CHARACTERIZING TYPE III CRISPR-CAS DEFENSE SYSTEMS

by

KAWANDA AMY FOSTER

(Under the Direction of Christine Szymanski)

ABSTRACT

CRISPR (Clustered Regularly Interspaced Palindromic Repeats)-Cas (CRISPR associated) is a robust immune system present in prokaryotes used to combat viruses, plasmids and other mobile genetic elements. The pathway to immunity occurs in three robust stages known as adaptation, crRNA biogenesis, and defense. CRISPR-Cas systems are very diverse and are currently organized into six overarching types containing close to thirty subtypes. The focus of this dissertation is the defense stage of subtype CRISPR-Cas systems III-A (Csm) and III-B (Cmr). Specifically, the requirements for successful anti-plasmid immunity and the roles of Csm6 and Csx1 in Type-III defense, was investigated. Several different technical approaches were utilized to study these systems, drawing upon elements of biochemistry, molecular biology, genetics, microbiology, and bioinformatics. Characterization of three Csm6 homologs from *Lactococcus lactis*, *Streptococcus thermophilus*, and *Staphylococcus epidermidis,* was performed using the model organism *Escherichia coli*. Csm6 was found to be a metal-independent single-stranded RNA-specific endoribonuclease. Cleavage activity of Csm6 was found to be dependent upon the C-terminal Higher Eukaryotes and Prokaryotes Nucleotide Binding (HEPN) motif that is common to a large family of ribonucleases.

Functioning of the defense response for all three Csm systems, *in vivo*, was found to be dependent upon the ribonuclease activity of Csm6. Characterization of the Type III-B system was conducted using the Cmr system of *Pyrococcus furiosus* (*Pfu*). In *Pfu*, the role of Csx1 in defense was explored by reconstituting the recently discovered cyclic oligoadenylate (cOA) production activity of Type III systems. Csx1 was found to be an important component of the Type III-B response of *Pfu*. Production of cOA molecules by the Cmr effector complex activates *Pfu* Csx1 endoribonuclease activity through first binding to the Cas Rossman Fold (CARF) domain. A component of the target RNA, the protospacer flanking sequence (PFS), was found to be essential for triggering both the DNA and RNA nuclease activities of the Type III-B system. The results presented in this dissertation provide important insight into the mechanism of action and regulation of Type III CRISPR-Cas systems and solidifies the role of highly related Csm6 and Csx1 ribonucleases in CRISPR-Cas immunity.

INDEX WORDS: CRISPR; Cas; Csm; Cmr; Csm6; Csx1; Type III; PFS; defense; targeting; invader silencing; RNA; cyclic oligoadenylate; nuclease; *Lactococcus lactis*; *Streptococcus thermophilus*; *Staphylococcus epidermidis*; *Pyrococcus furiosus*

CHARACTERIZING TYPE III CRISPR-CAS DEFENSE SYSTEMS

by

KAWANDA AMY FOSTER

B.S., University of Miami, 2011

M.P.H, Emory University, 2013

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

CHARACTERIZING TYPE III CRISPR-CAS DEFENSE SYSTEMS

by

KAWANDA AMY FOSTER

Major Professor: Christine Szymanski
Committee: Jorge C. Escalante
Robert J. Maier
Diana M. Downs

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

DEDICATION

This dissertation is dedicated to all the individuals who have supported me on this journey. You celebrated me during my highs and uplifted me during my lows. Your support and belief in me gave me the strength to keeping going, even when I did not believe in myself. Thank you.

ACKNOWLEDGEMENTS

I am a firm believer in the motto “It takes a village”. Thus, I would like to take a moment to acknowledge those who had a direct impact on my successful completion of this dissertation. I would like to express sincere gratitude to my scientific advisor Dr. Michael Terns for guiding and overseeing my academic endeavors and providing me with the resources and tools I needed to succeed. I would also like to thank my committee members Dr. Christine Szymanski, Dr. Jorge Escalante, Dr. Robert Maier, and Dr. Diana Downs for the helpful feedback, support, and encouragement over the years.

