VILNIUS UNIVERSITY
LIFE SCIENCES CENTRE

EGLĖ KUPČINSKAITĖ
Molecular Biotechnology study programme

Master Thesis

INVESTIGATION OF A TYPE I-F CRISPR-CAS INHIBITION BY ACRIF6 AND ACRIF9

VU LSC
Institute of Biotechnology
Department of Protein – DNA Interactions
Supervisor dr. Tomas Šinkūnas

Vilnius 2022

INVESTIGATION OF A TYPE I-F CRISPR-CAS INHIBITION BY ACRIF6 AND ACRIF9

Master thesis was prepared at Vilnius University Institute of Biotechnology Department of Protein – DNA Interctions (head of laboratory: professor Virginijus Šikšnys)

Student:
Eglė Kupčinskaitė _____
(Signature)

Supervisor:
Dr. Tomas Šinkūnas _____
(Signature)

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Aca	anti-CRISPR associated
Acr	anti-CRISPR
Cas	CRISPR-associated
Cascade	CRISPR-associated complex for antiviral defence
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
crRNA	CRISPR RNA
DNase	doexyribonuclease
IPTG	isopropyl beta-D-1-thiogalactopyranoside
LB	Luria-Bertani medium
PAM	protospacer adjacet motif
pre-crRNA	precursor CRISPR RNA
RNase	ribonuclease
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
tracrRNA	trans-activting CRISPR RNA

INTRODUCTION

Bacteriophages (phages) are a persistent threat to prokaryotic life. For this reason bacteria evolved to evade phage attacks by a number of strategies. One of such strategies is a prokaryotic system tht provides adaptive immunity against foreign nucleic acids known as CRISPR-Cas. This system coprises of CRISPR (Clustered Regurlarly Interspaced Short Palidromic Repeats) region and Cas (CRISPR-associated) protein encoding genes. CRISPR-Cas are divided into two classes, which are further divided into six types, each type comprises of several subtypes. Mechnism of action for these systems is composed of three closely linked stages: (i) adaptation where foreign nucleic acid fragment (spacer) is inserted into CRISPR region, (ii) crRNA biogenesis, where a trascript of CRISPR region is processed into small crRNA molecules and (iii) interference, where Cas-crRNA effector complex recognises DNA or RNA target and binds it. The target is cleaved by the effector complex or in some cases by an accessory Cas protein.

Due to evolutionary arms race bacteriophages coevolved and adapted strategies to evade bacterial defences. A fraction of phage genomes code for proteins, termed Anti-CRISPRs (Acrs), which inhibit CRISPR-Cas. Over 80 families of Acr proteins that are currently known are classified according to the type of the CRISPR-Cas that they inhibit. Most of the Acr proteins inhibit interference stage of CRISPR-Cas and work by inhibiting target binding or its cleavage.

Type I-F CRISPR-Cas encodes six Cas proteins (Cas1, Cas2/3, Casf5f, Cas6f, Cas7f, Cas8f). Four of these proteins (Cas5f, Cas6f, Cas7f, Cas8f) together with crRNA comprise an effector complex, termed Cascade. The complex specifically recognises DNA target with complementary sequence to crRNA (protospacer) and PAM (Protospacer Adjacent Motif). Upon crRNA recognition of DNA, a structure termed R-loop is formed, which engages Cas2/3 nuclease/helicase for the target DNA hydrolysis. Twenty four families of AcrIFs (AcrIF1-24) are currently known to inhinit type I-F CRISPR-Cas. The AcrIFs either interact with the Cascade and prevent DNA target binding or inhibit hydrolysis by Cas2/3.

CRISPR-Cas systems have became a significant molecular tool for genome editing and regulation of gene expression. Acr proteins could serve as a tool to control CRISPR-Cas, thus there is a signifiant amount of research dedicated to reveal mechanisms of action of these proteins. This work focuses on AcrIF6 and AcrIF9 and their impact on type I-F CRISPR-Cas from *Aggregtibacter Actinomycetemcomitans* D7S-1 (Aa-CIRSPR-Cas) bacterium.

AIMS AND OBJECTIVES

Research object:

Anti-CRISPR (Acr) proteins that inhibit type I-F CRISPR-Cas from *Aggregatibacter Actinomyvetemcomitans* D7S-1 (Aa-CRISPR-Cas).

Aim

Determine molecular mechanisms by which AcrIF6 and AcrIF9 inhibit type I-F Aa-CRISPR-Cas.

Objectives

1. Investigate the mode of action for AcrIF6 and AcrIF9 against Aa-CRISPR-Cas in vitro.
2. Crystallise and determine structures of the AcrIFs.
3. Conduct structure-guided mutagenesis and assay activity of the AcrIF mutants.

1. LITERATURE OVERVIEW

1.1 Evolutionary arms race between bacteria and phages

Bacteriophages pose a persistent threat to prokaryotic life. Consequently, bacteria have evolved numerous strategies to evade phage infection (Dy et al., 2014, Hampton et al., 2020). One of the best characterised antiphage defence systems includes R-M (restriction-modification), Abi (abortive infection) and CRISPR-Cas systems (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated). The extensive mutation rates enable phages to adapt counter defence strategies escaping prokaryotic defence (Samson et al., 2013; Dy et al., 2014, Hampton et al., 2020).

Anti-viral defence mechanisms can be divided into three large groups: (i) nucleic acid targeting systems (for example, R-M and CRISPR-Cas) (ii) Abi systems that lead to 'altruistic death' of host cell upon infection, and (iii) other systems. **R-M systems** destroy invading nucleic acids by employing restriction endonucleases (REases) that recognise specific sequences within phage genome and cleave them. To prevent self-targeting, epigenetic modification strategies of genome are used. R-M systems usually employ methyltransferases (MTases) to methylate cytosine and adenine at REase recognition sites only in host cell genome (Mruk & Kobayashi, 2014; Ershova et al., 2015). Group IV R-M systems are an exception because they target modified phage DNA while host cell DNA remains unmodified (Tock & Dryden, 2005). R-M systems are divided into four groups based on recognition sequence and mechanism of action and are found in more than 74% of prokaryotic genomes (Oliveira et al., 2014). **Abi** systems impede phage replication by committing 'altruistic death' of infected cell before phage reproductive cycle is completed. As a result, a phage is unable to replicate and non-infected cells survive. Abi systems are usually activated by a particular component, for example, a phage protein, a nucleic acid or cellular stress. Toxin-antitoxin systems, that have been shown to execute Abi, usually comprise of two proteins (i.e., toxin and antitoxin). When toxin and antitoxin molecules are bound, toxin is inactive. Upon a trigger, antitoxin molecule releases toxin and this results in cell death (Seed, 2015; Harms et al., 2018). **CRISPR-Cas** systems 'remember' invaders and prevent subsequent infections by employing ribonucleoprotein complexes (Hille et al., 2018). CRISPR-Cas comprise adaptive prokaryotic immunity. A detailed overview of CRISPR-Cas systems is provided in the following chapters.

Prokaryotic defence systems are clustered in genomic areas, termed defence islands (Makarova et al., 2013). Thorough genomic analysis of defence islands revealed more prokaryotic strategies for phage evasion in recent years. Some of these systems include BREX, pAgo and DISARM (Goldfarb et al., 2015; Gordeeva et al., 2019). It is assumed that the number of these systems in the defence islands is much greater. Additionally, new uncharacterised systems keep being reported. A recent

study has reported nine prokaryotic anti-phage defence systems, which have been named after protective deities from world mythology, including Gabija, Shedu, Zorya and others. Mechanisms of action for these systems remain to be disclosed (Doron et al., 2018).

The evidence shows that prokaryotic immunity comprises of a great variety of antiphage defences, some of which remain to be discovered. The research is accelerated by the known applications of R-M systems in molecular biology, and CRISPR-Cas in gene editing, diagnostics as well as therapeutics (Fellman et al., 2017; Felice et al., 2019; Javalkote et al., 2020). Potential of newly discovered systems serving as unique molecular tools has driven intensive research of prokaryotic immunity.

Rapid proliferation and high mutation rates of bacteriophages drive the evolution of numerous strategies to evade a broad arsenal of prokaryotic defence systems (Labrie et al., 2010; Samson et al., 2013). First, phages can escape R-M systems by acquiring mutations in R-M system recognition sites reducing the number of REase recognition sequences within phage genome. CRISPR-Cas system function can be abolished by mutations in PAM (Protospacer Adjacent Motif) or seed sequences of the protospacer. However, more robust means are required to escape prokaryotic defences. Therefore, phages code for anti-defence proteins, including anti-restriction proteins and anti-CRISPRs. These proteins are either injected into a cell with phage DNA or expressed early upon infection. The phage encoded MTases that methylate the phage genome prevent REase from cleaving or proteins that inhibit REases are introduced to the host cell (Labrie et al., 2010). CRISPR-Cas systems can be inhibited by small proteins, Anti-CRISPRs (Acrs), which will be reviewed in detail in the following chapters.

The antagonistic prokaryote and phage relationship has led to evolution of diverse antiphage defence systems. Evolvement of these mechanisms has been determined by constant coevolution of phages as they attempt to coexist. To survive, phages have employed countless strategies to overcome the vast diversity of bacterial defences.

1.2 CRISPR-Cas systems

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated) is a prokaryotic defence system that provides adaptive immunity against foreign nucleic acids (Barrangou et al., 2007; Terns & Terns, 2011; Wiedenheft et al., 2012; Hille et al, 2018). CRISPR-Cas immune systems are present in almost 90% archaeal and about 40% bacterial genomes (Makarova et al., 2020). In this chapter, CRISPR-Cas classification and mechanism of action will be reviewed.

CRISPR-Cas locus comprises of Cas protein encoding genes and CRISPR region (Haft et al., 2005). CRISPR region consists of short repeats interspaced by unique sequences (spacers) derived from phages or plasmids (Bolotin et al., 2005). The spacers encode information about encounters with invading nucleic acids and thus play a key role in adaptive immunity of CRISPR-Cas systems. Leader sequence in the CRISPR region has a regulatory function in adaptation and transcription initiation of CRISPR region (Alkhnbashi et al., 2016). *Cas* genes are usually organised in one operon, which is located nearby CRISPR locus (Figure 1.1.)

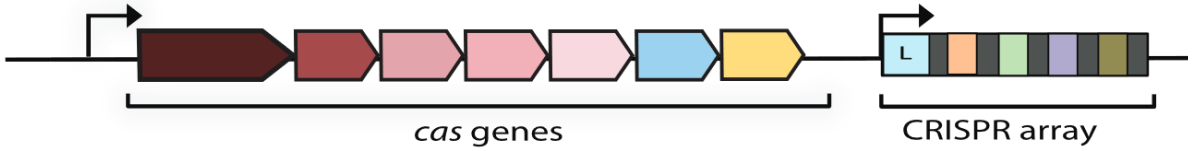

Figure 1.1. Organisation of CRISPR-Cas locus. CRISPR-Cas comprises of *cas* gene operon and CRISPR array. CRISPR region consists of spacers (marked as coloured rectangles) and repeats (marked as grey rectangles). Adapted from Mohanraju et al., 2016.

Classification of CRISPR-Cas systems is based on structural and functional differences between Cas proteins. CRISPR-Cas systems are divided into two large classes (class 1 and class 2) based on the organisation of effector complexes, which recognise and cleave foreign nucleic acids. In class 1 CRISPR-Cas, the effector complex comprises of several Cas protein subunits. By contrast, the effector of class 2 is a single multidomain protein (Makarova et al., 2020). Each class is divided into three types: I, III, IV and II, V, VI for classes 1 and 2, respectively. CRISPR-Cas types are characterised by unique Cas proteins and effector complexes. Five CRISPR-Cas types feature signature proteins: Cas3 for type I, Cas9 for type II, Cas10 for type III, Cas12 for type V and Cas13 for type VI (Figure 1.2).

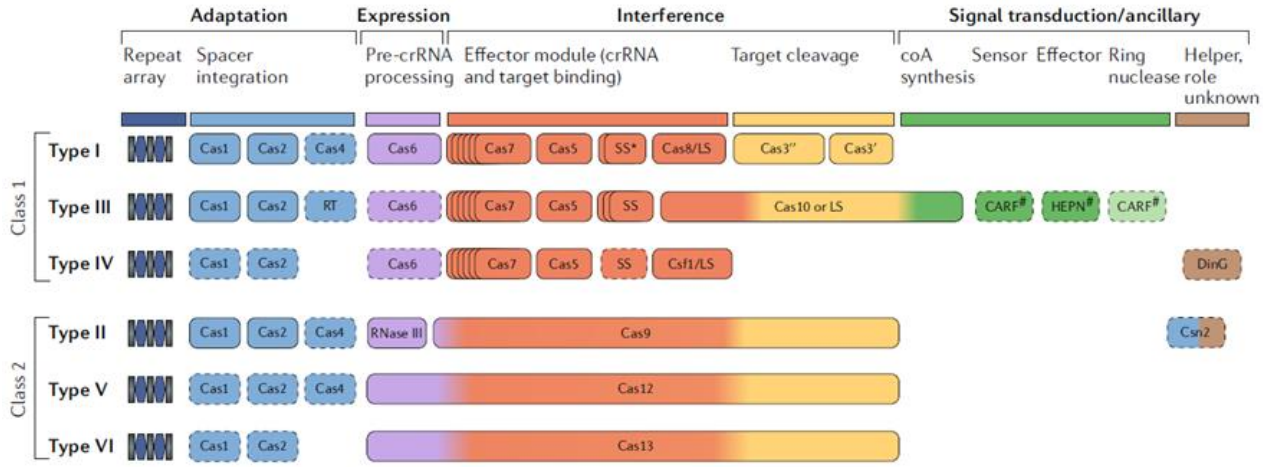

Figure 1.2. Classification of CRISPR-Cas systems. CRISPR-Cas systems are organized into two classes (1 and 2) and six types (I-VI). Division into classes is determined by differences in the structures of interference stage Cas proteins (Makarova et al., 2020).

Each CRISPR-Cas type includes several subtypes, among which differences are less pronounced and include, for example, unique locus organisation or subtype specific Cas proteins (Koonin & Makarova, 2019).

CRISPR-Cas mechanism of action comprises of three distinct but closely linked stages: (i) **adaptation**, (ii) **expression (crRNA biogenesis or maturation)** and (iii) **interference** (Mohanraju et al., 2016; Hille et al., 2018) (Figure 1.3).

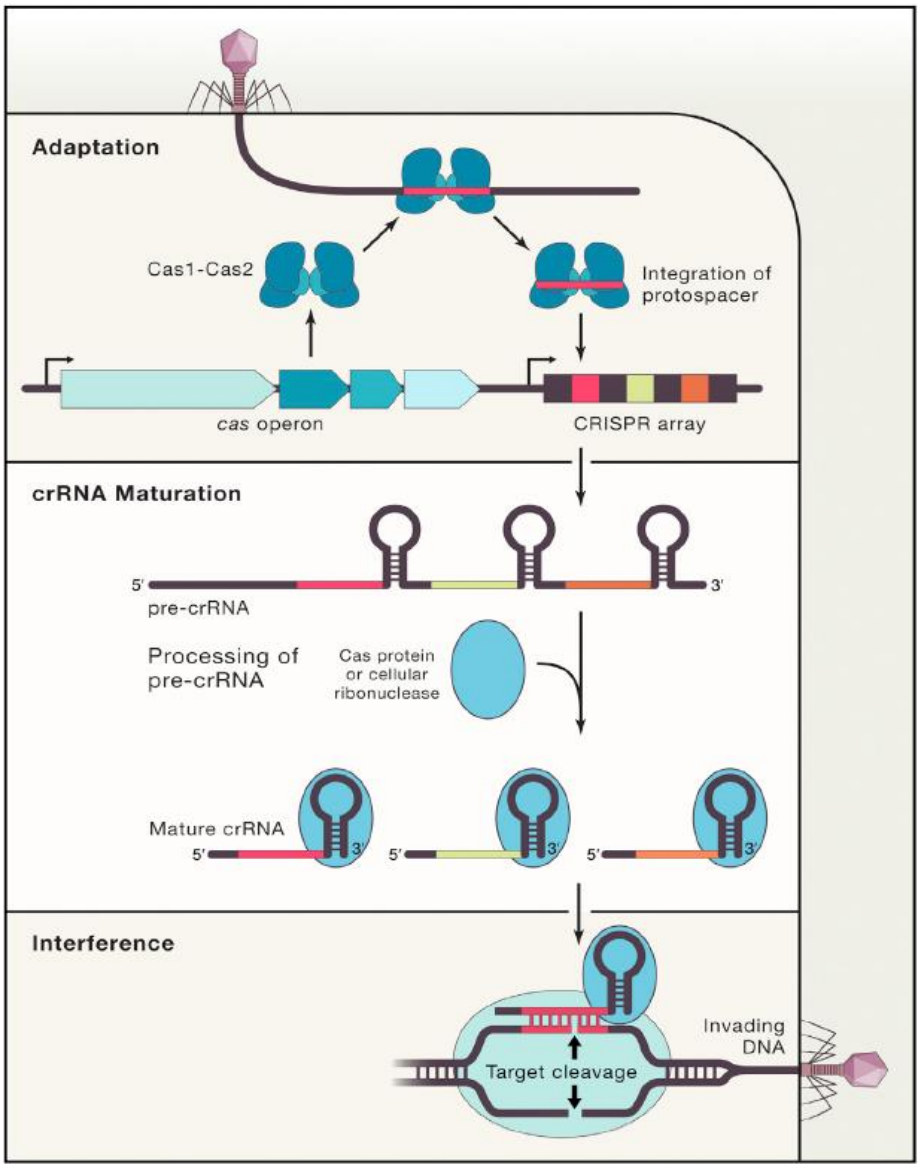

Figure 1.3. Three molecular stages of CRISPR-Cas immunity. During adaptation step, Cas1-Cas2 integrase excises a foreign nucleic acid fragment and inserts it into CRISPR array. In the following stage, crRNA biogenesis, pre-crRNA is transcribed and processed into crRNA. In the final stage, interference, Cas-crRNA surveillance complex recognises and destroys the invading nucleic acid (Hille et al., 2018).

