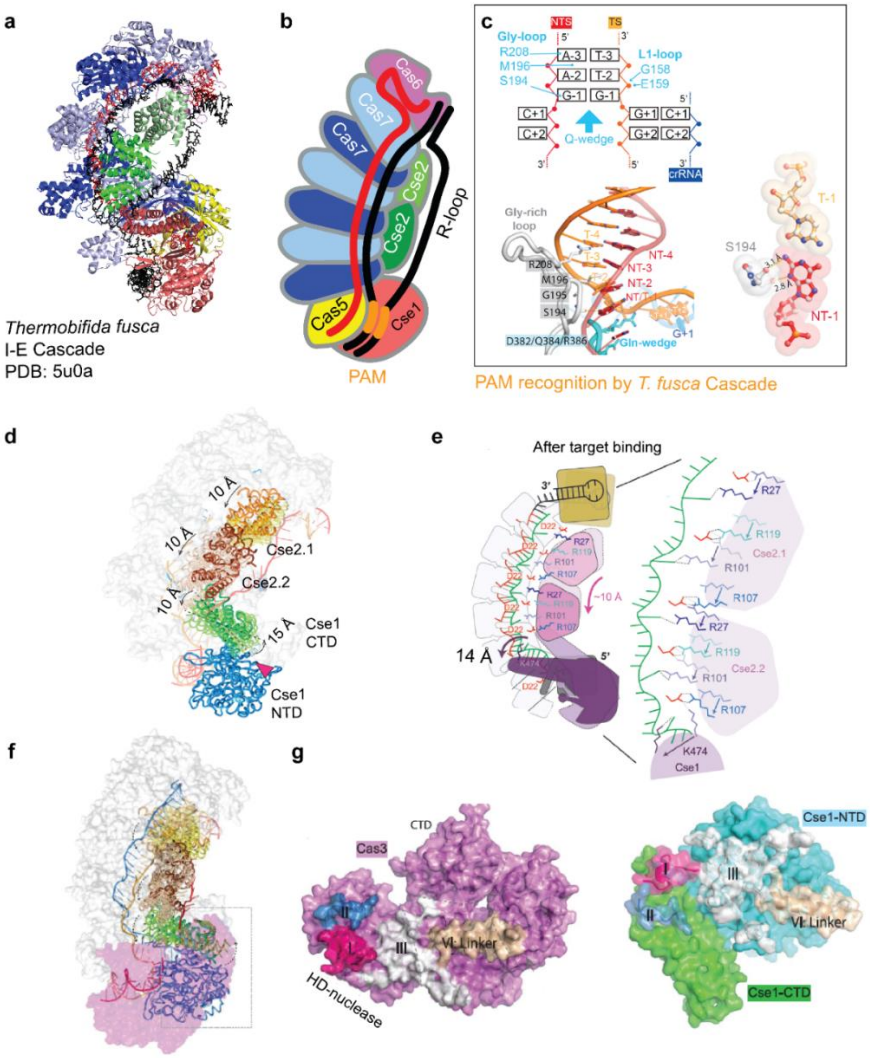


Figure 3. General structure and activation mechanism of I-E type Cascade complex. a. General "seahorse"-shaped structure of the *T. fusca* Cascade complex (PDB ID:5U0A). b. Schematic representation of the structure. PAM sequence shown in orange, crRNA shown in red and R-loop – in black. c. The PAM sequence recognition mechanism from minor groove side by *T. fusca* Cascade complex mediated by Gln-wedge and Gly-loop. Adapted from (Xiao et al., 2017). d. Structural rearrangements of I-E type Cascade complex upon the formation of a full R-loop. Small subunits Cse2 slide away from Cas6 closer to Cse1. CTD of Cse1 is tilted by 15Å. Adapted from (Xiao et al., 2017). e. Molecular mechanism of "target locking" by Cse2 and Cas7 in *E. coli* Cascade. Once full R-loop is formed, Arg from Cse2 forms salt-bridges to Asp22 from Cas7. Adapted from (Van Erp et al., 2015). f. Cas3 (purple) binding to full R-loop and Cas3 interactions with Cse1, adapted from (Xiao et al., 2018). g. The detected interaction surfaces, depicted I purple, II blue and III grey, respectively, of *T. fusca* Cas3 and Cse1, adapted from (Xiao et al., 2018).

Type I-F forms a four-component RNP termed Csy complex (Figure 4a). From the complex, Cas5 recognizes PAM from major groove side and Csy8f (large subunit) recognizes PAM from minor groove side (Figure 4c). The function of small subunits that is lacking in type I-F complexes is fulfilled by the CTD of Csy8f large subunit. Upon target recognition and R-loop formation, C-terminal helical bundle of Csy8f undergoes a drastic 180º degree rotation, mimicking R-loop locking by small subunits Cse2 in type I-E complex (Figure 4d) (Rollins et al., 2019). The R-loop activates Cas2/3 nuclease that is a fusion of Cas2 and Cas3 proteins in type I-F systems. Cas2/3 is negatively regulated by adaptation protein Cas1 and Cas2/3 HD-nuclease is activated once a functional R-loop is formed (Rollins et al., 2017). Once R-loop is not full, e. g. due to mutations in the protospacer, primed adaptation performed by Cas1-Cas2/3 takes place instead of the target degradation by Cas2/3.

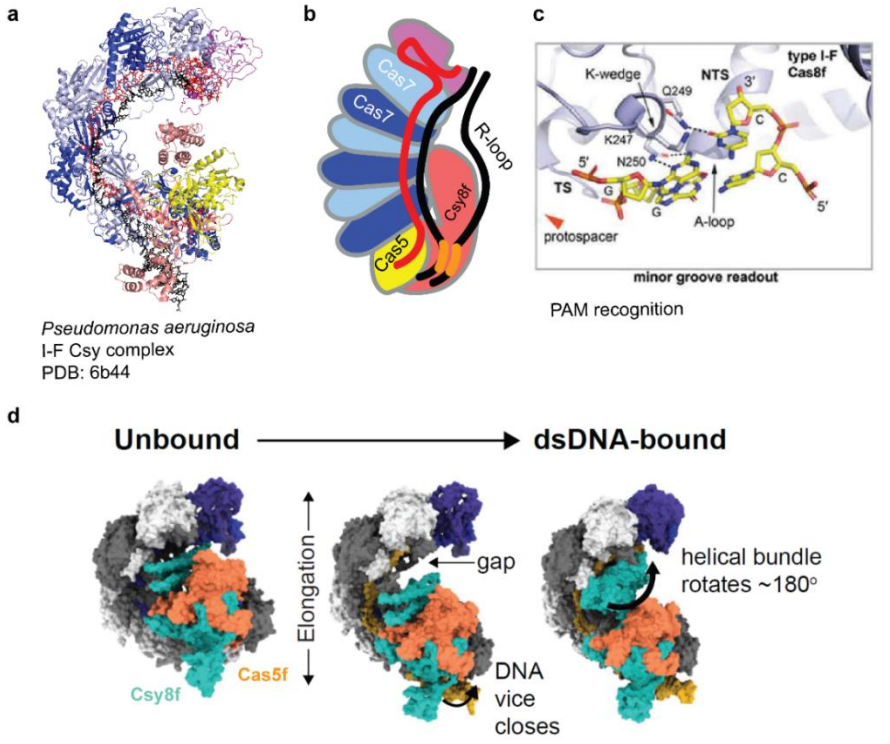


Figure 4. Structure and activation mechanism of I-F type Cascade complex. a. General structure of I-F Csy complex (PDB ID:6B44). The complex is composed of Cas6f, Cas7f, Cas5f and Csy8f (large subunit). b. Schematic representation of Csy complex, lack of Cse2 small subunits is compensated mainly by CTD of Csy8f. c. PAM recognition by I-F Csy complex large subunit Csy8f occurs from minor groove side, figure adapted from (Gleditzsch et al., 2019). d. Conformational changes in Csy8f large subunit that occur upon full target binding (R-loop locking). C-terminal

helical bundle rotates by 180º upon full R-loop recognition. Figure adapted from (Rollins et al., 2019).

Type III RNPs form similar, but more extended "worm-like" structures composed of five Cas proteins (Figure 5a). Type III RNPs usually lack Cas6 subunit and crRNA hairpin structure at 3′-end (Osawa et al., 2015; You et al., 2019). Type III systems act both on ssRNA and ssDNA bubbles. The type III complexes possess a set of enzymatic activities: trimming RNase cutting ssRNA target at every 6th base pair (3′-end of every flipped nucleotide in RNA:RNA duplex), located in Cas7 subunit (Csm3, Cmr4, respectively), ssRNA-activated ssDNase activity, located in large subunit Cas10 as a HD-domain (Csm1, Cmr2), and Palm-polymerase in Cas10, synthesizing signalling molecules, cyclic adenylates $cOA_{4\text{-}6}$, from ATP when ssRNA target is bound to the complex (Figure 5d, 5e) (Estrella et al., 2016; Hatoum-Aslan et al., 2014; Kazlauskiene et al., 2016, 2017; Niewoehner et al., 2017; Samai et al., 2015; Tamulaitis et al., 2014). Type III RNPs do not require PAM sequences on RNA for target binding, they use 5′-crRNA handle complementarities to the RNA target to distinguish self nucleic acids from non-self (You et al., 2019). Structures of multiprotein class 1 complexes are similar in terms of target recognition and crRNA-target duplex formation – Cas7 backbone subunits (Csm3, Csm4) insert a loop into RNA:RNA duplex, forming a "broken ladder" at every 6th base pair (Figure 5c) resembling type I complexes (You et al., 2019).

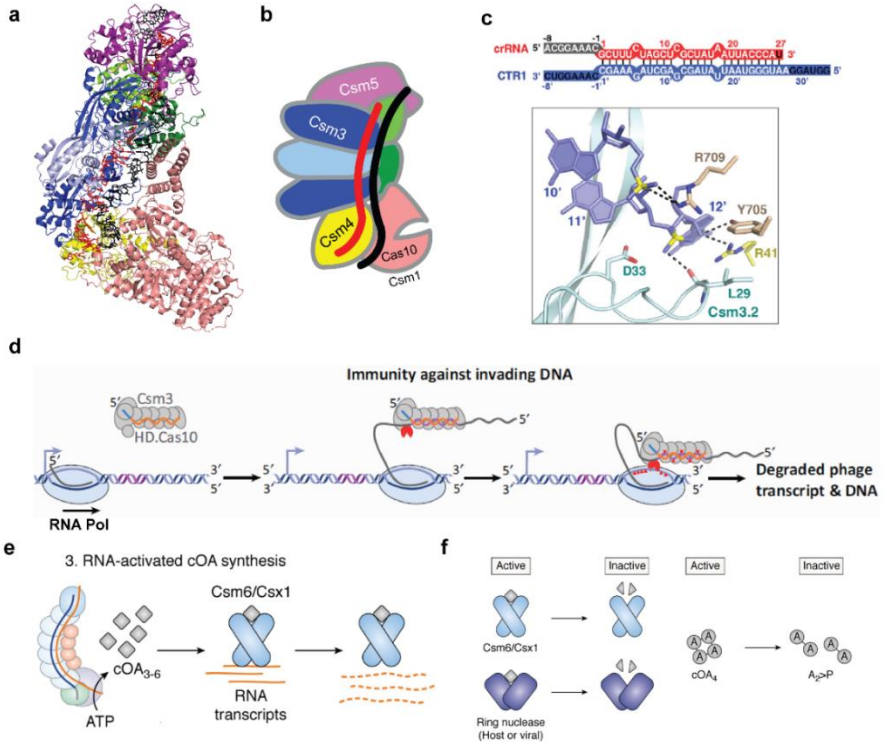


Figure 5. Structure and activity of type III Csm/Cmr complexes. a. Overall structure of Csm complex (PDB ID:6IFK) b. Schematic view of the composition of Csm complex, Csm3 is a homolog of Cas7, Csm4 (Cas5) and Csm1 (Cas10), respectively. c. Mechanism of ssRNA cleavage by Csm complex. crRNA:ssRNA duplex used for crystallization is depicted above. Csm3 (Cas7) catalytic finger with a conserved Asp33 (Asp33Asn mutant used for crystallization) and a scissile phosphate is shown, adapted from (You et al., 2019). d. Catalytic activity of Csm/Cmr complexes against RNA transcript and ssDNA bubble, adapted from (Tamulaitis et al., 2017). e. Catalytic activity of Csm/Cmr Palm-domain to produce cOA, activated by ssRNA binding, adapted from (Liu and Doudna, 2020). f. Catalytic activity of accessory Csm6/Csx1 or ring nucleases to neutralize cOA_{4-6} signalling molecules, adapted from (Liu and Doudna, 2020).

Cleavage activity of Cas7 (Cmr4) in Cmr complex is located in a palm structural motif and during target recognition the palm motif is sandwiched between thumb from Cmr4 and palm-loop from Cmr4 that forms “broken” RNA ladder and unpairs every 6th base forming complementary 5 nt fragments (Figure 5c) (Osawa et al., 2015). In Csm complexes thumb-like β-hairpin from Csm3 interspaces after every 5-nt in the target RNA unpairing 6th base, flipped bases stack to Arg41 from Csm2 and scissile bonds are placed near Asp44 from Csm3 (You et al., 2019). RNA binding to Csm/Cmr complexes, but not RNA-cleavage itself, activates a collateral ssDNA cleavage activity of DNA

bubbles by HD-domain located in the large subunit Cas10 (Kazlauskiene et al., 2016). Furthermore, RNA binding activates Palm-domain of Cas10 to produce cyclic signalling molecules $cOA_{4\text{-}6}$. The produced molecules activate CRISPR-Cas ancillary proteins (see Chapter 2.2), e.g. Csm6 or Csx1 by binding to CARF domain and activating mRNA cleaving HEPN domain (Figure 5e) (Han et al., 2017; Kazlauskiene et al., 2017; McMahon et al., 2020; Niewoehner et al., 2017). Cyclic signalling molecules can be neutralized either by active CARF, HEPN or RING-nucleases (Figure 5f) (Athukoralage et al., 2018; Brown et al., 2020; Smalakyte et al., 2020).

In **type I-D** systems complex is formed by four proteins. According to gene composition of type I-D CRISPR-Cas systems, these systems should lack small subunit, however, it was recently demonstrated that type I-D systems encode small subunits as an alternative open reading frame at the end of *cas10d* gene (Figure 6a, 6b) (McBride et al., 2020). Type I-D encodes an HD-domain in the large subunit Cas10d (Figure 6) (Makarova et al., 2019). Due to the lack of a functional Palm domain that is responsible for cOA synthesis in type III, Cas10d is classified as type I. However, due to the presence of a HD-domain in the large subunit that is unusual for type I, it may be considered as an evolutionary intermediate between type I and type III (Makarova et al., 2018). Furthermore, I-D system from *Sulfolobus islandicus LAL14/1* is shown to target dsDNA in a manner similar to I-E and I-F systems (required PAM on 3′-end of protospacer) and ssDNA in a manner similar to ssRNA trimming by type III systems (Figure 6d) (Lin et al., 2020). Activity of ssDNA cleavage depends on Cas7, reminiscent of type III RNA-cleavage performed by Cas7 homologs (Csm3, Cmr4), but the mechanism of it is unknown.

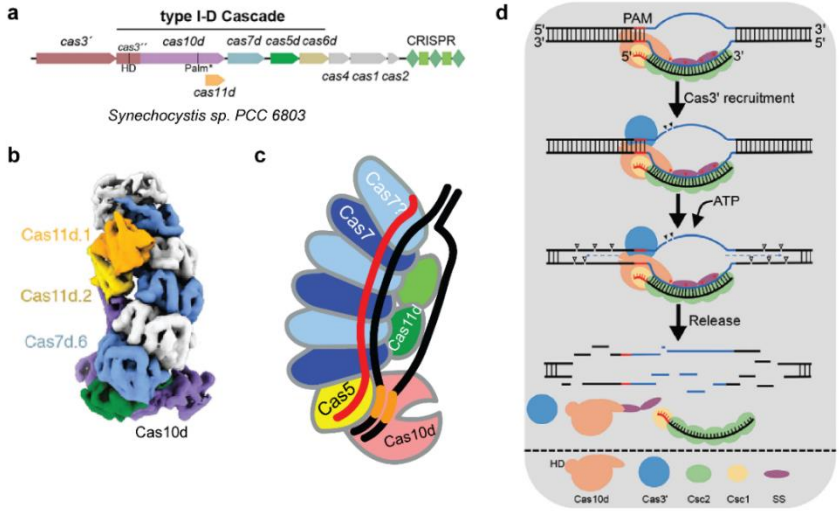


Figure 6. Type I-D CRISPR-Cas system. a. Genetic composition of I-D CRISPR-Cas system from *Synechocystis sp. PCC 6803*, encoding *Cas11d* small subunit as an alternative open reading frame of *Cas10d*, adapted from (McBride et al., 2020). b. Cryo-EM structure of I-D type Cascade complex from *Synechocystis sp. PCC 6803*, large subunit Cas10d depicted in purple, Cas5 depicted in green, Cas7 depicted in white and blue and small subunit Cas11 depicted in yellow. Adapted from (McBride et al., 2020). c. Schematic representation of type I-D complex with the R-loop. PAM sequence depicted in orange, crRNA depicted in red, dsDNA target is black d. Proposed mechanism of action based on *in vitro* activity of I-D type complex. Type I-D complexes recognize PAM sequence and form R-loop structures that mediate dsDNA cleavage by dCas10. Mulitprotein backbone formed by Cas7d subunit participates in PAM-independent ssDNA cleavage at multiple sites, adapted from (Lin et al., 2020).

Type IV lacks known nuclease domains and Cas1, Cas2 adaptation module and is known to restrict dsDNA plasmids by an unknown mechanism dependent on DinG helicase (Crowley et al., 2019; Pinilla-Redondo et al., 2020). By its structure, type IV complex best resembles type III-A Csm complex (a sea cucumber by its shape) with six Csf2 (Cas7-like) subunits forming a helical ‘‘backbone,’’ and five Cas11 subunits comprising a helical ‘‘belly’’. Due to similar structure, type IV complex may also possess RNA trimming activity, but further study is needed to show it (Zhou et al., 2021).

Type I-F variant transposon-encoded nuclease-deficient CRISPR-Cas systems have recently been shown to perform crRNA-guided transposition event that inserts DNA-cargo to 47-51 bp from 3′-crRNA end by cut-and-paste mechanism (Figure 7c) (Halpin-Healy et al., 2020; Klompe et al., 2019). The systems lack Cas3 nuclease, but encompass additional proteins: TniQ (transposase), TnsA, TnsB (both TnsA and TnsB cut DNA to form dsDNA breaks) and TnsC (integration ATPase). TnsA, TnsB, TnsC form a core transposase, while TnsQ interacts with Cascade Cas6 and Cas7 proteins at crRNA 3′-hairpin end (Figure 7a, 7b) (Halpin-Healy et al., 2020). Similarly to target recognition mechanism of I-E and I-E type effectors, transposition is sensitive to mismatches in seed region and removal of more than 4 nt from protospacer at the PAM-proximal site (Klompe et al., 2019). Cascade interacts with TniQ transposase through Cas7 and Cas6 and guides TniQ (Figure 7a, 7c). TniQ together with core transposase further mediates transposition event by cut-and-paste mechanism (Halpin-Healy et al., 2020; Klompe et al., 2019). Promiscuous PAM sequence recognition for I-F Csy and TniQ-Cascade complexes occurs from the minor dsDNA groove side (Halpin-Healy et al., 2020). 8 nt of seed sequence and 25-28 nt from PAM are shown to be critical regions for DNA integration (Klompe et al., 2019). R-loop formation mechanism of TniQ-Cascade with extended crRNA by +12 nt is reminiscent

of *S. thermophilus* I-E Cascade. The complex efficiently mediated transposition at WT target site was well as on the extended site shifted by 11-13 nt that corresponds well to the length change on the crRNA, confirming that WT-like target site recognition is sufficient to activate transposase, similarly to I-E type Cascade where WT-length target recognition triggers Cas3-mediated cleavage (**publication 2).**

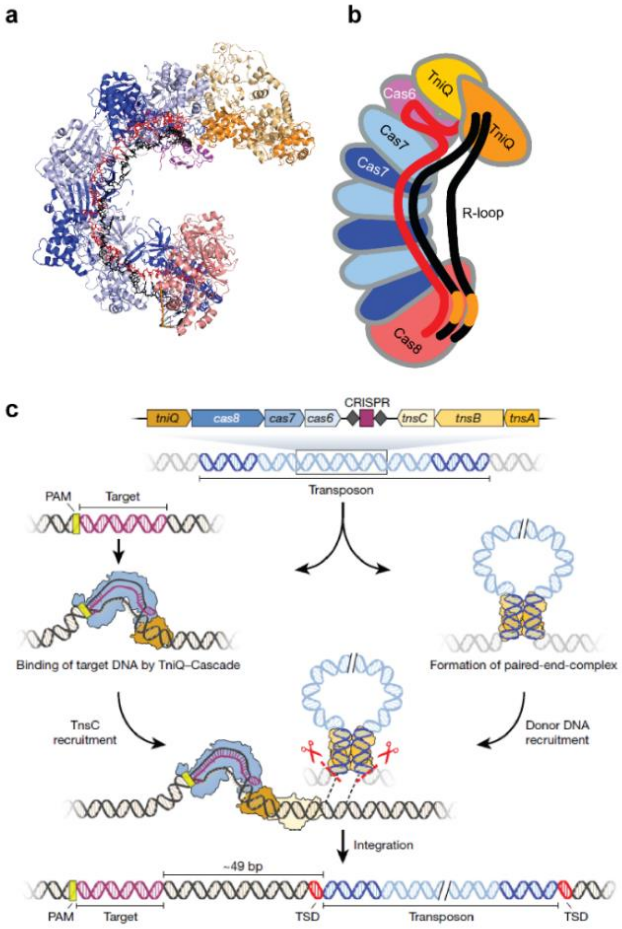


Figure 7. Structure and function of type I-F variant Cascade complex encoded close to transposon. a. Cryo-EM structure of Cascade-TniQ dimer (PDB ID: 6VBW). b. Schematic representation of transposon-encoded Cascade complex, Cas8 (large subunit) in this system is a fuse of Cas5-Cas8. c. Mechanism of action of transposon-encoded Cascade complex from *V. colerae*. Cascade complex with TniQ scans DNA for potential target sites. After full R-loop formation, TniQ presumably recruits TnsC, a non-specific DNA binding protein. TnsC is associated with TnsA and TnsB. Finally, excision of the transposon from its donor site and integration at a fixed length from Cascade-TniQ complex target site occurs by a cut-and-paste mechanism. Adapted from (Klompe et al., 2019).

2.4. Applications of class 1 CRISPR-Cas systems

Class 2 systems that have single protein effectors (Cas9, Cas12, Cas13) were shown to be efficient genome editing tools in a range of organisms from bacteria to human (Pickar-Oliver et al., 2019). Recently, CRISPR-Cas ancestor transposon nucleaseses (IsrB, IshB, IscB, TnpB) are also shown to mediate RNA-guided target DNA cleavage and potentially will be employed as a new generation genome engineering tools (Altae-Tran et al., 2021; Karvelis et al., 2021). Class 2 systems might be also employed for a range of applications, such as virus detection, gene silencing or activation (Hilton et al., 2015; Kellner et al., 2019; Thakore et al., 2015). Applications of class 1 systems as a potential alternative to class 2 are currently emerging. The major applications include gene silencing/activation or large chromosomal deletions. Type I complexes recognize longer targets, usually ranging from 31 bp to 35 bp and contain promiscuous PAM sequences, e.g. St-Cascade complex needs only single nucleotide (A) at 5′-end of the protospacer for DNA recognition and cleavage (Sinkunas et al., 2013). Also the class 1 dsDNA-targeting RNP complexes can be used as tools to target genomic DNA in order to interfere transcription or are employed as fusions to other known effector proteins/domains, such as FokI nuclease, transcription activation domains, activation-induced deaminase (AID) domains. In an elegant approach type I-E Cascade-FokI fusion was used for genome engineering in human cells. The study has shown that PAM-proximal 24-nt complementarity is critical for efficient editing (Cameron et al., 2019). Gene activation by I-E type system was performed in plants (corn *Zea mays)* with Cascade complex fused to transcription activator CBF1 (Young et al., 2019). When both the type I-E RNP complex and the accompanying nuclease Cas3 are used, deletions ranging from 100 bp up to 10 kb upstream to the target site are generated in human cells (Dolan et al., 2019; Morisaka et al., 2019). Even without knowledge about a biochemical mechanism of action, I-D type CRISPR-Cas system from *Microcystis aeruginosa* was used to introduce large genome deletions in tomatoes (Osakabe et al., 2020). Distinctly from I-E Cascade, that introduced large deletions upstream of the target site, I-D effectors introduced deletions on both sides (Osakabe et al., 2020). When I-E Cascade was directed to multiple promoter or ORF sequences, efficient gene silencing occurred in both *E. coli* and *S. typhimurium* (Rath et al., 2014). Downregulation of certain genes by type I-E complex caused dysregulation of metabolic pathways and accumulation of metabolites what can be later employed for metabolite production (Chang et al., 2016). I-E type systems were adapted for phage genome engineering by targeting certain genes of the phage and later

including target-directed sequences for repair (Kiro et al., 2014). A designed phage with a I-E type CRISPR-region was employed to guide Cascade complex to restrict antibiotic-resistance plasmids (Yosef et al., 2015). Type I-B complexes without Cas3-nuclease can also be reprogrammed for gene silencing (Pickar-Oliver et al., 2019). Minimal I-C type complex was employed for large deletions to reduce genome size of *Pseudomonas aeruginosa* (Csörgő et al., 2020). I-F type system associated with transposase enabled crRNA-guided up to 10 kb insertions into *E.coli* genome (Klompe et al., 2019). I-B type CRISPR-Cas system was employed for gene repression in archaea (Stachler and Marchfelder, 2016). Type III systems are under development as RNA interference tools, the RNAi activity has been demonstrated in zebrafish (Fricke et al., 2020).