To all Terns lab members past and present, thank you. Each one of you played a part in my learning and growth during this process and that helped me reach this point. I would especially like to thank the following members:

- Dr. Rebecca Terns: for your incredible mentorship and invaluable advice and feedback on data, presentations, and proposals.
- Dr. Caryn Hale: for discovering Cmr and all of your assay and protocol development.
- Dr. Joshua Elmore: for all your scientific contributions to *Pfu* work and Cmr and for being a great teacher, mentor, and friend. You have always helped me whenever I needed it and I will always appreciate you.
- Dr. Chris Cooper: for being a great trainer and early mentor. You saw my potential and gave me a great foundation to build on.

- Nolan Sheppard: for all your hard work with Csx1 that laid the foundation for my work and teaching me so many things and dealing with all of my shenanigans.
- Dr. Claiborne Glover: for your kindness and helpful feedback.
- Joshua Kalter: for being an amazing undergraduate research mentee and exceeding my expectations.
- Michael Ellis: for being a great teacher on protein techniques and for the numerous impromptu and delightful scientific chats.
- Dr. Yunzhou Wei: for all of your great cloning advice and helpful feedback.
- Walter Woodside: for your helpful discussions and feedback.
- Elizabeth Watts: for teaching me how to culture and utilize *Pfu* for my experiments.

Another group of individuals that I would like to thank are members of the scientific community. Thank you to Janet Westpheling for your wonderful scientific guidance and mentorship. And thank you to the CRISPR field for the amazing work which has made CRISPR-Cas an exciting area to study.

I would also like to thank all of my friends and family around the country who have loved and supported me throughout this. You all have laughed with me, cried with me, prayed for me, and kept me grounded. Thank you to my mom and Grandma Mills for being my backbones and instilling in me the values I needed to succeed in life. To (a few) of my best friends, Emily Carpinone, Ashley Hagen, Jenny Kim, and Ivette Nuñez, thank you for all of your emotional and mental support. Thank you to Blaze Foster and Whiskey Ladson for being the best cats a girl could ask for. Last but not least, thank you to Jeremy Ladson for being my biggest cheerleader and the best partner a girl could ask for. Your unconditional love and support has meant the world to me.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Introduction to CRISPR-Cas Systems

Bacteria and bacteriophages (phages) are involved in a perpetual and dynamic arms race for survival; over time each entity has evolved mechanisms to thwart the other [1,2]. Restriction modification systems, targeting of phage absorption receptors, toxin-antitoxin systems, and CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas (CRISPR associated) immune systems are all examples of defense mechanisms utilized by bacteria to defend against viral attacks [1,3-7]. CRISPR-Cas is a RNA-based adaptive and heritable immune system present in 83% of archaeal and 45% of sequenced bacterial genomes [8-10]. CRISPR-Cas systems have been found in viral genomes and on plasmids indicating that they are also actively exchanged through horizontal gene transfer [11,12].

CRISPR-Cas systems are composed of two components (Figure 1.1). The CRISPR component refers to the CRISPR locus. The CRISPR locus is composed of a leader sequence and alternating segments of repeat and spacer sequences. Repeat sequences are identical within a CRISPR locus. Spacer sequences are unique and serve as a memory log for invaders that a particular CRISPR-Cas system has interacted with. CRISPR-Cas immunity is executed in three overarching stages—adaptation, crRNA biogenesis, and defense (Figure 1.2). The process begins with adaptation. During

adaptation, foreign pieces of DNA are successfully incorporated into the CRISPR locus in between identical repeat sequences [4,13]. After successful incorporation of the newly acquired spacer into the CRISPR locus, the entire CRISPR locus is transcribed and the transcript is processed by an endonuclease during the crRNA biogenesis stage [14-17]. Resulting mature crRNAs combine with Cas protein(s) to form effector complexes during the final stage [14,18-20]. The defense stage commences with effector complexes degrading the nucleic acid of the invader. Specificity in this immune system is achieved via the existence and structure of the functional crRNA species which selectively recognize the target nucleic acid through base-pairing interactions.

A prominent contributing aspect of the diversity of CRISPR-Cas systems lies within the Cas component of the immune system. There are 13 core Cas proteins (Cas1-Cas13) [21]. Cas1 and Cas2 are viewed as universal Cas proteins and Cas3, Cas9, Cas10, Cas12, and Cas13 are viewed as signature Cas proteins that help define the six distinct types of CRISPR-Cas systems, respectively.