In **adaptation** stage, Cas protein complex integrates a fragment of foreign nucleic acid (protospacer) into the CRISPR locus. During this process acquired spacers keep memories of prior

infections (Sternberg et al., 2016; McGinn & Marraffini, 2019). In some CRISPR-Cas systems spacers can be acquired from RNA molecules by employment of a reverse transcriptase which is encoded within CRISPR-Cas locus (Silas et al., 2016; Toro et al., 2019). In **expression (crRNA biogenesis)** stage, CRISPR locus is transcribed as a single pre-crRNA molecule that is processed to form mature crRNAs (Haurwitz et al., 2010; Charpentier et al., 2015). In the final stage, **interference**, crRNA with Cas proteins form ribonucleoprotein complex that recognises invading nucleic acids. Target is cleaved and inactivated by a Cas nuclease, which is either part of the effector complex or a single protein recruited upon target recognition (Hille et al., 2018).

CRISPR-Cas diversity is at least partially determined by ongoing evolutionary arms race between bacteria and pages. This is confirmed by vast amount of anti-CRISPR proteins that are currently identified to supress CRISPR-Cas immunity (Shivram et al., 2021).

1.2.1 Adaptation

Adaptation is the first stage in CRISPR-Cas immunity, in which an invading nucleic acid fragment (protospacer) is integrated into CRISPR array in host cell genome to become a spacer (Figure 1.3: 'Adaptation'). Acquisition of spacers determines the adaptivity and heritability of CRISPR-Cas. Cas1 together with Cas2 comprise a heterohexameric complex that serves as a spacer integrase (Wright et al., 2017). The integrase inserts spacer to be in the beginning of CRISPR array, which results in a chronological array of spacers. Spacer acquisition from previously unencountered invaders is termed naïve adaptation. In type I CRISPR-Cas spacers can be acquired via priming pathway (see chapter 1.3). Depending on the type of CRISPR-Cas system, additional Cas or host cell proteins participate in the adaptation stage (Hille et al., 2018).

1.2.2 crRNA biogenesis

The transcript of CRISPR region is processed to form small crRNAs (Figure 1.3: 'crRNA maturation'). Initially, a single pre-crRNA is transcribed, which is further cleaved within repeat sequences to form RNA molecules. These RNAs either function as crRNA or are further processed by host cell ribonucleases. The process of crRNA biogenesis is different among class 1 and class 2 CRISPR-Cas.

CRISPR transcripts in type I and type III CRISPR-Cas (class 1) systems are processed by Cas6 endoribonuclease. In most cases, Cas6 cuts at the base of each hairpin structure formed by repeats within pre-crRNA (Kunin et al., 2007; Marraffini & Sontheimer, 2008; Charpentier et al., 2015). After cleavage, Cas6 remains bound to crRNA, which later serves as a scaffold for effector complex

assembly (Charpentier et al., 2015). In type I-A, I-B, and III, CRISPR transcripts lack palindromic sequences in repeats (Kunin et al., 2007). Thus, a Cas6 dimer or an independent pair of Cas6 has to bend RNA into a structure that is favourable for cleaving (Shao & Li, 2013; Shao et al., 2016). Then Cas6 releases crRNA and is not present in the effector complex (Carte et al., 2008).

In class 2 systems, effector complexes or other host cell proteins process crRNA (Charpentier et al., 2015). TracrRNA (trans-activating CRISPR RNA) is required for crRNA maturation in type II CRISPR-Cas. TracrRNA is partially complementary to repeat sequences of CRISPR transcript. Cas9 stabilises the duplex, which is cleaved by RNase III (Deltcheva et al., 2011; Karvelis et al., 2013). In type V (except V-B) and type VI crRNA maturing is independent of tracrRNA and CRISPR transcript is processed by the effector complexes. Interestingly, unprocessed CRISPR transcript in these complexes can act as functional crRNA (Hochstrasser & Doudna, 2015).

1.2.3 Interference

In the interference stage of CRISPR-Cas immunity, Cas-crRNA ribonucleoprotein complexes bind nucleic acids that are complementary to crRNA and cleave them (Figure 1.3: 'Interference'). The effector in class 2 is comprised of a single protein and guide RNA complex. In contrast, class 1 systems code for multi-protein effector complexes that bind crRNA.

Type I, II and V effector complexes target a double-stranded DNA, which has a protospacer sequence and PAM (protospacer adjacent motif). The PAM functions as a trigger for discrimination of non-self DNA as its absence in the CRISPR array prevents self-cleavage. The key players in type I CRISPR-Cas interference stage are surveillance complex Cascade and Cas3 nuclease/helicase, or in case of Type I-F, Cas2/3. Cascade binds the target via crRNA and dsDNA complementary base pairing. Hybridisation of crRNA with the target results in displacement of the non-complementary strand and formation of a structure, termed R-loop. Subsequently, Cas3 (or Cas2/3) is recruited to cleave the displaced strand of DNA target (Sinkunas et al., 2011; Westra et al., 2013).

In type II systems, the Cas9-tracrRNA-crRNA complex recognises the DNA target. Similarly to type I systems, R-loop structure formation is induced upon DNA target binding. R-loop triggers nuclease activity of Cas9. The HNH domain of Cas9 hydrolyses the crRNA-complementary strand of the target, while the crRNA noncomplementary strand is cleaved by the RuvC domain, resulting in blunt ends of the DNA (Jinek et al., 2012; Gasiunas et al., 2012). The type V effector protein Cas12 assembles with a single crRNA molecule. Cas12 is phylogenetically similar to Cas9; however, it has only the RuvC active site. Upon target binding, the RuvC induces double stranded DNA breaks in the target (Zetsche et al., 2015). Transcription dependent interference is carried out by type III effectors that are structurally similar to type I Cascade (Makarova et al., 2020). Although, unlike Cascade, type

III effectors possess RNase, DNase and adenyl cyclase activities (Elmore et al., 2016; Kazlauskiene et al., 2017). Type III effectors bind to RNA that is transcribed from the DNA target and cleave it every six nucleotides (Tamulaitis et al., 2014). DNase and cyclase are activated by the complementary binding of crRNA to the RNA transcript (Kazlauskienė et al., 2016). The DNase cleaves ssDNA in the transcription bubble. The cyclase synthesises cyclic oligonucleotides, which activate auxiliary Csm6 ribonuclease (Kazlauskiene et al., 2017; Niewoehner et al., 2017).

Type IV systems lack effector nucleases. Like other class 1 systems, type IV systems have an effector complex composed of multiple Cas proteins (Ozcan et al., 2019). The effector recognises nucleic acids, however, the mechanisms of target fate remain to be disclosed (Crowley et al., 2019).

Type VI CRISPR-Cas have Cas13 effector complex that targets RNA molecules (Abudayyeh et al., 2016), which has PFS (protospacer flanking sequences) near the target sequence (East-Seletsky et al., 2017; Smargon et al., 2017). Cas13 interaction with the target sequence unleashes Cas13 RNase activity, which degrades target RNA as well as non-specific RNAs (East-Seletsky et al., 2016).

1.3 Type I-F CRISPR-Cas

Type I CRISPR-Cas systems are the most diverse and prevalent among microbial communities. These systems are divided into seven subtypes (I-A, I-B, I-C, I-D, I-E, I-F and I-U) (Makarova et al., 2018). Type I-F is one of the best functionally and structurally characterised CRISPR-Cas system. Type I-F CRISPR-Cas codes for six *cas* genes (*cas1*, *cas5f*, *cas6f*, *cas7f*, *cas8f*, *cas2/3*). Cas5f, Cas6f, Cas8f and six Cas7f proteins together with a crRNA compose a ribonucleoprotein complex, termed Cascade (CRISPR-associated complex for antiviral defence). Cascade adapts a seahorse-like shape, where the 'head' is formed by the Cas6f bound to the 3' end of crRNA, six Cas7f subunits assemble along the spacer sequence to form 'backbone' and the 'tail' is composed of 5' end crRNA handle sandwiched by Cas5 and Cas8 (Zhang & Sontheimer, 2014; Jackson & Wiedenheft, 2015; Chowdhury et al., 2017). In type I-F systems, interference nuclease/helicase Cas3 is fused with adaptation protein Cas2, which in other type I systems are encoded by two separate genes (Makarova et al., 2020). Thus, adaptation complex comprises of Cas1-Cas2/3 complex (Rollins et al., 2017), which is responsible for naïve adaptation and mediate primed adaptation (priming). Primed adaptation relies on spacer acquisition from previously encountered nucleic acids that are recognised (primed) by a surveillance complex (Richter et al., 2014; Vorontsova et al., 2015; Fagerlund et al., 2017). It has been shown that priming can enhance spacer acquisition efficiency five hundred times over naïve adaptation (Staals et al, 2016). Due to primed adaptation escape mutants can be effectively captured by CRISPR-Cas.

It was demonstrated that Cascade scans invading nucleic acids for PAM sequences during interference (Rollins et al., 2015; Xue et al., 2017). Upon PAM recognition and binding by Cascade, crRNA hybridises along the length of target DNA protospacer forming R-loop structure (Szczelkun et al., 2014; Rutkauskas et al., 2015). Target DNA binding results in significant conformational changes that expose Cas8 for Cas2/3 nuclease/helicase recruitment (Xiao et al., 2018). Cas2/3, which in type I-F systems is a fusion of Cas2 and Cas3, is a signature protein of Type I-F CRISPR-Cas. Upon target recognition, Cas2/3 helicase/nuclease is recruited and invading nucleic acid hydrolysis is initiated (Rollins et al., 2017). Cas2/3 uses ATP energy and degrades ssDNA in 3'-5'direction. Final degradation of target is performed by single Cas2/3 proteins or host cell DNases (Sinkunas et al., 2011; Xiao et al., 2018; Tuminauskaite et al., 2020). Summary of type I-F CRISPR-Cas system composition and mechanism of action is provided in figure 1.4.

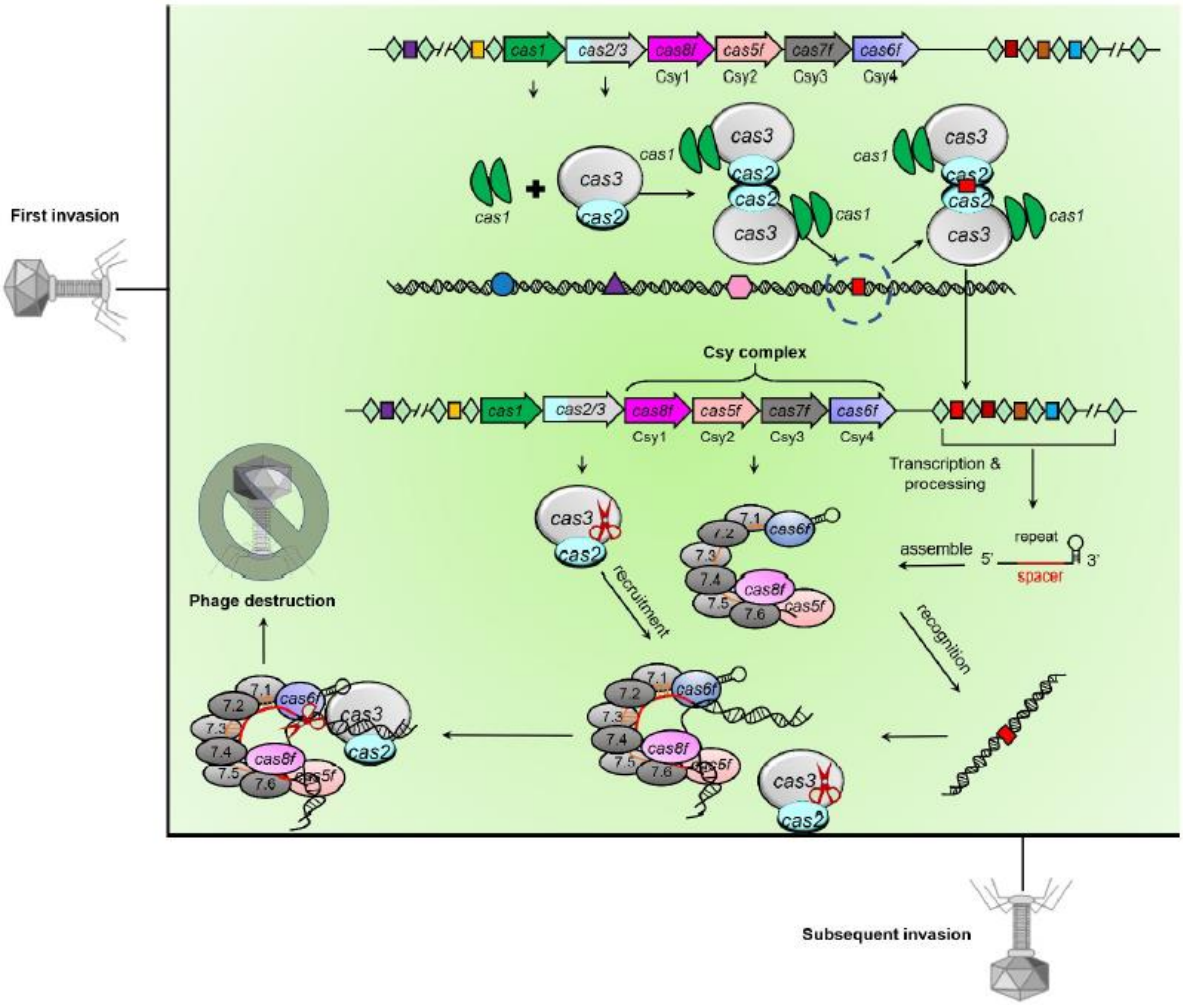

Figure 1.4. Structure and mechanism of action of type I-F CRISPR-Cas. Type I-F CRISPR-Cas system comprises of six Cas genes (*cas1*, *cas2/3*, *cas8f*, *cas5f*, *cas7f* and *cas6f*) and CRISPR array. During three stages of CRISPR-Cas immunity a new spacer is acquired from invading phage, Cascade complex is assembled and upon a subsequent phage invasion nucleic acid is recognised by the Cascade and cleaved by recruited Cas2/3 nuclease/helicase (Yang et al., 2021).

1.4 Inhibitors of CRISPR-Cas

CRISPR-Cas prokaryotic immunity allows bacteria and archaea to overcome phage attacks. However, phages can escape CRISPR-Cas targeting by a single nucleotide mutation. A point mutation in PAM or seed regions of a protospacer is sufficient to abolish CRISPR-Cas function even when there is a perfect crRNA complementarity to protospacer (Semenova et al, 2011). However, this strategy alone is insufficient for long term phage survival for multiple reasons. First, mutations have a fitness cost for phage survival. The fitness cost prevents the escapers to spread. Second, adaptive nature of CRISPR-Cas provides mutations to be ineffective via processes known as primed adaptation (see chapter 1.3). Primed adaptation allows bacteria to effectively counter nucleic acids that escape CRISPR-Cas interference (Datsenko et al., 2012; Fineran & Charpentier, 2012; Swarts et al., 2012; Sternberg et al., 2016; Jackson et al., 2017). Lastly, phages are distributed among diverse bacterial populations with vast diversity of acquired spacers. For phages to evade diverse CRISPR-Cas, it would require simultaneous mutations in all target sequences within phage genome (Li et al., 2020). Bacteriophages have evolved a broad spectrum of Anti-CRISPR proteins (Acrs) that are enable pages to overcome CRISPR-Cas immune pathway.

Acr proteins were first reported in a 2013 publication, which showed that several prophages encoded small anti-CRISPR proteins (50-150 residues), which inactivated the *Pseudomonas aeruginosa* type I-F CRISPR-Cas system (Bondy-Denomy et al., 2013). Subsequently, genes in the same locus were shown to inhibit type I-E CRISPR-Cas systems, another highly abundant CRISPR-Cas system in *P. aeruginosa* (Pawluk et al., 2014). Mature crRNA levels were not impacted by the AcrIF proteins (Bondy-Denomy et al., 2013), ruling out biogenesis inhibition. In vitro binding assays revealed that AcrIF1-4 each bind directly to different CRISPR-Cas components, preventing Cas proteins from binding or cleaving phage DNA (Bondy-Denomy et al., 2015). Since then anti-CRISPR arsenal has expanded significantly consisting of over 80 Acr families that inhibit type I, type II, type III, type V and type VI CRISPR-Cas (Bondy-Denomy et al., 2018; Marino et al., 2018). More recently discovered anti-CRISPR proteins as well usually tightly interact with Cas proteins.