2.5. Single molecule approaches to study CRISPR-Cas systems

Single-molecule techniques enable real-time observation of target search, base-pairing, DNA unwinding and R-loop formation. In this thesis single molecule magnetic tweezer method was used to gain insights into target recognition by I-E type Cascade complex, therefore this chapter shortly reviews the single molecule techniques used to study CRISPR-Cas systems and summarizes the major findings they revealed. Catalytic activity of Cas9, Cas12 (Cpf1), Cascade-Cas3, Csy complex depends on three critical steps (i) PAM recognition; (ii) R-loop formation (iii) target cleavage. These three steps have been studied by single molecule methods. Different single-molecule approaches that were used in the studies of CRISPR-Cas systems, including:

- **Magnetic tweezers.** Magnetic tweezer method has been employed to study type I-E CRISPR-Cas target recognition in this thesis (**publications 1 and 2**). DNA molecule of interest is attached to the surface by one end and to a magnetic bead by another end, and trapped by twisting magnets (Figure 8a). The force and torque applied to DNA molecule can be measured by the height of the bead (Rutkauskas et al., 2017). Each magnet rotation induces a turn on DNA molecule. From this basic measurement scheme, in which R-loop formation and dissolution can be seen as abrupt DNA length changes or as a shift of the rotation curve (Figure 8a). Magnetic tweezers revealed unidirectional target recognition mechanism starting from PAM by Cascade complex and Cas9 (Rutkauskas et al., 2015; Szczelkun et al., 2014a). For Cas9, the seed sequence (8-12 nt at PAM proximal site) is critical for R-loop formation while even 1-5 bp truncations from the

PAM-distal end are tolerated for the target cleavage (Szczelkun et al., 2014a). In contrast to the Cas9 complex for which unidirectional R-loop unwinding was enough to mediate a cleavage-competent conformation, Cascade performed additional "R-loop locking" step that was triggered by recognition of full R-loop (**publication 1**). Only "locked" conformation was able to attract Cas3 for target degradation (Rutkauskas et al., 2015).

- **DNA curtain assay.** DNA molecules are anchored to a fluid bilayer either by one or both ends and diffusion patterns of proteins are monitored by using total internal reflection fluorescence microscopy (TIRFM) (Figure 8b). Cas9 target recognition starts by screening of potential PAM sequence and is followed by adjacent target destabilization (Sternberg et al., 2014). For *E.coli* Cascade complex diffusion along phage λ DNA showed that Cascade complex find their target mostly by 3D diffusion versus 1D diffusion (Redding et al., 2015). PAM sequences increased the target binding efficiency, but protospacer lacking cognite PAM also can be bound (Redding et al., 2015; Szczelkun et al., 2014a). Noncanonical PAM attracts Cas1-Cas2 adaptation complex that mediates Cas3 interactions with *E. coli* Cascade (Redding et al., 2015). PAM-independent recognition mode as well as primed adaptation has not been demonstrated for I-E Cascade from *S. thermophilus* (unpublished data).
- **In single molecule FRET-TIRF assays** target binding/release is observed in real time by total internal reflection microscopy (TIRFM) employing Förster resonance energy transfer (FRET) between labelled molecules, mostly DNA or proteins. By this method, PAM recognition of Cas9 has been monitored. At least two distinct kinetic modes of PAM search are observed for *S. thermophilus (St)* St-Cas9. Cas9 can dissociate quickly (<0.5s) from PAM without further R-loop formation (quick 3D diffusion) or diffuse locally around PAM site (~2s) (local 1D diffusion) (Globyte et al., 2019). This is in line with the results for type I Cascade complex, obtained by DNA curtain assays confirming that effector complexes faviuor 3D diffusion versus 1D (Redding et al., 2015). In contrast to the Cascade complex, Cas9 does not show critical PAM-distal end verification or R-loop stabilization step, termed "target locking" though correct base pairing at PAM-distal end stabilizes HNH domain in a cleavage-competent conformation (Yang et al., 2018). Based on kinetic analysis, cleavage of the target strand by Cas9 HNH domain occurs first, followed by RuvC cleavage of non-target strand (Gong et al., 2018). Similarly to

Cas9, Cas12 binds target unidirectionally and PAM-distal mutations are not critical for target cleavage. However, distinctly from Cas9, it requires a 17 bp R-loop for both target binding and cleavage. Single molecule FRET assay revealed a PAM search mechanism by 1D diffusion and step-by-step target cleavage starting from the non-target strand (Jeon et al., 2018). Two different target binding modes ('adaptation' and 'interference' modes) by I-E Cascade and Cas3-mediated reeling by 3-nt steps of the non-target DNA strand were detected by FRET-TIRF (Blosser et al., 2015; Loeff et al., 2018) (Figure 8c). Cas3 nuclease-helicase remains bound to Cascade complex upon target DNA unwinding by 3-nt steps and degradation, in approx. 90 nt fragments (Loeff et al., 2018).

- **CHAMP-TIRF analysis** employs used MiSeq chips that are loaded by fluorescent oligonucleotides as markers for DNA cluster and labeled proteins. The mapping of the protein binding positions is detected by TIRFM (Figure 8d) (Jung et al., 2017). The analysis showed that *T. fusca* I-E Cascade complex binds extended PAM sequence up to 6 nt and in addition to tolerance to mismatches at every 6^{th} nt position in the target, can tolerate mismatches every 3^{rd} nt position in the protospacers more than any others (Jung et al., 2017).

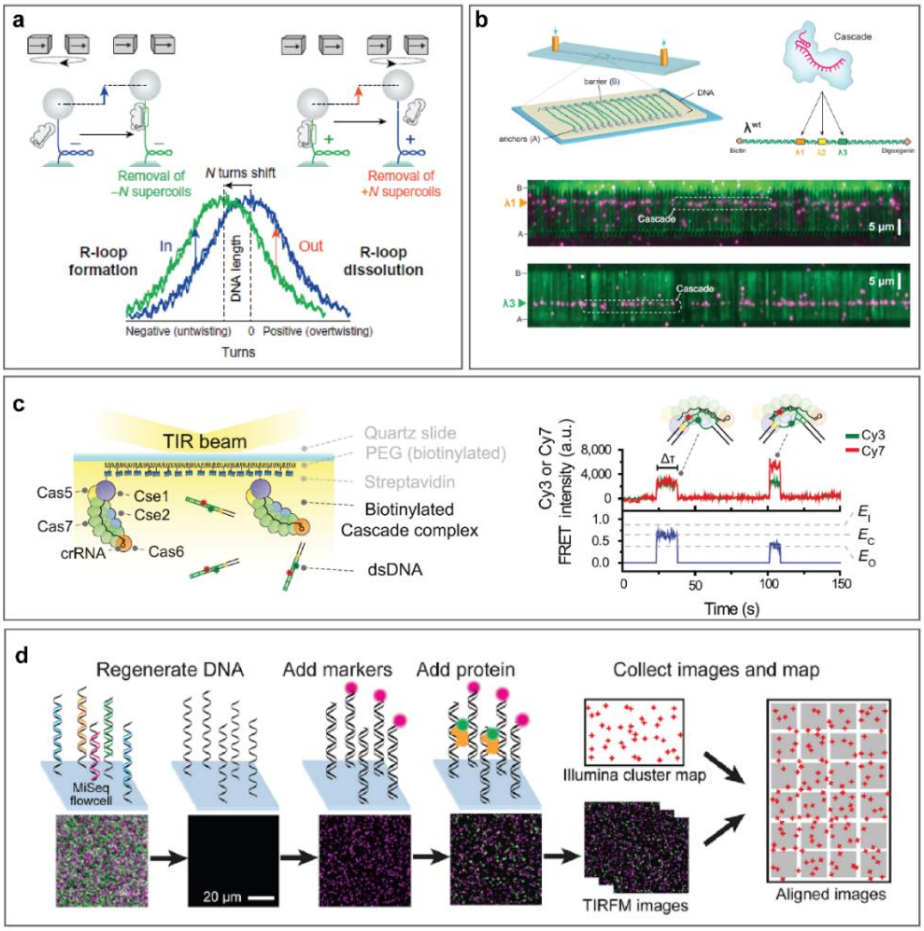


Figure 8. Single molecule techniques applied to study CRISPR-Cas systems. a. Magnetic tweezer assay and DNA length shift upon target unwinding and R-loop formation by Cascade complex, adapted from (Rutkauskas et al., 2017). b. DNA curtain assay with Cascade complex searching for a target on λ-phage sequence, adapted from (Redding et al., 2015). c. Single molecule FRET-TIRF assay for analysis of I-E type Cascade target recognition (Blosser et al., 2015). d. CHAMP-TIRF assay as an approach to study Cascade target recognition, adapted from (Jung et al., 2017).

3. R-loops in eukaryotes and bacteria

Class I CRISPR-Cas systems form R-loops on dsDNA prior to target destruction by Cas3 nuclease-helicase (Szczelkun et al., 2014a). Overall, R-loops are the structures that are formed when dsDNA is unwound and then displaced by RNA, forming a DNA:RNA hybrid and ssDNA (Figure 9a). R-loop structures have been detected in all organisms and play essential roles for gene regulation, immunoglobulin class switch or DNA replication (Groh and Gromak, 2014). I-E Cascade complexes employ an additional R-loop stabilization step, termed "locking" that has been extensively studied in this thesis (Figure 9b) (discussed in **publication 1 and 2**).

Timeline of R-loop research is depicted in Figure 9a. The presence of R-loops has been observed in 1976 by electron microscopy (EM). The revolution in the field was development of an antibody S9.6 that enabled detection of R-loops on microarrays (Figure 9a). Currently, the number of known R-loop functions is constantly increasing, including their role in CRISPR-Cas target recognition, histone modifications, DNA (de)methylation and genome instability (Niehrs and Luke, 2020). The best-characterized functions of R-loops are discussed below.

ColE1 plasmid replication relies on R-loop formation. RNAP forms a 550 bp length RNA II on the leading-strand DNA. RNA II is further processed by RNAse H1 and serves as a primer for DNA Pol I (Figure 9c) (Brantl, 2021). Mitochondrial DNA replication is also R-loop-dependent and is similar to ColE1 plasmid replication (Figure 9d) (reviewed by (Aguilera and García-Muse, 2012). Negative supercoiling arising from RNAP movement could lead to transient R-loop formation. ssDNA therefore can be more sensitive to damage or the R-loop may block further replication fork (RF) progression. R-loops or DNA damage block RF progression and further lead to homologous recombination (HR) or non-homologous end joining (NHEJ) (Aguilera and Gaillard, 2014). Last but not least, R-loops function in Ig class-switch recombination CSR (Figure 9e) in B-cells (Yu et al., 2003). CSR is facilitated by G-rich nature of the non-target strand of Ig switch region that leads to long (>1kb) and stable R-loop formation. R-loop formation induces activity of activation-induced cytidine deaminase (AID). AID deaminates dC to dU in the ssDNA region and so causes mutations and generates the variety of Ig (Figure 9e) (Groh and Gromak, 2014).

To conclude, the R-loop formation either *in cis* (during replication upstream of RNA Pol II) or *in trans* (replication independent) is important for genome stability and integrity. In the future, R-loop formation may be

employed for therapeutic purposes or gene regulation (Niehrs and Luke, 2020).

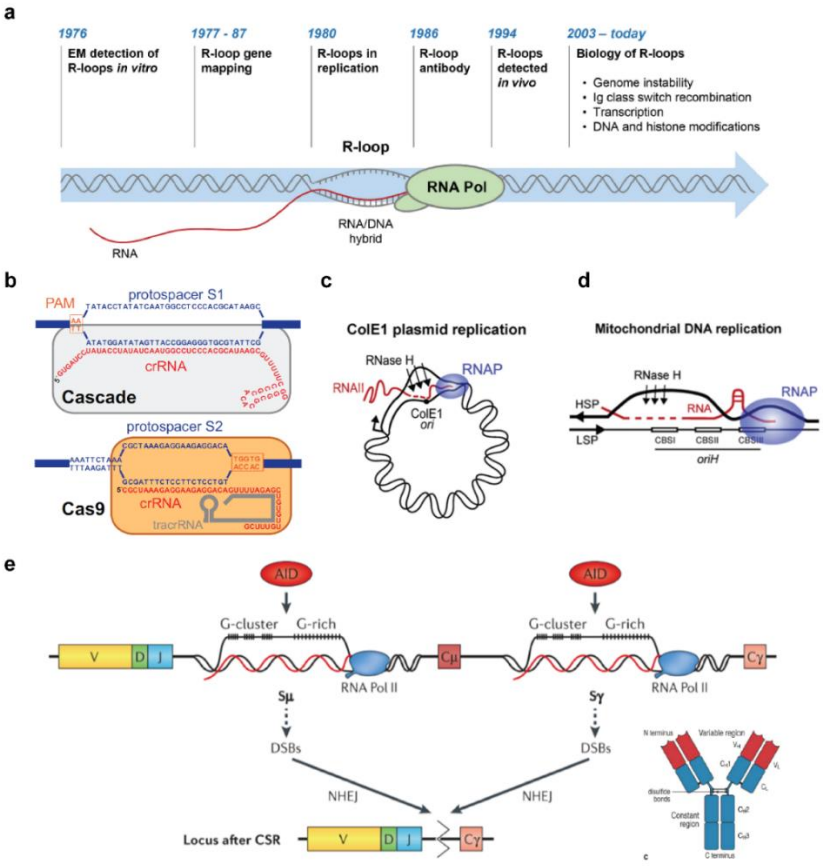


Figure 9. R-loops – history and main functions. a. A timeline of R-loop research, adapted from (Groh and Gromak, 2014). b. R-loop formed by type I-E Cascade complex and type II Cas9 complex, adapted from (Szczelkun et al., 2014a). c. ColE1 plasmid replication is mediated by R-loop formation in ColE1 *ori* region. RNAP produce 550 bp length RNA II on the leading strand. The RNA II is further processed by RNase H1 and serves as a primer, adapted from (Aguilera and García-Muse, 2012). d. Mitochondrial DNA replication mediated by R-loop formation, adapted from (Aguilera and García-Muse, 2012). e. R-loop mediated Ig class-switch repair, adapted from (Santos-pereira and Aguilera, 2015).

4. Classificiation and functions of toxins-antitoxins

Toxin-antitoxin (TA) systems are found in plasmids or chromosomes and are composed of stable toxic entity and its labile counteracting antitoxin (protein or RNA). The toxins interfere with key cellular processes, such as membrane homeostasis, replication, transcription or translation (Berghoff and Wagner, 2017; Hall et al., 2017; Harms et al., 2018). Currently, there are eight types of the TA systems proposed, with six major types (Figure 10) and two additional minor types that recently emerged (Song and Wood, 2020; Wang et al., 2020). The most abundant TA is type II where both toxin and the antitoxin are proteins forming a tight complex (Fraikin et al., 2020). TA systems are found as a part of mobilome and are frequently transfered by horizontal gene transfer (Goeders and Van Melderen, 2014). The biological function of most TA systems is unknown, though the stabilization of mobile elements through post-segregational killing, abortive infection has been observed or persister formation role has been proposed (Hall et al., 2017; Harms et al., 2018; Jurėnas and Van Melderen, 2020). RNases dominate among toxins and mRNA, tRNA, rRNA among the toxin-targets. RNA cleaving toxins may act through ribosome-dependant or independant pathways (Page and Peti, 2016). Types of TA systems and major mechanisms of their action are described below:

- **Type I** contains non-coding RNA antitoxin and protein toxin. Non-coding RNA binds to translation initation site of the toxin and thus through the base-pairing inhibits its translation. The most well known example is hok/sok TA system found on plasmid R1. The *hok* (*host killing*) expression depends on expression of *mok* (*modulation of killing*) mRNA. Sok (*supressor of killing*) antitoxin binds to *hok* mRNA and supress expression of both *mok* and *hok* (Thisted et al., 1994). This RNA duplex is later cleaved by RNase III (Berghoff and Wagner, 2017; Gerdes et al., 1992). Hok/sok TA system is active excluding T4 phage (Pecota and Wood, 1996).
- **Type II** TA systems are the most abundant in bacteria and most well studied among TA systems. They are encoded in the same operon and form protein-protein TA complexes where antitoxin inhibits the toxin. Majorly the toxins target RNA and impair protein translation. The most well studied example is mazEF TA pair with MazF toxin, that cleaves mRNA, tRNA or rRNA (Culviner and Laub, 2018; Schifano et al., 2013, 2014). Chromosomal MazF can protect cells from P1 phage infection (Kamruzzaman et al., 2021). VapC toxins cleave tRNAs or sarcin-ricin

loop in 23S rRNA (Winther et al., 2016). Type II systems are active against phages: RnlA and LsoA degrade most of the T4 mRNA to block phage propagation. To avoid TA systems, T4 phages developed an antitoxin-like protein that is able to block activity of RnlA and LsoA (Kamruzzaman et al., 2021). Most of type II antitoxins contain DNA-binding domain to regulate the expression of the TA operon (Jurėnas and Van Melderen, 2020).

- **Type III** systems encode pseudoknot RNA ToxI (*ToxN inhibitor*) that directly binds to toxin thus forming RNA-protein contacts and inhibiting the toxic endoRNase ToxN that prior process the ToxI. Three ToxI RNA molecules bind three ToxN proteins to form a heterhexameric triangular 3:3 complex (Blower et al., 2011; Short et al., 2013). ToxIN system protects against phages T4, T2, T5 and T6. ToxI gets degraded upon phage infection and so ToxN accumulates. ToxN cleaves phage mRNA at the later infection stage and thus protect neighbouring cells from phage burst (Guegler and Laub, 2021).
- In **type IV** systems toxin and antitoxin act though conteracting one another. The antitoxin blocks the toxin action on its target. YuuE (CbeA) antitoxin enhances cytoskeleton building antagonizing activity of YeeV. YeeV binds and inhibits polymeriziation of cell cytoskeleton proteins MreB and FtsZ (Masuda et al., 2012).
- In **type V** systems antitoxin is RNase that cleaves the toxin mRNA. Under stressful conditions, mRNA of the antitoxin is degraded and the toxin (a small hydrophobic peptide that interferes with cell membrane) is produced (Wang et al., 2012).
- **Type VI** systems are represented by SocA/SocB pair. The toxin SocB binds β sliding clamp DnaN and inhibits replication elongation. The antitoxin ensures the proteolytic degradation of the toxin by ClpXP protease (Aakre et al., 2013).
- **Type VII** systems represent TA systems where the antitoxin enzymatically inhibit the toxin, e.g. TomB antitoxin oxidizes Cys of the toxin Hha or antotoxin TakA phosphorylates toxin TglT thus inactivating it (recently reviewed by (Wang *et al*., 2020).
- **Type VIII** is the firt example where the both toxin and antitoxin are small RNA (Choi et al., 2018). sRNA SdsR mediates Hfq-dependant cell death while ryeA sRNA is encoded in opposite direction to SdsR and bloks SdsR by direct binding.

  Atypicl TA systems include tripartite TA systems composed of reverse transcriptase and retron msDNA as antitoxin and additional effector

domain as toxin, e.g. cold anaerobic protein (Bobonis et al., 2020a). Bacterial retrons are shown to function as antiphage defense systems (Bobonis et al., 2020b; Millman et al., 2020). Cascade with noncanonical crRNA termed creA was recently shown to block the promoter of the toxic small RNA creT. In creA-creT pair both toxin and antitoxin are small RNA. Small RNA CreA binds to Cascade complex and guides it to creT promoter to repress transcription of creT. Active toxin creT sequesters rare cellular tRNA-Arg (Li et al., 2021). Such a co-localization of the TA system with CRISPR-Cas locus is thought to safeguard the Cas genes or might be helpful in case of action of anti-CRISPR proteins that bind to Cascade complex. In such a case, the transcription of creT would not be blocked by Cascade and creA might induce dormancy or cell death (Li et al., 2021).

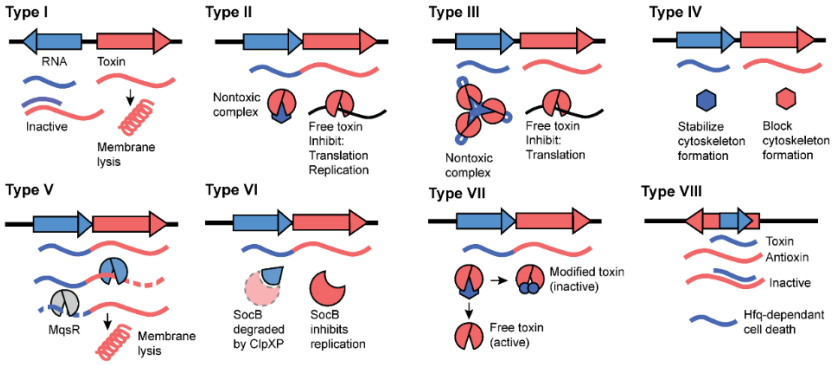


Figure 10. Classification of toxin-antitoxin systems.

5. AMPylation: from bacteria to eukaryotes

This chaper is presented in the thesis due to discovery of toxin-antitoxin system in I-D CRISPR-Cas operon that performs di-AMPylation to block the activity of HEPN RNase toxin (**publication 5**). Overall, covalent protein modifications may occur to block the activity of a new type VII toxins when antitoxin modifies the toxin so blocking its function, recently reviewed by (Wang *et al*., 2020).

Covalent modifications are versatile tools to regulate structural and functional flexibility of enzymes and therefore are present in all kingdoms of life.

AMPylation is a protein covalent modification that occurs when ATP molecule is used to transfer AMP group to -OH group of Ser, Thr or Tyr (Figure 11a). Currently known AMPylating or UMPylating proteins and their physiological functions are listed in the Table 3.

In late 1960s, AMPylation (adenylylation) has been discovered as a reversible post-transcriptional mechanism to regulate the activity of glutamine synthetase (GS) GlnA (Brown et al., 1971; Kingdon et al., 1967). GlnA is responsible for synthesis of L-glutamine from ammonia and glutamate. A regulating enzyme glutamine synthetase adenlytransferase GlnE (PII) attaches/removes AMP to/from tyrosine Tyr397 of GlnA, thus switching-off/on the activity of GlnA (Brown et al., 1971). This modification is flexible and can be removed by the NT-removing activity of GlnE as a response to nitrogen concentration. Moreover, GlnE itself is regulated by UMPylation performed by GlnD. The level of UMPylation is regulated by glutamate and α-ketoglutarate. UMPylated form of GlnE is inactive in AMPylation of GlnA (Figure 11b) (Jiang et al., 2007).

In 2009, AMPylation was “re-discovered” when it has been shown that bacterial pathogen *Vibrio parahaemolyticus* nucleotidyltransferase Fic domain containing protein VopS can AMPylate human Ras GTPase to block immune response. VopS prevents Rho GTPase interaction with downstream effectors thus leading to impaired actin cytoskeleton assembly and cell rounding (Yarbrough et al., 2009). AMPylation as toxic activity is observed in Fic family of toxins FicT. FicT AMPylates two crucial cellular proteins – ParE subunit of DNA topoisomerase and GyrB subunit of DNA gyrase (Engel et al., 2012; Harms et al., 2015). Fic (Fido) domains also perform UMPylation, phosphorylation and phosphocholination (Cruz and Woychik, 2014). Doc (*death on curing*) toxin exhibits Fic domain, but in contrast to FicT, Doc functions as a kinase and perfoms phosphorylation of elongation factor EF-Tu on a conserved Thr382 residue (Cruz et al., 2014).

Other pathogenic bacteria *Legionella pneumophila* creates replication vacuole in host cells. *Legionella* toxin DrrA (SidM) uses N-terminal Polβ-like nucleotidyltransferase (NT) domain to AMPylate Tyr of Rab1 cellular GTPase and to block interaction with GTPase activating proteins and trafficking to Golgi. AMPylation of Rab1 leads to constant activity of Rab1 (Muller et al., 2010). AMPylated Rab1 is later de-AMPylated by bacterial SidD phosphatase as the infection proceeds (Neunuebel et al., 2011). Another example of toxin-antitoxin activity includes type II Polβ-like nucleotidyltransferase bacterial toxin MenT that uses CTP or UTP to modify tRNA (Ser, Trp) 3′-acceptor stems thus further blocking aminoacylation (Cai et al., 2020).

In 2018, a pseudokinase selenoproetin-O SelO, was shown to AMPylate Tyr residues on glutaredoxin grxA and α-ketoglutarate dehydrogenase sucA, proposing that AMPylation may regulate redox homeostasis (Sreelatha et al., 2018).

To conclude, so far three families of protein domains are known to perform AMPylation: adenyltransferase domains (ATase) (Polβ-like superfamily enzymes, active site motif G-X_{11}-D-X-D), Fic domains (active site motif HPFx[D/E]GN[G/K]R) and recently discovered pseudokinases SelO (Casey and Orth, 2018; Sreelatha et al., 2018). Polβ-like superfamily enzymes include DNA X-family polymerases, primases, DNA ligases, DNA-dependent RNA polymerases, poly(A) polymerases, CCA-adding, mRNA capping enzymes and are involved in all fundamental cellular processes (Figure 11c) (Aravind and Koonin, 1999a). Nucleotidyltransferases adopt mixed β-sheet flanked by 4 α-helices with αβαβαβα topology and act without template (Figure 11d).

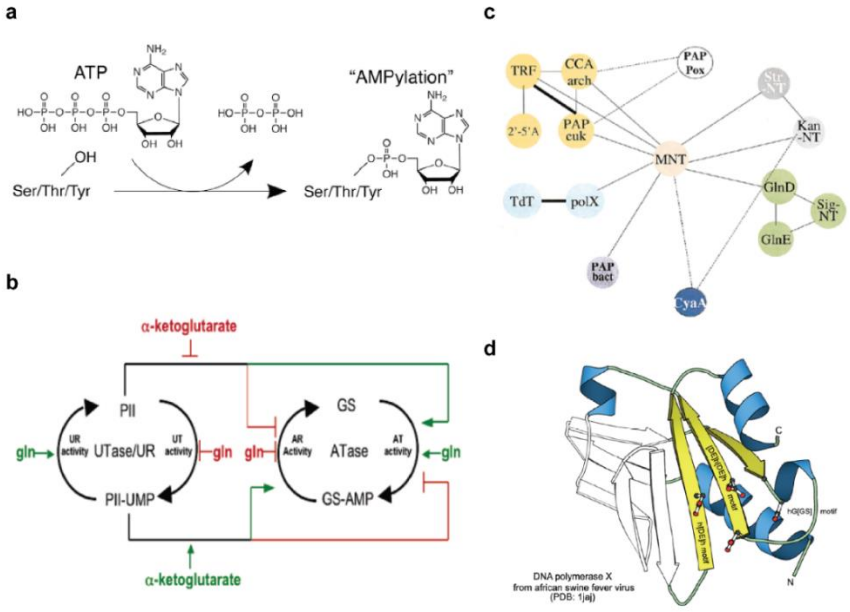


Figure 11. AMPylation and nucleotidyltransferases. a. General mechanism of AMPylation reaction. b. Bicylic regulation of glutamine synthetase activity by AMPylation regulated by glutamine and α-ketoglutarate, adapted from (Jiang et al., 2007). c. Polβ-like superfamily: MNT – minimal nucleotidyltransferase, CCA – CCA adding enzyme, TdT – terminal deoxynucleotidyltransferase, PAP – poly(A) polymerase, TRF – TRF proteins, 2′-5′ A – 2′-5′ oligoadenylate synthetase, adapted from (Aravind and Koonin, 1999a). d. Structural core (αβαβαβα topology) of nucleotidyltransferases, conserved amino acids are shown in sticks, adapted from (Kuchta et al., 2009).