Classification of CRISPR-Cas Systems

CRISPR-Cas systems are highly diverse. Thus, classifying them has become more difficult as new systems emerge. Currently, CRISPR-Cas systems are classified based the composition of their repeat sequences, *cas* gene content, and effector complex composition [22-24]. CRISPR-Cas systems are classified into two main classes—Class 1 and Class 2 [25]. Class 1 is characterized by containing systems that form multi-subunit effector complexes. Class 2 is characterized by containing systems that form single-subunit effector complexes. Each class is then segmented by system type. There are

currently 6 known types (I-IV). Class 1 includes Types I, III, and IV and class 2 includes Types II, V, and VI (Table1.1) [26]. The six system types are further broken down into subtypes. As shown in Table 1.1, there are currently 29 classified CRISPR-Cas subtypes [21]. Amongst CRISPR-Cas systems, Types I, II, and III are the most studied systems. Very little is known or understood about the Type IV CRISPR-Cas system beyond the mechanism of crRNA biogenesis; thus more investigation into Type IV systems is needed [21,25,27,28].

Great advancements in the practice of classifying and understanding CRISPR-Cas systems have occurred due to the combination of bioinformatic, biochemical, and genetic analyses that have been developed by researchers around the world. Cas1 and Cas2 are the most prominent Cas proteins and often regarded as the universal CRISPR systems that play a role in the CRISPR adaptation phase of the pathway [24,29]. However, as more functional CRISPR-Cas systems are discovered, more variation is seen in regards to the presence of both proteins. Some functionally active CRISPR-Cas systems are found in the absence of *cas1* and/or *cas2* and some systems depend on shared *cas1* and *cas2* genes located on different loci within the genome such as the Type III-A (Csm) system of *Lactococcus lactis* and the Type I-A (Csa) system of *Pyrococcus furiosus* [11,18]. Thus, in order to discover more functional CRISPR-Cas systems, alternative methods of identifying new systems are employed. One approach involves a two-step process of first creating a library of all *cas* genes to use as queries to search available genomes for highly similar genes. The architecture of the gene modules are then evaluated to identify signature and distinctive genes for classification [25]. Known Cas protein family sequences are often used to search thousands of genomes in the National

Center for Biotechnology Information (NCBI). A threshold value is utilized to assess similarity between known *cas* genes and putative *cas* genes. Gene neighborhoods are then analyzed for other components of a complete CRISPR-Cas system [25]. Multiple sequence alignments are also utilized to further assess classification and divergence. This and related bioinformatic approaches, have been highly effective in discovering new CRISPR-Cas systems.

Recombination events at CRISPR loci play a major role in the evolution of CRISPR-Cas systems [25,30,31]. Thus, several variants of CRISPR-Cas systems exist and can be more difficult to classify—especially in the absence of supporting biochemical data [21]. The Type III system of *Thermotoga lettingae* is an example of a variant system that has been difficult to classify [25,32]. The system has notable similarities to conventional Type III-A systems; however, it only encodes a single Cas7 protein family which is a prominent characteristic of Type IV systems [25,32]. As previously stated, CRISPR-Cas systems are present in approximately 50% of sequenced bacterial genomes and approximately 90% of sequenced archaeal genomes. Class 1 CRISPR-Cas systems dominate archaeal genomes in regards to representation [25]. Type I subtypes are the most abundant and complete subtypes in bacteria (60%) and archaea (64%). Type III subtypes compose about 34% of archaeal and about 25% of bacterial genomes. Despite their extraordinary utilization in the fields of biotechnology and genome engineering, archaea lack Type II systems and Type II subtypes are present in only about 13% of bacteria. Types IV, V, and VI are the rarest system types [25].

Overview of Adaptation in CRISPR-Cas Systems

Bacteria utilize several different processes to protect themselves against threats from phages [1]. CRISPR-Cas is an extremely useful mechanism amongst all of the systems due to its ability to adapt. Adaptation is the first stage of CRISPR-Cas and results in the incorporation of a small fragment of foreign invader DNA into the CRISPR locus (Figure 1.2) [33,34]. In order for adaptation to be successful, three main steps must be successfully executed: selection and processing of the protospacer, recognition of the CRISPR locus, and spacer integration into the CRISPR locus (Figure 1.3) [35]. Mechanisms of adaptation have predominantly been investigated in Type I and II systems.