1.4.1 Type I-F Anti-CRISPR (AcrIF) proteins

Three Anti-CRISPR proteins that inhibit Type I-F CRISPR-Cas from *Pseudomonas aeruginosa* were first in vitro characterised Acrs. AcrIF1 and AcrIF2 sterically block Cascade's interaction with the DNA target while AcrIF3 prevents DNA degradation by interacting with Cas3 nuclease and thus preventing its recruitment to the surveillance complex (Bondy-Denomy et al., 2015). Inhibition of target DNA binding and Cas3 nuclease recruitment remain most widely spread mechanisms of action

among Acrs to this day. Currently 24 AcrIF families (AcrIF1-AcrIF24) have been reported to inhibit type I-F CRISPR-Cas (Bondy-Denomy et al., 2018). Only a fraction of AcrIFs have been extensively studied. Recent studies reported that several AcrIFs employ unique strategies to supress CRISPR-Cas immunity, for example, AcrIF11 has an enzymatic activity (Niu et al., 2020) while AcrIF24 induces dimerization of the Cascade complex (Yang et al., 2022).

Most AcrIFs inhibit crRNA hybridisation with target DNA. AcrIF1 was one of the first discovered Acrs resulting in extensive amount of biochemical and structural studies to unravel the mechanism of action for this protein. Biochemical data showed that AcrIF1 interacts with the Cascade complex and prevents its binding to the target DNA. Structural and mutagenesis data revealed amino acid residues responsible for interactions with the Cascade complex. Three Cryo-EM structures demonstrated that AcrIF1 binds to the Cas7f backbone of the Cascade complex (Chowdhury et al., 2017; Guo et al., 2017; Peng et al., 2017). This confirmed that AcrIF1 indeed prevents the recognition and binding of DNA target by the surveillance complex.

AcrIF2 like AcrIF1 prevents DNA binding by the Cascade complex (Bondy-Denomy et al, 2015). The solved structure shows that it binds to the junction between Cas7.6f and Cas8f subunits (Chowdhury et al., 2017; Hong et al., 2017; Peng et al., 2017). However, AcrIF2 binding site only partially overlaps with that of DNA target as shown by solved structured of Cascade-AcrIF2 and Cascade-DNA complexes. AcrIF2 binding causes Cas8f hook to swing in the opposite direction to that seen with DNA binding (Guo et al., 2017). Thus, AcrIF2 prevents DNA binding through interacting with DNA binding surfaces of Cascade and pushing away the Cas8f hook from Cas7.6f. AcrIF7 has a negatively charged surface that interacts with Cas8f. Similarly to AcrIF2, AcrIF7 binds to Cas8f-Cas5f ‘tail’. AcrIF7 targets interaction surface of the Cascade that is important for PAM binding (Gabel et al., 2020; Zhang et al., 2020). Upon AcrIF10 binding, Cas8f swings towards Cas7.6f which also happens upon DNA binding. AcrIF10 interacts with residues on Cas7.6f, Cas8f, Cas5f surfaces that are involved in DNA binding (Guo et al., 2017).

AcrIF8 binding surfaces differ from those of AcrIF2/6/7/10. Nevertheless, AcrIF8 binds surfaces of Cas8f, Cas7.6f and Cas5f and prevents DNA binding. AcrIF8 also bind crRNA nucleotides that hybridise with DNA target. Thus, AcrIF8 acts as a DNA mimic by interacting with the Cas proteins of the Cascade and crRNA in this way preventing Cascade-crRNA complex hybridisation with DNA target (Zhang et al, 2020). AcrIF2, AcrIF6, AcrIF7, AcrIF8 and AcrIF10 are acidic proteins and have been shown to act as DNA mimics. They bind Cascade surfaces that are involved in interactions with DNA target thus preventing formation of Cascade-crRNA complexes and their base-pairing to DNA target.

AcrIF11 has been shown to be a potent CRISPR-Cas inhibitor (Pinilla-Redondo et al., 2020). Differently from the mechanisms described, AcrIF11 does not stably interact with the Cascade

proteins. AcrIF11 has been shown to have ADP-ribosylation activity. AcrIF11 ADP-ribosylates 'wedge' in Cas8f subunit, surface that serves in PAM recognition (Niu et al., 2020). Thus, AcrIF11 prevents DNA target binding through distinct mechanism of action.

Another less common mechanism of action for AcrIF proteins is DNA target cleavage inhibition. Initially this mechanism of action was described in AcrIF3. AcrIF3 homodimer binds to Cas2/3 and prevents its binding to the Cascade and DNA target (Bondy-Denomy et al., 2015; Wang et al., 2016). AcrIF3 keeps Cas2/3 in ADP bound state resulting in inactive nuclease/helicase. AcrIF3 is structurally similar to a domain of Cas8f subunit. Cas8f role in Cas3 recruitment to the surveillance complex suggests that AcrIF3 acts as a Cas8f mimic. AcrIF4 has been shown to inhibit CRISPR-Cas interference stage by interaction with Cas8f subunit of the surveillance complex. However, the interaction surface is distinct from that other AcrIFs bind to abolish PAM recognition and thus AcrIF4 does not compete with PAM recognition or crRNA hybridisation to dsDNA target. AcrIF4-Cascade complex is able to bind DNA (Gabel et al., 2021). AcrIF4 binds the Cascade and prevents conformational changes of Cas8f subunit that are required for Cas2/3 recruitment (Gabel et al., 2021; Rollins et al., 2019). Altogether, further studies are needed to elucidate mechanism of action for this protein. The mechanisms of action for AcrIF6, AcrIF9 and AcrIF14 are discussed in more detail in chapter 3.7. Mechanisms of action for AcrIFs reviewed in this work are summarised in figure 1.5.

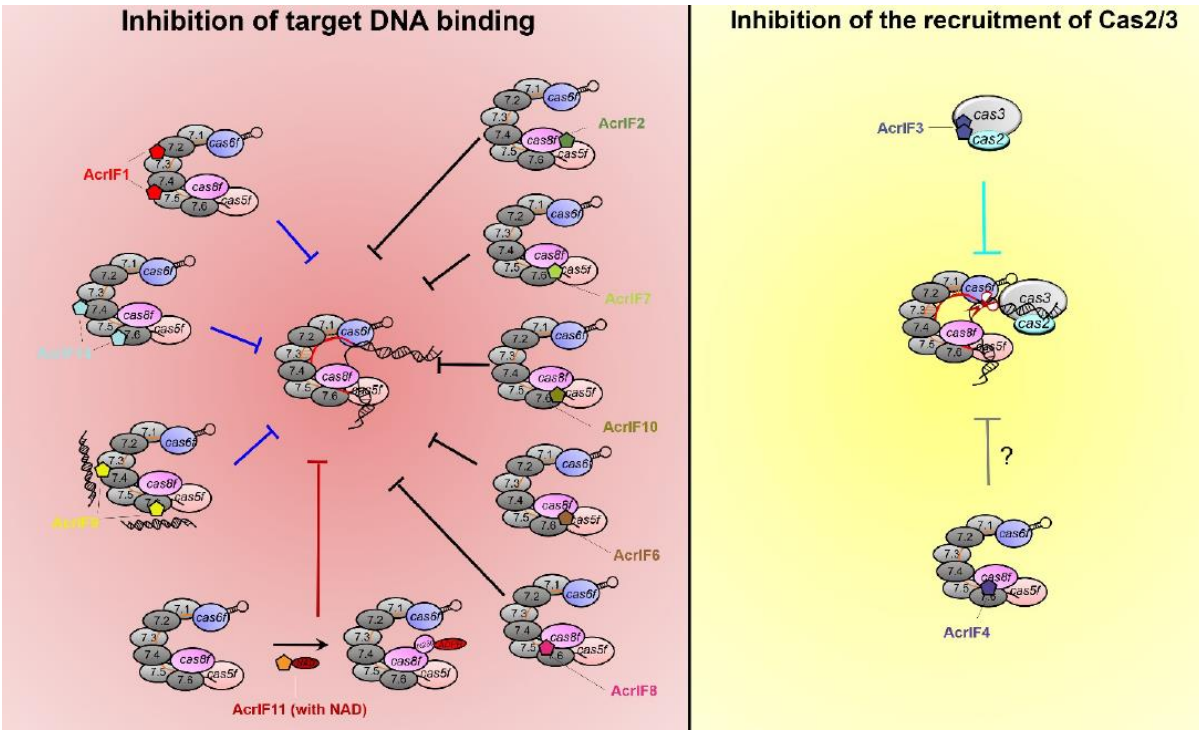

Figure 1.5. A summary of mechanisms of action of AcrIF proteins. AcrIFs work by two main mechanisms of action, they either inhibit DNA target binding or its cleavage (Yang et al., 2021).

2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Reagents

Amresco:	Tris (tris(hydroxymethyl)aminomethane), SDS (sodium dodecyl sulfate), APS (ammonium persulfate), 2-mercaptoethanol.
BIO-RAD:	Ethidium bromide, TEMED (tetramethylethylenediamine).
Hartman analytic:	γ-^{32}P ATP.
Roth:	Chloramphenicol, Kanamycin, Streptomycin, PMSF (Phenylmethylsulfonyl fluoride), acrylamide/N,N'-methylenebis(acrylamide) (29:1) 40% (w/V), L(+)-arabinose, HCl, NaCl.
Sigma-Aldrich:	Orange G dye, NaOH, LiCl, imidazole.
Thermo Fisher Scietific:	ATP, TopVision agarose, IPTG (Isopropyl β-D-1-thiogalactopyranoside), dNTP, Bacto™ tryptone, Bacto™ Yeast Extract, DTT (dithiothreitol), glycerol.
FORMEDIUM:	Agar.

2.1.2 Enzymes, other proteins and kits

Thermo Fisher Scientific products used in enzymatic reactions:

- Bovine Serum Albumin (BSA);
- DreamTaq™ DNA polymerase and Phusion™ HF DNA polymerase;
- FastAP Thermosensitive Alkaline Phosphatase;
- FastDigest™ restriction enzymes NcoI, BamHI, XhoI;
- Rapid DNA Ligation Kit;
- T4 Polynucleotide Kinase.

Buffer solutions for enzymatic reactions (Thermo Fisher Scientific):

- 10× FastDigest™ Green buffer (used in restriction enzyme reactions);
- DreamTaq™ Green Buffer (10×);
- Phusion™ HF Buffer (5×).

Nuleic acid and protein ladders (Thermo Fisher Scientific):

- O‘geneRuler DNA Ladder Mix ready to use;
- PageRuler^{TM} Plus Prestained Protein Ladder. Sizes (kDa): 250, 130, 100, 70, 55, 35, 25, 15, 10.

Kits (Thermo Fisher Scientific):

- GeneJet Plasmid miniprep Kit;
- PCR Purification Kit;
- GeneJET Gel Extraction Kit;
- Qubit™ dsDNA HS Assay Kit.

Buffers used for protein – nucleic acid interactions assays:

- EMSA: TAE buffer, 10% glycerol, 0.1 mg/ml BSA, 200 mM NaCl;
- Nuclease activity assays: 20 mM HEPES pH 7.3, 0.1 mg/ml BSA, 2 mM ATP, 200 mM NaCl, 1.5 mM $CoCl_2$.

Other solutions

- Nucleic acid dye SybrGold (Invitrogen)

2.1.3 Buffers and other solutions

Protein crystallisation kits:

- Commercial Hampton Research kits: Crystal Screen, Index, PEG/Ion, SaltRx, PEGRx;
- Produced in Institute of Biotechnology: Morpheus (Gorrec, 2009).

AcrIF9 protein crystallisation buffer solution:

- 0.1 M Tris-HCl (pH 8.5), 3.0 M NaCl.

E. coli competent cell preparation solutions:

- „Na^+“ solution: 5 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0 (25 °C);
- „Ca^{2+}“ solution: 5 mM Tris-HCl, 100 mM $CaCl_2$, 5 mM $MgCl_2$, pH 8.0 (25 °C).

Other solutions:

- 4× protein dye: 200 mM Tris-HCl, 400 mM DTT. 8 % SDS (w/V), 0.4 % bromophenol blue (w/V), 40 % (V/V) glycerol, pH 6.8 (25 °C);
- Acrylamide gel dye: PageBlue^{TM} Protein Staining Solution (Thermo Scientific);
- 3× stop solution (75 mM EDTA, 0.025% OrangeG, 0.5% SDS (sodium dodecyl sulphate), 50% (V/V) glycerol, pH 8.0 (25°C)

DNA electrophoresis buffers:

- NBE buffer: 0.1 M H_3BO_3, 0.2 mM EDTA, 0.5 μg/ml ethidium bromide, pH 8.3 (25 °C);
- TAE buffer: 40 mM Tris-acetate, 1 mM EDTA, pH 8.3 (25 °C);
- TBE buffer: 89 mM Tris-HCl, 89 mM H_3BO_3, 2 mM EDTA, pH 8.3 (25 °C).

2.1.4 Bacterial growth media

- LB broth: 1.0 % (w/V) tryptone, 0.5 % (w/V) yeast extract, 0.5 % (w/V) NaCl, pH 7.0 (25°C);
- Agar LB broth: 1.5 % (w/V) agar, 1.0 % (w/V) tryptone, 0.5 % (w/V) yeast extract, 0.5 % (w/V) NaCl, pH 7.0 (25°C).

2.1.5 Protein purification solutions

Ni^{2+} affinity chromatography buffer solutions:

- Cell sonication buffer: 20 mM Tris-HCl (pH = 8.0 (25°C)), 500 mM NaCl, 5mM 2-mercaptoethanol, 25 mM imidazole, 0.2 mM PMSF.
- A buffer: 20 mM Tris-HCl (pH = 8 (25°C)), 500 mM NaCl, 5mM 2-mercaptoethanol.
- B buffer: 20 mM Tris-HCl (pH = 8 (25°C)), 500 mM NaCl, 5 mM 2-mercaptoethanol, 500 mM imidazole.

Size exclusion chromatography buffer solution:

- 20 mM TrisHCl (pH = 8.0 (25°C)), 250 mM NaCl, 5 mM 2-mercaptoethanol.

Dialyis buffer solution:

- 20 mM Tris-HCl (pH = 8.0 (25°C)), 300 mM NaCl, 2mM DTT, 50% (V/V) glycerol.

2.1.6 Protein electrophoresis solutions

- 4 % stacking polyacrylamide gel: 4 % acrylamide/N,N' – methylenebis(acrylamide) (ratio 37,5:1), 125 mM Tris-HCl, 0,1 % SDS, pH 6,8 (25°C);
- 15 % resolving polyacrylamide gel: 15 % acrylamide/N,N'– methylenebis(acrylamide) (ratio 37,5:1), 375 mM Tris-HCl, 0.1 % (w/V) SDS, pH 8.8 (25°C);
- Electrophoresis buffer solution: 25 mM Tris-HCl, 190 mM glycine, 0.1 % (w/V) SDS, pH 8.3 (25°C).

2.1.7 Bacteria strains

- *E. coli* BL21-AI (DE3): F- ompT hsdSB (rB- mB-) gal dcm araB::T7RNAP-tetA.
- *E. coli* DH5α: F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ–.

2.1.8 AcrIF sequences

Table 2.1. *AcrIF* gene sequences used in this work.

Anti-CRISPR	DNA sequence
Os-AcrIF6	atgactatctatctcagcaacgctatcgaaaacgccacttccattgaacaagtcgttgaactgatcaatgaaggaacttgcgagggaatggaaggcattgagttcactgccgatatacttgctggtcaatacgcctggagtgccgctaaagatttgttgattcggcaggcatgacagctgatgacctagaagcccagcttgagttcctaagcgccgccggcgcagtattcagcgagagcgtcgcagtagatcatggcttgaaactggccgccgccgacaccgaataa
Aa1-AcrIF9	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttgaaagcgcagactcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcttccaagggaccacgttaaagattggaacgatctactcactgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttac
AcrIF9-Aa1 mutant	DNA sequence
Aa1-AcrIF9-K24A/R37A/K72A	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttggcggcgcagactcttgcggatgccaaacgcgaagcgtcggcgcgtcagtgcttccaaggtaccacgttaaagattggaacgatctactcactgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaagcgtgggtcgatcgttac

Table 2.1. AcrIF gene sequences used in this work [continued].