Table 3. AMPylators and UMPylators and their targets.

AMP(UMP)ylator	**Active site**	**Bacteria**	**AMPylation/UMPylation target**	**Function**	**Ref.**
GlnE	Polβ-like	*E. coli and others*	Glutamine synthetase GlnA	On/Off states of GS-ATase activity	(Brown et al., 1971; Kingdon et al., 1967)
DrrA (SidM)	Polβ-like	*Legionella pneumophila*	Rab1 GTPase Y77	Block cellular vesicle trafficking	(Muller et al., 2010)
FicT toxin	Fic	*Bartonella schoenbuchensis, Yersinia enterocolitica, Pseudomonas aeruginosa*	ParE topoisomerase, GyrB gyrase	Toxin-antitoxin system	(Harms et al., 2015)
VopS	Fic	*Vibrio parahaemolyticus*	Rho GTPase Cdc42	Impair cytoskeleton assembly and cause cell rounding, block immune response	(Yarbrough et al., 2009)
IbpA	Fic	*Histophilus somni*	Rho GTPases (RhoA, Rac, Cdc42)	Causes collapse of host cell cytoskeleton	(Worby et al., 2009)

HYPE	Fic	*Homo sapiens*	BiP (ER chaperone)	Increases BiP activity. BiP refolds unfolded proteins in ER.	(Sanyal et al., 2015)
Bep2	Fic	*Bartonella rochalimae*	AMPylation of filament protein vimentin	Possibly block immune response	(Woolery et al., 2010)
GlnD	Polβ-like	*E. coli and others*	UMPylation of GlnE	Regulation of glutamine synthetase	(Brown et al., 1971)
AvrAC	Fic	*Xanthomonas campestris*	UMPylation of BIK and RIBK in *Arabidopsis thaliana*	Inhibit plant immunity	(Feng et al., 2012)
SelO	Pseudokinase	*Pseudomonas syringae, E. coli*	AMPylation of GrxA glutaredoxin, SucA α-ketoglutarate dehydrogenase	Regulation of redox homeostasis	(Sreelatha et al., 2018)
YdiU	Pseudokinase	*Salmonella*	UMPylation of DnaK and GroEL Self-AMPylation	UMPylation of chaperones negatively regulate their activity	(Yang et al., 2020)

6. Regulation of HEPN RNase domains

This chapter is presented in this thesis due to analysis the HEPN toxin, present near I-D type CRISPR-Cas operon (**publication 5**). We aimed to understand the mechanism that switch-on activity of HEPN RNase and therefore general mechanisms known to regulate HEPN RNase activity are reviewed here.

HEPN (**H**igher **E**ukaryotes and **P**rokaryotes **N**ucleotide Binding) domains are present in proteins in all kingdoms of life. The active HEPN domains form dimers and contain the conserved $R\Phi x_{4\text{-}6}H$ (Φ – polar amino acid, x – any amino acid) active site motif (Anantharaman et al., 2013a). HEPN domains perform metal-independent catalysis leaving 5′-OH and 2′,3′-cyclic phosphate (Pillon and Stanley, 2018). Most of the details on the enzymatic action of HEPN domains came from structural studies of type VI CRISPR-Cas proteins Cas13 (also known as C2c2) that contain two HEPN domains: the first HEPN domain acts in pre-crRNA maturation to functional crRNA and the second HEPN performs ssRNA *cis*- and *trans*-cleavage once protospacer containing ssRNA is bound (Figure 12a) (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Smargon et al., 2017; Yan et al., 2018). crRNA recognise ssRNA by base-complementarity forming a continuous A-form RNA:RNA duplex (Liu et al., 2017b; Zhang et al., 2018). Some of Cas13 (Cas13a, b) need PFS- protospacer flanking sequences at 3′-end of ssRNA to discriminate self nucleic acids from non-self, while Cas13d and *Leptotrichia bucallis* LbuCas13a does not have such sequence (Konermann et al., 2018; Liu et al., 2017b; Zhang et al., 2019). Extended complementarity between crRNA and target ssRNA at its 3′-end abolishes the active site formation and cleavage (Wang et al., 2021). The Cas13 proteins have central ssRNA seed regions and 20 nt ssRNA is needed to activate *trans*-RNase (Liu et al., 2017b). Activation occurs by introducing drastic or slight structural rearrangements of HEPN active site, e.g. His477 shifting by 12 Å and Arg472 by 6 Å to form an active site cleft in LbuCas13 (Figure 12b). Slight rearrangements (2-3 Å shift of catalytic helix) are observed in *Eubacterium siraeum* EsCas13d (Figure 12c) (Zhang et al., 2018).

Other HEPN proteins are generally regulated by ligand binding. CRISPR-Cas ancillary proteins Csm6 and Csx1 contain CARF-ligand cOA binding and HEPN RNase domains (Kazlauskiene et al., 2017; Niewoehner et al., 2017). Biochemical studies revealed ligand-mediated HEPN RNase activation, though structural studies are still unable to resolve HEPN domain in on- and off- states (Garcia-Doval et al., 2020; Molina et al., 2019). HEPN

domain containing protein Ire1 is located in the membrane of endoplasmic reticulum and senses misfolded proteins. Misfolded proteins induce *trans*-autophosphorylation of Ire1 followed by HEPN RNase activation (Pillon et al., 2020a). HEPN protein RNase L dimerizes upon binding of 2′-5′ oligoadenylate synthetized by 2′-5′ oligoadenylate synthetase induced by interferon signaling (Huang et al., 2014). Grc3/Las1 performs pre-rRNA maturation from single 60S pre-rRNA transcript to 7S pre-rRNA and 26S pre-rRNA (Pillon et al., 2017). Grc3/Las1 is a heterotetramer of dimeric Grc3 polynucleotide kinase and dimer of Las1 HEPN ribonuclease. Grc3 and Las1 works sequentially: Las1 cleaves 35S pre-rRNA to leave 5′-OH on 26S pre-rRNA that is later phosphorylated by Grc3 (Pillon et al., 2018). Histidine and conserved threonine in the active site provide a switch between active and inactive forms of Grc3/Las1 (Pillon et al., 2018, 2020b). HEPN domains are also present in type II TA system RnlA/RnlB that restricts *dmd*$^-$ mutant of T4 phage. Toxin RnlA contains HEPN and TBP-domains and is inhibited by RnlB antitoxin directly binding a canonical HEPN dimer. Yet, RnlB does not directly bind an active site groove of RnlA, it is positioned adjacent to it and blocks entrance/exit of the RNA substrate (Garcia-Rodriguez et al., 2021).

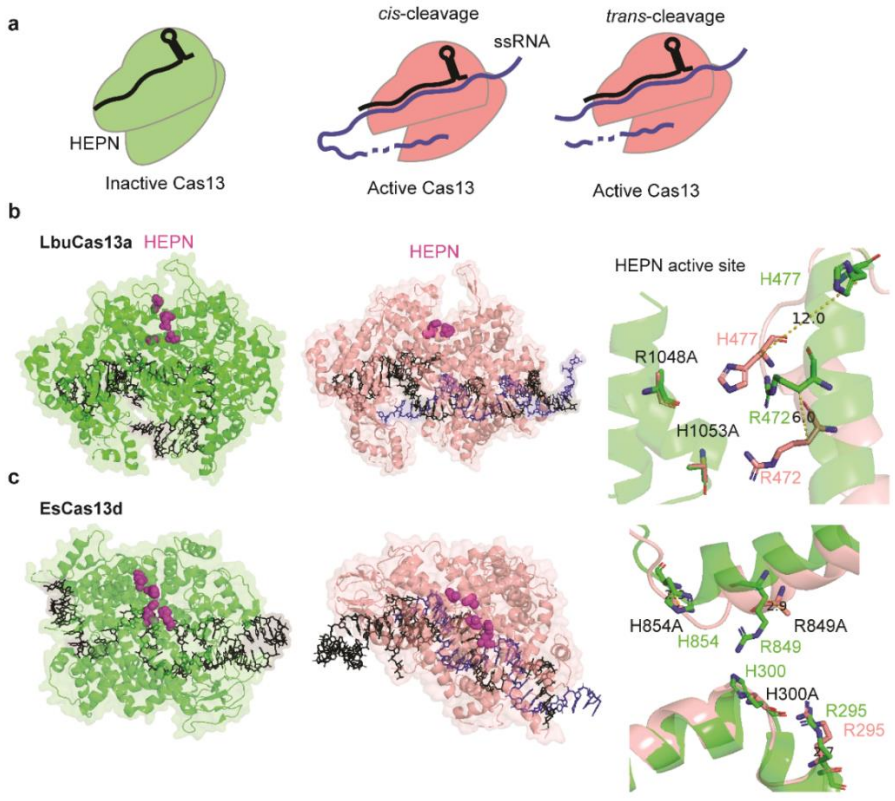


Figure 12. a. *Cis*- and *trans*- ssRNA cleavage activities of Cas13. Cas13 cleavage is activated by target RNA binding to crRNA and might be executed by cleaving the same RNA molecule (in *cis*) or other RNA target (in *trans*). b. LbuCas13a conformational rearrangement of HEPN domain in inactive state (green, PDB ID:5XWY) versus active state (salmon, ssRNA-target bound PDB ID: 5XWP). c. EsCas13d conformational rearrangements of HEPN domain in inactive state (green, PDB ID: 6E9E) versus active state (salmon, ssRNA-bound, PDB ID:6E9F).

7. Results

7.1. Unidirectional R-loop zipping by I-E type Cascade (publication 1)

Based on phage escape mutants and previous single molecule analysis (Redding et al., 2015; Semenova et al., 2011; Szczelkun et al., 2014a), unidirectional model of R-loop formation starting from the PAM sequence has been proposed. This model might explain the importance of the seed sequence and elucidate Cascade dissociation events from mutated targets. This may allow us hypothesize how phages overcome CRISPR-Cas immunity and give useful insights for applications of type I Cascade complexes. To monitor the target recognition pathway and elucidate separate single molecule target-recognition events, we applied previously developed magnetic tweezers-based DNA supercoiling assay (Szczelkun et al., 2014a). Initially, we applied the assay (described in Chapter 2.5, Figure 8a) to monitor R-loop formation on protospacers bearing single point mutations of the WT target. This included mutations in the seed region at positions 2, 4, and 8 as well as in the remaining protospacer sequence at positions 15, 18, 19, 20, and 22 bp from 5′-PAM (later labeled as M2, M8, etc.). Changes in supercoiling accompanying formation of full R-loop formation vs mismatch position increased linearly with PAM and mismatch distance, thus confirming that R-loops are initiated at the PAM position and expand unidirectionally towards the protospacer end. Shorter R-loops M15, M18 collapsed without overcoming mismatch position, while M20 or M22 formed full R-loops without dissociation. To test how length-dependent dissociation of the R-loops enables more efficient search of matching targets, we employed competition-based EMSA assays with labelled WT target and unlabelled competitor target with a mismatch. Seed mutants M2, M4 and M8 are the least efficient competitors to WT target, while mutations further to the PAM sequence, both single mismatches (M15, M18, M19, M22) or sequence truncations at protospacer end (termed E24, E26, E28, E30), are more efficient competitors. R-loop “locking” is abolished for E24, E26 and E28, therefore competitive binding may occur through formation of intermediate R-loops. Single molecule experiments with non-functional PAM sequence CC proved that both fully formed locked R-loop and PAM verification are crucial for Cas3-mediated target cleavage. DNA cleavage in bulk of supercoiled and linear DNA was investigated at 25°C and 37°C. Both WT targets were cleaved, E26 target with abolished R-loop locking was not cleaved. Supercoiled, but not linear, forms with M2 and M22 mutations were cleaved at 37°C, minor degradation was observed for M22 plasmid at 25°C.

Negative supercoiling or elevated temperature enabled R-loop formation for mismatch targets.

Main findings of the publication 1:

- I-E type Cascade forms R-loop unidirectionally starting from PAM
- R-loop collapse probability is higher when mismatch is closer to PAM sequence
- Target cleavage requires a correct PAM sequence and locked R-loop configuration (verification of PAM-distal end)

7.2. Target recognition by I-E type Cascade complexes (publication 2)

Cascade complexes contain modular architecture, thus indicating that the composition of the complexes might be altered. As the 3′-end of the protospacer is crucial for the target locking and further degradation, next we asked wheater this locking position might be shifted by altering the length of the crRNA. In this work, we initially raised the question whether Cascade complex is able to accommodate altered length crRNA and how the complexes form the R-loops. To this aim, we produced *S. thermophilus* (St) St-Cascade complexes containing crRNA spacers of altered length ranging from 15–57 nt and monitored R-loop formation using magnetic tweezer assay described in **publication 1**. The results showed that the complexes with crRNA of altered length can successfully form R-loops that extended over the full protospacer. Complexes for which the spacer length differed by even multiples of 6 nt from the WT (i.e., -12, +12, and +24 nt) were found to be stable and homogeneous according to native mass spectrometric analysis. In contrast to reports for Ec-Cascade (Kuznedelov et al., 2016; Luo et al., 2016), complexes for which the crRNA spacer length differed by odd multiples of 6 nt (i.e., -18, -6, +6, and +18) were found to be unstable, with the dominating population being of significantly reduced molecular weight. The central aim of this study was to elucidate the target recognition and R-loop locking by these complexes. R-loop locking and Cas3-mediated target cleavage occurred when an R-loop could extend over more than 28 bp independent of the spacer length of the complex. Despite the common minimum R-loop length that triggered locking for all complexes, R-loop stability was found to be considerably increased for elongated spacers compared with the WT spacer. However, no clear correlation of R-loop stability with the R-loop or the spacer length was observed. This provides evidence that RNA:DNA base-pairing is

not the major determinant of R-loop stabilization for I-E type St-Cascade complex.

Main findings of the publication 2:

- I-E type St-Cascade forms stable complexes on crRNA that differs by ±12 nt from WT crRNA-length, but not by ± 6 nt
- The altered complexes form R-loops and trigger DNA-cleavage by Cas3 once WT-length R-loop is formed
- RNA:DNA length is not a major determinant of R-loop stability in I-E type St-Cascade complex

7.3. Complex formation of I-E type CRISPR-Cas system (publication 3)

Even though the composition of the stable I-E type Cascade complexes from *E. coli* with WT or altered length crRNA has been studied and the data showed that Ec-Cascade can form stable complexes on both ±6 nt and ±12 nt crRNA (Luo et al., 2015), St-Cascade complexes were stable only when ±12 nt increment was introduced. Little was known about the assembly and composition of the Cascade complexes, therefore, we addressed this question by native mass spectrometry. Here, we analysed the same Cascade complexes with spacers ranging from –18 nt to +24 nt from WT crRNA used in **publication 2** for single molecule study. Combining the results from various complex species, we were able to propose the mechanism of the assembly pathway for I-E type St-Cascade complex. The major peak distribution in the native mass spectrum of the wild-type complex corresponds to a mass of 422 242 Da (Figure 3a) representing the intact St-Cascade complex with a stoichiometry of $(Cse1)_1(Cse2)_2(Cas7)_6(Cas5)_1(Cas6)_1(crRNA)_1$, same composition has been observed for homologous I-E Cascade complex from *E. coli* (Hayes et al., 2016; Jackson, 2014; Mulepati et al., 2014). St-Cascade complexes assembled on crRNAs that differ in 12 nt units when compared with the wild-type complex, i.e., crRNA spacer length of +12, +24, and −12 nt. For these complexes incorporation or loss of two Cas7 and one Cse2 copies per 12 nucleotides was suggested (Luo et al., 2015). The native mass spectrum of the St-Cascade complex assembled on 12 nt shorter crRNA showed three species: one major species, which comprises the expected stoichiometry of $(Cse1)_1(Cse2)_1(Cas7)_4(Cas5)_1(Cas6)_1(crRNA)_1$ and two subcomplexes thereof, which lost the peripheral large subunit Cse1 as well as Cas7 and Cse2 subunits. Loss of Cse1 and Cse2 subunits was also observed for subcomplexes of the wild-type St-Cascade complex. In agreement to that, St-Cascade

complexes that were assembled on crRNA which is 12 nt longer than wild-type crRNA, one Cse2 and two Cas7 subunits were incorporated. The mass spectra of St-Cascade complexes assembled on crRNAs with decreased or increased spacer length (−12, +12, and +24 nucleotides) confirmed the incorporation or loss of the $(Cas7)_2(Cse2)$ module for each 12 nt unit. Increasing the crRNA spacer length by 6 nt yielded heterogeneous populations of St-Cascade complexes presumably due to space constraints on the crRNA. The preferred assembly included incorporation of one copy Cse2 and one copy Cas7 accompanied by the loss of Cse1. Cse1 is responsible for PAM recognition, therefore these complexes showed lower activity in molecular tweezer and biochemical assays (**publication 2**).

Major findings of the publication 3:

- Native mass spectrometry allows characterization of heterogeneous populations of CRISPR-Cas ribonucleoprotein complexes
- The $(Cas7)_2(Cse2)_1$ module is preferably incorporated or lost per 12 nt increment or decrement in crRNA
- When elongating/shortening crRNA by 6 nt, a stable Cse2 backbone cannot be formed causing dissociation of other protein subunits, preferably Cse1, what correlated with reduced activity of the Cascade complexes to form R-loops

7.4. R-loop formation and interference by I-F type Csy complex (publication 4)

Here, we examined the mechanism of DNA interference provided by the type I-F1 system from *Aggregatibacter actinomycetemcomitans D7S-1* (Aa). I-F type CRISPR-Cas systems differ from I-E by the absence of the small subunits that have critical role for R-loop locking in type I-E (Chapter 2, Figure 3, Figure 4). The presence of a single large subunit Csy8f and its 180º C-terminal rotation, indicated that R-loop locking may occur by a different mechanism in comparison to I-E type Cascade (Figure 4d). Therefore, we have constructed and purified I-F Csy complexes that contained 14-144 nt spacers in crRNA and checked their ability to form R-loops. In line with type I-E complex formation, we show that functional type I-F Aa-Cascade complexes can be assembled not only with WT spacer of 32 nt but also with shorter or longer (14–176 nt) crRNA. All complexes guided by the crRNA bind to the target DNA sequence (protospacer) forming an R-loop when a single C nucleotide or CT dinucleotide motif (PAM) is present

immediately upstream of the protospacer. The length of crRNA determined the length of the formed R-loop, however, same as observed for I-E type Cascade complex, WT protospacer length was a minimal sufficient R-loop length to provide Cas3-guided target interference. Next the I-F Cascade complexes with or without Cas2/3-nuclease were checked for their ability to restrict phage M13mp18 or multicopy pUC plasmids with protospacer and functional PAM. The complexes that contained WT-length crRNA or longer crRNA restricted phage by similar order of magnitude. When crRNA was shortened by 6 nt or 12 nt, the ability of I-F Cascade+Cas2/3 to restrict plasmid was impaired, however, type I-F Csy complex could still provide phage interference by two orders of magnitude. Therefore, we have proposed that I-F type Cascade employs a similar mode of action as I-E type St-Cascade complex, where PAM-distal end is important for target verification and triggers Cas3-mediated degradation.

Major findings of the publication 4:

- I-F type Csy complex is able to form R-loops ranging from 14 nt to 144 nt
- I-F type complex verifies its target by PAM-distal mode
- I-F type Csy complexes with shorter crRNA (−6 and −12 nt) are able to restrict phages, but not a multicopy plasmid

7.5. OligoAMPylation in HEPN-MNT toxin-antitoxin system (publication 5)

Cyanobacteria *Aphanizomoneon flos-aquae* encodes a single I-D CRISPR-Cas system and two type III-B Cmr systems (Šulčius et al., 2015). We noticed that the operon of type I-D CRISPR-Cas system encodes a putative type II toxin-antitoxin system. Toxin-antitoxin TA systems are abundant in defense islands, though co-operation between TA and CRISPR-Cas system has not been demonstrated yet. Initial bioinformatical analysis of the TA system has detected minimal nucleotidyltransferase Polβ-like domain and HEPN domain that is found in RNases (Makarova et al., 2009). Furthermore, it has been experimentally demonstrated that MNT-HEPN pair from *Shewanella oneidensis* functions as a proposed type II TA system (Yao et al., 2015). Polβ-like superfamily enzymes include DNA polymerases, primases, DNA ligases, DNA-dependent RNA polymerases, poly(A) polymerases, CCA-adding, mRNA capping enzymes and are involved in all fundamental cellular processes (Aravind and Koonin, 1999a). HEPN domains contain the conserved $R\Phi x_{4\text{-}6}H$ (Φ – polar amino acid, x – any amino acid)

active site motif and HEPN proteins typically implement suicidal or dormancy induced strategies, are involved in eukaryotic immunity and employed for foreign nucleic acid cleavage, e.g. RNase L in eukaryotes (Huang et al., 2014; Pillon et al., 2018, 2020a). The domains might have been encoded by fast-evolving mobile elements involved in biological conflicts in prokaryotes, they are commonly found associated with Abi, CRISPR-Cas, restriction-modification or Pgl systems (Anantharaman et al., 2013). The molecular mode of action of HEPN-MNT TA has not been known, as well as its possible co-operation mechanisms with CRISPR-Cas systems. We initially aimed to characterize the putative TA system biochemically. First, we cloned HEPN toxin and MNT antitoxin with His-tags and tried to co-purify them. Both HEPN and MNT purified as standalone proteins, but not as a complex as observed for *Shewanella oneidensis* MNT-HEPN system (Yao et al., 2015). Mass-spectrometric analysis of HEPN toxin indicated its mass increase by 659 Da, which was consistent with a potential di-AMPylation of the toxin. The modification was later confirmed by mutagenesis and structural analysis of HEPN toxin. The structural studies revealed how the di-AMPylation interfered with the activity of the toxin. MNT modified conserved Tyr close to the HEPN RNase active site motif and by doing so changed conformation of the active site loop. At the end, we aimed to elucidate the cellular targets of the active HEPN toxin. To address this question, we expressed the unmodified toxin in *E. coli* cell-free *in vitro* translation mixture and sequenced the RNA from the mixture. The analysis revealed that unmodified HEPN cleaved 4 nt from tRNA 3′-end.

Major findings of publication 5:

- MNT nucleotidyltransferease neutralizes HEPN RNase toxin by covalent di-AMPylation
- Di-AMPylation changes conformation of an active site loop
- HEPN cleaves 4 nt from 3′-end of tRNA

OVERVIEW OF THE RESULTS

In this thesis, class 1 type I CRISPR-Cas systems and related (ancillary) genes forming a toxin-antitoxin system were studied. Class 1, type I systems are different from class 2 due to formation of multiprotein effector complexes that target invading nucleic acids, dsDNA. Furthermore, class1, type I systems recognize longer dsDNA targets in comparison to class 2 therefore they might be expoited for a more precise gene editing or silencing. This work started by a single molecule analysis of I-E type Cascade target recognition. The analysis revealed the target-end validation as a main mechanism for a precise target recognition and further cleavage by Cas3 nuclease. Target-end validation mechanism is a special feature of type I complexes, but not Cas9 protein (Szczelkun et al., 2014b). Further, artificial Cascade complexes containg crRNA of different length were constructed and analyzed in terms of their target recognition. Cascade complexes bearing shorter or longer crRNA were able to form R-loops, however, only the complexes containing WT or longer crRNA mediated the target cleavage by Cas3. The target-end validation mechanism occured in a similar manner for WT or longer crRNA containing Cascade complexes, indicating that longer crRNA and formation of longer R-loops has not provided a more precise target recognition in comparison to WT Cascade complexes. In line to that, the formation of type I-E Cascade complexes has been monitored by native mass spectrometry. Incoropration or loss of $(Cas7)_2(Cse2)_1$ protein module was observed upon addiotion/loss of every 12 nt to crRNA. Dissotation of the large subunit Cse1 monitored in some cases revealed the reason why some complexes formed heterogenious inactive species observed on single molecule assays. Based on structural studies, Cse2 together with Cse1 are thought to be major players for full-target recognition and stabilization for Cas3-mediated cleavage (Chapter 2.3). In contrast to I-E type systems, type I-F systems do not encode small subunits (Cse2) and perform the target end validation by 180º rotation of large subunit Csy8f C-terminal helical boundle. Therefore, we asked how longer crRNA containing type I-F complexes recognise their targets and employ the target end validation. The results showed the target recognition by WT-like mode even when crRNAs were longer. Overall, this indicates that both type I-E and I-F complexes adapted to their crRNA length and crRNA length changes have not provided higher sequence specificity for longer DNA targets.