Cas1 and Cas2 are often regarded as the workhorses of adaptation due to their almost universal existence amongst CRISPR-Cas systems and essentiality in the adaptation process (Figure 1.3) [36-39]. Cas1 is the integrase enzyme that catalyzes the spacer integration reaction and that Cas2 is a partner protein that appears to play a structural role that helps guide accurate integration of invader DNA fragments at CRISPR repeats. In the Type I system of *E. coli*, Cas1 and Cas2 are sufficient for executing adaptation [36,37]. However, for other subtypes, additional Cas proteins are needed for adaptation events. Investigation into the Type II systems of *Streptococcus pyogenes* and *Streptococcus thermophilus* have revealed that in addition to Cas1 and Cas2, Cas9 and Csn2 (Type II specific Cas proteins) are important components of the adaptation process [40-43]. The Type I-D system of *Synechocystis sp* was found to require Cas1, Cas2, and Cas4 [44]. In the CRISPR-Cas system of *Pyrococcus furiosus* which harbors two Type I subtypes (I-A and I-G) and one Type III subtype (II-B), two distinct Cas4 proteins have

been shown to be required for adaptation along with Cas1 and Cas2 [45,46]. Furthermore, analysis conducted in the Type I-F system of *Pseudomonas aeruginosa* discovered that Cas1, Cas2, and the entire I-F effector complex is needed for adaptation [47].

A prominent question within the field of adaptation has focused on the process of protospacer generation. Evidence has shown that double-stranded (ds)DNA is the preferred substrate of the adaptation complex [48]. Thus, the origins and specificity of this chosen dsDNA has been an emerging area of exploration. Recent findings have connected cellular processes to protospacer generation such as replication and restriction enzyme activity that produce double-stranded breaks (DSB) in the DNA (Figure 1.4) [35]. In *E. coli,* replication forks are repaired via the RecBCD DNA repair complex. The DSBs are recognized by the RecBCD helicase-nuclease complex which rapidly unwinds and degrades the DNA [49]. This process continues until RecBCD reaches a crossover hotspot instigator (Chi) site. Chi sites are recombination hot spots; they are sequences that are recognized by the repair complex to attenuate activity and trigger recombination. Chi sites sequences vary between organisms but an example sequence in *E. coli* is 5′-GCTGGTGG-3′. Analysis conducted in the Type I-E system found that major hotspots for spacer uptake mapped back to regions in which DSBs had occurred. These hotspots were additionally bound by Chi sites, further tying RecBCD degradation byproducts to targeted spacers by CRISPR-Cas systems [34]. Another aspect of protospacer selection pertains to recognition of a short DNA sequence adjacent to the protospacer region called the protospacer adjacent motif or PAM. A PAM sequence is a critical component of substrate capture by the adaptation complex and is typically 2 to 5 base pairs in length (Figure 1.3) [50-53]. Prior to integration into the CRISPR array, the

protospacers are processed into suitable spacers for the adaptation complex. The lengths of spacers integrated into CRISPR arrays vary amongst systems. In *P. furiosus*, two distinct Cas4 proteins are responsible for processing protospacers prior to integration by the Cas1-Cas2 complex [46]. More exploration into protospacer processing is needed for systems with and without Cas4.

Successful adaptation can be categorized into naïve or primed. Naïve adaptation occurs when a new spacer is incorporated into the CRISPR array from a unique source [54]. During naïve adaptation, self-originated protospacers are primarily generated by stalled replication forks that occur during the replication process and highly transcribed regions [34,55]. The primary mechanisms of non-self (foreign) generated protospacer origins are still widely unknown and the subject of much investigation. During primed adaptation, existing spacers within the CRISPR array are repurposed to trigger adaptation of protospacers of similar sequence. Primed adaptation is an essential form of adaptation due to the ability of viruses to mutate resulting in CRISPR-escaper mutants. Priming allows the CRISPR-Cas system to take advantage of existing spacers to overcome the escaper phenomenon and still defend [56-58]. An additional advantage of priming is the increase in efficiency of successful adaptation. With priming, systems have been found to enhance spacer acquisition 10 to 20-fold compared to naïve conditions [58,59]. During the naïve state, the system's effector complex may or may not be required for spacer acquisition depending on the subtype under study. However, during the primed state, the effector complex is required [57,59-61].