Aa1-AcrIF9-K24E	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttggaagcgcaga ctcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcttccaaggtaccacgttaaagattggaacgatctactc actgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttac
Aa1-AcrIF9-R37E	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttgaaagcgcagac tcttgcggatgccaaacgcgaagcgtcggaacgtcagtgcttccaaggtaccacgttaaagattggaacgatctactc actgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttac
Aa1-AcrIF9-F41A	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttgaaagcgcagac tcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcgcccaaggtaccacgttaaagattggaacgatctactc actgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttac
AcrIF9-Aa1-N12A/N13A/N15A/Y17A	atgacgaacgtagtttattactttacggaaactgcggcgatcgcggctgcggcgactgcggaggctttgaaagcgca gactcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcttccaaggtaccacgttaaagattggaacgatcta ctcactgaacagtgacgggcttctggtagacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttac
AcrIF9-Aa1-Δ(L54-V60)	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttgaaagcgcagac tcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcttccaagggaccacgttaaagattggaacgatctactc agacgagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttactaa
AcrIF9-Aa1-I51-D61->VNGL	atgacgaacgtagtttattactttacggaaactaataacatcaacgcttatgcgactgcggaggctttgaaagcgcagac tcttgcggatgccaaacgcgaagcgtcgcgtcgtcagtgcttccaagggaccacgttaaagattggaacggtaaacg gactggagatcaccagtaaagaagatgggaaaaaatgggtcgatcgttactaa

2.1.9 Oligonucleotides

Table 2.2. Oligonucleotide sequences used in this work.

Primer	Sequence 5'→3'	Purpose
Cloning		
TS918	GAAGTGCCATTCCGCCTGACC	Amplification of *acrIF* genes from synthetic fragment; forward primer
TS919	CACTGAGCCTCCACCTAGCCT	Amplification of *acrIF* genes from synthetic fragment; reverse primer
EMSA		
Lgur	GCGAGGAAGCGGAAGAGCGCCC	SP1 or NS 420 bp substrate obtained by PCR-amplification from pSP or pNS plasmids as templates, respectively.
pUC57-For	GCCAGGGTTTTCCCAGTCACGA	
Cas2/3 cleavage		
GG222	CTTGAGATCCTTTTTTTCTGC	SP2 1444 bp substrate obtained by PCR-amplification from pSP plasmid as a template.
TS759	GAGCAGACAAGCCCGTCAG	
Competitive DNA binding		
TS931*	GCTCGGTACCCGACCACCCTTTTTGATATCC	SP* 272 bp substrate obtained by PCR-amplification from pSP plasmid as a template.
TS930	GCGGGCAGTGAGCGCAACG	

2.1.10 Plasmids

Table 2.3. Plasmids used in this work.

Plasmid	Description	Cloning sites	Primers	Source
Cloning				
pCDF-HS [Str]	T7 promoter based expression vector. AcrIF cloning.	N/A	N/A	Tuminauskaite et al., 2020
Expression				
pCd [Cm]	Cascade cassette (*cas8f-cas5f-cas7f-cas6f*). Cas-AcrIF complex purification for crystallisation.	N/A	N/A	Tuminauskaite et al., 2020
pCd-H [Str]	Cascade cassette (*cas8f-cas5f-cas7f* and *cas6f* with N-terminal His_6 sequence). Cascade complex purification.	N/A	N/A	Tuminauskaite et al., 2020
pCas2/3-H [Amp]	*cas2/3* gene in pET-SH; C-terminus of Cas2/3 is fused with His_6 tag. Cas2/3 purification.	N/A	N/A	Tuminauskaite et al., 2020
pCR [Cm]	CRISPR region; spacer targeting sequence in pSP plasmid. In was used for expression of Cascade complex, which was used for EMSA and cleavage assay.	N/A	N/A	Tuminauskaite et al., 2020
pCDF-Os-AcrIF6 [Str]	*acrIF6* from *O. smirnovii*	NcoI/XhoI	TS918/TS919	This study
pCDF-Aa1-AcrIF9 [Str]	*acrIF9* from *A. actinomycetemcomitans serotype* e str. SC1083	NcoI/XhoI	TS918/TS919	This study
pCDF-His-Os-AcrIF6 [Str]	*AcrIF6-Osm* with N-terminal His_6 sequence	BamHI/XhoI	TS918/TS919	This study
pCDF-His-Aa1-AcrIF9 [Str]	*Aa1-AcrIF9* with N-terminal His_6 sequence	BamHI/XhoI	TS918/TS919	This study
pCDF-His-Aa1-AcrIF9-K24A/R37A/K72A [Str]	*AcrIF9-Aa1-K24A/R37A/K72A* with N-terminal His_6 sequence; putative DNA interaction surface [M1]	BamHI/XhoI	TS918/TS919	This study
pCDF-His-Aa1-AcrIF9-K24E [Str]	*AcrIF9-Aa1-K24E* with N-terminal His_6 sequence; putative DNA interaction surface [M2]	BamHI/XhoI	TS918/TS919	This study

Table 2.3. Plasmids used in this work [continued].

pCDF-His-Aa1-AcrIF9-R37E [Str]	*AcrIF9-Aa1- R37E* with N-terminal His_6 sequence; putative DNA interaction surface [M3]	BamHI/XhoI	TS918/TS919	This study
pCDF-His-Aa1-AcrIF9-N12A/N13A/N15A/Y17A [Str]	*AcrIF9-Aa1-N12A/N13A/N15A/Y17A* with N-terminal His_6 sequence; putative Cascade interaction surface [M4]	BamHI/XhoI	TS918/TS919	This study
pCDF-His-Aa1-AcrIF9-F41A [Str]	*AcrIF9-Aa1- F41A* with N-terminal His_6 sequence; putative Cascade interaction surface [M5]	BamHI/XhoI	TS918/TS919	This study
pCDF-His-AcrIF9-Aa1-Δ(L54-V60) [Str]	*AcrIF9-Aa1-Δ(L54-V60)* with N-terminal His_6 sequence; loop deletion [M6]	BamHI/XhoI	TS918/TS919	This study
pCDF-His-AcrIF9-Aa1-I51-D61->VNGL [Str]	*AcrIF9-Aa1-I51-D61->VNGL* with N-terminal His6 sequence; loop replacement with sequence of Vp-AcrIF9 loop [M7]	BamHI/XhoI	TS918/TS919	This study
Substrate production				
pSP [Amp]	CC PAM and target protospacer sequences in pUC19. Template for production of SP1, SP2 and SP* DNA substrates.	N/A	N/A	Sinkunas et al., 2013
pNS [Amp]	AA PAM and non-target protospacer sequences in pUC19. Template for production of NS DNA substrate.	N/A	N/A	Sinkunas et al., 2013

2.2 Methods

2.2.1 Cloning, expression and purification of proteins

Cascade complex and Cas2/3 protein were obtained as described previously (Tuminauskaite et al., 2020). All plasmids used in this study are provided in table 2.3.

Synthetic DNA fragments of *acrIFs* were obtained from Twist Bioscience (sequences of cloned genes are provided in table 2.2.) The fragments were inserted into either NcoI/XhoI or BamHI/XhoI sites of pCDF-HS expression vector obtaining AcrIF proteins either without tag or fused to N-terminal His_6-tag, respectively (table 2.1.)

E. coli BL21-AI strain containing AcrIF and Cascade cassette vectors (for preparation of samples for protein crystallisation) or AcrIF vector was grown in LB medium supplemented with respective antibiotics until it reached an optical density of ~0.5 (OD600nm). Then protein expression was induced for 4 h at 37°C by addition of 1 mM IPTG (isopropylthiogalactoside) and 0.2% (w/V) arabinose. Cells were collected by centrifugation and resuspended in buffer containing 20 mM Tris-HCl, 500 mM NaCl, 5 mM imidazole, 5 mM 2-ME (2-mercaptoethanol) and 1 mM PMSF (phenylmethylsulphonyl fluoride). Cells were lysed by sonication and cell debris was removed by centrifugation. The resulting supernatant was loaded on Ni^{2+}-charged HiTrap column (GE Healthcare) and eluted with a linear gradient of increasing imidazole. The fractions containing AcrIF or AcrIF-Cas complex were pooled and loaded on Superdex 200 (HiLoad 16/600; GE Healthcare) column for separation by gel filtration. The fractions containing AcrIF or Cas-AcrIF complex were dialysed into 20 mM Tris-HCl (pH 8), 300 mM NaCl, 50% (V/V) glycerol, 2mM DTT (dithiothreitol) and stored at -20°C. The concentrations of AcrIFs were measured by UV (280 nm) absorbance.

2.2.2 Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)

Gel comprising of 4% stacking and 15% resolving layers was prepared. Protein samples were mixed with 4× protein dye and incubated at 100 °C for 5 minutes allowing denaturation of the proteins. The samples then were loaded in wells formed in the staking layer of the gel. Electrophoresis was conducted at 15V/cm voltage. After electrophoresis the gel was stained with Page Blue dye according to the manufacturer recommendations.

2.2.3 Electrophoretic mobility shift assay (EMSA)

Two component mixing approaches were used to conduct EMSA: (i) 30 nM or 100 nM of Cascade complex was pre-bound with 20 nM DNA substrate (SP1 or NS; table 2.2.) for 10 minutes forming the R-loop then AcrIF6 or AcrIF9 was added and incubated for additional 20 minutes; (ii) 30 nM or 100 nM of Cascade complex was pre-incubated with AcrIF6, AcrIF9 or respective AcrIF9 mutant (30 nM, 300 nM, 3000 nM, 20000 nM) for 10 minutes then mixed with 20 nM DNA substrate (SP1 or NS) and incubated for additional 20 minutes. The incubations were performed at RT in 1× TAE (Invitrogen) buffer supplemented with 200 mM NaCl, 10% (V/V) glycerol and 0.1 mg/mL BSA (bovine serum albumin). The samples were analysed on 1% (w/V) agarose gel and visualised by SYBR-Gold (Invitrogen) staining.

2.2.4 Cas2/3 cleavage assay

Prior Cas2/3 cleavage initiation, the reaction components were mixed by two approaches: (i) the Cascade was mixed with DNA target forming the R-loop then AcrIF was introduced; (ii) AcrIF was pre-incubated with Cascade before adding the DNA target. The reactions conducted at 37°C in a buffer containing 20 mM HEPES (pH 7.2 at at 25°C), 120 mM NaCl, 10% glycerol, 0.1 mg/mL BSA, 2 mM CoCl2, 2 mM ATP, 5 nM SP2 substrate (table 2.1.), 20 nM Cascade, 200 nM Cas2/3 and AcrIF6, AcrIF9 or respective AcrIF9 mutant (5, 50, 500, and 5000 nM). After incubation for 1 h, the reactions were terminated by adding 3× stop solution (75 mM EDTA, 0.025% OrangeG, 0.5% SDS (sodium dodecyl sulphate), 50% (V/V) glycerol (pH 8 at 25°C)) followed by heating at 75°C for 10 minutes. Nucleic acids were separated from proteins by phenol-chloroform extraction. The reaction products were analysed on 0.8% (w/V) agarose gel and visualised by ethidium bromide staining.

2.2.5 Preparation of radioactively labelled DNA substrate

SP* substrate (Table 2.2.) was used to investigate AcrIF9-Cascade interaction in competitive DNA binding experiment. Oligonucleotide TS931 was radioactively labelled at 5’ end using γ-^{32}P ATP and T4 PNK. The reaction was conducted according to manufacturer recommendations. The labelled oligonucleotide together with TS930 was used for PCR to obtain SP* substrate. Reaction product was purified using GeneJet PCR Purification Kit.

2.2.6 Competitive DNA binding

Two component mixing approaches were used to assay competitive DNA binding: (i) the Cascade was mixed with DNA target forming the R-loop then AcrIF9 together with competing DNA was introduced; (ii) AcrIF9 was pre-incubated with Cascade before adding the DNA target mixed with competing DNA. The binding reactions were conducted at 37°C in binding buffer (1× TAE buffer (Invitrogen), 200 mM NaCl, 10% (V/V) glycerol and 0.1 mg/mL BSA) supplemented with 100 nM Cascade, 20 µM AcrIF9, 20 nM SP* target DNA (table 2.2.) and 0.05 nM, 0.5 nM or 5 nM of competetive λ DNA (~48 kbp). The samples were analysed in 1% (w/V) agarose gel and visualised by autoradiography.

2.2.7 SEC pull-down assay

Cascade complex was incubated with respective AcrIFs (AcrIF6 or AcrIF9) at the molar ratio of 1:100 in 1× TAE buffer at 37°C for 30 minutes. The samples were fractionated by Superdex 200 (HiLoad 16/600; GE Healthcare) column. Proteins in the collected fractions were precipitated using TCA (trichloroacetic acid) then analysed by SDS-PAGE and visualised by Coomassie blue staining.

2.2.8 Protein stability assay

The stabilities of the Aa1-AcrIF9 and its mutants were assayed by nanoscale differential scanning fluorimetry (Prometheus, Nanotemper). The capillaries were loaded with protein samples of 0.5 mg/ml concentration in 20 mM Tris-HCl, 300 mM NaCl and 25 % (V/V) glycerol. The denaturation curves were recorded increasing temperature from 15 to 95ºC at 1 ºC/min rate. The onset and inflection point temperatures were determined from three separate measurements.

2.2.9 Protein crystallization

The crystals of AcrIFs were obtained by sitting drop vapor diffusion method at 19 °C by mixing AcrIF with reservoir solutions containing various buffers from protein crystallisation kits (see chapter 2.1.3: ‘Protein crystallisation kits’).

2.2.10 Protein structure determination

AcrIF9 structure was determined by dr. Giedrė Tamulaitienė. For data collection the crystals were shortly dipped into 2.9 M Na-malonate (pH 7.0) and flash-freezed. The X-ray diffraction dataset was collected at the EMBL/DESY Petra III P13 beamline (Germany) at 100 K. XDS (Kabsch, 2010), SCALA and TRUNCATE (CCP4, 1994) were used for data processing.. Homology model of Aa1-AcrIF9 prepared by Swiss-Model server (https://swissmodel.expasy.org/) (Waterhouse et al., 2018) using Pa-AcrIF9 (PDB ID 6VQV chain A) as a template was used for molecular replacement in MOLREP (Vagin and Teplyakov, 2010). Manual model rebuilding of the models was performed in COOT (Emsley and Cowtan, 2004) and the structure was refined with phenix.refine-1.12-2829 (Afonine et al., 2012). All molecular scale representations were prepared using Pymol (Schrodinger, LLC (2015), The PyMOL Molecular Graphics System, Version 1.8.).

3. RESULTS

3.1 Object of the research

CRISPR-Cas provides an adaptive prokaryotic immunity against invading phages and plasmids. Type I-F CRISPR-Cas surveillance complex (Cascade) recognises and binds DNA target in a presence of PAM sequence, while Cas2/3 helicase/nuclease cleaves the target. Anti-CRISPRs (Acrs) are bacteriophage encoded typically small proteins that inhibit CRISPR-Cas and enable phage survival and replication. In this work AcrIF6 and AcrIF9 interplay with type I-F CRISPR-Cas from *Aggregatibacter actinomycetemcomitans* D7S-1 (Aa-CRISPR-Cas) was studied (Figure 3.1).

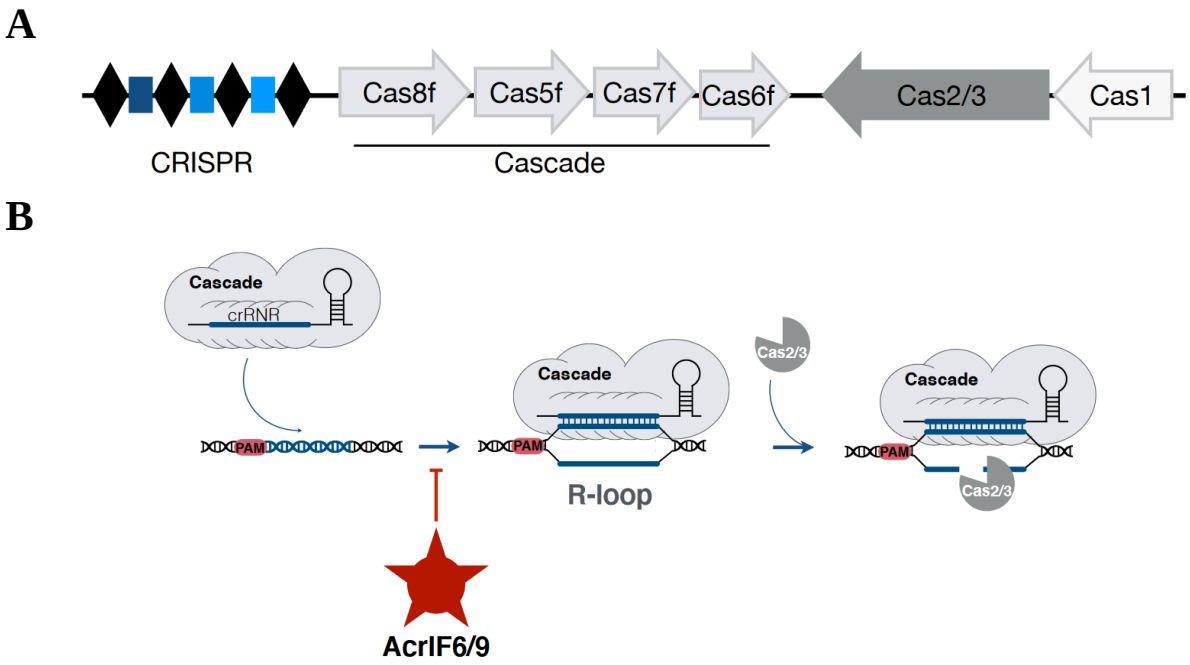

Figure 3.1. Type I-F CRISPR-Cas from *Aggregatibacter actinomycetemcomitans* (Aa-CRISPR-Cas) and its Acr inhibitors (AcrIFs) investigated in this study. (A) Aa-CRISPR-Cas locus is comprised of CRISPR region, *cas8f*, *cas5f*, *cas7f* and *cas6f* genes that encode Cascade complex as well as *cas2/3* and *cas1* that encode interference and adaptation stage proteins. (B) AcrIF6 and AcrIF9 act in interference stage of CRISPR-Cas immunity and hinder DNA target binding.