Ancillary genes of CRISPR-Cas systems are used as an additional layer for antiphage defense. In this thesis TA system located in the same operon as I-D type CRISPR-Cas effector complex genes were studied. It was shown that

a minimal nucleotidyltransferase antitoxin MNT uses ATP to block activity of toxic RNase HEPN. Presumably low ATP concentrations can lead to release of active HEPN toxin. The presence of ATPase domains in a wide range of mobile genetic elements or phage genomes indicate that the preditors extensively use host ATP and may create an intracellular „ATP sink“ (Shim et al., 2021), therefore the presence of a systems that detects the decrease of cellullar ATP level in case of phage infection might be a logical evolutionary move. Cas3 helicase that uses ATP upon the phage dsDNA unwinding and destruction in type I-D CRISPR-Cas system might also be responsible for a substantial decline of cellular ATP levels. In other systems with HEPN domains, such as Cas13 or Csm6/Csx1 proteins, HEPN domains contribute to non-specific ssRNA degradation, such as mRNA. However, in HEPN-MNT TA system, HEPN specifically cleaves 4 nt from tRNA 3′-stem loop. Notably, a tRNA variant truncated at 3′-end by 4 nt can not be repaired by CCA-adding polymerase that requires the 4th „discriminatory“ nucleotide for adding a 3′-tail (Rammelt and Rossmanith, 2016). This activity of HEPN may be compared to other tRNA modifying and protein synthesis inhibiting toxins: cleavage by VapC (Winther and Gerdes, 2011), acetylation of attached amino acid by GNAT toxins (Jurenas et al., 2017), inactivation of tRNA by addition of UTP, CTP by MenT toxin (Cai et al., 2020) or pyrophosphorylation by RelA/SpoT homologs RSH toxins (Kurata et al., 2021). The factors that contribute to activation mechanism of HEPN toxin remain unknown. One of the possibilities is that the formed HEPN-MNT complex is unstable in the absence of ATP and therefore unmodified active HEPN is released. However, special enzymes dedicated for de-AMPylation of modified HEPN might also exist. Furthermore, how these ancillary TA genes contribute to antiphage defense and if (or how) TA system co-operates with CRISPR-Cas action also remains to be identified.

CONCLUSIONS

1. Type I-E Cascade complex recognizes its target by unidirectional zipping mechanism starting from the PAM sequence (**publication 1**)
2. The proximal side of the R-loop induces the conformational change in the Cascade complex termed "target locking" that triggers target degradation by Cas3 nuclease (**publication 1**)
3. Type I-E Cascade complex is able to accommodate crRNA of altered length and forms shorter or longer R-loops (ranging from 15 bp to 57 bp, respectively) depending on the spacer length (**publication 2 and 3**)
4. Regardless of the crRNA length and the length of the formed R-loop, I-E Cascade or I-F Csy complexes recognize their targets and induce target degradation by the same mechanism as WT complexes using 3′-end validation step for dsDNA cleavage (**publication 2 and 4**)
5. HEPN-MNT from *Aphanizomenon flos-aqaue* I-D CRISPR-Cas operon form a toxin-antitoxin system where HEPN RNase is neutralized by di-AMPylation (**publication 5**)
6. Unmodified HEPN cleaves tRNA removing 4 nt from the 3′-end: CCA and "discriminatory" nucleotide (**publication 5**)

THESIS MATERIAL WAS PRESENTED IN THE FOLLOWING CONFERENCES[1]:

1. **Songailienė I**, Rutkauskas M, Sinkunas T, Seidel R, Siksnys V.

"Target recognition by Cascade complexes bearing altered length crRNA" poster in FEBS 2017 and oral presentation in FEBS Young Scientist Forum, 2017, Israel.

2. **Songailienė I**, Sinkunas T, Rutkauskas M, Seidel R, Siksnys V. "Reconstitution of Cascade complex in vitro and R-loop formation". Poster in "CRISPR 2016", Israel.
3. **Songailiene I**, Sinkunas T, Rutkauskas M, Seidel R ir Siksnys V, "Cascade complex reconstitution in vitro and R-loop formation by single molecule experiments", poster presentation, Open Readings 2016, Lithuania
4. **Songailienė I.** EMBO Practical course "Non-coding RNR in infection", 2014, Wurzbug, Germany, oral presentation "CRISPR-Cas systems in bacterial defense".
5. **Songailienė I**, Juozapaitis J, Tamulaitiene G, Sasnauskas G, Venclovas Č, Siksnys V "HEPN-MNT toxin-antitoxin as a bacterial ATP sensor". 2021 Open Readings, oral presentation, Lithuania.
6. **Songailienė I**, Juozapaitis J, Tamulaitiene G, Sasnauskas G, Venclovas Č, Siksnys V "HEPN-MNT toxin-antitoxin as a proposed bacterial ATP sensor". 2021 EMBO meeting "New approaches and concepts in microbiology", 2021 June 7-9, oral and poster presentation, virtual meeting.

[1] All presentations in the conferences were presented by the first authors.

ACKNOWLEDGEMENTS

At first, I am very glad and thankful for my PhD supervisor prof. dr. Virginijus Šikšnys for the opportunity to study CRISPR-Cas systems, intriguing scientific discussions, shared readings, paper editing, trust and countless support over the years.

Second, I am very pleased to have my colleague dr. Tomas Šinkūnas who taught me a lot regarding the CRISPR-Cas experiments at the very start of my PhD. The start of my PhD would not have been possible without my undergrad supervisor dr. Giedrė Tamulaitienė who gave me strong fundamentals in biochemistry and a wide range of experimental work. Experiments may be scarce without bioinformatic models or hypothesis, therefore dr. Česlovas Venclovas and his group gave extra boost to my PhD project.

Moreover, I would like to thank prof. Ralf Seidel and Marius Rutkauskas who both taught me the basics of single molecule experiments during my visit and who are our successful collaboration partners.

I am glad to know dr. Carla Schmidt, dr. Sabine Witting and Marie Babel Barth who have successfully performed native mass spectrometric analysis of Cascade samples.

I thank dr. Dukas Jurėnas and dr. Julija Armalytė for the insightful revisions that helped to improve this thesis.

I am also very grateful to my colleagues dr. Elena Manakova, dr. Giedrius Sasnauskas, dr. Arūnas Šilanskas, dr. Mindaugas Zaremba, dr. Gintautas Tamulaitis, dr. Tautvydas Karvelis, dr. Miglė Kazlauskienė, dr. Paulius Toliušis, dr. Giedrius Gasiūnas, dr. Saulius Gražulis, dr. Evelina Zagorskaitė, Mantvyda Grušytė, Greta Bigelytė, Irmantas Mogila, dr. Gediminas Drabavičius, Edvardas Golovinas for a number of lessons, shared plasmids and joyful coffee or lunch-times. Last, but not least, are my supervised students Jonas Juozapaitis, Eglė Lelešiūtė, Matas Jakubėlis, Matas Grigaliūnas, Dainius Tautvaišas, Ieva Juozapaitytė and Raminta Mineikaitė who taught me a lot about teamwork.

Finally, I would like to thank for love and care given by my family that I will always be thankful for.

REFERENCES

1. Aakre, C.D., Phung, T.N., Huang, D., and Laub, M.T. (2013). A Bacterial Toxin Inhibits DNA Replication Elongation through a Direct Interaction with the β Sliding Clamp. Mol. Cell *52*, 617–628.
2. Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*.
3. Aguilera, A., and Gaillard, H. (2014). Transcription and recombination: when RNA meets DNA. Cold Spring Harb. Perspect. Biol. *6*, a016543-.
4. Aguilera, A., and García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. Mol. Cell *46*, 115–124.
5. Al-Shayeb, B., Sachdeva, R., Chen, L.-X., Ward, F., Munk, P., Devoto, A., Castelle, C.J., Olm, M.R., Bouma-Gregson, K., Amano, Y., et al. (2020). Clades of huge phages from across Earth's ecosystems. Nature *578*, 425–431.
6. Altae-Tran, H., Kannan, S., Demircioglu, F.E., Oshiro, R., Nety, S.P., McKay, L.J., Dlakić, M., Inskeep, W.P., Makarova, K.S., Macrae, R.K., et al. (2021). The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. Science.
7. Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E. V., and Aravind, L. (2013a). Comprehensive analysis of the HEPN superfamily: Identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct *8*.
8. Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E.V., and Aravind, L. (2013b). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct *8*, 15.
9. Aravind, L., and Koonin, E. V (1999a). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. *27*, 1609–1618.
10. Aravind, L., and Koonin, E.V. (1999b). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. *27*, 1609–1618.

11. Athukoralage, J.S., Rouillon, C., Graham, S., Grüschow, S., and White, M.F. (2018). Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. Nature *562*, 277–280.
12. Athukoralage, J.S., Graham, S., Rouillon, C., Grüschow, S., Czekster, C.M., and White, M.F. (2020). The dynamic interplay of host and viral enzymes in type III CRISPR-mediated cyclic nucleotide signalling. ELife *9*, e55852.
13. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *1709*.
14. Berghoff, B.A., and Wagner, E.G.H. (2017). RNA-based regulation in type I toxin–antitoxin systems and its implication for bacterial persistence. Curr. Genet. *63*, 1011–1016.
15. Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. Nat. Rev. Microbiol. *18*, 113–119.
16. Blosser, T.R., Loeff, L., Westra, E.R., Vlot, M., Künne, T., Sobota, M., Dekker, C., Brouns, S.J.J., and Joo, C. (2015). Two Distinct DNA Binding Modes Guide Dual Roles of a CRISPR-Cas Protein Complex. Mol. Cell 1–11.
17. Blower, T.R., Pei, X.Y., Short, F.L., Fineran, P.C., Humphreys, D.P., Luisi, B.F., and Salmond, G.P.C. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. Nat. Struct. Mol. Biol. *18*, 185–190.
18. Bobonis, J., Mitosch, K., Mateus, A., Kritikos, G., Elfenbein, J.R., Savitski, M.M., Andrews-Polymenis, H., and Typas, A. (2020a). Phage proteins block and trigger retron toxin/antitoxin systems (Microbiology).
19. Bobonis, J., Mateus, A., Pfalz, B., Garcia-Santamarina, S., Galardini, M., Kobayashi, C., Stein, F., Savitski, M.M., Elfenbein, J.R., Andrews-Polymenis, H., et al. (2020b). Bacterial retrons encode tripartite toxin/antitoxin systems (Microbiology).
20. Brantl, S. (2021). Plasmid Replication Control by Antisense RNAs. 19.
21. 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., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.
22. Brown, M.S., Segal, A., and Stadtman, E.R. (1971). Modulation of Glutamine Synthetase Adenylylation and Deadenylylation Is Mediated by Metabolic Transformation of the PI,-Regulatory Protein. Proc Nat Acad Sci USA 5.

23. Brown, S., Gauvin, C.C., Charbonneau, A.A., Burman, N., and Lawrence, C.M. (2020). Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type III CRISPR–Cas that degrades the second messenger cA4. J. Biol. Chem. *295*, 14963–14972.
24. Cai, Y., Usher, B., Gutierrez, C., Tolcan, A., Mansour, M., Fineran, P.C., Condon, C., Neyrolles, O., Genevaux, P., and Blower, T.R. (2020). A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems. Sci. Adv. *6*, eabb6651.
25. Cameron, P., Coons, M.M., Klompe, S.E., Lied, A.M., Smith, S.C., Vidal, B., Donohoue, P.D., Rotstein, T., Kohrs, B.W., Nyer, D.B., et al. (2019). Harnessing type I CRISPR–Cas systems for genome engineering in human cells. Nat. Biotechnol. *37*, 1471–1477.
26. Carte, J., Pfister, N.T., Compton, M.M., Terns, R.M., and Terns, M.P. (2010). Binding and cleavage of CRISPR RNA by Cas6. RNA N. Y. N *16*, 2181–2188.
27. Casey, A.K., and Orth, K. (2018). Enzymes Involved in AMPylation and deAMPylation. Chem. Rev. *118*, 1199–1215.
28. Chang, Y., Su, T., Qi, Q., and Liang, Q. (2016). Easy regulation of metabolic flux in Escherichia coli using an endogenous type I-E CRISPR-Cas system. Microb. Cell Factories *15*, 195–195.
29. Charpentier, E., Richter, H., van der Oost, J., and White, M.F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol. Rev. *39*, 428–441.
30. Choi, J.S., Kim, W., Suk, S., Park, H., Bak, G., Yoon, J., and Lee, Y. (2018). The small RNA, SdsR, acts as a novel type of toxin in *Escherichia coli*. RNA Biol. *15*, 1319–1335.
31. Chowdhury, S., Carter, J., Rollins, M.F., Golden, S.M., Jackson, R.N., Hoffmann, C., Nosaka, L., Bondy-Denomy, J., Maxwell, K.L., Davidson, A.R., et al. (2017). Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. Cell *169*, 47-57.e11.
32. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP–AMP signalling protects bacteria against viral infection. Nature *574*, 691–695.
33. Crowley, V.M., Catching, A., Taylor, H.N., Borges, A.L., Metcalf, J., Bondy-Denomy, J., and Jackson, R.N. (2019). A Type IV-A CRISPR-Cas System in *Pseudomonas aeruginosa* Mediates RNA-Guided Plasmid Interference *In Vivo*. CRISPR J. *2*, 434–440.

34. Cruz, J.W., and Woychik, N.A. (2014). Teaching Fido New ModiFICation Tricks. PLoS Pathog. *10*.
35. Cruz, J.W., Rothenbacher, F.P., Maehigashi, T., Lane, W.S., Dunham, C.M., and Woychik, N.A. (2014). Doc toxin is a kinase that inactivates elongation factor Tu. J. Biol. Chem. *289*, 7788–7798.
36. Csörgő, B., León, L.M., Chau-Ly, I.J., Vasquez-Rifo, A., Berry, J.D., Mahendra, C., Crawford, E.D., Lewis, J.D., and Bondy-Denomy, J. (2020). A compact Cascade–Cas3 system for targeted genome engineering. Nat. Methods *17*, 1183–1190.
37. Culviner, P.H., and Laub, M.T. (2018). Global Analysis of the E. coli Toxin MazF Reveals Widespread Cleavage of mRNA and the Inhibition of rRNA Maturation and Ribosome Biogenesis. Mol. Cell *70*, 868-880.e10.
38. Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.
39. Dolan, A.E., Hou, Z., Xiao, Y., Gramelspacher, M.J., Heo, J., Howden, S.E., Freddolino, P.L., Ke, A., and Zhang, Y. (2019). Introducing a Spectrum of Long-Range Genomic Deletions in Human Embryonic Stem Cells Using Type I CRISPR-Cas. Mol. Cell *74*, 936-950.e5.
40. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. Science *359*, eaar4120.
41. East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature *538*, 270–273.
42. Estrella, M.A., Kuo, F.-T., and Bailey, S. (2016). RNA-activated DNA cleavage by the Type III-B CRISPR–Cas effector complex. Genes Dev. *30*, 460–470.
43. Feng, F., Yang, F., Rong, W., Wu, X., Zhang, J., Chen, S., He, C., and Zhou, J.-M. (2012). A Xanthomonas uridine 5′-monophosphate transferase inhibits plant immune kinases. Nature *485*, 114–118.
44. Figure, S. Nature Structural and Molecular Biology : doi : 10 . 1038 / nsmb . 2875 Nature Structural and Molecular Biology : doi : 10 . 1038 / nsmb . 2875. *8*.
45. Fraikin, N., Goormaghtigh, F., and Van Melderen, L. (2020). Type II Toxin-Antitoxin Systems: Evolution and Revolutions. J. Bacteriol. *202*, e00763-19, /jb/202/7/JB.00763-19.atom.

46. Fricke, T., Smalakyte, D., Lapinski, M., Pateria, A., Weige, C., Pastor, M., Kolano, A., Winata, C., Siksnys, V., Tamulaitis, G., et al. (2020). Targeted RNA Knockdown by a Type III CRISPR-Cas Complex in Zebrafish. CRISPR J. *3*, 299–313.
47. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., and Zhang, F. (2020). Diverse enzymatic activities mediate antiviral immunity in prokaryotes. 9.
48. Garcia-Doval, C., Schwede, F., Berk, C., Rostøl, J.T., Niewoehner, O., Tejero, O., Hall, J., Marraffini, L.A., and Jinek, M. (2020). Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. Nat. Commun. *11*, 1596.
49. Garcia-Rodriguez, G., Charlier, D., Wilmaerts, D., Michiels, J., and Loris, R. Alternative dimerization is required for activity and inhibition of the HEPN ribonuclease RnlA. 15.
50. Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. *109*, 2579–2586.
51. Gerdes, K., Nielsen, A., Thorsted, P., and Wagner, E.G.H. (1992). Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable Hok, SrnB and PndA effector messenger RNAs. J. Mol. Biol. *226*, 637–649.
52. Gleditzsch, D., Pausch, P., Müller-Esparza, H., Özcan, A., Guo, X., Bange, G., and Randau, L. (2019). PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. RNA Biol. *16*, 504–517.
53. Globyte, V., Lee, S.H., Bae, T., Kim, J., and Joo, C. (2019). CRISPR /Cas9 searches for a protospacer adjacent motif by lateral diffusion. EMBO J. *38*.
54. Goeders, N., and Van Melderen, L. (2014). Toxin-Antitoxin Systems as Multilevel Interaction Systems. Toxins *6*, 304–324.
55. Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-amikam, Y., Afik, S., Ofir, G., and Sorek, R. (2014). BREX is a novel phage resistance system widespread in microbial genomes. 1–16.
56. Gong, S., Yu, H.H., Johnson, K.A., and Taylor, D.W. (2018). DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. Cell Rep. *22*, 359–371.
57. Gootenberg, J.S., Abudayyeh, O.O., Lee, J.W., Essletzbichler, P., Dy, A.J., Joung, J., Verdine, V., Donghia, N., Daringer, N.M., Freije, C.A.,

et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. Science *356*, 438–442.
58. Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics *8*, 172.
59. Groh, M., and Gromak, N. (2014). Out of balance: R-loops in human disease. PLoS Genet. *10*, e1004630–e1004630.
60. Grüschow, S., Athukoralage, J.S., Graham, S., Hoogeboom, T., and White, M.F. (2019). Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. Nucleic Acids Res. *47*, 9259–9270.
61. Guegler, C.K., and Laub, M.T. (2021). Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. Mol. Cell S1097276521002227.
62. Guo, T.W., Bartesaghi, A., Yang, H., Falconieri, V., Rao, P., Merk, A., Eng, E.T., Raczkowski, A.M., Fox, T., Earl, L.A., et al. (2017). Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex. Cell *171*, 414-426.e12.
63. Hale, C.R., Cocozaki, A., Li, H., Terns, R.M., and Terns, M.P. (2014). Target RNA capture and cleavage by the Cmr type III-B CRISPR–Cas effector complex. Genes Dev. *28*, 2432–2443.
64. Hall, A.M., Gollan, B., and Helaine, S. (2017). Toxin–antitoxin systems: reversible toxicity. Curr. Opin. Microbiol. *36*, 102–110.
65. Halpin-Healy, T.S., Klompe, S.E., Sternberg, S.H., and Fernández, I.S. (2020). Structural basis of DNA targeting by a transposon-encoded CRISPR–Cas system. Nature *577*, 271–274.
66. Han, W., Pan, S., López-Méndez, B., Montoya, G., and She, Q. (2017). Allosteric regulation of Csx1, a type IIIB-associated CARF domain ribonuclease by RNas carrying a tetraadenylate tail. Nucleic Acids Res. *45*, 10740–10750.
67. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., Montoya, G., and She, Q. (2018). A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. Nucleic Acids Res.
68. Harms, A., Stanger, F.V., Scheu, P.D., de Jong, I.G., Goepfert, A., Glatter, T., Gerdes, K., Schirmer, T., and Dehio, C. (2015). Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. Cell Rep. *12*, 1497–1507.

69. Harms, A., Brodersen, D.E., Mitarai, N., and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. Mol. Cell *70*, 768–784.
70. Harrington, L.B., Burstein, D., Chen, J.S., Paez-Espino, D., Ma, E., Witte, I.P., Cofsky, J.C., Kyrpides, N.C., Banfield, J.F., and Doudna, J.A. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. Science *362*, 839–842.
71. Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L.A. (2014). Genetic Characterization of Antiplasmid Immunity through a Type III-A CRISPR-Cas System. J. Bacteriol. *196*, 310–317.
72. Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B.G., Rajashankar, K., Bailey, S., Wiedenheft, B., and Ke, A. (2016). Structural basis for promiscuous PAM recognition in type I–E Cascade from E. coli. Nature 1–16.
73. Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. *33*, 510–517.
74. Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E., and Doudna, J. a (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. Proc. Natl. Acad. Sci. *111*, 6618–6623.
75. Hochstrasser, M.L., Taylor, D.W., Kornfeld, J.E., Nogales, E., and Doudna, J.A. (2016). DNA Targeting by a Minimal CRISPR RNA-Guided Cascade. Mol. Cell *63*, 840–851.
76. Huang, H., Zeqiraj, E., Dong, B., Jha, B.K., Duffy, N.M., Orlicky, S., Thevakumaran, N., Talukdar, M., Pillon, M.C., Ceccarelli, D.F., et al. (2014). Dimeric Structure of Pseudokinase RNase L Bound to 2-5A Reveals a Basis for Interferon-Induced Antiviral Activity. Mol. Cell *53*, 221–234.
77. Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, M.D., Zhou, S., Rajashankar, K., Kurinov, I., et al. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. Nat. Struct. Mol. Biol. *21*, 771–777.
78. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. *169*, 5429–5433.
79. Jackson, R.N. (2014). No Title. *1473*.
80. Jeon, Y., Choi, Y.H., Jang, Y., Yu, J., Goo, J., Lee, G., Jeong, Y.K., Lee, S.H., Kim, I.-S., Kim, J.-S., et al. (2018). Direct observation of DNA

target searching and cleavage by CRISPR-Cas12a. Nat. Commun. *9*, 2777.

81. Jia, N., Mo, C.Y., Wang, C., Eng, E.T., Marraffini, L.A., and Patel, D.J. (2019). Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. Mol. Cell *73*, 264-277.e5.
82. Jia, X., Yao, J., Gao, Z., Liu, G., Dong, Y., Wang, X., and Zhang, H. (2018). Structure–function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN–MNT toxin–antitoxin. J. Biol. Chem. jbc.RA118.002421-jbc.RA118.002421.
83. Jiang, P., Mayo, A.E., and Ninfa, A.J. (2007). Escherichia coli glutamine synthetase adenylyltransferase (ATase, EC 2.7.7.49): Kinetic characterization of regulation by PII, PII-UMP, glutamine, and α-ketoglutarate. Biochemistry *46*, 4133–4146.
84. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science *337*, 816–821.
85. Jung, C., Hawkins, J.A., Jones, S.K., Xiao, Y., Rybarski, J.R., Dillard, K.E., Hussmann, J., Saifuddin, F.A., Savran, C.A., Ellington, A.D., et al. (2017). Massively Parallel Biophysical Analysis of CRISPR-Cas Complexes on Next Generation Sequencing Chips. Cell *170*, 35-47.e13.
86. Jurėnas, D., and Van Melderen, L. (2020). The Variety in the Common Theme of Translation Inhibition by Type II Toxin–Antitoxin Systems. Front. Genet. *11*, 262.
87. Jurenas, D., Chatterjee, S., Konijnenberg, A., Sobott, F., Droogmans, L., Garcia-Pino, A., and Van Melderen, L. (2017). AtaT blocks translation initiation by N-acetylation of the initiator tRNAfMet. Nat. Chem. Biol. *13*, 640–646.
88. Jurėnas, D., Chatterjee, S., Konijnenberg, A., Sobott, F., Droogmans, L., Garcia-Pino, A., and Van Melderen, L. (2017). AtaT blocks translation initiation by N-acetylation of the initiator tRNAfMet. Nat. Chem. Biol. *13*, 640–646.
89. Ka, D., Oh, H., Park, E., Kim, J.-H., and Bae, E. (2020). Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD+ degradation. Nat. Commun. *11*, 2816.
90. Kamruzzaman, M., Wu, A.Y., and Iredell, J.R. (2021). Biological Functions of Type II Toxin-Antitoxin Systems in Bacteria. Microorganisms *9*, 1276.
91. Karvelis, T., Bigelyte, G., Young, J.K., Hou, Z., Zedaveinyte, R., Budre, K., Paulraj, S., Djukanovic, V., Gasior, S., Silanskas, A., et al. (2020).

PAM recognition by miniature CRISPR–Cas12f nucleases triggers programmable double-stranded DNA target cleavage. Nucleic Acids Res. *48*, 5016–5023.

92. Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., Kazlauskas, D., Venclovas, Č., and Siksnys, V. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. Nature.
93. Kazlauskiene, M., Tamulaitis, G., Georgij, K., Venclovas, Č., and Siksnys, V. (2016). Spatio temporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. Mol. Cell *62*, 295–306.
94. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science *357*, 605–609.
95. Kellner, M.J., Koob, J.G., Gootenberg, J.S., Abudayyeh, O.O., and Zhang, F. (2019). SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat. Protoc. *14*, 2986–3012.
96. Kingdon, H.S., Shapiro, B.M., and Stadtman, E. (1967). Regulation of glutamine synthetase, viii. ATP: Glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase. Proc. Natl. Acad. Sci. U. S. A. *58*, 1703–1710.
97. Kiro, R., Shitrit, D., and Qimron, U. (2014). Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. RNA Biol. *11*, 1–3.
98. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S., and Sternberg, S.H. (2019). Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. Nature *571*, 219–225.
99. Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell *173*, 665-676.e14.
100. Koonin, E.V., Makarova, K.S., and Wolf, Y.I. (2017). Evolutionary Genomics of Defense Systems in Archaea and Bacteria. Annu. Rev. Microbiol. *71*, 233–261.
101. Kronheim, S., Daniel-Ivad, M., Duan, Z., Hwang, S., Wong, A.I., Mantel, I., Nodwell, J.R., and Maxwell, K.L. (2018). A chemical defence against phage infection. Nature *564*, 283–286.
102. Kuchta, K., Knizewski, L., Wyrwicz, L.S., Rychlewski, L., and Ginalski, K. (2009). Comprehensive classification of nucleotidyltransferase fold

proteins: Identification of novel families and their representatives in human. Nucleic Acids Res. *37*, 7701–7714.