We previously tested ten AcrIF families (AcrIF1-10) and showed that one AcrIF6 and four AcrIF9 homologues inhibit type I-F Aa-CRISPR-Cas in vivo (Kupcinskaite, 2020). In this work, detailed molecular inhibition mechanisms of Os-AcrIF6 (AcrIF6 from *Oceanimonas smirnovii*) and Aa1-AcrIF9 (AcrIF9 from *Aggregatibacter actinomycetemcomitans* serotype e str. SC1083) were assayed. For simplicity, these proteins are referred to as AcrIF6 and AcrIF9 further in the text.

3.2 In vitro characterisation of AcrIF6 and AcrIF9

Majority of Anti-CRISPR proteins supress CRISPR-Cas immunity by direct binding to CRISPR-Cas effector complex. AcrIFs either bind to the Cascade complex and hinder its interactions with DNA target or interact with Cas2/3 helicase/nuclease and prevent hydrolysis of the target. To assess detailed inhibitory action of AcrIF6 and AcrIF9 on Aa-CRISPR-Cas, in vitro experiments including EMSA, nuclease activity and size exclusion chromatography pull down assays were conducted.

To evaluate the effect of the AcrIFs on DNA binding by Cascade dual strategy was employed: AcrIF6 or AcrIF9 was mixed with Cascade-DNA complex (preformed R-loop), or free DNA target before the addition of Cascade. Analogous mixing order was also selected for a non-target DNA (DNA with absent protospacer lacking complementarity to crRNA). Additionally, AcrIFs were mixed with DNA to check whether DNA binding by AcrIFs occurs (Figure 3.2. A).

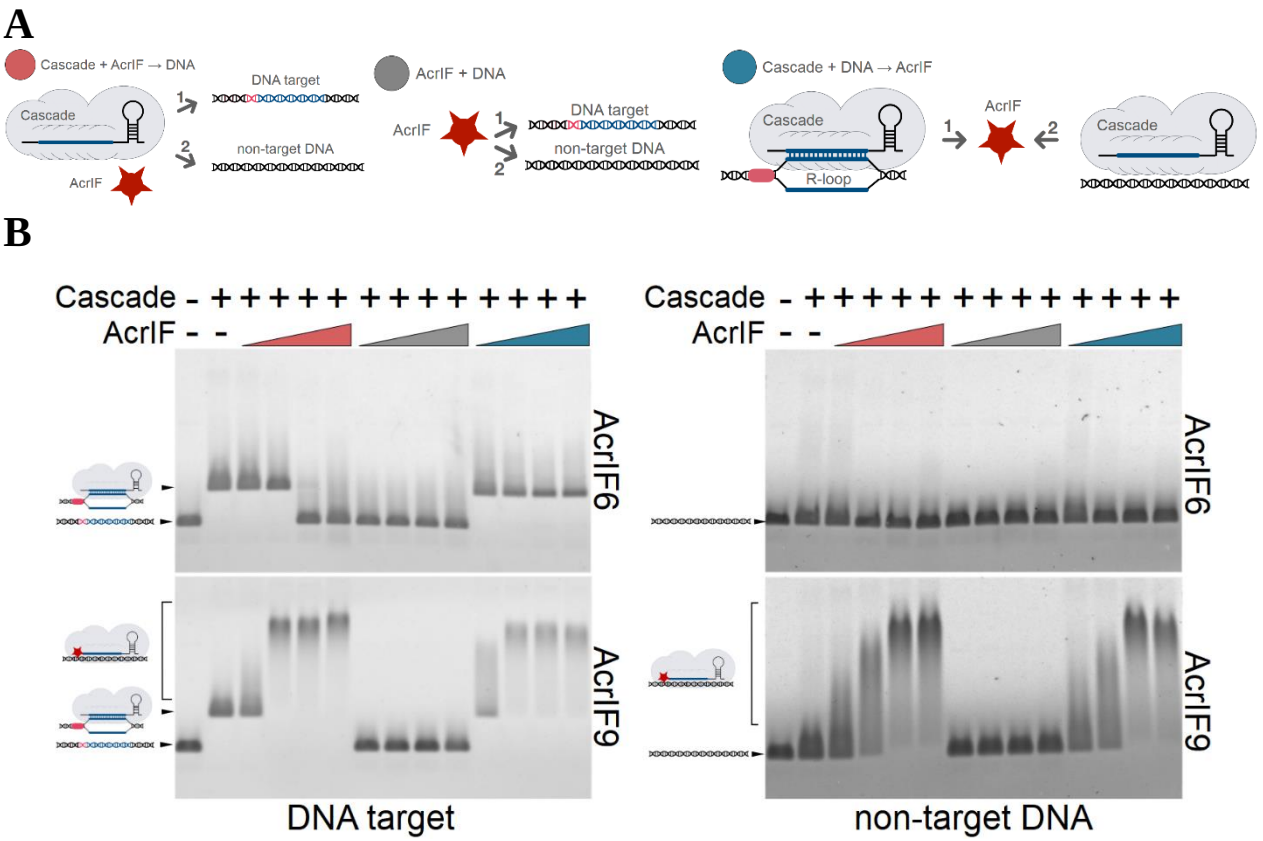

Figure 3.2. Influence of AcrIF6 and AcrIF9 on Cascade binding to target and non-target DNA. Scheme and results of the EMSA experiment. (A) Different mixing orders of reaction components were employed to determine whether AcrIF6/9 could (i) inhibit R-loop formation (red), (ii) directly interact with DNA (grey) or (iii) disassemble R-loop (blue). AcrIF6/9 protein was introduced either before (red) or after (blue) DNA addition. The Cascade binds target DNA sequence forming R-loop structure, while non-target DNA remains unbound. The proteins were mixed with either target DNA (1) or non-targer DNA (2). (B) EMSA results using target-DNA (left pane) and non-target DNA (right pane) mixed with Cascade and AcrIF6/9 protein. AcrIF6/9 (30, 300, 3000, 20000 nM) and/or 100 nM Cascade were introduced to reaction mixture as indicated in (A).

Cascade preincubation with AcrIF6 led to unbound DNA visible in agarose gel. This indicates that AcrIF6 hinders DNA binding by Cascade and R-loop formation. However, R-loop remains intact when AcrIF6 is introduced to reaction mixture with preformed R-loop. Interestingly, Cascade-DNA interaction is altered both when AcrIF9 is added before and after R-loop formation. Addition of AcrIF9 by both mixing orders resulted in smeared DNA bands with decreased mobility. Furthermore, smeared DNA bands were observed when identical experiment was conducted with non-specific DNA. Meanwhile, no DNA band smearing was present when AcrIF6 was added to the reaction mixtures (Figure 3.2. B).

Next, the AcrIF9 mediated Cascade binding to the target DNA was tested in the context of the competitive DNA (Figure 3.3). In this experiment, two mixing approaches were used: (i) Cascade and AcrIF9 protein mixture was added to radioactively labelled target DNA and unlabelled non-target DNA mixture, (ii) preincubated radioactively labelled target DNA and Cascade mixture was added to reaction solution containing unlabelled non-target DNA (Figure 3.3 A). The non-target DNA displaced the radioactively labelled target DNA in the presence of AcrIF9, which is seen from reduced DNA band smearing on the agarose gel when non-target DNA concentration is increased (Figure 3.3. B). Overall, these data suggest that AcrIF9-Cascade complex forms non-specific interactions with DNA.

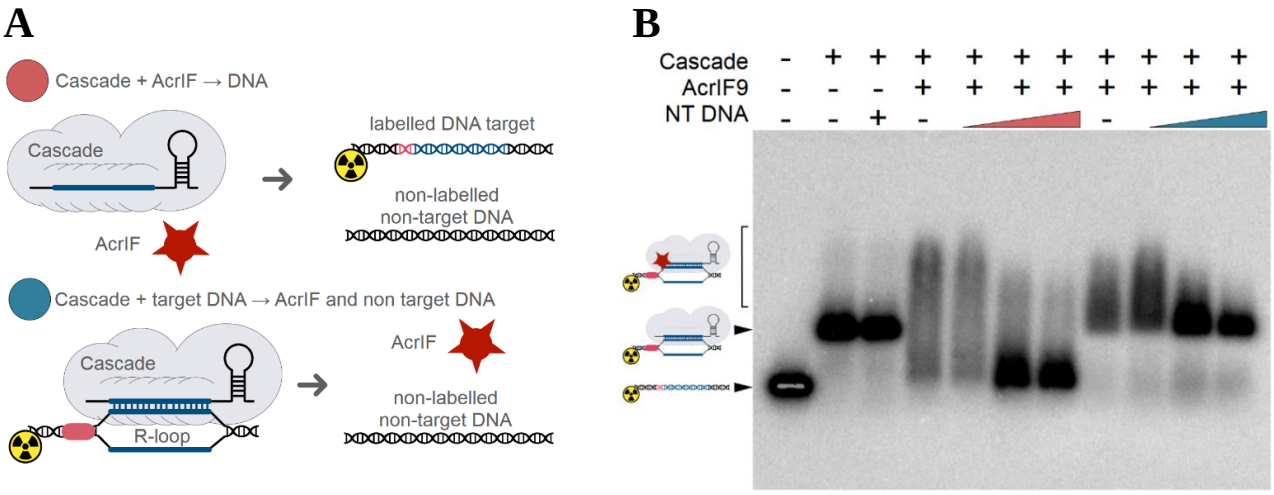

Figure 3.3. Competitive DNA binding experiment. The AcrIF9-mediated Cascade binding to target DNA in the presence of non-specific DNA. (A) Mixing order of reaction components. Two approaches were used for competitive (non-target) DNA addition: (i) AcrIF9 protein was incubated with Cascade then added to a mixture of the labelled target DNA and non-labelled competitive DNA containing no target sequence (red); (ii) the preformed R-loop was introduced to the mixture of AcrIF9 and non-labelled competitive DNA (blue). (B) EMSA results of binding reactions. Increasing concentrations of competitive ~48 kbp length λ phage genomic DNA (0.05, 0.5, and 5 nM) were introduced to binding reactions of 20 μM AcrIF9, 100 nM Cascade and 20 nM labelled target DNA (~0.3 kbp length) as indicated in (A).

The R-loop structure is necessary for Cas2/3 mediated DNA target cleavage. Cas2/3 nuclease activity was assessed in the presence of AcrIF6 and AcrIF9. The cleavage assays showed that AcrIF6 or AcrIF9 prevent DNA hydrolysis. However, neither AcrIF6 nor AcrIF9 influence Cascade-DNA complex integrity when R-loop is pre-formed, thus cleavage occurs. In agreement with EMSA, Cascade and DNA interactions remain in tact when AcrIFs are added to a mixture where R-loop is present (Figure 3.4.).

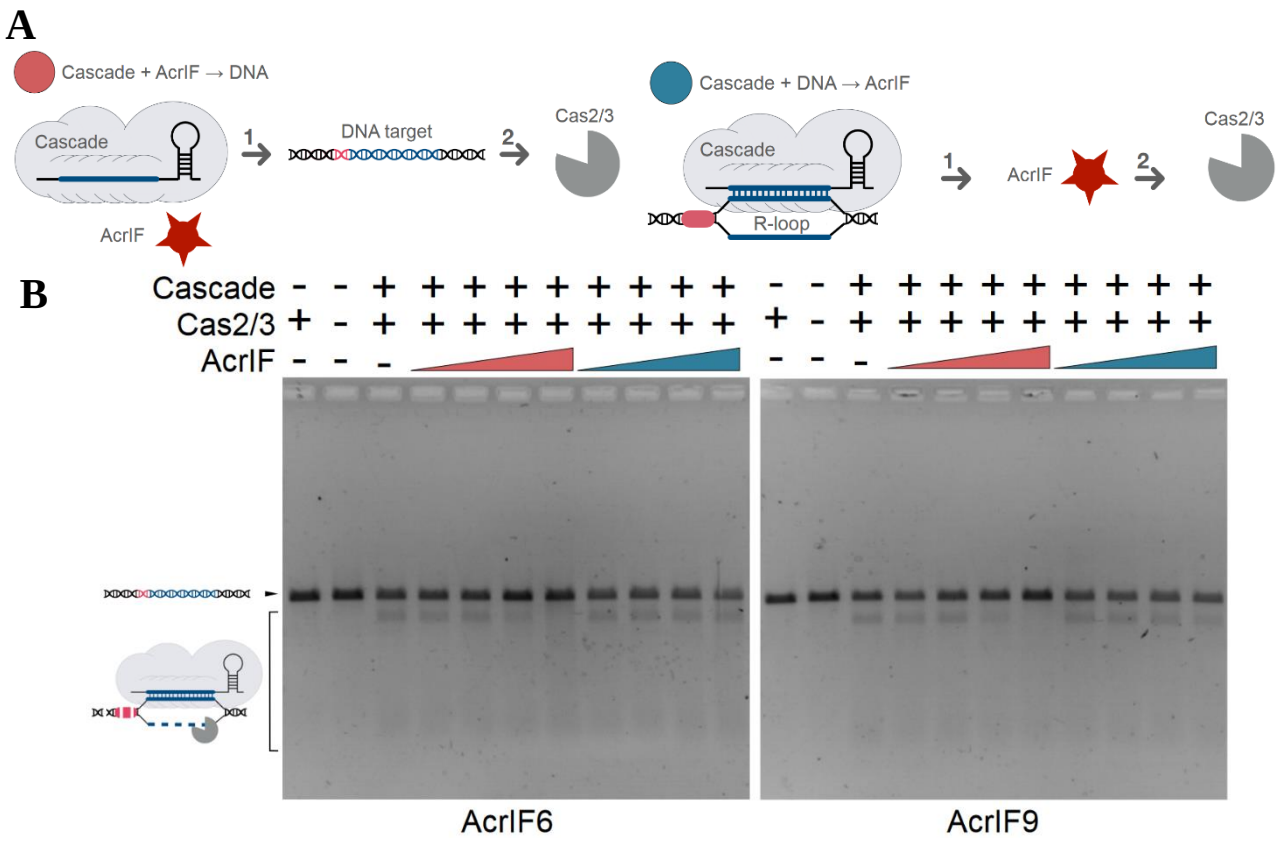

Figure 3.4. AcrIF6 and AcrIF9 influence on nuclease activity of CRISPR-Cas. (A) Mixing order of reaction components. AcrIF6/9 protein was introduced either before (red) or after (blue) the R-loop assembly (1). Then cleavage was initiated by Cas2/3 addition (2). (B) The Cas2/3-mediated cleavage. Increasing AcrIF concentrations (5, 50, 500, 5000 nM) were introduced to the reaction mixtures as pictured in A.

Therefore, AcrIF6 and AcrIF9 both inhibit specific Cascade binding to DNA target thus preventing target cleavage. In addition, AcrIF9 promotes Cascade binding to non-specific DNA.

3.3 AcrIF6 and AcrIF9 directly interact with the Cascade complex

To confirm AcrIF6 and AcrIF9 binding to the Cascade, size exclusion chromatography (SEC) assay was conducted. Purified samples of Cascade and AcrIF6/9 protein were preincubated allowing proteins to bind. The samples then were loaded on a Superdex 200 SEC column to separate unbound

AcrIF6/9 (Figure 3.5. A and C). Elution peaks were analysed by SDS-PAGE (Figure 3.5. A, B, C and D).

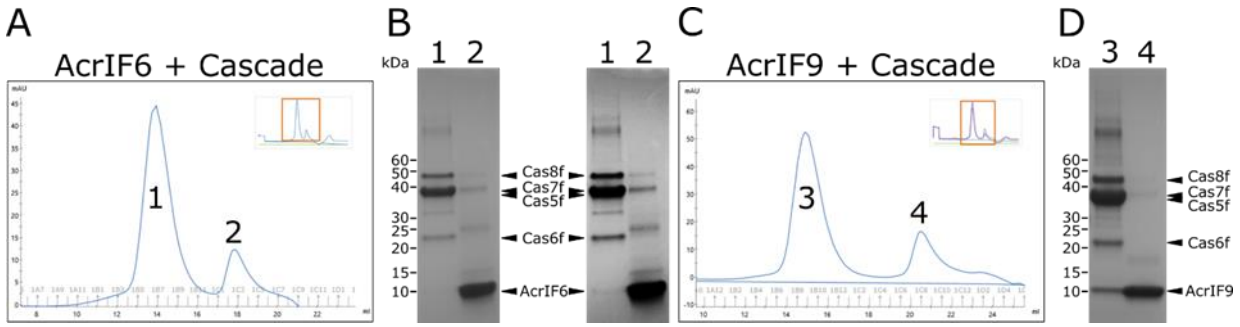

Figure 3.5. SEC pull down assay of AcrIF6, AcrIF9 and Cascade. (A) After incubation**,** Cascade and AcrIF6 elution peaks (1 and 2 respectively) were analysed by SDS-PAGE. (B) The brightness and contrast of the gel picture on the right was adjusted to sharpen the poorly stained band of AcrIF6, which eluted together with Cascade. (C) Elution peaks 3 and 4 in chromatogram of AcrIF9 and Cascade incubation were analysed by SDS-PAGE (D) SDS-PAGE gel of AcrIF9 and Cascade. Expected molecular weights for Cas8f, Cas5f, Cas7f, Cas6f, AcrIF6 and AcrIF9 proteins were approximately 52, 36, 38, 23, 11.6, and 10.3 kDa, respectively.