103. Kurata, T., Brodiazhenko, T., Alves Oliveira, S.R., Roghanian, M., Sakaguchi, Y., Turnbull, K.J., Bulvas, O., Takada, H., Tamman, H., Ainelo, A., et al. (2021). RelA-SpoT Homolog toxins pyrophosphorylate the CCA end of tRNA to inhibit protein synthesis. Mol. Cell S1097276521004524.
104. Kuznedelov, K., Mekler, V., Lemak, S., Tokmina-Lukaszewska, M., Datsenko, K.A., Jain, I., Savitskaya, E., Mallon, J., Shmakov, S., Bothner, B., et al. (2016). Altered stoichiometry Escherichia coli Cascade complexes with shortened CRISPR RNA spacers are capable of interference and primed adaptation. Nucleic Acids Res. *44*, 10849–10861.
105. Li, M., Gong, L., Cheng, F., Yu, H., Zhao, D., Wang, R., Wang, T., Zhang, S., Zhou, J., Shmakov, S.A., et al. (2021). Toxin-antitoxin RNA pairs safeguard CRISPR-Cas systems. Science *372*, eabe5601.
106. Li, S.-Y., Cheng, Q.-X., Liu, J.-K., Nie, X.-Q., Zhao, G.-P., and Wang, J. (2018). CRISPR-Cas12a has both cis- and trans-cleavage activities on single-stranded DNA. Cell Res. *28*, 491–493.
107. Lin, J., Fuglsang, A., Kjeldsen, A.L., Sun, K., Bhoobalan-Chitty, Y., and Peng, X. (2020). DNA targeting by subtype I-D CRISPR–Cas shows type I and type III features. Nucleic Acids Res. gkaa749.
108. Liu, T.Y., and Doudna, J.A. (2020). Chemistry of Class 1 CRISPR-Cas effectors: Binding, editing, and regulation. J. Biol. Chem. *295*, 14473–14487.
109. Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017a). Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. Cell *168*, 121-134.e12.
110. Liu, L., Li, X., Ma, J., Li, Z., You, L., Wang, J., Wang, M., Zhang, X., and Wang, Y. (2017b). The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. Cell *170*, 714-726.e10.
111. Loeff, L., Brouns, S.J.J., and Joo, C. (2018). Repetitive DNA Reeling by the Cascade-Cas3 Complex in Nucleotide Unwinding Steps. Mol. Cell *70*, 385-394.e3.
112. Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. Annu. Rev. Virol. *7*, annurev-virology-011620-040628.
113. Luo, M.L., Mullis, A.S., Leenay, R.T., and Beisel, C.L. (2015). Repurposing endogenous type i CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res. *43*, 674–681.

114. Luo, M.L., Jackson, R.N., Denny, S.R., Tokmina-Lukaszewska, M., Maksimchuk, K.R., Lin, W., Bothner, B., Wiedenheft, B., and Beisel, C.L. (2016). The CRISPR RNA-guided surveillance complex in Escherichia coli accommodates extended RNA spacers. Nucleic Acids Res. *44*, 7385–7394.
115. Luria, S.E., and Human, M.L. A NONHREDITARY, HOST-INDUCED VARIATION OF BACTERIAL VIRUSES'. *64*, 13.
116. Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2009). Comprehensive comparative-genomic analysis of Type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biol. Direct *4*.
117. Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2018). Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? CRISPR J. *1*, 325–336.
118. 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., et al. (2019). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol.
119. 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., et al. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. *18*, 67–83.
120. Masuda, H., Tan, Q., Awano, N., Wu, K.-P., and Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in Escherichia coli: YeeU enhances bundling of MreB and FtsZ filaments. Mol. Microbiol. *84*, 979–989.
121. McBride, T.M., Schwartz, E.A., Kumar, A., Taylor, D.W., Fineran, P.C., and Fagerlund, R.D. (2020). Diverse CRISPR-Cas complexes require independent translation of small and large subunits from a single gene (Molecular Biology).
122. McMahon, S.A., Zhu, W., Graham, S., Rambo, R., White, M.F., and Gloster, T.M. (2020). Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. Nat. Commun. *11*, 500.
123. Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., and Sorek, R. (2020). Bacterial retrons function in anti-phage defense (Microbiology).
124. Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., López-Méndez, B., She, Q., and Montoya, G. (2019). Structure of

Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. Nat. Commun. *10*, 4302.

125. Morisaka, H., Yoshimi, K., Okuzaki, Y., Gee, P., Kunihiro, Y., Sonpho, E., Xu, H., Sasakawa, N., Naito, Y., Nakada, S., et al. (2019). CRISPR-Cas3 induces broad and unidirectional genome editing in human cells. Nat. Commun. *10*, 5302.
126. Mulepati, S., and Bailey, S. (2013). In vitro reconstitution of an Escherichia coli RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA Target. J. Biol. Chem. *288*, 22184–22192.
127. Mulepati, S., Héroux, A., and Bailey, S. (2014). Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. Science 1–9.
128. Muller, M.P., Peters, H., Blumer, J., Blankenfeldt, W., Goody, R.S., and Itzen, A. (2010). The Legionella Effector Protein DrrA AMPylates the Membrane Traffic Regulator Rab1b. Science *329*, 946–949.
129. Neunuebel, M.R., Chen, Y., Gaspar, A.H., Backlund, P.S., Yergey, A., and Machner, M.P. (2011). De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila. Science *333*, 453–456.
130. Niehrs, C., and Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. Nat. Rev. Mol. Cell Biol. *21*, 167–178.
131. Niewoehner, O., Garcia-Doval, C., Rostøl, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., and Jinek, M. (2017). Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature *548*, 543–548.
132. Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S., and Sorek, R. (2018). DISARM is a widespread bacterial defence system with broad anti-phage activities. Nat. Microbiol. *3*, 90–98.
133. Ogura, T., and Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc. Natl. Acad. Sci. *80*, 4784–4788.
134. Osakabe, K., Wada, N., Miyaji, T., Murakami, E., Marui, K., Ueta, R., Hashimoto, R., Abe-Hara, C., Kong, B., Yano, K., et al. (2020). Genome editing in plants using CRISPR type I-D nuclease. Commun. Biol. *3*, 648.
135. Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. Mol. Cell 1–13.

136. Page, R., and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. Nat. Chem. Biol. *12*, 208–214.
137. Pausch, P., Müller-Esparza, H., Gleditzsch, D., Altegoer, F., Randau, L., and Bange, G. (2017). Structural Variation of Type I-F CRISPR RNA Guided DNA Surveillance. Mol. Cell *67*, 622-632.e4.
138. Pecota, D.C., and Wood, T.K. (1996). Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. J. Bacteriol. *178*, 2044–2050.
139. Petassi, M.T., Hsieh, S.-C., and Peters, J.E. (2020). Guide RNA Categorization Enables Target Site Choice in Tn7-CRISPR-Cas Transposons. Cell *183*, 1757-1771.e18.
140. Peters, J.E., Makarova, K.S., Shmakov, S., and Koonin, E.V. Recruitment of CRISPR-Cas systems by Tn7-like transposons. 40.
141. Pickar-Oliver, A., Black, J.B., Lewis, M.M., Mutchnick, K.J., Klann, T.S., Gilcrest, K.A., Sitton, M.J., Nelson, C.E., Barrera, A., Bartelt, L.C., et al. (2019). Targeted transcriptional modulation with type I CRISPR–Cas systems in human cells. Nat. Biotechnol.
142. Pillon, M.C., and Stanley, R.E. (2018). Nuclease integrated kinase super assemblies (NiKs) and their role in RNA processing. Curr. Genet. *64*, 183–190.
143. Pillon, M.C., Sobhany, M., Borgnia, M.J., Williams, J.G., and Stanley, R.E. (2017). Grc3 programs the essential endoribonuclease Las1 for specific RNA cleavage. Proc. Natl. Acad. Sci. *114*, E5530–E5538.
144. Pillon, M.C., Sobhany, M., and Stanley, R.E. (2018). Characterization of the molecular crosstalk within the essential Grc3/Las1 pre-rRNA processing complex. RNA *24*, 721–738.
145. Pillon, M.C., Gordon, J., Frazier, M.N., and Stanley, R.E. (2020a). HEPN RNases – an emerging class of functionally distinct RNA processing and degradation enzymes. Crit. Rev. Biochem. Mol. Biol. 1–30.
146. Pillon, M.C., Goslen, K.H., Gordon, J., Wells, M.L., Williams, J.G., and Stanley, R.E. (2020b). It takes two (Las1 HEPN endoribonuclease Domains) to cut RNA correctly. J. Biol. Chem. jbc.RA119.011193.
147. Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R.A., Randau, L., Sørensen, S.J., and Shah, S.A. (2020). Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. Nucleic Acids Res. *48*, 2000–2012.
148. Rammelt, C., and Rossmanith, W. (2016). Repairing tRNA termini: News from the 3’ end. RNA Biol. *13*, 1182–1188.
149. Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P.R., and Lundgren, M. (2014). Efficient programmable gene silencing by Cascade. Nucleic Acids Res. *43*, 237–246.

150. Redding, S., Sternberg, S.H., Wiedenheft, B., Jennifer, A., Greene, E.C., Redding, S., Sternberg, S.H., Marshall, M., Gibb, B., Bhat, P., et al. (2015). Supplemental Information: Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System. Cell *163*, 1–12.
151. Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Wilkinson, R.A., Bondy-Denomy, J., Lander, G.C., and Wiedenheft, B. (2017). Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. Proc. Natl. Acad. Sci. 201616395.
152. Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Miettinen, H.M., Santiago-Frangos, A., Faith, D., Lawrence, C.M., Lander, G.C., and Wiedenheft, B. (2019). Structure Reveals a Mechanism of CRISPR-RNA-Guided Nuclease Recruitment and Anti-CRISPR Viral Mimicry. Mol. Cell *74*, 132-142.e5.
153. Rostøl, J.T., Xie, W., Kuryavyi, V., Maguin, P., Kao, K., Froom, R., Patel, D.J., and Marraffini, L.A. (2021). The Card1 nuclease provides defence during type-III CRISPR immunity. Nature.
154. Rouillon, C., Athukoralage, J.S., Graham, S., Grüschow, S., and White, M.F. (2018). Control of cyclic oligoadenylate synthesis in a type III CRISPR system. ELife *7*, e36734.
155. Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V., and Seidel, R. (2015). Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. Cell Rep. *10*, 1534–1543.
156. Rutkauskas, M., Krivoy, A., Szczelkun, M.D., Rouillon, C., and Seidel, R. (2017). Single-Molecule Insight Into Target Recognition by CRISPR–Cas Complexes. In Methods in Enzymology, (Elsevier), pp. 239–273.
157. Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. Cell 1–11.
158. Santos-pereira, J.M., and Aguilera, A. (2015). REVIEWS R loops : new modulators of genome dynamics and function. Nat. Publ. Group *16*, 583–597.
159. Sanyal, A., Chen, A.J., Nakayasu, E.S., Lazar, C.S., Zbornik, E.A., Worby, C.A., Koller, A., and Mattoo, S. (2015). A novel link between fic (filamentation induced by cAMP)-mediated adenylylation/AMPylation and the unfolded protein response. J. Biol. Chem. *290*, 8482–8499.
160. Sasnauskas, G., and Siksnys, V. (2020). CRISPR adaptation from a structural perspective. Curr. Opin. Struct. Biol. *65*, 17–25.

161. Schifano, J.M., Edifor, R., Sharp, J.D., Ouyang, M., Konkimalla, A., Husson, R.N., and Woychik, N.A. (2013). Mycobacterial toxin MazF-mt6 inhibits translation through cleavage of 23S rRNA at the ribosomal A site. Proc. Natl. Acad. Sci. *110*, 8501–8506.
162. Schifano, J.M., Vvedenskaya, I.O., Knoblauch, J.G., Ouyang, M., Nickels, B.E., and Woychik, N.A. (2014). An RNA-seq method for defining endoribonuclease cleavage specificity identifies dual rRNA substrates for toxin MazF-mt3. Nat. Commun. *5*, 3538.
163. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., and Westra, E.R. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *108*, 10098–10103.
164. Semenova, E., Kuznedelov, K., Datsenko, K.A., Boudry, P.M., Savitskaya, E.E., Medvedeva, S., Beloglazova, N., Logacheva, M., Yakunin, A.F., and Severinov, K. (2015). The Cas6e ribonuclease is not required for interference and adaptation by the E . coli type I-E CRISPR-Cas system. 1–13.
165. Shim, H., Shivram, H., Lei, S., Doudna, J.A., and Banfield, J.F. (2021). Diverse ATPase Proteins in Mobilomes Constitute a Large Potential Sink for Prokaryotic Host ATP. Front. Microbiol. *12*, 691847.
166. Short, F.L., Pei, X.Y., Blower, T.R., Ong, S.-L., Fineran, P.C., Luisi, B.F., and Salmond, G.P.C. (2013). Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. Proc. Natl. Acad. Sci. *110*, E241–E249.
167. Silva, M.B. A nucleotide-sensing endonuclease from the Gabija bacterial defense system. 23.
168. Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J. *30*, 1335–1342.
169. Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO J. 1–10.
170. Smalakyte, D., Kazlauskiene, M., F. Havelund, J., Rukšėnaitė, A., Rimaite, A., Tamulaitiene, G., Færgeman, N.J., Tamulaitis, G., and Siksnys, V. (2020). Type III-A CRISPR-associated protein Csm6 degrades cyclic hexa-adenylate activator using both CARF and HEPN domains. Nucleic Acids Res. *48*, 9204–9217.

171. Smargon, A.A., Cox, D.B.T., Pyzocha, N.K., Zheng, K., Slaymaker, I.M., Gootenberg, J.S., Abudayyeh, O.A., Essletzbichler, P., Shmakov, S., Makarova, K.S., et al. (2017). Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Mol. Cell *65*, 618-630.e7.
172. Song, S., and Wood, T.K. (2020). Toxin/Antitoxin System Paradigms: Toxins Bound to Antitoxins Are Not Likely Activated by Preferential Antitoxin Degradation. Adv. Biosyst. *4*, 1900290.
173. Songailiene, I., Rutkauskas, M., Sinkunas, T., Manakova, E., Wittig, S., Schmidt, C., Siksnys, V., and Seidel, R. (2019). Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length. Cell Rep. *28*, 3157-3166.e4.
174. Songailiene, I., Juozapaitis, J., Tamulaitiene, G., Ruksenaite, A., Šulčius, S., Sasnauskas, G., Venclovas, Č., and Siksnys, V. (2020). HEPN-MNT Toxin-Antitoxin System: The HEPN Ribonuclease Is Neutralized by OligoAMPylation. Mol. Cell S1097276520308340.
175. Sreelatha, A., Yee, S.S., Lopez, V.A., Park, B.C., Kinch, L.N., Pilch, S., Servage, K.A., Zhang, J., Jiou, J., Karasiewicz-Urbańska, M., et al. (2018). Protein AMPylation by an Evolutionarily Conserved Pseudokinase. Cell *175*, 809-821.e19.
176. Stachler, A.-E., and Marchfelder, A. (2016). Gene Repression in Haloarchaea Using the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas I-B System. J. Biol. Chem. *291*, 15226–15242.
177. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.
178. Šulčius, S., Alzbutas, G., Kvederavičiūtė, K., Koreivienė, J., Zakrys, L., Lubys, A., and Paškauskas, R. (2015). Draft Genome Sequence of the Cyanobacterium Aphanizomenon flos-aquae Strain 2012/KM1/D3, Isolated from the Curonian Lagoon (Baltic Sea). Genome Announc. *3*, e01392-14, /ga/3/1/e01392-14.atom.
179. Swarts, D.C., Jore, M.M., Westra, E.R., Zhu, Y., Janssen, J.H., Snijders, A.P., Wang, Y., Patel, D.J., Berenguer, J., Brouns, S.J.J., et al. (2014). DNA-guided DNA interference by a prokaryotic Argonaute. Nature *507*, 258–261.
180. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., and Karvelis, T. (2014a). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. 6–11.

181. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014b). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc. Natl. Acad. Sci. *111*, 9798–9803.
182. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Nwokeoji, A.O., Dickman, M.J., Horvath, P., and Siksnys, V. (2014). Article Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. 1–12.
183. Tamulaitis, G., Venclovas, Č., and Siksnys, V. (2017). Type III CRISPR-Cas Immunity: Major Differences Brushed Aside. Trends Microbiol. *25*, 49–61.
184. Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods *12*, 1143–1149.
185. Thisted, T., Sørensen, N.S., Wagner, E.G., and Gerdes, K. (1994). Mechanism of post-segregational killing: Sok antisense RNA interacts with Hok mRNA via its 5′-end single-stranded leader and competes with the 3′-end of Hok mRNA for binding to the mok translational initiation region. EMBO J. *13*, 1960–1968.
186. Tuminauskaite, D., Norkunaite, D., Fiodorovaite, M., Tumas, S., Songailiene, I., Tamulaitiene, G., and Sinkunas, T. (2020). DNA interference is controlled by R-loop length in a type I-F1 CRISPR-Cas system. BMC Biol. *18*, 65.
187. Van Erp, P.B.G., Jackson, R.N., Carter, J., Golden, S.M., Bailey, S., and Wiedenheft, B. (2015). Mechanism of CRISPR-RNA guided recognition of DNA targets in Escherichia coli. Nucleic Acids Res. *43*, 8381–8391.
188. Wang, B., Zhang, T., Yin, J., Yu, Y., Xu, W., Ding, J., Patel, D.J., and Yang, H. (2021). Structural basis for self-cleavage prevention by tag:anti-tag pairing complementarity in type VI Cas13 CRISPR systems. Mol. Cell S1097276520309540.
189. Wang, X., Lord, D.M., Cheng, H.-Y., Osbourne, D.O., Hong, S.H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., et al. (2012). A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. Nat. Chem. Biol. *8*, 855–861.
190. Wang, X., Yao, J., Sun, Y.-C., and Wood, T.K. (2020). Type VII Toxin/Antitoxin Classification System for Antitoxins that Enzymatically Neutralize Toxins. Trends Microbiol. S0966842X20303164.
191. Westra, E.R., van Erp, P.B.G., Künne, T., Wong, S.P., Staals, R.H.J., Seegers, C.L.C., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., et

al. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol. Cell *46*, 595–605.

192. 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., et al. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature *567*, 194–199.
193. Winther, K.S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc. Natl. Acad. Sci. *108*, 7403–7407.
194. Winther, K., Tree, J.J., Tollervey, D., and Gerdes, K. (2016). VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation. Nucleic Acids Res. *44*, 9860–9871.
195. Wittig, S., Songailiene, I., and Schmidt, C. (2020). Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry. J. Am. Soc. Mass Spectrom. *31*, 538–546.
196. Woolery, A.R., Luong, P., Broberg, C.A., and Orth, K. (2010). Ampylation: Something old is new again. Front. Microbiol. *1*, 2–7.
197. Worby, C.A., Mattoo, S., Kruger, R.P., Corbeil, L.B., Koller, A., Mendez, J.C., Zekarias, B., Lazar, C., and Dixon, J.E. (2009). The Fic Domain: Regulation of Cell Signaling by Adenylylation. Mol. Cell *34*, 93–103.
198. Xiao, Y., Luo, M., Hayes, R.P., Kim, J., Ng, S., Ding, F., Liao, M., and Ke, A. (2017). Structure Basis for Directional R-loop Formation and Substrate Handover Mechanisms in Type I CRISPR- Cas System Cryo-EM structures of type I CRISPR-Cas system resolve the mechanisms governing the PAM-dependent R-loop formation, Cas3 recruitment, and subst. Cell *170*, 48-60.e11.
199. Xiao, Y., Xiao, Y., Luo, M., Dolan, A.E., Liao, M., and Ke, A. (2018). Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *0839*, 1–12.
200. Yan, W.X., Chong, S., Zhang, H., Makarova, K.S., Koonin, E.V., Cheng, D.R., and Scott, D.A. (2018). Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. Mol. Cell *70*, 327-339.e5.
201. Yang, M., Peng, S., Sun, R., Lin, J., Wang, N., and Chen, C. (2018). The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET. Cell Rep. *22*, 372–382.

202. Yang, Y., Yue, Y., Song, N., Li, C., Yuan, Z., Wang, Y., Ma, Y., Li, H., Zhang, F., Wang, W., et al. (2020). The YdiU Domain Modulates Bacterial Stress Signaling through Mn2+-Dependent UMPylation. Cell Rep. *32*, 108161.
203. Yao, J., Guo, Y., Zeng, Z., Liu, X., Shi, F., and Wang, X. (2015). Identification and characterization of a HEPN-MNT family type II toxin-antitoxin in Shewanella oneidensis. Microb. Biotechnol. *8*, 961–973.
204. Yarbrough, M., Yan, L., Kinch, L., Grishin, N., Ball, H., and Orth, K. (2009). AMPylation of Rho GTPases by. *323*, 269–272.
205. Yosef, I., Manor, M., Kiro, R., and Qimron, U. (2015). Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. Proc. Natl. Acad. Sci. *112*, 7267–7272.
206. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., and Wang, Y. (2019). Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. Cell *176*, 239-253.e16.
207. Young, J.K., Gasior, S.L., Jones, S., Wang, L., Navarro, P., Vickroy, B., and Barrangou, R. (2019). The repurposing of type I-E CRISPR-Cascade for gene activation in plants. Commun. Biol. *2*, 383.
208. Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. *4*, 442–451.
209. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell *163*, 759–771.
210. Zhang, B., Ye, Y., Ye, W., Perčulija, V., Jiang, H., Chen, Y., Li, Y., Chen, J., Lin, J., Wang, S., et al. (2019). Two HEPN domains dictate CRISPR RNA maturation and target cleavage in Cas13d. Nat. Commun. *10*, 2544.
211. Zhang, C., Konermann, S., Brideau, N.J., Lotfy, P., Wu, X., Novick, S.J., Strutzenberg, T., Griffin, P.R., Hsu, P.D., and Lyumkis, D. (2018). Structural Basis for the RNA-Guided Ribonuclease Activity of CRISPR-Cas13d. Cell *175*, 212-223.e17.
212. Zhao, H., Sheng, G., Wang, J., Wang, M., Bunkoczi, G., Gong, W., Wei, Z., and Wang, Y. (2014). Crystal structure of the RNA-guided immune surveillance Cascade complex in Escherichia coli. Nature *515*, 147–150.
213. Zhou, Y., Bravo, J.P.K., Taylor, H.N., Steens, J.A., Jackson, R.N., Staals, R.H.J., and Taylor, D.W. (2021). Structure of a type IV CRISPR-Cas ribonucleoprotein complex. IScience *24*, 102201.

214. Zhu, X., and Ye, K. (2014). Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. Nucleic Acids Res. *43*, 1257–1267.
215. Zhu, W., McQuarrie, S., Grüschow, S., McMahon, S.A., Graham, S., Gloster, T.M., and White, M.F. (2021). The CRISPR ancillary effector Can2 is a dual-specificity nuclease potentiating type III CRISPR defence. Nucleic Acids Res. gkab073.
216. Zhu, W., McQuarrie, S., Graham, S., Gloster, T.M., and White, M.F. The CRISPR ancillary effector Can2 is a dual-specificity nuclease potentiating type III CRISPR defence. 13.

SUPPLEMENTARY TABLE

Table S1. The cryo-EM aqnd X-ray structures solved for class 1 systems. The structures solved by X-ray are in bold.

Type	**Name**	**PDB ID**	**Organism**	**Structure with**	**Ref.**
I-E	Cascade	**4TVX** **4U7U** **4QYZ** **5CD4** **5H9E** **5H9F** 5U07 5U0A 6C66	*E. coli* *Thermobifida fusca* 	+crRNA +crRNA +crRNA + ssDNA +crRNA +crRNA+dsDNA +crRNA+dsDNA +crRNA +dsDNA target (R-loop) +crRNA+dsDNA+Cas 3 *(pre-nicking)*	(Figure; Hayes et al., 2016; Jackson, 2014; Mulepati et al., 2014; Van Erp et al., 2015; Xiao et al., 2017, 2018)
	Cas3	**4QQX** **4QQY** **4QQZ** **4QQW** **4Q2C** **4Q2D**	*Thermobifida fusca*	+ATP +ADP +AMPPNP +ssDNA +dATP	
I-F	Csy complex	6B44 6B45 6B47 6B48 5UZ9 6VQV 6VQW 6VQX 5XLO 5XLP 6NE0	*Pseudomonas aeruginosa*	+crRNA+dsDNA +crRNA +crRNA+AcrF2 +crRNA+AcrF10 +crRNA+AcrF1/2 + crRNA+AcrF9 + crRNA+AcrF8 +crRNA+AcrF6 +crRNA+AcrF1/2 +crRNA+AcrF1/2 +crRNA+dsDNA	(Chowdhury et al., 2017; Guo et al., 2017; Rollins et al., 2019)
	Cas3	5B7I			
I-Fv		**5O6U** **5O7H**	*Shewanella putrefaciens*	+crRNA+dsDNA	(Pausch et al., 2017)
I-D	Cascade complex			+crRNA	(McBride et al., 2020)
III-A	Csm complex	6MUR 6MUT 6IFK 6IFL		+ssRNA +ssRNA +ssRNA +AMPPNP	(Jia et al., 2019; You et al., 2019)

		6O7I		+ssRNA+cA4	
III-B	Cmr complex	3X1L		+ssRNA	(Osawa et al., 2015)
I-F variant	TniQ-Cascade TniQ	6PIF 6PIG 6PIJ 6LNB 6LNC 6V9Q 6VBW 6UVN **6LND** **6V9P**	*Vibrio cholerae* *Vibrio cholerae*	+crRNA (close) +crRNA (open) +crRNA (close) +crRNA+dsDNA	(Halpin-Healy et al., 2020)

CURRICULUM VITAE

Inga Songailienė

RESEARCH EXPERIENCE

2020 May-December	*Supervisor* of the project „Investigation of I-D type CRISPR-Cas effector complex“ founded by Vilnius University intramural grant
2017-present	*Junior scientist* Structural and functional characterization of CRISPR-Cas systems
2013-2017	*Research assistant*
2010-2013	*Research technician* Structural and functional investigation of restriction endonucleases.
2013 February-April	Professional practice at research lab founded by EU.
2011 November-2012 January	Professional practice founded by Research Council of Lithuania.
2010 June-October	ERASMUS laboratory practice at the University of Bristol, School of Biomedical Sciences, supervisor prof. H. Mellor.

EDUCATION

2013 October - present	*PhD studies* on CRISPR-Cas systems, supervisor prof. Virginijus Siksnys
2011-2013	MSc in Biochemistry *magna cum laude,* Vilnius University**.**
2007-2011	BSc in Biochemistry *cum laude***,** Vilnius University

METHODS

Cloning and expression in *E. coli*, DNA, RNA electrophoresis, radioactive labeling of nucleic acids, SDS-PAGE, protein purification by liquid chromatography, HPLC, mass spectrometry, Western blot, Northern blot, RNA extraction, RNA-seq library preparation, native and denatured PAGE electrophoresis, protein crystallization

MAJOR AWARDS

2013	The best Master thesis defended in Life Sciences in Lithuania. „Characterization of restriction endonuclease MnlI N-terminal domain structure with specific DNA sequence“, supervisor dr. G. Tamulaitienė.
2020	The best life sciences publication of the year at Vilnius university.