AcrIF6 and AcrIF9 eluted together with the Cascade complex indicating their direct interactions. This is in agreement with EMSA experiments and confirms that AcrIF6 and AcrIF9 indeed bind to the Cascade complex. Further, structure-based studies were employed.

3.4 Crystallisation of AcrIF6 and AcrIF9

Sitting drop vapour diffusion method was used for protein crystallisation screening. We hoped to obtain crystals of AcrIF6 and AcrIF9 in apo form and bound to Cas proteins. For this, AcrIF6 and AcrIF9 were co-expressed with Cas proteins of the Cascade and the complexes were purified.

Number of screened conditions for each AcrIF protein or Cas-AcrIF6/9 complex is provided in Table 3.1.

Table 3.1 The numbers of screened crystallisation conditions.

Proteins	Number of conditions screened
AcrIF6 + Cas	576
AcrIF6	576
AcrIF9	1152
AcrIF9 + Cas	576
Total	2880

Initially 576 crystal trials in Crystal Screen, Index, Morpheus, SaltRx, PEGRx, PEGIon kits were tested for each AcrIF protein or AcrIF-Cas complexes. Some protein crystallization hits were

identified for AcrIF9 (Figure 3.6. A). Unfortunately, no crystals of AcrIF6 in apo form, AcrIF6 or AcrIF9 bound to Cas proteins were obtained during this screening (Figure 3.6. B, C and D). Protein precipitates, thin needle shaped crystals and spherulites that are not suitable for X-ray diffraction analysis were obtained.

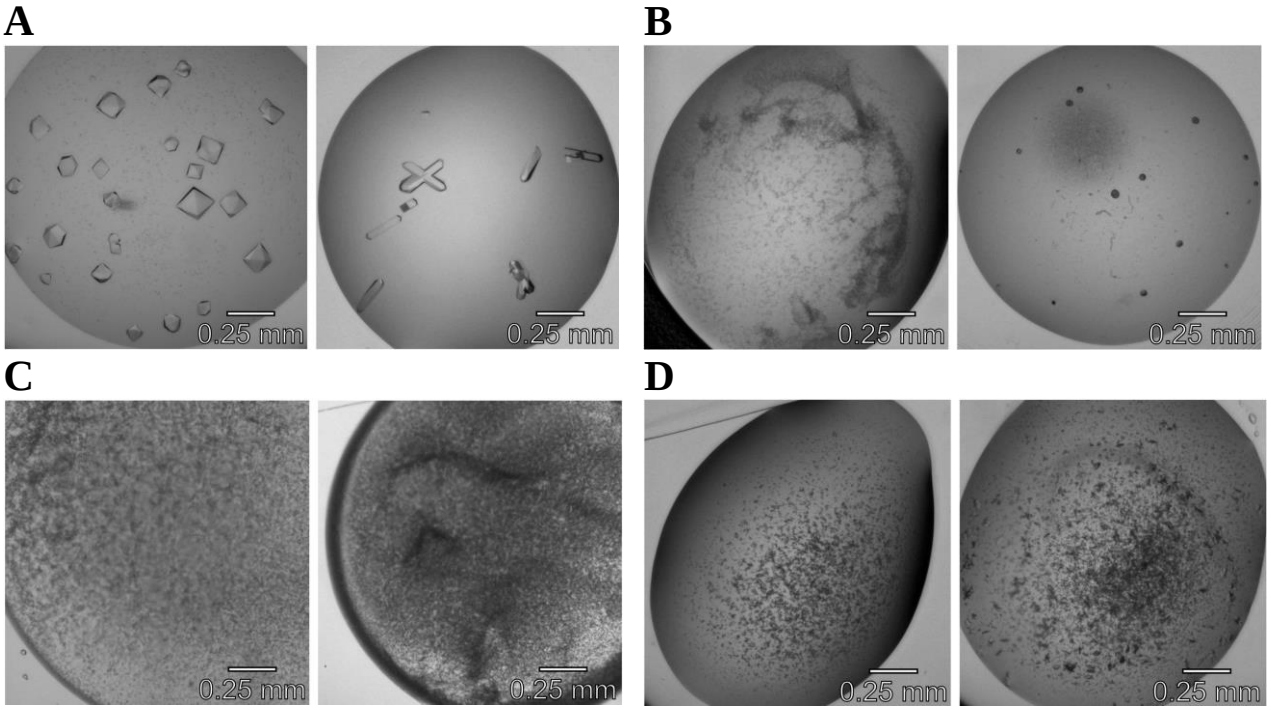

Figure 3.6. A fraction of screened protein crystallisation results. (A) AcrIF9 crystals. (B) AcrIF9-Cas crystallisation resulted in precipitate and spherulites. (C) AcrIF6 protein precipitate. (D) AcrIF6-Cas overnucleated crystals and precipitate.

Protein crystal formation of AcrIF9 was favoured at high salt concentrations (for example, 3 M NaCl). Further optimisation for AcrIF9 included changing buffer solution and protein solution ratio in the droplet as well as replicating crystallisation conditions that already retrieved good form crystals to obtain reproducible results. In the end, AcrIF9 crystals that diffracted X-ray were collected and shipped to EMBL/DESY Petra synchrotron (Germany) for X-ray diffraction analysis. The tertiary structure of AcrIF9 protein (fom *A. actinomycetemcomitans* prophage) was solved by dr. Giedrė Tamulaitienė. Thus, for further analysis only AcrIF9 was selected.

3.5 AcrIF9 mutant selection

Aa-AcrIF9 structure was solved at 2.3 Å resolution (PDB ID: 7BB5). The overall structure of Aa-AcrIF9 protein is similar to other AcrIF9 homologues, which were solved in a complex with the Cascade; however, Aa-AcrIF9 possesses a longer loop (L3-4), which could serve a functionally important role. The cryo-EM structures of *Pseudomonas aeruginosa* and *Proteus panneri* AcrIF9 (Pa-AcrIF9 and Pp-AcrIF9) bound to the Pa-Cascade complex indicate that two AcrIF9 molecules

bind to the thumbs of Cas7.4f and Cas7.6f subunits and the two AcrIF9 binding sites are similar (Hirschi et al., 2020; Zhang et al., 2020). The Pp-AcrIF9 also interacts with additional DNA molecules (Hirschi et al., 2020). The structures were superimposed and the Aa1-AcrIF9 surfaces potentially involved in the interactions with Cas7f subunits and DNA were identified (Figure 3.7.)

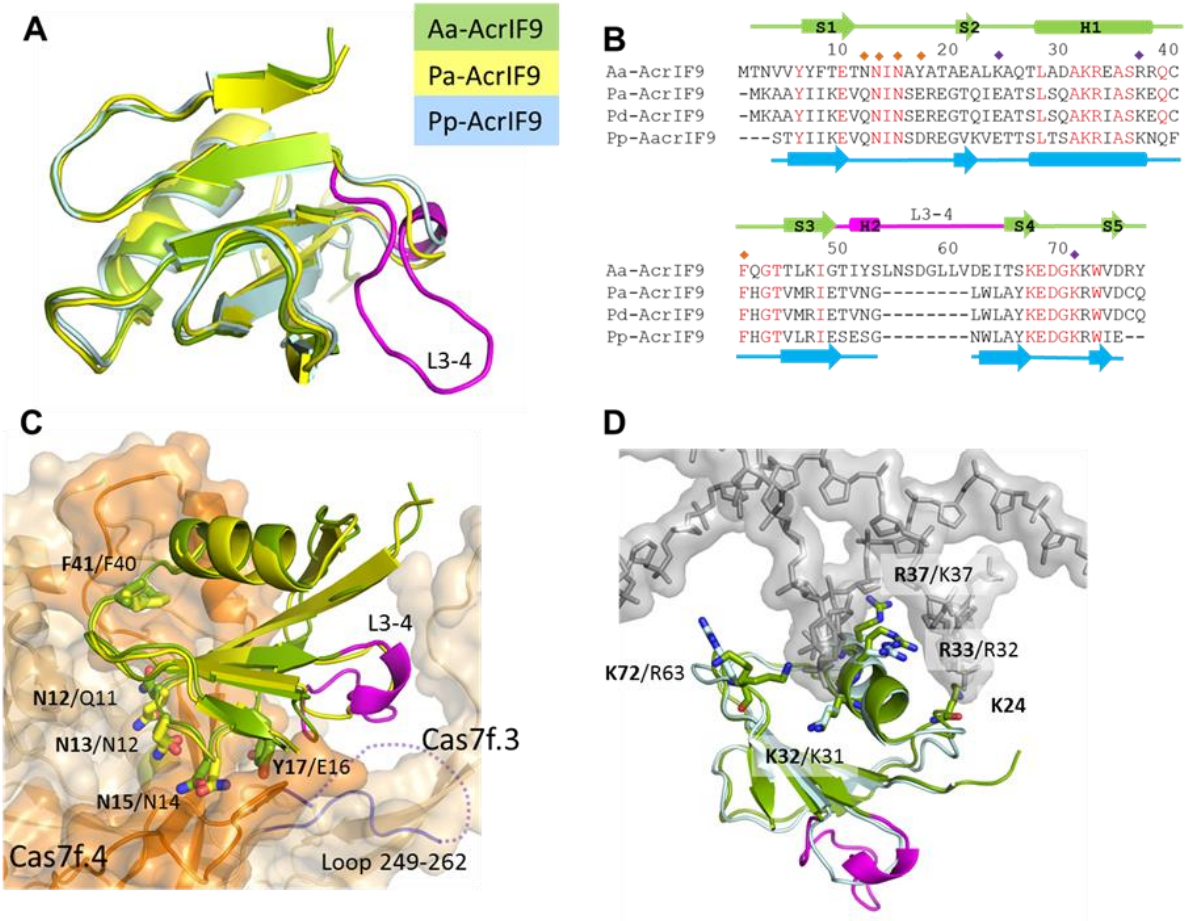

Figure 3.7. Aa1-AcrIF9 structural superimpositions. (A) Superposition of Aa1-AcrIF9 (green) with Pa-AcrIF9 (PDB ID 6VQV, chain A, yellow) and Pp-AcrIF9 (PDB ID 6W1X, chain J, light blue). (B) Structure based sequence alignment of AcrIF9 homologs. Identical residues are red. Aa1-AcrIF9 residues presumably involved in contacts with Cas7f and DNA are marked by orange and purple diamonds, respectively. (C) Model of Aa1-AcrIF9 binding to the Pa-Cascade complex. Aa1-AcrIF9 (green) was superimposed with the chain A of Pa-AcrIF9-Cascade complex (PDB ID 6VQV). Cas7.4f and Cas7.3f subunits are coloured orange and light orange, respectively. Residues of Aa1-AcrIF9 interacting with Cas7f.4 are shown in stick representation and labelled. Aa1-AcrIF9 residues are labelled in bold. Loop 249-262 of Cas7.3f presumably interacting with L3-4 is coloured blue. (D) Model of DNA binding by Aa1-AcrIF9. Aa1-AcrIF (green) was superimposed with the chain J of Pp-AcrIF9-Cascade-DNA complex (PDB ID 6WHI). Residues of Aa1-AcrIF9 interacting with DNA are shown in stick representation and labelled. Aa1-AcrIF9 residues are labelled in bold.

To test the possible interaction surfaces, seven vectors encoding AcrIF9 mutants were constructed. Mutations K24A/R37A/K72A (mutant termed M1), K24E (M2), and R37E (M3) that may be responsible for interaction with DNA were introduced. Additionally, mutations N12A/N13A/N15A/Y17A (M4) and F41A (M5) were made to evaluate their effect on AcrIF9-

Cascade interactions. Lastly, the L3-4 loop that may be involved in the interactions with the Cas7f subunits was deleted ΔL54-V60 (M6) or replaced I51-D61→VNGL (M7; the residues from the Pa-AcrIF9).

3.6 In vitro characterisation of AcrIF9 mutants

After identifying possible AcrIF9 interaction surfaces and constructing seven AcrIF9 mutant sequences, *acrIF9* mutants were expressed and the proteins were purified (Figure 3.8.)

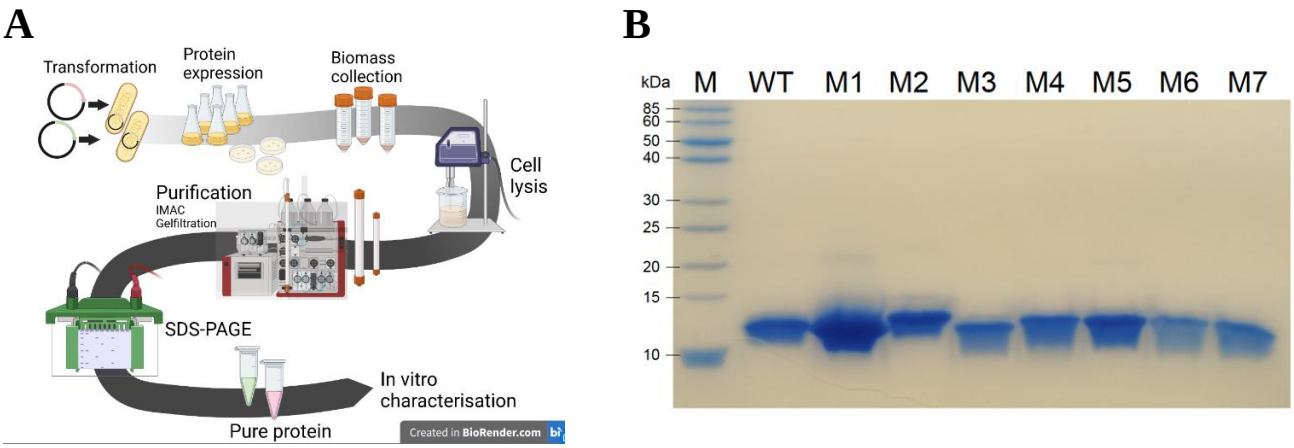

Figure 3. 8. Expression and purification of AcrIF9 mutants. (A) Flow of *acrIF9* mutant expression and protein purification. (B) SDS-PAGE assay of the purified proteins. Wild type AcrIF9 and seven AcrIF9 mutant proteins were purified: M1 (K24A/R37A/K72A), M2 (K24E), M3 (R37E), M4 (N12A/N13A/N15A/Y17A), M5 (F41A), M6 (ΔL54-V60), M7 (I51-D61->VNGL).

After successfully purifying the AcrIF9 mutants, in vitro assays were performed in order to determine significance of the mutations for AcrIF9 interactions with DNA and Cascade (Figure 3.9.) The M1, M2 and M3 mutations resulted in obstruction of AcrIF9 dependent non-specific DNA binding, consequently smeared DNA bands are absent in lanes containing the three mutants (Figure 3.9. B).