MAJOR CONFERENCES

2021	**Songailiene I**, Juozapaitis J, Tamulaitiene G, Sasnauskas G, Venclovas Č, Siksnys V “HEPN-MNT

	toxin-antitoxin as a bacterial ATP sensor". 2021 Open Readings, oral presentation, Lithuania
2017	**Songailiene I**, Rutkauskas M, Šinkunas T, Seidel R, Siksnys V "Target recognition by Cascade complexes bearing altered length crRNA" Poster at FEBS 2017 and oral presentation at FEBS Young Scientist Forum, 2017, Israel.
2016	**Songailiene I**, Sinkunas T, Rutkauskas M, Seidel R ir Siksnys V, "Cascade complex reconstitution in vitro and R-loop formation by single molecule experiments", poster presentation, Open Readings 2016, Lithuania
2016	**Songailiene I**, T. Šinkunas, M. Rutkauskas, R. Seidel, V. Šikšnys "Reconstitution of Cascade complex *in vitro* and R-loop formation" Poster "CRISPR2016", Israel
2014	EMBO Practical course "Non-coding RNR in infection", oral presentation "CRISPR-Cas systems in bacterial defense".

COURSES

2016	Non-coding RNA in infection, EMBO practical course
2014	Fundamentals of modern methods of biocrystallography Biocrys2, Oreiras, Portugal.
2012	Insubria International Summer School for Heath and Biosciences.

RESEARCH INTERNSHIP

2015 December	Single molecule experiments with CRISPR-Cas systems, under supervision of prof. Ralf Seidel, University of Leipzig, founded by Research Council of Lithuania

TEACHING

2017-present	„CRISPR-Cas biology" course for MSc students. Responsible for practical work.

MAJOR INTERESTS

Non-coding RNA, CRISPR-Cas, phage-bacteria defense systems, covalent modifications, signaling molecules

FOREIGN LANGUAGES: English (fluent), German (basics)

PUBLICATIONS

The list of publications and contributions forming the basis of this thesis:

1. Rutkauskas M, Sinkunas T, **Songailiene I**, Tikhomirova MS, Siksnys V, Seidel R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep.* 2015 Mar 10;10(9):1534-1543. doi: 10.1016/j.celrep.2015.01.067.
   Performed experiments in bulk, analyzed the results and prepared the results for publication.
2. **Songailiene I*,** Rutkauskas M*, Sinkunas T, Manakova E, Wittig S, Schmidt C, Siksnys V, Seidel R. Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length. *Cell Rep.* 2019 Sep 17;28(12):3157-3166.e4. doi: 10.1016/j.celrep.2019.08.033
   Designed the study, performed biochemical and initial single molecule experiments, wrote/edited the publication.
3. Wittig S, **Songailiene I**, Schmidt C. Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry. *J Am Soc Mass Spectrom.* 2020 Mar 4;31(3):538-546. doi: 10.1021/jasms.9b00011.
   Purified and characterized Cascade complexes, edited publication draft.
4. Tuminauskaite D*, Norkunaite D*, Fiodorovaite M, Tumas S, **Songailiene I**, Tamulaitiene G, Sinkunas T. DNA interference is controlled by an R-loop length in a type I-F1 CRISPR-Cas system. *BMC biology.* 2020 Jun 15;18(1):65. doi: 10.1186/s12915-020-00799-z.
   Performed experiments in vitro and in vivo with I-F Csy complexes.
5. **Songailiene I***, Juozapaitis J*, Tamulaitiene G, Ruksenaite A, Sulčius S, Sasnauskas G, Venclovas Č, Siksnys V. HEPN-MNT toxin-antitoxin system: the toxin is neutralized by OligoAMPylation. Mol Cell, 2020 Dec 17;80(6):955-970.e7. doi: 10.1016/j.molcel.2020.11.034. Epub 2020 Dec 7.
   Designed the study and performed in vitro experiments, wrote and revised the publication.

Publication 1

It is an open access publication on Cell Reports, accessible by doi: 10.1016/j.celrep.2015.01.067.

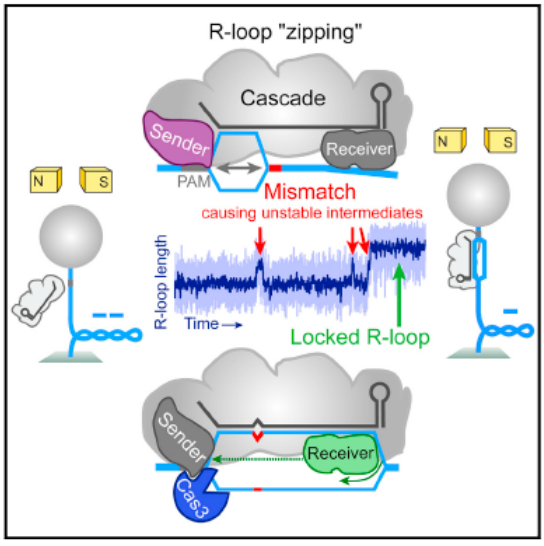
Rutkauskas M, Sinkunas T, **Songailiene I**, Tikhomirova MS, Siksnys V, Seidel R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep.* 2015 Mar 10;10(9):1534-1543. doi: 10.1016/j.celrep.2015.01.067.

Article

Cell Reports

Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection

Graphical Abstract



Authors

Marius Rutkauskas, Tomas Sinkunas, ..., Virginijus Siksnys, Ralf Seidel

Correspondence

siksnys@ibt.lt (V.S.),
ralf.seidel@uni-muenster.de (R.S.)

In Brief

The CRISPR-Cas surveillance complex Cascade targets invading DNA for destruction with the help of a short crRNA. Rutkauskas et al. use real-time supercoiling experiments on single DNA molecules to directly reveal the impact of mismatches during the directional zipping of the crRNA on a DNA target.

Publication 2

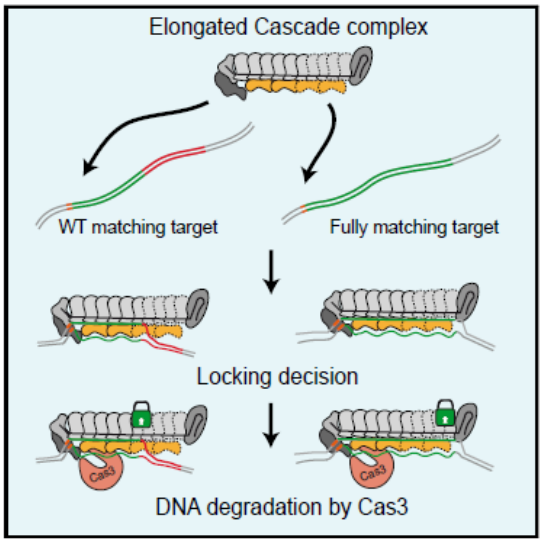
It is an open access publication on Cell Reports, accessible by
doi: 10.1016/j.celrep.2019.08.033
Songailiene I*, Rutkauskas M*, Sinkunas T, Manakova E, Wittig S, Schmidt C, Siksnys V, Seidel R. Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length. *Cell Rep.* 2019 Sep 17;28(12):3157-3166.e4. doi: 10.1016/j.celrep.2019.08.033

Article

Cell Reports

Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length

Graphical Abstract



Authors

Inga Songailiene, Marius Rutkauskas, Tomas Sinkunas, ..., Carla Schmidt, Virginijus Siksnys, Ralf Seidel

Correspondence

siksnys@ibt.lt (V.S.),
ralf.seidel@physik.uni-leipzig.de (R.S.)

In Brief

Songailiene et al. show that engineered versions of Cascade with elongated crRNAs can base-pair with target DNA over 57 bp. However, target recognition only requires wild-type-sized base-pairing of ~30-bp length. These findings define constraints for improving the specificity of these complexes in biotechnological applications.

Publication 3

It is an open access publication on Journal of the American Society for Mass Spectrometry, accessible by doi: 10.1021/jasms.9b00011.

Wittig S, **Songailiene I**, Schmidt C. Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry. *J Am Soc Mass Spectrom*. 2020 Mar 4;31(3):538-546. doi: 10.1021/jasms.9b00011.

Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry

Sabine Wittig, Inga Songailiene, and Carla Schmidt*

Cite This: *J. Am. Soc. Mass Spectrom.* 2020, 31, 538−546 Read Online

ABSTRACT: The adaptive immune system of bacteria and archaea against viral DNA is based on clustered, regularly interspaced, short palindromic repeats (CRISPRs) which are encoded in the host genome and translated into CRISPR RNAs (crRNAs) containing single spacer sequences complementary to foreign DNA. crRNAs assemble with CRISPR-associated (Cas) proteins forming surveillance complexes that base-pair with viral DNA and mediate its degradation. As specificity of degradation is provided by the crRNA spacer sequence, genetic engineering of the CRISPR system has emerged as a popular molecular tool, for instance, in gene silencing and programmed DNA degradation. Elongating or shortening the crRNA spacer sequence are therefore promising ventures to modify specificity toward the target DNA. However, even though the stoichiometry of wild-type complexes is well established, it is unknown how variations in crRNA spacer length affect their stoichiometry. The CRISPR-associated antiviral defense surveillance complexes of *Streptococcus thermophilus* (StCascade complexes) contain crRNA and five protein subunits. Using native mass spectrometry, we studied the formation and stoichiometry of StCascade complexes assembled on a set of crRNAs with different spacer lengths. We assigned all relevant complexes and gained insights into the stoichiometry of the complexes as well as their preferred assembly. We found that stable complexes, which incorporate or lose a $(Cas7)_2(Cse2)_1$-module, assemble on crRNA varied in length by 12-nucleotide units, while varying crRNA length in six-nucleotide units results in heterogeneous mixtures of complexes. Combining our results from the various variants, we generated an assembly pathway revealing general features of I-E type Cascade complex formation.

KEYWORDS: *CRISPR, type I-E Cascade complex, crRNA, stoichiometry, complex heterogeneity, native mass spectrometry*

Publication 4

It is an open access publication on BMC Journal, accessible by doi: 10.1186/s12915-020-00799-z.
Tuminauskaite D*, Norkunaite D*, Fiodorovaite M, Tumas S, **Songailiene I**, Tamulaitiene G, Sinkunas T. DNA interference is controlled by an R-loop length in a type I-F1 CRISPR-Cas system. *BMC biology*. 2020 Jun 15;18(1):65. doi: 10.1186/s12915-020-00799-z.

DNA interference is controlled by R-loop length in a type I-F1 CRISPR-Cas system

Donata Tuminauskaite[1†], Danguole Norkunaite[1†], Marija Fiodorovaite[1], Sarunas Tumas[1,2], Inga Songailiene[1], Giedre Tamulaitiene[1] and Tomas Sinkunas[1*]

Abstract

Background: CRISPR-Cas systems, which provide adaptive immunity against foreign nucleic acids in prokaryotes, can serve as useful molecular tools for multiple applications in genome engineering. Diverse CRISPR-Cas systems originating from distinct prokaryotes function through a common mechanism involving the assembly of small crRNA molecules and Cas proteins into a ribonucleoprotein (RNP) effector complex, and formation of an R-loop structure upon binding to the target DNA. Extensive research on the I-E subtype established the prototypical mechanism of DNA interference in type I systems, where the coordinated action of a ribonucleoprotein Cascade complex and Cas3 protein destroys foreign DNA. However, diverse protein composition between type I subtypes suggests differences in the mechanism of DNA interference that could be exploited for novel practical applications that call for further exploration of these systems.

Results: Here we examined the mechanism of DNA interference provided by the type I-F1 system from *Aggregatibacter actinomycetemcomitans* D7S-1 (Aa). We show that functional Aa-Cascade complexes can be assembled not only with WT spacer of 32 nt but also with shorter or longer (14–176 nt) spacers. All complexes guided by the spacer bind to the target DNA sequence (protospacer) forming an R-loop when a C or CT protospacer adjacent motif (PAM) is present immediately upstream the protospacer (at −1 or −2,−1 position, respectively). The range of spacer and protospacer complementarity predetermine the length of the R-loop; however, only R-loops of WT length or longer trigger the nuclease/helicase Cas2/3, which initiates ATP-dependent unidirectional degradation at the PAM-distal end of the WT R-loop. Meanwhile, truncation of the WT R-loop at the PAM-distal end abolishes Cas2/3 cleavage.

Conclusions: We provide a comprehensive characterisation of the DNA interference mechanism in the type I-F1 CRISPR-Cas system, which is different from the type I-E in a few aspects. First, DNA cleavage initiation, which usually happens at the PAM-proximal end in type I-E, is shifted to the PAM-distal end of WT R-loop in the type I-F1. Second, the R-loop length controls on/off switch of DNA interference in the type I-F1, while cleavage initiation is less restricted in the type I-E. These results indicate that DNA interference in type I-F1 systems is governed through a checkpoint provided by the Cascade complex, which verifies the appropriate length for the R-loop.

Keywords: CRISPR-Cas immunity, CRISPR protection, type I-F, Cascade, Csy, R-loop, Cas3 nuclease/helicase, Cas2/3, DNA interference, DNA degradation

Publication 5

It is a publication on Molecular Cell, accessible by
doi: 10.1016/j.molcel.2020.11.034. Epub 2020 Dec 7.
If you can not access the publication, please contact *inga.songailiene@bti.vu.lt*

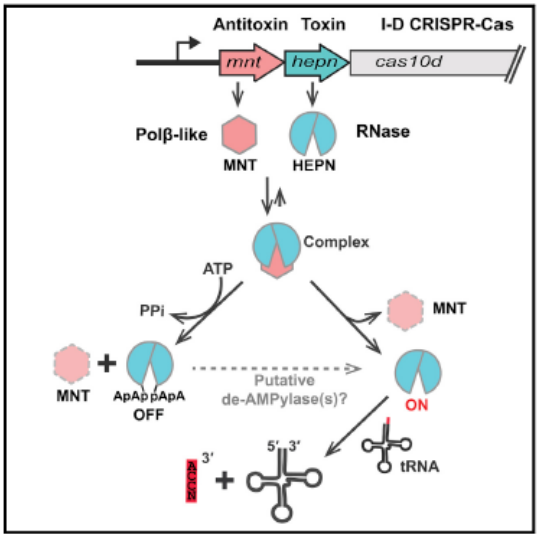
Songailiene I*, Juozapaitis J*, Tamulaitiene G, Ruksenaite A, Sulčius S, Sasnauskas G, Venclovas Č, Siksnys V. HEPN-MNT toxin-antitoxin system: the toxin is neutralized by OligoAMPylation. Mol Cell, 2020 Dec 17;80(6):955-970.e7. doi: 10.1016/j.molcel.2020.11.034. Epub 2020 Dec 7.

Article

Molecular Cell

HEPN-MNT Toxin-Antitoxin System: The HEPN Ribonuclease Is Neutralized by OligoAMPylation

Graphical Abstract



Authors

Inga Songailiene, Jonas Juozapaitis, Giedre Tamulaitiene, ..., Giedrius Sasnauskas, Česlovas Venclovas, Virginijus Siksnys

Correspondence

siksnys@ibt.lt

In Brief

Songailiene et al. report the mechanism of *A. flos-aquae* HEPN-MNT toxin-antitoxin system. HEPN RNase is a toxin, cutting 4 nt from the tRNA acceptor stem. The minimal nucleotidyltransferase (MNT) inactivates HEPN through oligoAMPylation. The system may act as a cellular sensor of ATP homeostasis, releasing active HEPN at low ATP.

SANTRAUKA

1. Santrumpų sąrašas

(d)NTP	(deoksi)nukleozidų trifosfatai
Abi	abortatyvi infekcija
Cas	CRISPR-asocijuotas
Cascade	CRISPR-asocijuotas kompleksas priešvirusiniam atsakui
cOA	cikliniai oligoadenilatai
CRISPR	reguliariai sutelkti palindrominiai pasikartojimai
crRNR	CRISPR asocijuota maža RNR
dg	dvigrandininė
EMSA	elektroforetinio mobilumo tyrimas
PAM	prie proskirtuko esantis sekos motyvas
PDB	baltymų struktūrų duomenų bankas
pre-crRNR	crRNA prieš jai subręstant
RNaze	ribonukleazė
RNP	ribonukleobaltyminis kompleksas
vg	viengrandininė
TA	toksinas-antitoksinas
WT	laukinis tipas

2. Įvadas

Nuo žarnyno bakterijų iki ligų sukėlėjų – mikrooganizmai yra mūsų gyvenimo dalis. Šiuo metu antibiotiktų atsparumas yra viena iš opiausių problemų medicinoje. Kovai su antibiotikams atspariomis bakterijomis pasitelkiami vis kitokie įrankiai, tokie kaip antimikrobiniai peptidai, antikūnai, o tarp jų, kaip alternatyva, bakterijų virusų (fagų) terapija. Norint sėkmingai pritaikyti šį metodą, visų pirma reikia suprasti bakterijų ir virusų sąveiką bei natūralius bakterijų kovos su virusais būdus. Bakterijos su virusais dažniausiai kovoja blokuodamos ląstelės paviršiaus receptorius arba degraduodamos į ląstelės vidų patekusias fagų nukleorūgštis (DNR arba RNR), kam panaudojama daugybė įvairiausių sistemų. Bakterijų priešvirusinės apsaugos sistemų tyrinėjimas padėjo atrasti tokius genomo inžinerijos įrankius kaip restrikcijos endonukleazės, transpozonai arba CRISPR-Cas sistemos (Bernheim and Sorek, 2020). Būtent CRISPR-Cas sistemos, sutinkamos virš 40% bakterijų ir beveik 90% archėjų genomų

(Grissa et al., 2007), sukėlė proveržį šiuolaikinėje genų inžinerijoje. CRISPR-Cas sistemų veikiantieji komponentai – efektoriniai baltymų ir trumpų RNR (crRNR) kompleksai. Šie kompleksai sugeba aptikti savo taikinį svetimoje (fago) nukleorūgštyje dėka crRNR, kuri veikia kaip vedlys (gidas). Tokiu būdu, keičiant crRNR lengvai galima sistemas perprogramuoti ir nukreipti prie kito taikinio. CRISPR-Cas sistemų efektoriniai kompleksai sudaryti iš vieno arba kelių Cas baltymų ir crRNR. 1 klasei priskiriami daugiabaltyminiai kompleksai ir crRNR, o 2 klasę sudaro pavienis baltymas (Cas9, Cas12, Cas13) su crRNR (Makarova et al., 2019). Klasės, reminatis genų organizacija, smulkiau skirstomos į tipus. I tipo CRISPR-Cas sistemos veikia prieš dgDNR arba vgDNR (išimtinai tik I-D tipo kompleksai). Šių sistemų efektorinius RNP kompleksus sudaro 3-5 baltyminiai komponentai ir crRNR. Jie atpažįsta trumpą nukleotidų motyvą ant dgDNR, vadinamą PAM seka ir tuomet praskiriant dgDNR grandines vieną iš jų išstumia ir sudaro DNR:RNR hibridą, vadinamą R-kilpa. Pilnos R-kilpos susidarymas pritraukia Cas3 nukleazę-helikazę, vykdančią taikinio degradaciją. Kartu su CRISPR-Cas sistemomis dažnai veikia papildomi baltymai, papildantys šių sistemų poveikį ir gebantys sulėtinti ląstelinius procesus aptikus fago infekciją (abortyvinė infekcija) arba sukelti programuojamą ląstelės žūtį. Tai ypač būdinga III tipo sistemoms, kurios po fago infekcijos specialių signalinių molekulių ciklinių adenilatų cOA dėka sugeba aktyvuoti pagalbinius baltymus, koduojamus CRISPR-Cas sistemos operone, bet nesudarančius efektorinių kompleksų. Šie baltymai sukelia abortyvinę infekciją (Abi) – sulėtina ląstelės metabolizmą taip, kad fagas nebegali pasidauginti (Lopatina et al., 2020). Paprastai šie baltymai veikia skaldydami RNR arba DNR (Athukoralage et al., 2018; Kazlauskiene et al., 2017; McMahon et al., 2020; Niewoehner et al., 2017; Rostøl et al., 2021; Zhu et al., 2021).

Šios disertacijos pagrindiniai tyrimo objektai – pirmos klasės, I-E, I-F ir I-D tipo CRISPR-Cas sistemos ir su I-D tipo sistema asocijuota toksino-antitiksino sistema HEPN-MNT. Gilinamos žinios apie I-E ir I-F tipo CRISPR-Cas sistemų taikinio atpažinimo mechanizmą pavienių molekulių ir biocheminiais metodais. Charakterizuojamas HEPN-MNT toksino-antitoksino sistemos veikimas ir bandoma jį susieti su tame pačiame operone esančia I-D tipo CRISPR-Cas sistema.

Pagrindinis **darbo tikslas** yra suprasti I tipo CRISPR-Cas sistemų ir su jomis asociuojtos toksino-antitoksino sistemos HEPN-MNT veikimą biocheminiais ir pavienių molekulių metodais. Tam pasiekti išsikelti šie **uždaviniai:**

1. Charakterizuoti I-E tipo Cascade komplekso taikinio atpažinimo mechanizmą;

2. Nustatyti, kaip I-E ir I-F tipo sisitemos su ilgesnėmis ar trumpesnėmis crRNR, geba atpažinti savo taikinius;
3. Nustatyti, kaip formuojasi I-E tipo Cascade efektronis kompleksas;
4. Charakterizuoti I-D tipo sistemos operone esančius papildomus *mnt* ir *hepn* genus, kurie spėjamai sudaro toksino-antitoksino sistemą;

Mokslinis naujumas ir aktualumas: Jau parodyta, jog pirmos klasės, I tipo CRISPR-Cas sistemos gali būti naudojamos prokariotų ir eukariotų genomo redagavimui. Supratimas, kaip šios sistemos komponentai atpažįsta savo takinius pavienių molekulių lygyje, leidžia geriau paaiškinti tokius *in vivo* stebimus efektus, kaip stipresnis genų raiškos slopinimas ar tam tikrais atvejais pagerėjęs genų redagavimo efektyvumas. Be to, žinios apie tai, kokių minimalių komponentų reikia sistemos veikimui, leidžia tikslingai planuoti genomo redagavimą ar genų nutildymą. R-kilpos formavimasis yra svarbus žingsnis CRISPR-Cas interferencijoje, tad šio proceso tyrinėjimas leidžia suprasti, kaip ir kodėl fagai išvengia CRISPR-Cas sistemų poveikio. Septinto tipo toksino-antitoksino (TA) sistema, charakterizuota šiame darbe, yra viena pirmųjų sistemų prisikriamų šiam tipui. Šią sistemą sudaro minimali nukleotidiltransferazė MNT ir HEPN ribonuklazė. Priešingai, negu prieš tai manyta, ši sistema nepriklauso II tipo TA sistemoms, kurių veikimas paremtas toksino blokavimu pasitelkiant tiesioginę baltymų sąveiką. HEPN-MNT sistemos veikimas paremtas tuo, jog antitoksinas veikia fermentiškai, perkeldamas dvi ATP molekules ant konservatyvios HEPN tirozino aminorūgšties, taip prie baltymo prijungiant di-AMP grandinę. Struktūriniai tyrimai atskleidė, jog di-AMP prijungimas pakeičia toksino aktyvaus centro konformaciją ir aktyvaus centro konservatyvios histidino aminorūgštys nutolsta viena nuo kitos per 18Å. Darbe taip pat analizuoti HEPN ribonukleazės taikiniai. Išanalizavus HEPN toksino aktyvumą *E. coli* beląsteliniame ekstrakte panaudojant aktyvų HEPN toksiną arba jo mutantus ir atlikus mažųjų RNR sekoskaitą, paaiškėjo, kad esant aktyviam HEPN toksinui, nemaža dalis tRNR molekulių nebeturi 4 nt jų 3′-gale (ant akceptorinio stiebo). Toks ribonukleazinis aktyvumas nėra būdingas iki šiol aprašytoms RNazėms.

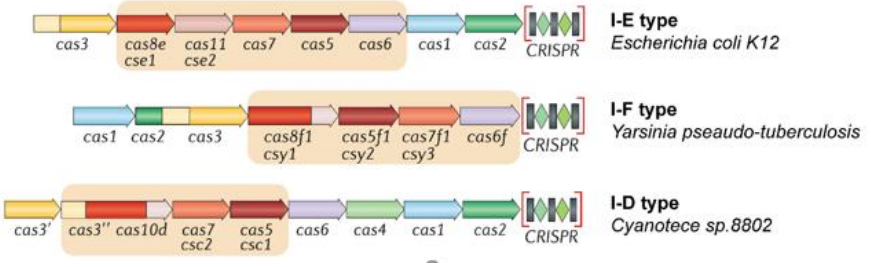
3. Metodai

Šiame darbe panaudoti įvairūs metodai – taikytas pavienių molekulių magnetinių pincetų metodas, natyvios masių spektrometrijos ir biocheminiai metodai I-tipo CRISPR-Cas sistemų taikinio atpažinimui tyrinėti. Pavienių molekulių magentinių pincetų metodas leidžia nukleotido tikslumu aptikti R-kilpos susiformavimą ir stebėti šių kilpų ilgį. Natyvios masių spektrometrijos

metodas pritaikytas suprasti, kaip formuojasi I-E tipo sistemos efektorinis kompleksas Cascade. Masių spektrometrijos ir struktūriniai metodai pritaikyti analizuojant toksino-antitoksino HEPN-MNT sistemos veikimą. Visi tyrinėti baltymai, išskyrus aktyvią HEPN ribonukleazę, buvo gaunami iš *E. coli* bakterijų, prieš tai genus klonuojant į plazmides ir pasitelkiant indukuojamas baltymų gamybos sistemas. Naudoti baltymai gryninti skysčių chromatografijos metodais. HEPN ribonukleazės veikimas analizuotas beląsteliniame *E. coli* ekstrakte, pasitekliant mažųjų RNR sekoskaitos metodą. Naudoti metodai, oligonukleotidai ir buferiai detaliai aprašyti atskirose publikacijose.

4. Darbo objektai

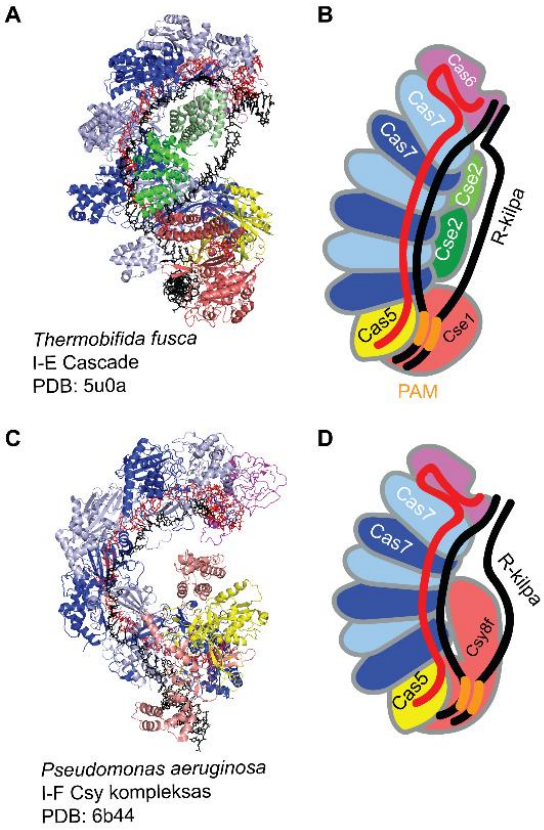
Šis darbas atliktas su pirmos klasės, I tipo CRISPR-Cas sistemų kompleksais ir VII tipo toksino-antitoksino sistema HEPN-MNT. Bendra šių I tipo CRISPR-Cas sistemų genų organizacija pavaizduota pav. 1.



Paveikslas 1. Tiriamų 1 klasės, I-E, I-F ir I-D tipo sistemų bendra genų organizacija. I-E tipo sistemos efektorinį kompleksą sudaro penki baltymai (apvesti rudai): Cse1, Cse2, Cas7, Cas5 ir Cas6. I-F tipo sistemos kompleksą sudaro keturi baltymai: Csy1, Csy2, Csy3 ir Cas6f. Iš bioinformatinių duomenų buvo spėjama, kad I-D tipo kompleksą sudaro trys komponentai: Cas10d, Csc2, Csc1, tačiau dabar žinoma, kad Cas10d turi alternatyvų skaitymo rėmelį ir nuo jo sintetinamas mažasis subvienetas Cas11d, tad kompleksą viso sudaro keturi komponentai (McBride et al., 2020). Paveikslas adaptuotas iš (Makarova et al., 2019).

Besiformuojant 1 klasės, I tipo, efektoriniam kompleksui, crRNR 5′-galas yra atpažįstamas Cas5 baltymo, o crRNR 3′-galo plaukų segtukas po pre-crRNR brendimo surišamas Cas6 baltymo (Carte et al., 2010). Cas7 baltymas nusėda ant crRNR molekulės kaip karkasas (Pav. 2a, 2b). Prie Cas5, prisijungia sistemos didysis subvienetas, Cas8e (Cse1), atsakingas už PAM sekos atpažinimą (Hayes et al., 2016; Mulepati et al., 2014; Xiao et al., 2018). I tipo sistemos tarpusavyje skiriasi tuo, ar turi mažąjį subvienetą (Cse2, Cas11d). I-E tipo sistema šį subvienetą koduoja kaip atskirą geną Cse2, tačiau I-F tipo sistemoje jis sulietas su didžiuoju subvienetu (Pav. 2c, 2d) (Rollins et

al., 2019). Kartais funkcionalūs I tipo kompleksai susidaro ir nesant Cas6 (Lin et al., 2020; Semenova et al., 2015). I-E tipo sistema, naudota pavienių molekulių tyrimams šiame darbe, originaliai yra koduojama *Streptoccocus thermophilus* bakterijose (Sinkunas et al., 2013). I-F tipo sistema yra koduojama *Aggregatibacter actinomycetemcomitans D7S-1* patogeninėje bakterijoje (Tuminauskaite et al., 2020). I-D tipo CRISPR-Cas sistema ir HEPN-MNT toksino-antitoksino (TA) sistema aptinkama melsvabakterėse *Aphanizomenon flos-aquae*. HEPN-MNT sistema aptikta tame pačiame genų operone, kaip sutinkama ir I-D tipo CRISPR-Cas sistema ir *mnt-hepn* genų operonas persidengia su sistemos didžiuojo subvieneto *Cas10d* skaitymo rėmeliu. I-D tipo sistema ypatinga tuo, jog HD-nukleazinis domenas, priešingai negu būdinga I-E ir I-F tipo sistemoms, yra lokalizuotas komplekso didžiajame subvienete Cas10d. Neseniai charakterizuotas I-D tipo sistemos veikimas iš archėjų *Sulfolobus islandicus,* parodė, kad sistemos atpažįsta PAM seką, sudaro R-kilpą ir degraduoja dgDNR padedant Cas3 helikazei, be to, turi ir nuo Cas3 helikazės nepriklausomą aktyvumą, – degraduoja vgDNR (Lin et al., 2020). I-D tipo sistemos iš melsvabakterės *Aphanizomenon flos-aquae* atveju, be I-D tipo komplekso, tame pačiame skaitymo rėmelyje aptikta ir TA sistema, kas nėra dažna tarp CRISPR-Cas sistemų. Kadangi jau buvo nustatytas I-D tipo sistemos aktyvumas (Lin et al., 2020), tačiau nebuvo parodyta, kad TA sistemos ir CRISPR-Cas sistemos galėtų kooperuoti, susidomėta šios TA sistemos veikimo mechanizmu ir funkcija. Bioinformatiškai charakterizavus šios sistemos komponentus, pastebėta, kad sistemą sudaro Polβ-like nukleotidiltransferazė MNT ir potenciali ribonukleazė HEPN, kur abu baltymai turi konservatyvias aktyviesiems centrams būdingas aminorūgštis. Panaši charakterizuota sistema iš bakterijų *Shewanella oneidensis*, sudarė baltyminį kompleksą, kuris, kaip pasiūlyta, yra reikalingas toksino blokavimui (Jia et al., 2018). Polβ-like nukleotidiltransferazės atlieka įvairias funkcijas ląstelėse ir dažnai veikia kaip nuo matricos nepriklausomos polimerazės, pvz. CCA-polilmerazė, poli(A)-polimerazė, kanamicino nukleotidiltransferazė (Aravind and Koonin, 1999b).



Paveikslas 2. I-E tipo Cascade ir I-F tipo Csy kompleksų struktūrinė organizacija. A. *Thermobifida fusca* I-E tipo Cascade kristalinė struktūra (PDB ID:5u0a). B. I-E tipo Cascade struktūros schematinis vaizdas išvardinant skirtingus sistemą sudarančius baltymus. R-kilpa pavaizduota juodai, PAM seka – oranžinė. C. I-F tipo Csy komplekso kristalinė struktūra (PDB ID:6b44). D. Csy komplekso struktūros schematinis vaizdas. Csy kompleksas nebeturi mažųjų subvienetų Cse2, kuriuos turi I-E tipo Cascade, tačiau mažųjų suvienetų atitikmuo koduojamas kaip C-galinis didžiojo subvieneto Csy8f domenas.

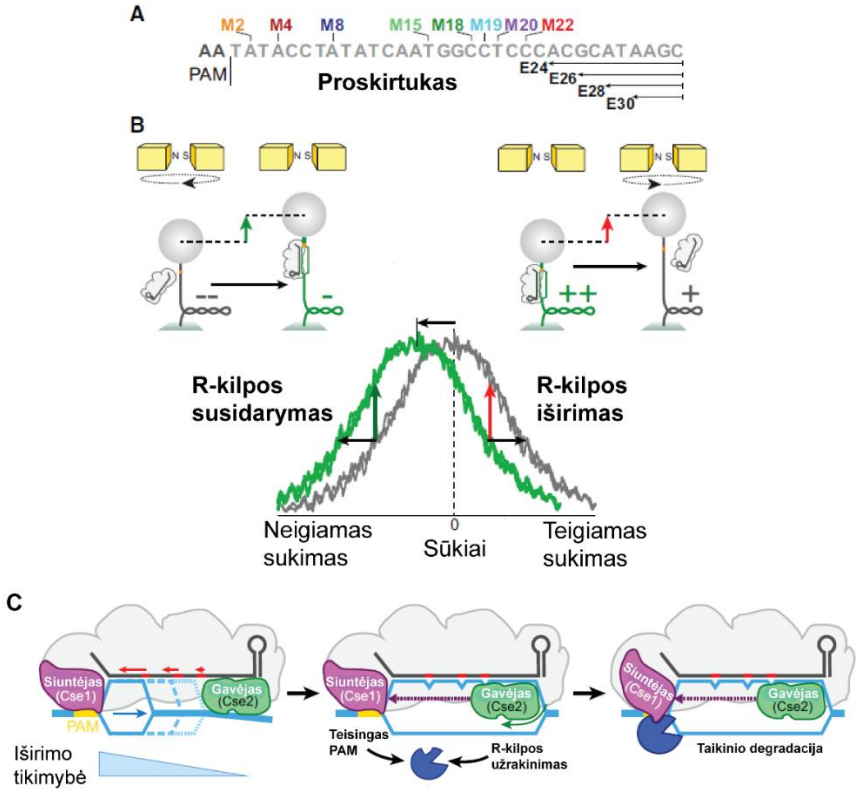
5. Rezultatai

I-E tipo Cascade taikinio atpažinimas – vienkryptis *(publikacija 1)*

Remiantis tuo, kad Cascade taikinio atpažinimui yra būtina PAM seka ir tuo, kad šalia PAM esančios sekos turi daugiausiai įtakos išvengiant Cascade kompleksų vykdomos taikinio degradacijos, buvo spėjama, kad Cascade kompleksas atpažįsta savo taikinį kryptingai, pradedant nuo PAM sekos (Semenova et al., 2011; Szczelkun et al., 2014a). Šiai hipotezei patikrinti, sukonstruoti keli skirtingi 33 bp proskirtukai, turintys mutacijas skirtingai nutolusias nuo PAM sekos. Tai buvo pavienių nukleotidų mutacijos (M2, M4, M8, etc) arba kelių nukleotidų delecijos sekos pabaigoje (E24, E26, etc) (Pav. 3a). Su šiais substratais atlikti pavienių molekulių magnetinių pincetų (eksperimento schema pavaizduota Pav. 3b) ir biocheminiai tyrimai. Paaiškėjo, kad kuo toliau nuo PAM sekos yra įvesta mutacija, tuo ilgiau I-E tipo Cascade kompleksas išsilaiko ant savo taikinio nedisocijavęs. Atpažinus taikinį pilnai buvo stebimas papildomas taikinio stabilizacijos žingsnis, pavadintas „R-kilpos užrakinimu“. Taikinio užrakinimas nebevyko, pašalinus daugiau negu kad 6 nt nuo taikinio pabaigos (E24, E26 mutantai). Siekiant patikrinti, kaip realiomis sąlygomis vyksta konkurencija tarp skirtingų taikinių, pasinaudota elektroforetinio mobilumo poslinkio (EMSA) eksperimentais. Eksperimento esmė – turint vienodą koncentraciją laukinio tipo žymėto taikinio, pridedama vis didėjanti koncentracija konkuruojančios DNR sekos, turinčios mutaciją proskirtuke. Tokiu būdu atskleista, jog taikiniai, turintys mutaciją toliau nuo PAM sekos, yra geresni konkurentiniai inhibitoriai sėkmingam Cascade prisirišimui prie nemutuoto taikinio.

Pagrindiniai publikacijos 1 teiginiai:

- I-E tipo Cascade kompleksas atpažįsta taikinį kryptingai, pradedant nuo PAM sekos;
- R-kilpos išsiardymo tikimybė tuo didesnė, kuo mutacija yra arčiau PAM sekos;
- Taikinio kirpimui yra būtina tiksli PAM seka ir R-kilpos užrakinimas, kuris vyksta atpažinus nuo PAM nutolusią taikinio dalį.



Paveikslas 3. Magnetinių pincetų eksperimentas su I-E tipo Cascade kompleksu ir Cascade veikimo modelis. A. Mutacijų, įvestų į taikinio proskirtuką, schema. M žymimos taškinės mutacijos, E žymimi ilgesni sekos pakeitimai taikinio pabaigoje B. Pavienių molekulių magnetinių pincetų eksperimento schema. DNR išsukimas į neigiamos superspiralizacijos pusę palengvina R-kilpos susidarymą ir taip molekulės rotacinė kreivė pasislenka į kairę. Sukant molekulę link teigiamos superspiralizacijos, tam tikrame etape Cascade kompleksas disocijuoja, taip rotacijos kreivė pasislenka į dešinę. C. I-E tipo Cascade taikinio atpažinimo modelis. Cse1 didysis subvienetas (siuntėjas) atpažįsta PAM seką ir inicijuoja R-kilpos susidarymą. Kuomet susidaro pilno ilgio R-kilpa, Cse2 mažasis subvienetas (gavėjas), įvykdo R-kilpos užrakinimą. Tam įvykus, pritraukiama Cas3 nukleazė-helikazė, kuri vykdo taikinio degradaciją. Paveikslai adaptuoti iš (Rutkauskas et al., 2015).

Taikinio atpažinimas, panaudojant skirtingus I-E tipo Cascade kompleksus

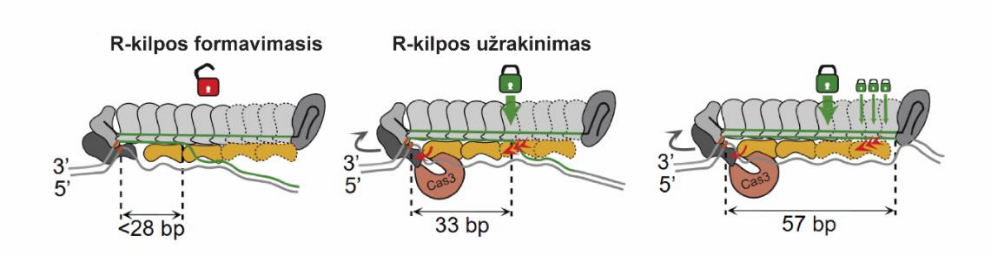
(publikacija 2)

Pastebėję, jog taikinio R-kilpos užrakinimui yra būtinas nuo PAM nutolusios taikinio dalies atpažinimas, toliau norėjome ištirti, ar ši taikinio užrakinimo pozicija gali būti pakeista, jeigu pasikeistų R-kilpos ilgis. Kaip jau

prieš tai minėta, Cascade kompleksai yra sudaryti iš pasikartojančių Cas7 subvienetų, susirinkusių ant crRNR molekulės ir atpažįstančių po 6 nt ant jos, tad galimai keičiant crRNR ilgį, galėtų keistis ir komplekso sudėtis ir jo taikinio atpažinimo mechanizmas. Pradžioje tikriname, ar šie kompleksai galėtų prisirišti prie skirtigo ilgio crRNR ir sudaryti funkcionalias R-kilpas. crRNR skirtuko ilgis buvo keičiamas nuo 15 nt iki 57 nt ir R-kilpos susidarymas stebėtas, panaudojant magnetinių pincetų metodą, taikytą ir publikacijoje 1. Rezultatai parodė, jog kompleksai geba formuoti funkcionalias R-kilpas, tačiau priešingai negu tikėtasi, R-kilpos užrakinimo žingsnis vyko toje pačioje taikinio dalyje, kaip ir naudojant laukinio tipo Cascade kompleksą (bent 28 nt nuo PAM sekos). Dėl to pasiūlytas taikinio atpažinimo modelis, kuomet ilgesnės negu 28 nt ilgio R-kilpos susiformavimas yra lemiantis veiksnys Cas3 nukleazės-helikazės aktyvacijai (Pav. 4). Norint ištirti, ar susidaro homogeniški Cascade kompleksai, panaudotas natyvios masių spektrometrijos metodas. Kompleksai, kurių crRNR skyrėsi po 12 nt nuo WT crRNR (-12, +12, +24 nt), buvo homogeniški, tačiau esant 6 nt skirtumui nuo originalaus crRNR ilgio, daugiausiai formavosi subkompleksai ir buvo prarandamas didysis Cse1 subvienetas. Nepaisant to, kad R-kilpos užrakinimo žingsnis išliko ten pat, kaip ir WT komplekso atveju, ilgesnes crRNR turintys kompleksai sudarė stabilesnes R-kilpas, tačiau R-kilpos stabilumas ir crRNR ilgis tarpusavyje nekoreliavo, taip parodant, jog ne DNR:RNR hibrido ilgis yra svarbiausias nulemiant R-kilpos stabilumą.

Pagrindiniai publikacijos 2 teiginiai:

- Cas baltymai geba formuoti stabilius Cascade kompleksus, kai crRNR ilgis varijuoja 12 nt nuo laukinio tipo crRNR ilgio;
- Ilgesnes nei WT ilgio crRNR turintys Cascade kompleksai geba formuoti R-kilpas, kurių užrakinimas visuomet vyksta toje pačioje dalyje kaip ir WT tipo komplekso, ir inicijuoti taikinio degradaciją po įvykusio R-kilpos užrakinimo;
- RNR:DNR duplekso ilgis nėra pagrindinis veiksnys, nulemiantis R-kilpų stabilumą.



Paveikslas 4. R-kilpos užrakinimo modelis, kuomet kompleksui suformuoti naudojama ilgesnė crRNR molekulė. Kol nesusidaro 28 bp arba ilgesnė R-kilpa, kilpos užrakinimas ir Cas3 nukleazės-helikazės pritraukimas nevyksta (tai pavaizduota kairėje). Susidarius ilgesnei negu 28 bp R-kilpai (pvz. 33 bp, kas būdinga WT Cascade kompleksui), vyksta sėkmingas taikinio užrakinimas, Cas3 pritraukimas ir taikinio degradacija (vidurinis paveikslas). Susidarant ilgesnėms R-kilpoms, pagrindinis R-kilpos stabilizavimo etapas vis tiek išlieka toks, koks būdingas WT kompleksui (už 33 bp), o ilgesnės R-kilpos susiformavimas tik papildomai stabilizuoja kompleksą, tačiau įtakos taikinio kirpimui nedaro (paveikslas dešinėje). Pritaikyta iš (Songailiene et al., 2019).

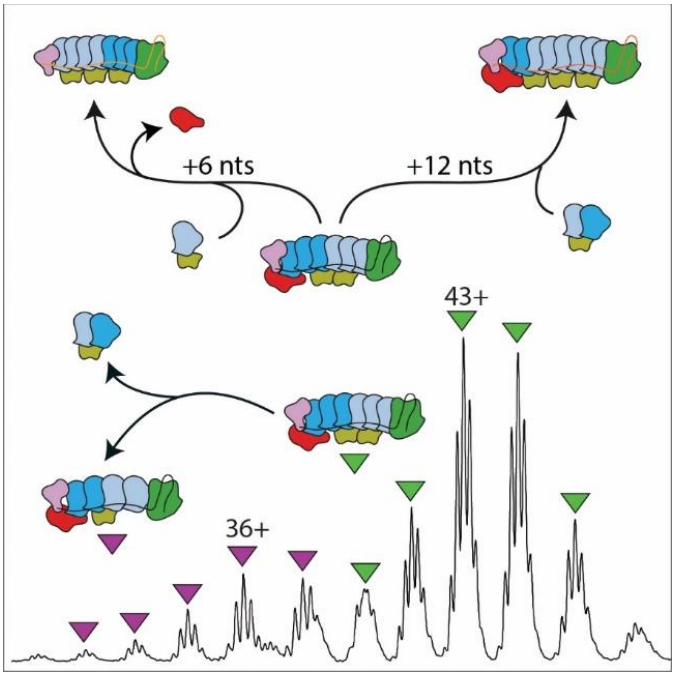
Komplekso susiformavimas I-E tipo CRISPR-Cas sistemose *(publikacija 3)*

Nepaisant to, kad homogeninio Cascade komplekso iš *E. coli* su pakeisto ilgio crRNR stabilumo tyrimas parodė, kad jis sudaro stabilius komplekus su įvairiaus ilgio crRNR (Luo et al., 2016), *S. thermophilus* Cascade kompleksas funkcionalias R-kilpas formavo kuomet naudotos ±12 nt skirtumą nuo WT turinčios crRNR (publikacija 2). Tai iškėlė klausimą, kaip susidaro Cascade kompleksai ir kodėl tam tikri kompleksai nėra funkcionalūs. Atsakyti į šį klausimą pasitelkta natyvi masių spektrometrija. Pradžioje tirtas WT kompleksas, kuris buvo 422 242 Da dydžio ir turėjo stechiometriją $(Cse1)_1(Cse2)_2(Cas7)_6(Cas5)_1(Cas6)_1(crRNA)_1$, būdingą ir homogeniniam *E. coli* kompleksui (Mulepati et al., 2014). Kuomet crRNR prailginama/sutrumpinama 12 nt, papildomų dviejų $(Cas7)_2$ ir vieno $(Cse2)_1$ atsiradimas/praradimas komplekse buvo pasiūlytas Luo et al., 2016. *S. thermophilus* (St) Cascade komplekso natyvi masių spektrometrija su 12 nt trumpesne crRNR parodė, kad mėginį sudaro šios dalelės: pagrindinė dalis, turinti stechiometriją $(Cse1)_1(Cse2)_1(Cas7)_4(Cas5)_1(Cas6)_1(crRNA)_1$ ir du subkompleksai, kurių vienas neturėjo didžiojo subvieneto Cse1, o kiti dar $(Cas7)_2$ ir $(Cse2)_1$. Sekant tendenciją, St-Cascade kompleksas, surinktas ant 12 nt ilgesnės crRNA negu WT, turėjo du papildomus Cas7 ir vieną Cse2. Siūloma Cascade kompleksų susidarymo schema pavaizduota Pav. 4. Cse1 didžiojo subvieneto, atpažįstančio PAM seką, praradimas stebimas ir WT tipo komplekse, taip parodant, jog šios sąveikos komplekse yra silpniausios. Labiausiai Cse1 praradimas buvo būdingas kompleksams, kurių ilgiai nuo crRNR skyrėsi po 6 nt. Šie kompleksai atitinkamai 6 nt prailginus/sutrumpinus crRNR prisijungė/prarado po vieną Cas7 ir Cse2 subvienetą (Pav. 5), be to, prarasdavo Cse1, kuris yra būtinas PAM sekos atpažinimui. Šie natyvios masių spektrometrijos rezultatai paaiškina, kodėl funkcionalias R-kilpas galėjo sudaryti tik kompleksai, besiskiriantys po 12 nt nuo laukinio tipo kompleksų ir turintys Cse1, o 6 nt nuo WT crRNR

besiskiriančių kompleksų, kurie prarasdavo Cse1, R-kilpos nebuvo homogeniškos ir funkcionalios.

Pagrindiniai publikacijos 3 teiginiai:

- Natyvi masių spektrometrija leidžia chakaterizuoti Cascade I-E tipo kompleksų heterogeniškumą;
- $(Cas7)_2(Cse2)_1$ baltymai yra įterpiami/prarandami į/iš Cascade komplekso, kuomet crRNR prailgėja/sutrumpėja per 12 nt;
- Kuomet crRNR yra prailginama arba sutrumpinama 6 nt, nebegali susiformuoti stabilus Cse2 karkasas, kas greičiausiai nulemia Cse1 didžiojo subvieneto disociaciją iš komplekso.



Paveikslas 5. I-E tipo Cascade kompleksų susidarymo mechanizmas, keičiant crRNR ilgį. Prailginus crRNR 6 nt, prie komplekso prisijungia papildomai Cas7 (mėlynas) ir Cse2 (gelsvas), tačiau disocijuoja Cse1 didysis subvienetas (raudonas). Prailginus crRNR 12 nt, prie komplekso prisijungia papildomai du Cas7 subvienetai ir vienas Cse2, o Cse1 nedisocijuoja, susidarant funkcionaliam Cascade kompleksui. Pritaikyta pagal (Wittig et al., 2020).

Taikinio atpažinimas, panaudojant skirtingus I-F tipo Csy kompleksus *(publikacija 4)*

Šioje publikacijoje yra aprašytas *Aggregatibacter actinomycetemcomitans D7S-1* (Aa) I-F tipo Csy komplekso taikinio atpažinimo tyrimas. I-F tipo kompleksai skiriasi nuo I-E tipo tuo, jog neturi mažųjų subvienetų, kurie I-E tipo komplekso taikinio atpažinimo mechanizme atlieka R-kilpos užrakinimą. Norėdami patikrinti taikinio užrakinimo modelį I-F tipo sistemoje, mes sukonstravome Csy kompleksus, kurie turėtų 14-144 nt ilgio crRNR skirtuką ir tikrinome, ar šie kompleksai geba suformuoti R-kilpas. Panašiai kaip ir I-E tipo Cascade kompleksai, I-F tipo sistema suformavo funkcionalius kompleksus ir R-kilpas. Toliau mes tikrinome, ar šie kompleksai geba užtikrinti ląstelių atsparumą prieš fagą M13mp18, turintį taikinio proskirtuką ir ar geba apsaugoti nuo plazmidžių pUC19 patekimo į ląsteles. Stebima, jog Csy kompleksas geba riboti fagus arba plazmides, kuomet susidaro WT arba ilgesnė R-kilpa. Plazmidžių ribojimas nebevyksta tada, kai R-kilpa yra trumpesnė negu būdinga WT kompleksui (<32 bp). Tuo tarpu su trumpesnėmis R-kilpomis fagų ribojimas dar vyksta 100 kartų geriau lyginant su proskirtuko neturinčiu variantu, taip parodant, jog Cas2/3 taikinio degradacija nėra būtina fagų DNR ribojimui, o užteka Csy komplekso susirišimo.

Pagrindiniai publikacijos 4 teiginiai:

- I-F tipo Csy kompleksai geba formuoti R-kilpas nuo 14 iki 144 nt ilgio;
- I-F tipo kompleksai, taip pat kaip ir I-E tipo kompleksai, patikrina savo taikinį susidarant pilno WT ilgio R-kilpoms. Ilgesnės R-kilpos susidarymas nėra kritinis taikinio atpažinimui ir degradacijai;
- Csy komplekso prisijungimas prie fago M13mp18 nukleorūgšties be Cas2/3 kirpimo geba riboti fago daugininimąsi, tačiau negali riboti pUC19 plazmidžių patekimo į ląstelės, nes tam būtina Cas2/3 degradacija.