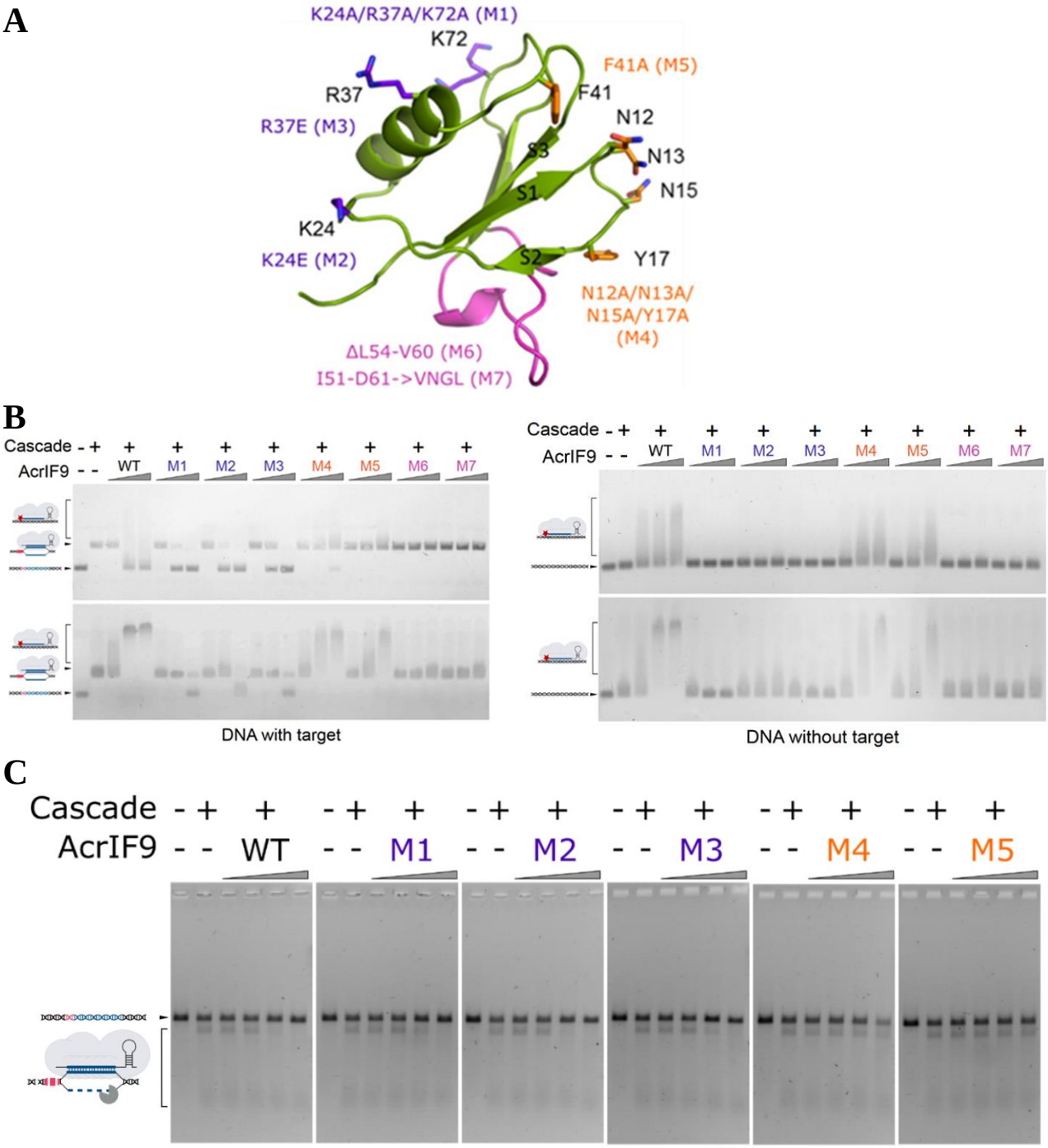

Figure 3.9. Overall structure of AcrIF9 with selected mutations indicated and in vitro assays of AcrIF9 mutants. (A) Cascade and DNA interaction surfaces. Residues of Aa1-AcrIF9 potentially interacting with Cascade and DNA are shown in stick representation and coloured orange and purple, respectively. The loop L3-4 is coloured in magenta. Mutants of the corresponding residues are indicated. (B) Cascade binding to the DNA target (left pane) and non-specific DNA (right pane) in the presence of the AcrIF9 mutants. The increasing concentrations of the AcrIF9 mutant proteins (50, 500, 5000 nM) incubated with either 20 nM (top) or 100 nM (bottom) Cascade then were introduced to the binding buffer containing 20 nM DNA. Protein interactions with the DNA were assayed by EMSA in agarose gel. (C) Cas2/3-mediated cleavage inhibition by AcrIF9 mutants. Increasing concentrations (5, 50, 500, 5000 nM) of respective AcrIF9 mutant were pre-incubated with 20 nM Cascade then mixed with 5 nM target DNA (SP3). The DNA degradation was initiated by addition of 200 nM Cas2/3. Positions of substrate and degradation products are indicated on the left of agarose gel.

Although non-specific binding was interrupted, the M1, M2, M3 mutants remained bound to the Cascade complex leading to hindrance of Cas2/3 mediated DNA target cleavage, hence the uncleaved substrate in the agarose gel (Figure 3.9 C).

In the presence of M4 and M5 mutations, Cascade remained bound to the target DNA and R-loop formation was not disturbed (Figure 3.9. B). Consequently, Cas2/3 recruitment was permitted resulting in cleavage (Figure 3.9. C). Nevertheless, M4 and M5 mutants showed reduced non-specific DNA binding as seen from reduced DNA band smearing (Figure 3.9. B). Regardless, DNA-Cascade interaction was not fully disrupted by M4 and M5 mutants. The lack of effect of Cascade's binding capacity in presence of M4 and M5 could be explained by the longer loop in Aa1-AcrIF9 structure that may form additional interactions with Cas7f subunits and thus stabilise AcrIF9-Cascade complex. M6 and M7 AcrIF9 mutants contain alterations in the L3-4 loop, which result in hindering of Cascade-AcrIF9 interactions and thus DNA cleavage (Figure 3.9.) In addition, the stabilities of WT and mutant AcrIF9s were tested (Figure 3.10.)

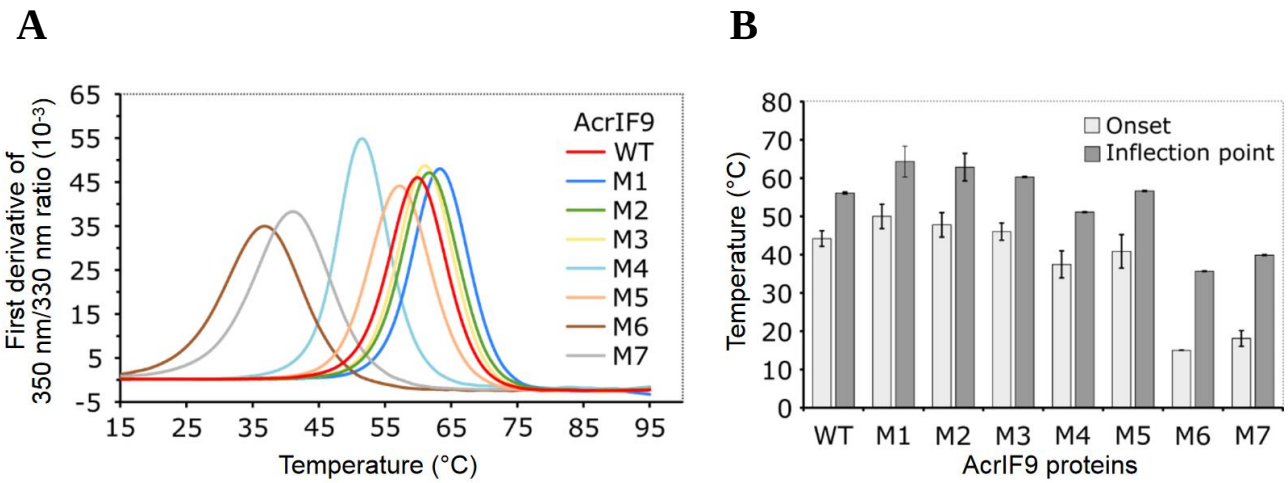

Figure 3.10. The stability of the Aa1-AcrIF9 proteins. The samples of Aa1-AcrIF9 (WT) and its K24A/R37A/K72A (M1), K24E (M2), R37E (M3), N12A/N13A/N15A/Y17A (M4), F41A (M5), ΔL54-V60 (M6), I51-D61->VNGL (M7) mutants were assayed by nanoscale differential scanning fluorimetry. (A) Denaturation curves are provided as the first derivative of 350 nm/330 nm ratio. (B) Temperatures of the denaturation curve onset and inflection points. Error bars represent standard deviations of average in at least three separate measurements.

M6 and M7 mutants showed notably lower protein stability than WT AcrIF9 or other mutants (M1 – M5). Therefore, mutations in M6 and M7 show that L3-4 loop structure is important for AcrIF9 structural integrity and may play a role in Cascade-AcrIF9 complex assembly.

3.7 Discussion

Previously we tested AcrIF1-10 inhibitory effect on Aa-CRISPR-Cas in vivo and showed that two AcrIF families (AcrIF6 and AcrIF9) supress activity of Aa-CRISPR-Cas (Kupcinskaite, 2020). The aim of this work was to unravel molecular mechanisms of Aa-CRSIPR-Cas inhibition by AcrIF6 and AcrIF9. A spectrum of in vitro experiments as well as structural and structure-guided mutagenesis studies were conducted.

This work demonstrates that AcrIF6 and AcrIF9 directly bind to Aa-Cascade complex and supress CRISPR-Cas immunity. These AcrIF proteins hinder DNA target binding by Cascade, consequently this prevents DNA cleavage by Cas2/3 helicase/nuclease. However, AcrIF6/9 cannot disassemble R-loop and once it is formed, AcrIF6 and AcrIF9 are inactive against Aa-CRISPR-Cas. Zhang and colleagues obtained structures of AcrIF6 and AcrIF9 bound to *Pseudomonas aeruginosa* Cascade (Pa-Cascade) complex, which revealed AcrIF6 and AcrIF9 binding sites (Zhang et al., 2020). AcrIF9 has been shown to interact with Cas7.4f and Cas7.6f subunits. AcrIF9 binding to the Cascade blocks PAM recognition and R-loop formation. Meanwhile, AcrF6 binds at the junction between Cas7.6f and Cas8f hindering PAM recognition. AcrIF6 forms interactions with Cas8f residues that have been shown to be important for target DNA duplex splitting via a process termed ‘wedging’ (Rollins et al., 2019). This suggests that AcrIF6 interferes with ‘wedging’ and thus acts as a DNA mimic. Overall, sterical blocking of hybridization between the complementary DNA strand and the crRNA is a general inhibition approach for type I-F Acrs (AcrIF1/2/6/7/8/9/10) (Li et al., 2020; Yang et al., 2021).

In addition to inhibition of DNA target recognition, AcrIF9 promotes non-specific interactions with DNA, which was studied in detail during this work. AcrIF9 has a positively charged surface that could mediate these interactions. This guided a mutational study of AcrIF9. Three AcrIF9 mutants (K24A/R37A/K72A, K24E, and R37E) were constructed to verify the putative surfaces for interaction with DNA. Here, it was confirmed that the positively charged residues promote interactions with DNA since replacing them resulted in a decreased AcrIF9-mediated binding to non-specific DNA. The AcrIF9 surfaces interacting with Cascade were as well identified. The N12A/N13A/N15A/Y17A and F41A mutations alleviated Cascade binding to the target indicating reduced inhibitory activity of AcrIF9. Overall, the solved structure of Aa1-AcrIF9 is similar to other available homologous structures (Hirschi et al., 2020; Kim et al., 2020; Zhang et al., 2020), however, it has a unique loop which serves a role in structural stability of Aa1-AcrIF9 and may be important for additional interactions with Cascade.

The question of function for non-specific DNA binding by AcrIF9-Cascade complex remains. This work demonstrates that AcrIF9 binding to Cascade is sufficient to abolish CRISPR-Cas

functionality. It is unlikely that the AcrIF9 surface interacting with DNA is an evolutionary remnant, which does not play a role in CRISPR-Cas immunity inhibition. AcrIF9 homologues retained this surface. The AcrIF14 family, has also been shown to mediate non-specific DNA binding via Cascade complex. AcrIF14 binds to Cas7f subunits inhibiting Cascade-DNA interactions while its N-terminal domain together with Cas8f PAM recognition loop is essential for non-specific DNA binding (Liu et al., 2021). Some of class II CRISPR-Cas inhibitors also contain DNA binding domains. DNA binding does not directly inhibit CRISPR-Cas but is a part of a bifunctional strategy employed by these Acrs.

AcrIIA1 has a two-domain architecture with its C-terminal domain (CTD) serving as an Acr protein and its N-terminal domain (NTD) as a transcriptional repressor. When Cas9 expression is high, AcrIIA1 binds and inhibits Cas9 and de-represses its own promoter (Osuna et al., 2020a; Osuna et al 2020b). Other HTH-Acr fusions with autorepression capabilities were also recently reported for AcrIIA13-AcrIIA15 found in several *Staphylococcus* genomes (Watters et al., 2018). Additionally, AcrIIA6 and AcrVA4 possess HTH domains with unknown functions (Hynes et al., 2018; Zhang et al., 2019). Therefore, it is possible that AcrIF9-Cascade complex could have some sequence preference. Therefore, the function of non-specific DNA binding mediated by AcrIF9 and potentially other Acrs remains to be disclosed.

To summarize, AcrIF6 and AcrIF9 proteins bind to the conserved surfaces of the Cascade and hinder its binding to the DNA target. This work also provides evidence for the AcrIF9-mediated non-specific DNA binding, which might facilitate CRISPR-Cas inhibition.

CONCLUSIONS

1. AcrIF6 and AcrIF9 inhibit CRISPR-Cas by hindering target DNA recognition. Additionally, AcrIF9 mediates non-specific DNA binding by Cascade.
2. AcrIF9 protein was successfully crystallised, and the protein structure was solved at 2.3 Å resolution.
3. AcrIF9 surfaces responsible for interaction with Cascade and non-specific DNA were identified.

PUBLICATIONS

Data in this thesis was presented at these conferences:

Kupcinskaite E., Tutkus M., Kopustas A., Asmontas S., Jankunec M., Zaremba M., Tamulaitiene G., and **Sinkunas T.** Disarming of CRISPR-Cas system by anti-CRISPR proteins AcrIF6 and AcrIF9. EMBO|EMBL Symposium: New Approaches and Concepts in Microbiology. 2021 07 7-9 (digital poster presentation).

Kupcinskaite E., Tamulaitiene G. and Sinkunas T. Inhibition of Cascade complex by the AcrIF6 and AcrIF9 proteins. CRISPR 2021, virtual conference. 2021 11 16-18 (digital poster presentation).

Kupcinskaite E, Tamulaitiene G, **Sinkunas T.** Inhibition of crRNA-guided effector complex by the AcrIF proteins. The EMBO | EMBL Symposium: The Complex Life of RNA, virtual conference. 2020 10 7-9 (digital poster presentation).

Manuscript with data presented in this thesis was submitted to Scientific Reports and currenly is under revision stage:

Kupcinskaite E, Tutkus M, Kopustas A, Asmontas S, Jankunec M, Zaremba M, Tamulaitiene G, Sinkunas T. Disarming of type I-F CRISPR-Cas surveillance complex by anti-CRISPR proteins AcrIF6 and AcrIF9.

VILNIUS UNIVERSITY
LIFE SCIENCES CENTRE
Eglė Kupčinskaitė
Master thesis

INVESTIGATION OF A TYPE I-F CRISPR-CAS INHIBITION BY ACRIF6 AND ACRIF9

SUMMARY

CRISPR-Cas is a diverse prokaryotic system that provides adaptive immunity against foreign nucleic acids. Bacteriophages have developed a vast variety of strategies to escape CRISPR-Cas protection in an evolutionary bacterial-phage arms race. One of the strategies is phage-encoded small proteins, named anti-CRISPRs (Acrs), that inhibit CRISPR-Cas protection and enable phage evasion. In this work, type I-F CRISPR-Cas from *Aggregatibacter actinomycetecomitans* D7S-1 bacteria (Aa-CRIRSPR-Cas) was employed to investigate AcrIF6 and AcrIF9 inhibitory effects. In Aa-CRISPR-Cas, the effector complex (Cascade) binds the DNA target and triggers its hydrolysis by Cas2/3 helicase/nuclease. AcrIF6 and AcrIF9 bind to Cascade and hinder DNA target recognition. In addition, AcrIF9 promotes non-specific interactions with DNA, which may facilitate CRISPR-Cas inhibition. Further, the AcrIF9 crystal structure was solved at 2.3 Å resolution and structure-based mutagenesis was employed. Putative AcrIF9 interaction surfaces with DNA and Cascade were identified, following the construction of seven AcrIF9 mutant vectors. The AcrIF9 mutans were expressed and purified. Finally, the significance for Cascade and DNA binding of the predicted interaction surfaces was confirmed by biochemical assays.

ACKNOWLEDGEMENTS

My academic journey would not be possible without encouragement from my mentor dr. Tomas Šinkūnas, family and friends.

I am grateful to my supervisor dr. Tomas Šinkūnas for sharing knowledge, providing help and guidance during preparation of this thesis.

I thank dr. Giedrė Tamulaitienė for the AcrIF9 structure.

I thank professor Virginijus Šikšnys for opportunity to work at department of Protein – DNA interactions and the whole BNSTS team for productive and friendly work environment.

Finally, I am extremely grateful for my family and friends for emotional support.

This work was supported by Research Council of Lithuania (S-MIP-20-39) to dr. Tomas Šinkūnas.

REFERENCE LIST

1. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DBT, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016 Aug 5;353(6299):aaf5573.
2. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica Section D, Biological crystallography 68, 352-367.
3. Alkhnbashi OS, Shah SA, Garrett RA, Saunders SJ, Costa F, Backofen R. Characterizing leader sequences of CRISPR loci. Bioinformatics. 2016 Sep 1;32(17):i576–85.
4. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science. 2007 Mar 23;315(5819):1709–12.
5. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005 Aug 1;151(8):2551–61.
6. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 2013 Jan 17;493(7432):429–32.
7. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. 2008 Dec 15;22(24):3489–96.
8. CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50, 760-763.
9. Charpentier E, Richter H, van der Oost J, White MF. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiology Reviews. 2015 May 1;39(3):428–41.
10. Chowdhury S, Carter J, Rollins MF, Golden SM, Jackson RN, Hoffmann C, et al. Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. Cell. 2017 Mar;169(1):47-57.e11.
11. Chowdhury S, Carter J, Rollins MF, Golden SM, Jackson RN, Hoffmann C, et al. Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. Cell. 2017 Mar;169(1):47-57.e11.
12. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011 Mar;471(7340):602–7.
13. Di Felice F, Micheli G, Camilloni G. Restriction enzymes and their use in molecular biology: An overview. J Biosci. 2019 Jun;44(2):38.
14. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. Science 359, eaar4120 (2018)
15. Dy RL, Richter C, Salmond GPC, Fineran PC. Remarkable Mechanisms in Microbes to Resist Phage Infections. Annu Rev Virol. 2014 Nov 3;1(1):307–31.
16. East-Seletsky A, O'Connell MR, Burstein D, Knott GJ, Doudna JA. RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. Molecular Cell. 2017 May;66(3):373-383.e3.
17. East-Seletsky A, O'Connell MR, Knight SC, Burstein D, Cate JHD, Tjian R, et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature. 2016 Oct;538(7624):270–3.

18. Elmore JR, Sheppard NF, Ramia N, Deighan T, Li H, Terns RM, et al. Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. Genes Dev. 2016 Feb 15;30(4):447–59.
19. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta crystallographica Section D, Biological crystallography *60*, 2126-2132.
20. Ershova AS, Rusinov IS, Spirin SA, Karyagina AS, Alexeevski AV. Role of restriction-modification systems in prokaryotic evolution and ecology. Biochemistry Moscow. 2015 Oct;80(10):1373–86.
21. Fagerlund RD, Wilkinson ME, Klykov O, Barendregt A, Pearce FG, Kieper SN, et al. Spacer capture and integration by a type I-F Cas1–Cas2-3 CRISPR adaptation complex. Proc Natl Acad Sci USA [Internet]. 2017 Jun 27 [cited 2022 Apr 7];114(26). Available from: https://pnas.org/doi/full/10.1073/pnas.1618421114
22. Fellmann C, Gowen BG, Lin P-C, Doudna JA, Corn JE. Cornerstones of CRISPR–Cas in drug discovery and therapy. Nat Rev Drug Discov. 2017 Feb;16(2):89–100.
23. Gabel C, Li Z, Zhang H, Chang L. Structural basis for inhibition of the type I-F CRISPR–Cas surveillance complex by AcrIF4, AcrIF7 and AcrIF14. Nucleic Acids Research. 2021 Jan 11;49(1):584–94.
24. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci USA [Internet]. 2012 Sep 25 [cited 2022 Apr 19];109(39). Available from: https://pnas.org/doi/full/10.1073/pnas.1208507109
25. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak-Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. EMBO J. 2015 Jan 14;34(2):169–83.
26. Gordeeva J, Morozova N, Sierro N, Isaev A, Sinkunas T, Tsvetkova K, et al. BREX system of Escherichia coli distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Research. 2019 Jan 10;47(1):253–65.
27. Gorrec F. The MORPHEUS protein crystallization screen. J Appl Crystallogr. 2009 Dec 1;42(6):1035–42.
28. Guo TW, Bartesaghi A, Yang H, Falconieri V, Rao P, Merk A, et al. Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex. Cell. 2017 Oct;171(2):414-426.e12.
29. Haft DH, Selengut J, Mongodin EF, Nelson KE. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. Eisen JA, editor. PLoS Comput Biol. 2005 Nov 11;1(6):e60.
30. Hampton HG, Watson BNJ, Fineran PC. The arms race between bacteria and their phage foes. Nature. 2020 Jan 16;577(7790):327–36.
31. Harms, A., Brodersen, D. E., Mitarai, N. & Gerdes, K. Toxins, targets, and triggers: an overview of toxin-antitoxin biology. Mol. Cell 70, 768–784 (2018).
32. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. Science. 2010 Sep 10;329(5997):1355–8.
33. Hille F, Richter H, Wong SP, Bratovič M, Ressel S, Charpentier E. The Biology of CRISPR-Cas: Backward and Forward. Cell. 2018 Mar;172(6):1239–59.
34. Hirschi M, Lu WT, Santiago-Frangos A, Wilkinson R, Golden SM, Davidson AR, et al. AcrIF9 tethers non-sequence specific dsDNA to the CRISPR RNA-guided surveillance complex. Nat Commun. 2020 Dec;11(1):2730.
35. Hochstrasser ML, Doudna JA. Cutting it close: CRISPR-associated endoribonuclease structure and function. Trends in Biochemical Sciences. 2015 Jan;40(1):58–66.
36. Hong S, Ka D, Yoon SJ, Suh N, Jeong M, Suh JY, et al. CRISPR RNA and anti-CRISPR protein binding to the Xanthomonas albilineans Csy1-Csy2 heterodimer in

the type I-F CRISPR-Cas system. Journal of Biological Chemistry. 2018 Feb;293(8):2744–54.
37. Jackson RN, Wiedenheft B. A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. Molecular Cell. 2015 Jun;58(5):722–8.
38. Javalkote VS, Kancharla N, Bhadra B, Shukla M, Soni B, Sapre A, et al. CRISPR-based assays for rapid detection of SARS-CoV-2. Methods. 2020 Oct;S1046202320302176.
39. Jia N, Patel DJ. Structure-based functional mechanisms and biotechnology applications of anti-CRISPR proteins. Nat Rev Mol Cell Biol. 2021 Aug;22(8):563–79.
40. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. 2012;337:7.
41. Kabsch, W. (2010). Xds. Acta crystallographica Section D, Biological crystallography *66*, 125-132.
42. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V. crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. RNA Biology. 2013 May;10(5):841–51.
43. Kazlauskiene M, Kostiuk G, Venclovas Č, Tamulaitis G, Siksnys V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science. 2017 Aug 11;357(6351):605–9.
44. Kim GE, Lee SY, Park HH. A high-resolution (1.2 Å) crystal structure of the anti-CRISPR protein AcrIF9. FEBS Open Bio. 2020 Dec;10(12):2532–40.
45. Koonin EV, Makarova KS. Origins and evolution of CRISPR-Cas systems. Phil Trans R Soc B. 2019 May 13;374(1772):20180087.
46. Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. Genome Biol. 2007;8(4):R61.
47. Kupcinskaite E. Investigation of Anti-CRISPR Proteins of a Type I-F CRISPR-Cas [bachelor's thesis]. Vilnius: Vilnius University; 2020.
48. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. Nat. Rev. Microbiol. 8, 317–327 (2010).
49. Li Y, Bondy-Denomy J. Anti-CRISPRs go viral: The infection biology of CRISPR-Cas inhibitors. Cell Host & Microbe. 2021 May;29(5):704–14.
50. Liu, X., Zhang, L., Xiu, Y., Gao, T., Huang, L., Xie, Y., Yang, L., Wang, W., Wang, P., Zhang, Y., et al. (2021). Insights into the dual functions of AcrIF14 during the inhibition of type I-F CRISPR-Cas surveillance complex. Nucleic acids research 49, 10178-10191.
51. Lu WT, Trost CN, Müller-Esparza H, Randau L, Davidson AR. Anti-CRISPR AcrIF9 functions by inducing the CRISPR–Cas complex to bind DNA non-specifically. Nucleic Acids Research. 2021 Apr 6;49(6):3381–93.
52. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, et al. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol. 2020 Feb;18(2):67–83.
53. Makarova KS, Wolf YI, Koonin EV. Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? The CRISPR Journal. 2018 Oct;1(5):325–36.
54. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Research. 2013 Apr;41(8):4360–77.
55. Makarova KS, Zhang F, Koonin EV. SnapShot: Class 1 CRISPR-Cas Systems. Cell. 2017 Feb;168(5):946-946.e1.
56. Makarova KS, Zhang F, Koonin EV. SnapShot: Class 2 CRISPR-Cas Systems. Cell. 2017 Jan;168(1–2):328-328.e1.

57. Marino ND, Zhang JY, Borges AL, Sousa AA, Leon LM, Rauch BJ, et al. Discovery of widespread type I and type V CRISPR-Cas inhibitors. Science. 2018 Oct 12;362(6411):240–2.
58. Marraffini LA, Sontheimer EJ. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science. 2008;322(5909):1843–5.
59. McGinn J, Marraffini LA. Molecular mechanisms of CRISPR–Cas spacer acquisition. Nat Rev Microbiol. 2019 Jan;17(1):7–12.
60. Mohanraju P., Makarova K.S., Zetsche B., Zhang F., Koonin E.V., van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science. 2016; 353:aad5147.
61. Mruk I, Kobayashi I. To be or not to be: regulation of restriction–modification systems and other toxin–antitoxin systems. Nucleic Acids Research. 2014 Jan 1;42(1):70–86.
62. Niewoehner O, Garcia-Doval C, Rostøl JT, Berk C, Schwede F, Bigler L, et al. Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature. 2017 Aug;548(7669):543–8.
63. Niu Y, Yang L, Gao T, Dong C, Zhang B, Yin P, et al. A Type I-F Anti-CRISPR Protein Inhibits the CRISPR-Cas Surveillance Complex by ADP-Ribosylation. Molecular Cell. 2020 Nov;80(3):512-524.e5.
64. Oliveira, P. H., Touchon, M. & Rocha, E. P. C. The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. Nucleic Acids Res. 42, 10618–10631 (2014).
65. Osuna BA, Karambelkar S, Mahendra C, Christie KA, Garcia B, Davidson AR, et al. Listeria Phages Induce Cas9 Degradation to Protect Lysogenic Genomes. Cell Host & Microbe. 2020a Jul;28(1):31-40.e9.
66. Osuna BA, Karambelkar S, Mahendra C, Sarbach A, Johnson MC, Kilcher S, et al. Critical Anti-CRISPR Locus Repression by a Bi-functional Cas9 Inhibitor. Cell Host & Microbe. 2020 Jul;28(1):23-30.e5.
67. Özcan A, Pausch P, Linden A, Wulf A, Schühle K, Heider J, et al. Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. Nat Microbiol. 2019 Jan;4(1):89–96.
68. Pawluk A, Bondy-Denomy J, Cheung VHW, Maxwell KL, Davidson AR. A New Group of Phage Anti-CRISPR Genes Inhibits the Type I-E CRISPR-Cas System of Pseudomonas aeruginosa. Hendrix R, editor. mBio. 2014 May;5(2):e00896-14.
69. Pawluk A, Staals RHJ, Taylor C, Watson BNJ, Saha S, Fineran PC, et al. Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. Nat Microbiol. 2016 Aug;1(8):16085.
70. Peng R, Xu Y, Zhu T, Li N, Qi J, Chai Y, et al. Alternate binding modes of anti-CRISPR viral suppressors AcrF1/2 to Csy surveillance complex revealed by cryo-EM structures. Cell Res. 2017 Jul;27(7):853–64.
71. Pinilla-Redondo,R., Shehreen,S., Marino,N.D., Fagerlund,R.D., Brown,C.M., Sørensen,S.J., Fineran,P.C. and Bondy-Denomy,J. (2020) Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. Nat. Commun., 11, doi:10.1038/s41467-020-19415-3.
72. Richter C, Dy RL, McKenzie RE, Watson BNJ, Taylor C, Chang JT, et al. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. Nucleic Acids Res. 2014 Jul 29;42(13):8516–26
73. Rollins MF, Chowdhury S, Carter J, Golden SM, Wilkinson RA, Bondy-Denomy J, et al. Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. Proc Natl Acad Sci USA [Internet]. 2017 Jun 27 [cited 2022 Apr 7];114(26). Available from: https://pnas.org/doi/full/10.1073/pnas.1616395114

74. Rollins MF, Schuman JT, Paulus K, Bukhari HST, Wiedenheft B. Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from Pseudomonas aeruginosa. Nucleic Acids Research. 2015 Feb 27;43(4):2216–22.
75. Rollins,M.C.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.
76. 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 Reports. 2015 Mar;10(9):1534–43.
77. Samson JE, Magadán AH, Sabri M, Moineau S. Revenge of the phages: defeating bacterial defences. Nat Rev Microbiol. 2013 Oct;11(10):675–87.
78. Seed KD. Battling Phages: How Bacteria Defend against Viral Attack. Miller VL, editor. PLoS Pathog. 2015 Jun 11;11(6):e1004847.
79. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci USA. 2011 Jun 21;108(25):10098–103.
80. Shao Y, Li H. Recognition and Cleavage of a Nonstructured CRISPR RNA by Its Processing Endoribonuclease Cas6. Structure. 2013 Mar;21(3):385–93.
81. Shao Y, Richter H, Sun S, Sharma K, Urlaub H, Randau L, et al. A Non-Stem-Loop CRISPR RNA Is Processed by Dual Binding Cas6. Structure. 2016 Apr;24(4):547–54.
82. Shivram H, Cress BF, Knott GJ, Doudna JA. Controlling and enhancing CRISPR systems. Nat Chem Biol. 2021 Jan;17(1):10–9.
83. Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, et al. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase–Cas1 fusion protein. Science. 2016 Feb 26;351(6276):aad4234.
84. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J. 2011;30(7):1335–42.
85. Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P. & Siksnys, V. (2013) In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus, The EMBO journal. 32, 385-94.
86. Smargon AA, Cox DBT, Pyzocha NK, Zheng K, Slaymaker IM, Gootenberg JS, et al. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Molecular Cell. 2017 Feb;65(4):618-630.e7.
87. Staals RHJ, Jackson SA, Biswas A, Brouns SJJ, Brown CM, Fineran PC. Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR–Cas system. Nat Commun. 2016 Nov;7(1):12853.
88. Sternberg S.H., Richter H., Charpentier E., Qimron U. Adaptation in CRISPR-Cas systems. Mol. Cell. 2016; 61:797–808.
89. Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P, et al. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc Natl Acad Sci USA. 2014 Jul 8;111(27):9798–803.
90. Tamulaitis G, Kazlauskiene M, Manakova E, Venclovas C, Nwokeoji AO, Dickman MJ, et al. Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. Molecular Cell. 2014 Nov;56(4):506–17.
91. Terns M.P., Terns R.M. CRISPR-based adaptive immune systems. Curr. Opin. Microbiol. 2011; 14:321–327
92. Tock MR, Dryden DT. The biology of restriction and anti-restriction. Current Opinion in Microbiology. 2005 Aug;8(4):466–72.

93. Toro N, Mestre MR, Martínez-Abarca F, González-Delgado A. Recruitment of Reverse Transcriptase-Cas1 Fusion Proteins by Type VI-A CRISPR-Cas Systems. Front Microbiol. 2019 Sep 13;10:2160.
94. Tuminauskaite D, Norkunaite D, Fiodorovaite M, Tumas S, Songailiene I, Tamulaitiene G, et al. DNA interference is controlled by R-loop length in a type I-F1 CRISPR-Cas system. BMC Biol. 2020 Dec;18(1):65.
95. Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta crystallographica Section D, Biological crystallography 66, 22-25.
96. Vorontsova D, Datsenko KA, Medvedeva S, Bondy-Denomy J, Savitskaya EE, Pougach K, et al. Foreign DNA acquisition by the I-F CRISPR–Cas system requires all components of the interference machinery. Nucleic Acids Res. 2015 Dec 15;43(22):10848–60.
97. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research 46, W296-W303.
98. Watters KE, Fellmann C, Bai HB, Ren SM, Doudna JA. Systematic discovery of natural CRISPR-Cas12a inhibitors. Science. 2018 Oct 12;362(6411):236–9.
99. Westra ER, van Erp PBG, Künne T, Wong SP, Staals RHJ, Seegers CLC, et al. CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. Molecular Cell. 2012 Jun;46(5):595–605.
100. Wiedenheft B., Sternberg S.H., Doudna J.A. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012; 482:331–338.
101. Wright AV, Liu J-J, Knott GJ, Doxzen KW, Nogales E, Doudna JA. Structures of the CRISPR genome integration complex. Science. 2017 Sep 15;357(6356):1113–8.
102. Xiao Y, Luo M, Dolan AE, Liao M, Ke A. Structure basis for RNA-guided DNA degradation by Cascade and Cas3. Science. 2018 Jul 6;361(6397):eaat0839.
103. Xue C, Zhu Y, Zhang X, Shin Y-K, Sashital DG. Real-Time Observation of Target Search by the CRISPR Surveillance Complex Cascade. Cell Reports. 2017 Dec;21(13):3717–27.
104. Yang L, Zhang L, Yin P, Ding H, Xiao Y, Zeng J, Wang W, Zhou H, Wang Q, Zhang Y, Chen Z, Yang M, Feng Y. Insights into the inhibition of type I-F CRISPR-Cas system by a multifunctional anti-CRISPR protein AcrIF24. Nat Commun. 2022 Apr 11;13(1):1931. doi: 10.1038/s41467-022-29581-1. PMID: 35411005.
105. Yang L, Zhang Y, Yin P, Feng Y. Structural insights into the inactivation of the type I-F CRISPR-Cas system by anti-CRISPR proteins. RNA Biology. 2021 Nov 12;18(sup2):562–73.
106. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell. 2015 Oct;163(3):759–71.
107. Zhang K, Wang S, Li S, Zhu Y, Pintilie GD, Mou TC, et al. Inhibition mechanisms of AcrF9, AcrF8, and AcrF6 against type I-F CRISPR–Cas complex revealed by cryo-EM. Proc Natl Acad Sci USA. 2020 Mar 31;117(13):7176–82.
108. Zhang Y, Sontheimer EJ. Cascading into focus. Science. 2014 Sep 19;345(6203):1452–3.