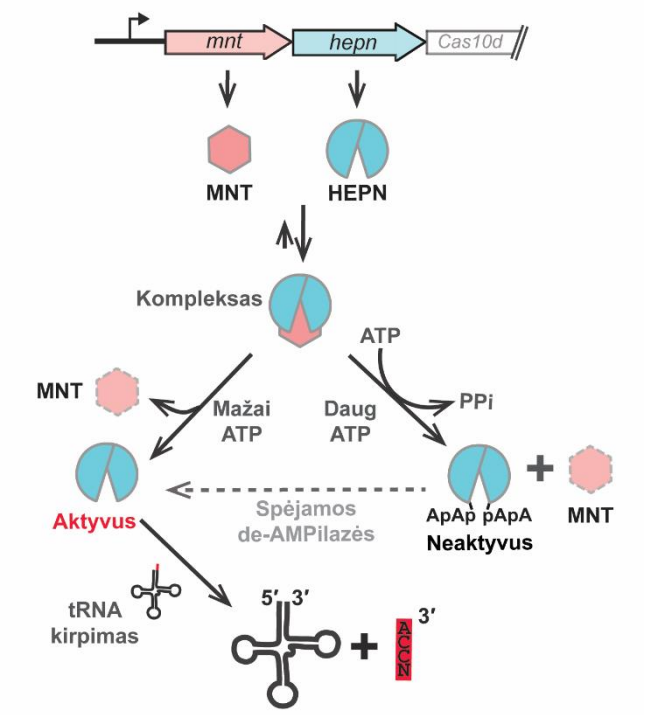
OligoAMPilinimas HEPN-MNT toksino-antitoksino poroje *(publikacija 5)*

Melsvabakterės *Aphanizomenon flos-aquae* I-D tipo CRISPR-Cas sistemos operone aptikta toksino-antitoksino sistema. Ją charakterizavus bioinformatiniais metodais paaiškėjo, kad antitoksinas yra homologiškas į

Polβ-panašioms nukleotidiltransferazėms, o toksinas turi HEPN RNazės aktyvųjį centrą. Prieš tai homologiška sistema buvo aprašyta kaip II tipo TA sistema iš bakterijų *Shewanella oneidensis* (Jia et al., 2018). Darbe norėjome patikrinti *Aphanizomenon flos-aquae* HEPN-MNT sistemos veikimą biocheminiais metodais ir nustatyti ribonukleazės blokavimo mechanizmą. Tam pradžioje buvo siekiama išgryninti šių baltymų kompleksą, naudojant afininius His_6-inkarus. Pastebėta, jog HEPN ir MNT baltymai nesudaro stabilaus komplekso ir juos galima išgryninti kaip pavienius. Masių spektrometrijos metodu nustačius baltymų molekulinį svorį, paaiškėjo, kad HEPN molekulinis svoris yra apie 659 Da didesnis negu turėtų būti pagal teorinius skaičiavimus. Kadangi jau anksčiau buvo hipotezių, kad HEPN gali būti neutralizuojamas prie jo prijungiant AMP (Anantharaman et al., 2013b), paskaičiavome, kad prijungta molekulinė masė atitinka di-AMP grandinę. Tuomet buvo ieškota aminorūgšties pozicijos, prie kurios prijungta ši grandinė. Tam sukonstruotas HEPN sekų palyginys, kuris leido aptikti konservatyvią tirozino aminorūgštį, esančią šalia HEPN aktyviojo centro. Kadangi aminorūgštis nepriklauo aktyviajam centrui, tačiau yra konservatyvi, buvo padaryta išvada, jog tai galėtų būti potenciali modifikacijos pozicija. Norint sukonstruoti HEPN mutantus, kuomet tirozinas pakeistas į fenilalaniną Y109F, kad modifikavimas nevyktų, to nepavyko padaryti dėl jų toksiškumo, tad papildomai sukonstruotas aktyviojo centro Arg102Ala mutantas ($M_{w\ teor}$ +659 Da) ir dvigubas R102A+Y109F mutantas, kurių masės nustatinėtos masių spektrometrija. Šio R102A+Y109F mutanto molekulinė masė atitiko teorinius skaičiavimus, taip įrodant, kad būtent Tyr109 aminorūgštis yra reikalinga HEPN di-AMPilinimui, o Y109F mutacija panaikina –OH grupę, reikiamą modifikacijos prijungimui.

Toliau buvo vykdomi HEPN ribonukleazės aktyvumo tyrimai kuomet tirta, kokius taikinius kerpa aktyvus HEPN toksinas. Negavus toksino raiškos *E. coli* ląstelėse, jo gavimui panaudotas beląstelinis *E. coli* ekstraktas. Pastebėta, kad aktyvus toksinas kerpa šiame ekstrakte esančias mažas RNR (trumpesnes negu 200 nt), tad panaudojus mažųjų RNR sekoskaitos eksperimentus, nustatinėta, kokius taikinius kerpa HEPN. Lyginant aktyvų ir neaktyvų toksiną turinčius mėginius, nustatyta, kad HEPN pagrindinai pašalina 4 nt nuo tRNR 3′-galo stiebo: CCA ir ketvirtą nukleotidą, vadinamą diskriminuojančiu. Toks RNR kirpimo fermentinis aktyvumas iki šiol nebuvo aptiktas. Remiantis HEPN-MNT poros veikimu, pasiūlytas inaktyvacijos mechanizmas, kurio veikimas priklauso nuo ATP koncentracijos ląstelėje (Pav. 6). Esant pakankamai ATP koncentracijai, MNT geba užmodifikuoti HEPN toksiną ir užblokuoti jo aktyvumą prieš tRNR. Kuomet ATP

koncentracija krenta, pvz. esant streso sąlygoms arba fago infekcijai, HEPN nebemodifikuojamas ir geba kirpti tRNR.



Paveikslas 6. Toksino-antitoksino sistemos HEPN-MNT siūlomas veikimo mechanizmas. Pradžioje TA pora HEPN-MNT sudaro kompleksą. Esant pakankamam ATP kiekiui ląstelėje, toksinas di-AMPilinamas išjungiant jo aktyvumą. Esant mažai ląstelinei ATP koncentracijai, toksinas lieka nemodifikuotas ir geba kirpti 4 nt nuo tRNR 3′-stiebo galų. Paveikslas adaptuotas iš (Songailiene et al., 2020).

REZULTATŲ APTARIMAS

Šioje disertacijoje tirtos I tipo CRISPR-Cas sistemos ir su I-D tipo CRISPR-Cas sistema asocijuota toksino-antitoksino pora. Priešingai negu 2 klasės sistemos, kurias sudarytos iš vieno baltymo efektorius (Cas9, Cas12, Cas13), 1 klasės sistemos formuoja daugiabaltyminį kompleksą, kuris atpažįsta taikinio dgDNR ir formuoja ant jos R-kilpą. Susiformavusi R-kilpa pritraukia nukleazę-helikazę Cas3 taikinio degradacijai. 1 klasės kompleksai lyginant su 2 klasės sistemomis atpažįsta ilgesnius taikinius, tad galėtų būti panaudojami tikslesniam genomo redagavimui. Šis darbas pradėtas nuo pavienių molekulių tyrimų, kaip I-E tipo Cascade kompleksai atpažįsta savo taikinį. Analizė atskleidė, kad taikinys atpažįstamas kryptingai nuo PAM sekos susidarant R-kilpai. Taikinio atpažinimo tikslumas tikrinamas priešingoje pusėje nuo PAM sekos ir esant komplementarumui toje srityje, įvyksta „R-kilpos užrakinimas“. Šis mechanizmas būdingas I-E tipo Cascade kompleksui, tačiau nebūdingas Cas9 (Szczelkun et al., 2014a). Toliau siekiant patikrinti, ar galima keisti taikinio užrakinimo poziciją ir susidriusios R-kilpos ilgį, sukonstruotos crRNR, turinčios ilgesnius arba trumpesnius skirtukus ir stebėtas šių kompleksų taikinių atpažinimas. Parodyta, jog keičiantis crRNR ilgiui, susidaro skirtingo ilgio R-kilpos, tačiau taikinio užrakinimo pozicija lieka toje pačioje vietoje kaip ir laukinio tipo komplekse. Tai parodo, kad didesnio taikinio atpažinimo tikslumo su I-E tipo Cascade keičiant R-kilpos ilgį nepavyko pasiekti. Kadangi I-F tipo kompleksas nekoduoja atskiro mažojo subvieneto, kilo klausimas, kokiu mechanimu jis atpažįsta savo taikinius ir ar šis kompleksas gebėtų sudaryti skirtingo ilgio R-kilpas. Parodyta, kad naudojant tokio ilgio crRNR, kurios būdingos WT, arba ilgesnes, I-F tipo Csy kompleksas geba riboti fagų arba plazmidžių nukleorūgščių patekimą į ląsteles. Ribojimo efektyvumas paklaidos ribose nesiskyrė. Tai rodo, kad WT-ilgio crRNR yra pilnai pakankamos I-F tipo komplekso veikimui ir crRNR ilginimas įtakos komplekso veikimui nepadaro, plazmidžių ar fagų ribojimo efektyvumo nepagerina. Apibendrinant, šie rezultatai parodo, kad tiek I-E, tiek I-F tipo efektoriniai kompleksai prisitaikė atpažinti savo taikinius su CRISPR regiono koduojamų skirtukų ilgiu ir crRNR ilgio pakeitimas neužtikrina geresnio efektyvumo ir tikslumo atpažįstant taikinius.

CRISPR-Cas sistemų veikimui kartais panaudojami papildomi pagalbiniai baltymai, kurie geba sustiprinti CRISPR-Cas sistemų poveikį išvengiant bakteriofagų infekcijos. Šiame darbe aptikta tame pačiame operone kaip I-D tipo CRISPR-Cas genai koduojama toksino-antitoksino sistema

HEPN-MNT. Atskleista, kad šioje sistemoje toksino neutralizavimui antitoksinas panaudoja fermentinį aktyvumą – di-AMPilina kilpą, kurioje yra HEPN RNazės aktyviojo centro aminorūgštys, taip pašalindamas toksino žalingą poveikį ir pakeisdamas RNazės aktyvaus centro kilpos konformaciją. Aktyvi HEPN RNazė veikia kirpdama 4 nt nuo 3′-tRNA stiebų. Daugumos toksinų taikiniai yra tRNR ir jie veikia tiek modifikuodami tRNR, tiek ją kirpdami. GNAT toksinai aminoacilina fMet-tRNR, kad ši nebegalėtų patekti į baltymų sintezės iniciacijos kompleksą (Jurėnas et al., 2017). Dalis VapC toksinų kerpa iniciatorinę tRNR arba kitas tRNR šalia antikodono kilpos ir taip blokuoja transliaciją (Winther and Gerdes, 2011; Winther et al., 2016). MenT toksinas, veikiantis kaip nukleotidiltransferazė, perneša CTP arba UTP molekules ant 3′-tRNA stiebų (Cai et al., 2020), dėl ko jos nebegali būti užkraunamos reikiama aminorūgštimi. RelA/SpoT baltymų homologas, toksinas RSH, veikia pirofosforilindamas 3′-tRNA stiebus (Kurata et al., 2021). Tai rodo, kad transliacijos inhibicija, pasitelkiant įvairius mechanizmus nukreiptus prieš tRNR, yra dažna strategija toksinų veikimui. Nėra aišku, kas tiksliai aktyvuoja HEPN toksiną – galimai dėl ATP trūkumo HEPN-MNT kompleksas tampa nestabilus ir atsiskiria laisvas HEPN toksinas, tačiau gali būti ir specialių fermentų – de-AMPilazių, kurie nuima nuo neaktyvaus HEPN modifikaciją, taip paleisdami aktyvų toksiną. Kodėl būtent ši TA pora yra randama tame pačiame operone ir kaip I-D CRISPR-Cas sistemos buvimas šalia šios TA sistemos yra susiję, lieka neatsakytas klausimas. Galimai Cas3′-helikazė, įsijungus I-D efektorinio kompleko veikimui kuomet šis susijungia su dgDNR proskirtuku, staigiai sunaudoja ląstelinį ATP ir HEPN toksinas nebegali būti efektyviai modifikuojamas, tačiau ši hipotezė dar nėra patikrinta eksperimentiškai.

IŠVADOS

1. I-E tipo Cascade kompleksas atpažįsta savo taikinį kryptingai, pradedant nuo PAM sekos (**publikacija 1**);
2. Labiausiai nuo PAM nutolusi taikinio dalis yra atsakinga už taikinio stablizaciją ir tolesnį taikinio nukreipimą degradacijai, vykdomai Cas3 nukleazės-helikazės (**publikacija 1 ir 2**);
3. I-E tipo ir I-F tipo kompleksai geba formuotis ant kitokio nei WT ilgio crRNR molekulių ir sudaryti R-kilpas (**publikacija 2, 3 ir 4**);
4. Netgi ir skirtingo ilgio R-kilpas sudarantys I-E ir I-F tipo kompleksai, naudoja taikinio atpažinimo mechanizmą, būdingą laukinio tipo kompleksams (**publikacija 2 ir 4**);
5. Toksino-antitoksino poroje HEPN-MNT, MNT veikia kaip nukleotidiltransferazė, modifikuojanti (di-AMPilinanti) HEPN toksiną ir taip užblokuodama šio toksino aktyvumą (**publikacija 5**);
6. Aktyvus HEPN toksinas nukerpa 4 nt nuo tRNR 3′-stiebo (**publikacija 5**).

CURRICULUM VITAE

Inga Songailienė

Tyrimų patirtis	
2020 Gegužė-Gruodis	Projekto „I-D tipo CRISPR-Cas sistemos tyrimai“ finansuojamo Vilniaus universiteto Mokslo fondo vadovė
2017-dabar	*Jaunesnioji mokslo darbuotoja*
2013-2017	*Biologė tyrėja*
2010-2013	*Laborantė*
2013 Vasaris-Rugpjūtis	Profesnė praktika mokslinių tyrimų laboratorijoje finansuojama EU.
2011 Lapkritis-2012 Sausis	Profesinė praktika mokslinių tyrimų laboratorijoje finansuojama Lietuvos mokslo tarybos.
2010 Birželis-Spalis	ERASMUS praktika laboratorijoje, Bristolio universitete, vadovas prof. H. Mellor.

Išsilavinimas	
2013 spalis - 2017	Doktorantūros studijos “CRISPR-Cas sistemų charakterizavimas”, vadovas prof. Virginijus Šikšnys
2011-2013	Magistrantūros biochemijos studijos *magna cum laude,* Vilniaus universitetas
2007-2011	Bakalauro biochemijos studijos *cum laude,* Vilniaus universitetas

Metodai

Klonavimas ir baltymų ekspresija *E. coli*, RNR, DNR elektroforezė, radioaktyvus nukleorūgščių žymėjimas, NDS-PAGE, baltymų skysčių chormatografija, aukšto slėgio skysčių chromatografija, masių spektrometrija, Western blotas, Northern blotas, RNR išskyrimas, RNR-sekoskaitos bibliotekos paruošimas, baltymų kristalizacija

Pagrindiniai apdovanojimai	
2013	Geriausias magsitro darbas, apgintas 2013m. biomokslų srityje „MnlI restrikcijos endonukleazės N-domeno sąveikos su DNR tyrimas“, vadovė dr. G. Tamulaitienė.
2020	Geriausia Vilniaus univeristeto gamtos mokslų metų publikacija.

Pagrindinės konferencijos	
2021	**Songailiene I**, Juozapaitis J, Tamulaitiene G, Sasnauskas G, Venclovas Č, Siksnys V “HEPN-MNT toxin-antitoxin as a bacterial ATP sensor”. 2021 Open Readings, žodinis pranešimas, Lietuva

2021	**Songailiene I**, Juozapaitis J, Tamulaitiene G, Sasnauskas G, Venclovas Č, Siksnys V "HEPN-MNT toxin-antitoxin as a bacterial ATP sensor". 2021 EMBO meeting "New approaches and concepts in microbiology", žodinis ir stendinis pranešimas, virtuali konferencija
2017	**Songailiene I**, Rutkauskas M, Šinkunas T, Seidel R, Siksnys V "Target recognition by Cascade complexes bearing altered length crRNA" plakatas FEBS 2017 ir žodinis pranešimas FEBS Jaunųjų mokslininkų formume, 2017, Izrealis
2016	**Songailiene I**, Sinkunas T, Rutkauskas M, Seidel R ir Siksnys V, "Cascade complex reconstitution in vitro and R-loop formation by single molecule experiments", plakatas Open Readings 2016, Lietuva
2016	**Songailiene I**, T. Šinkunas, M. Rutkauskas, R. Seidel, V. Šikšnys "Reconstitution of Cascade complex *in vitro* and R-loop formation", plakatas "CRISPR 2016", Izraelis
2014	EMBO Practical course "Non-coding RNR in infection", žodinis pranešimas "CRISPR-Cas systems in bacterial defense".
Kursai	
2016	Nekoduojanti RNR infekcijose, EMBO praktinis kursas, Viurzburgas, Vokietija
2014	Modernių metodų biokristalografijoje pagrindai Biocrys2, Oreiras, Portugalija
2012	Insubria tarptautinė mokykla sveikatai ir biomokslams, Como, Italija.
Moksliniai vizitai	
2015 Gruodis	Paveinių molekulių magnetinių pincetų eksperimentai, vadovas prof. Ralf Seidel, Leipcigo universitetas, finansuojama Lietuvos mokslo tarybos
Mokymas	
2017-dabar	„CRISPR-Cas biologija" kursas biochemijos ir medicinos genetikos magistro studentams. Atsakinga už praktinius darbus ir seminarus.
Domėjimosi sritys	

Nekoduojančios RNR, signalinės molekulės, kovalentinės modifikacijos, baterijų apsaugos nuo virusų sistemos, CRISPR-Cas tyrimai ir šių sistemų panaudojimas

Užsienio kalbos: Anglų (sklandžiai), Vokiečių (pagrindai)

PUBLIKACIJŲ SĄRAŠAS

Publikacijų, kurių pagrindu parengta ši disertacija, sąrašas:

1. Rutkauskas M, Sinkunas T, **Songailiene I**, Tikhomirova MS, Siksnys V, Seidel R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep*. 2015 Mar 10;10(9):1534-1543. doi: 10.1016/j.celrep.2015.01.067.
   Atlikau biocheminius EMSA eksperimentus ir parengiau jų rezultatus publikavimui.
2. **Songailiene I*,** Rutkauskas M*, Sinkunas T, Manakova E, Wittig S, Schmidt C, Siksnys V, Seidel R. Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length. *Cell Rep*. 2019 Sep 17;28(12):3157-3166.e4. doi: 10.1016/j.celrep.2019.08.033
   Parengiau tyrimo planą ir gavau pradinius pavienių molekulių eksperimentų rezultatus, parengiau publikacijos rankraštį ir jį redagavau.
3. Wittig S, **Songailiene I**, Schmidt C. Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry. *J Am Soc Mass Spectrom*. 2020 Mar 4;31(3):538-546. doi: 10.1021/jasms.9b00011.
   Išgryninau ir biochemiškai charakterizavau Cascade kompleksus, redagavau publikacijos rankraštį.
4. Tuminauskaite D*, Norkunaite D*, Fiodorovaite M, Tumas S, **Songailiene I**, Tamulaitiene G, Sinkunas T. DNA interference is controlled by an R-loop length in a type I-F1 CRISPR-Cas system. *BMC biology*. 2020 Jun 15;18(1):65. doi: 10.1186/s12915-020-00799-z.
   Atlikau in vitro ir in vivo eksperimentus su I-F Csy kompleksais.
5. **Songailiene I***, Juozapaitis J*, Tamulaitiene G, Ruksenaite A, Sulčius S, Sasnauskas G, Venclovas Č, Siksnys V. HEPN-MNT toxin-antitoxin system: the toxin is neutralized by OligoAMPylation. Mol Cell, 2020 Dec 17;80(6):955-970.e7. doi: 10.1016/j.molcel.2020.11.034. Epub 2020 Dec 7.
   Parengiau tyrimo planą, atlikau in vitro eksperimentus, rašiau ir redagavau publikacijos rankraštį.

SANTRAUKOS LITRATŪROS SĄRAŠAS

1. Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E.V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct 8, 15.
2. Aravind, L., and Koonin, E.V. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. 27, 1609–1618.
3. Athukoralage, J.S., Rouillon, C., Graham, S., Grüschow, S., and White, M.F. (2018). Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. Nature 562, 277–280.
4. Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. Nat. Rev. Microbiol. 18, 113–119.
5. Cai, Y., Usher, B., Gutierrez, C., Tolcan, A., Mansour, M., Fineran, P.C., Condon, C., Neyrolles, O., Genevaux, P., and Blower, T.R. (2020). A nucleotidyltransferase toxin inhibits growth of Mycobacterium tuberculosis through inactivation of tRNA acceptor stems. Sci. Adv. 6, eabb6651.
6. Carte, J., Pfister, N.T., Compton, M.M., Terns, R.M., and Terns, M.P. (2010). Binding and cleavage of CRISPR RNA by Cas6. RNA N. Y. N 16, 2181–2188.
7. Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics 8, 172.
8. Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B.G., Rajashankar, K., Bailey, S., Wiedenheft, B., and Ke, A. (2016). Structural basis for promiscuous PAM recognition in type I–E Cascade from E. coli. Nature 1–16.
9. Jia, X., Yao, J., Gao, Z., Liu, G., Dong, Y., Wang, X., and Zhang, H. (2018). Structure–function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN–MNT toxin–antitoxin. J. Biol. Chem. jbc.RA118.002421-jbc.RA118.002421.
10. Jurėnas, D., Chatterjee, S., Konijnenberg, A., Sobott, F., Droogmans, L., Garcia-Pino, A., and Van Melderen, L. (2017). AtaT blocks

translation initiation by N-acetylation of the initiator tRNAfMet. Nat. Chem. Biol. 13, 640–646.

11. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science 357, 605–609.
12. Kurata, T., Brodiazhenko, T., Alves Oliveira, S.R., Roghanian, M., Sakaguchi, Y., Turnbull, K.J., Bulvas, O., Takada, H., Tamman, H., Ainelo, A., et al. (2021). RelA-SpoT Homolog toxins pyrophosphorylate the CCA end of tRNA to inhibit protein synthesis. Mol. Cell S1097276521004524.
13. Lin, J., Fuglsang, A., Kjeldsen, A.L., Sun, K., Bhoobalan-Chitty, Y., and Peng, X. (2020). DNA targeting by subtype I-D CRISPR–Cas shows type I and type III features. Nucleic Acids Res. gkaa749.
14. Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. Annu. Rev. Virol. 7, annurev-virology-011620-040628.
15. Luo, M.L., Jackson, R.N., Denny, S.R., Tokmina-Lukaszewska, M., Maksimchuk, K.R., Lin, W., Bothner, B., Wiedenheft, B., and Beisel, C.L. (2016). The CRISPR RNA-guided surveillance complex in Escherichia coli accommodates extended RNA spacers. Nucleic Acids Res. 44, 7385–7394.
16. 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., et al. (2019). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol.
17. McBride, T.M., Schwartz, E.A., Kumar, A., Taylor, D.W., Fineran, P.C., and Fagerlund, R.D. (2020). Diverse CRISPR-Cas complexes require independent translation of small and large subunits from a single gene (Molecular Biology).
18. McMahon, S.A., Zhu, W., Graham, S., Rambo, R., White, M.F., and Gloster, T.M. (2020). Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. Nat. Commun. 11, 500.
19. Mulepati, S., Héroux, A., and Bailey, S. (2014). Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. Science 1–9.
20. Niewoehner, O., Garcia-Doval, C., Rostøl, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., and Jinek, M. (2017). Type III

CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature 548, 543–548.

21. Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Miettinen, H.M., Santiago-Frangos, A., Faith, D., Lawrence, C.M., Lander, G.C., and Wiedenheft, B. (2019). Structure Reveals a Mechanism of CRISPR-RNA-Guided Nuclease Recruitment and Anti-CRISPR Viral Mimicry. Mol. Cell 74, 132-142.e5.
22. Rostøl, J.T., Xie, W., Kuryavyi, V., Maguin, P., Kao, K., Froom, R., Patel, D.J., and Marraffini, L.A. (2021). The Card1 nuclease provides defence during type-III CRISPR immunity. Nature.
23. Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V., and Seidel, R. (2015). Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. Cell Rep. 10, 1534–1543.
24. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., and Westra, E.R. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. 108, 10098–10103.
25. Semenova, E., Kuznedelov, K., Datsenko, K.A., Boudry, P.M., Savitskaya, E.E., Medvedeva, S., Beloglazova, N., Logacheva, M., Yakunin, A.F., and Severinov, K. (2015). The Cas6e ribonuclease is not required for interference and adaptation by the E . coli type I-E CRISPR-Cas system. 1–13.
26. Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO J. 1–10.
27. Songailiene, I., Rutkauskas, M., Sinkunas, T., Manakova, E., Wittig, S., Schmidt, C., Siksnys, V., and Seidel, R. (2019). Decision-Making in Cascade Complexes Harboring crRNAs of Altered Length. Cell Rep. 28, 3157-3166.e4.
28. Songailiene, I., Juozapaitis, J., Tamulaitiene, G., Ruksenaite, A., Šulčius, S., Sasnauskas, G., Venclovas, Č., and Siksnys, V. (2020). HEPN-MNT Toxin-Antitoxin System: The HEPN Ribonuclease Is Neutralized by OligoAMPylation. Mol. Cell S1097276520308340.
29. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., and Karvelis, T. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. 6–11.
30. Tuminauskaite, D., Norkunaite, D., Fiodorovaite, M., Tumas, S., Songailiene, I., Tamulaitiene, G., and Sinkunas, T. (2020). DNA

interference is controlled by R-loop length in a type I-F1 CRISPR-Cas system. BMC Biol. 18, 65.

31. Winther, K.S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc. Natl. Acad. Sci. 108, 7403–7407.
32. Winther, K., Tree, J.J., Tollervey, D., and Gerdes, K. (2016). VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation. Nucleic Acids Res. 44, 9860–9871.
33. Wittig, S., Songailiene, I., and Schmidt, C. (2020). Formation and Stoichiometry of CRISPR-Cascade Complexes with Varying Spacer Lengths Revealed by Native Mass Spectrometry. J. Am. Soc. Mass Spectrom. 31, 538–546.
34. Xiao, Y., Xiao, Y., Luo, M., Dolan, A.E., Liao, M., and Ke, A. (2018). Structure basis for RNA-guided DNA degradation by Cascade and Cas3. 0839, 1–12.
35. Zhu, W., McQuarrie, S., Grüschow, S., McMahon, S.A., Graham, S., Gloster, T.M., and White, M.F. (2021). The CRISPR ancillary effector Can2 is a dual-specificity nuclease potentiating type III CRISPR defence. Nucleic Acids Res. gkab073.

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