Once the Cas1-Cas2 centric adaptation complex successfully binds its substrate, the CRISPR array must be recognized. In *E. coli*, Cas1 and Cas2 are facilitated through this

process with the assistance of host proteins such as integration host factor (IHF) [62,63]. IHF aides in specific CRISPR locus recognition by bending the leader portion by 120° to direct spacer integration at the leader-proximal repeat with high specificity [62,64]. The bending by IHF also increases specificity of integration into the leader-proximal repeat portion of the CRISPR array. In the absence of IHF, about 35% of integration events occurred at the leader-proximal repeat region *in vitro*; however, in the presence of IHF, that percentage increased to 80% *in vitro* [62]. *In vivo*, IHF appears to be essential for integration and this occurs virtually exclusively at the leader-proximal repeat. Many CRISPR containing organisms do not contain IHF. Thus, alternative mechanisms for specificity of CRISPR array recognition exist. In the Type II-A systems of *S. pyogenes* and *S. thermophilus* short (less than 10 base pairs in length) sequences within the repeat adjacent end of the leader region are critical for specific integration into the CRISPR array by Cas1 and Cas2 [41,42,65].

Adaptation commences upon successful integration of the spacer into the CRISPR array. Integration of the new spacer primarily occurs at the leader-repeat end of the locus and involves a two-step transesterification reaction via the 3'-OH groups of the spacer (Figure 1.5). For most systems, specific integration into the CRISPR locus occurs with directionality. In the first round of attack, the leader-proximal repeat undergoes a nucleophilic attack on the plus strand of the leader-repeat junction resulting in a half-site intermediate. During the second attack, the minus strand of the locus undergoes a nucleophilic attack at the repeat-spacer junction of the same repeat resulting in full-site integration [65-67]. The remaining gaps are then filled by DNA polymerase and DNA ligase and the immune process continues [48].

Overview of crRNA Biogenesis in CRISPR-Cas Systems

In the CRISPR-Cas pathway, crRNA biogenesis is a crucial step in the process to produce functional crRNAs that assemble with Cas proteins to form effector complexes (Figure 1.2). Without the generation of the appropriate crRNA component, the effector complexes cannot form or function and CRISPR-Cas defense will not occur [14,68,69]. The structure of a mature or functional crRNA differs based on the system and the process of producing functional crRNAs occurs in two to three main steps. The process begins with transcription of the CRISPR locus to produce a long precursor crRNA (pre-crRNA). This pre-crRNA is then cleaved within the repeat regions to produce crRNAs that contain full length spacer regions flanked by repeat sequences of variable length. In some systems, this is the final step and the crRNA is functional. However, some systems undergo a third step to further process the crRNA to remove additional repeat and/or spacer sequence [69]. Figure 1.6 provides a summary of functional crRNA species structures for Types I - VI.

Class 1 CRISPR systems overwhelmingly utilize the endonuclease activity of the Cas6 superfamily of proteins to produce crRNAs [15,17,27,28], except for the I-C subtype which utilizes Cas5 for pre-crRNA processing [70,71]. After production of the pre-crRNA, each spacer unit in the CRISPR array is released via the nuclease activity of Cas6 or Cas5 (Type I-C). The cleavage occurs within the repeat regions which leaves each spacer unit flanked by repeat sequence on its 5' and 3' ends for some Class 1 systems such as Type I-E and Type IV systems [14,15,28,72,73]. Alternatively, many Class 1 CRISPR systems process the crRNAs even further to remove the repeat sequences from the 3' end such as Type I-A and Type III systems via an additional unknown nuclease [16,73-

75]. Cas6 is a divalent metal independent nuclease; thus, it uses acid-base chemistry to cleave the primary transcript via a transesterification reaction [76]. A nucleophilic attack from a 2' hydroxyl group located upstream from a scissile phosphate catalyzes the cleavage reaction by Cas6 [76]. Cleavage occurs on the 3' end of the repeat regions as a result of the stem loops that form due to the palindromic nature of the repeat regions [14,15,69]. In systems with non-palindromic repeats and thus unstable stem loops, such as I-A and I-B, Cas6 dimerizes and restructures the repeat to form the pre-requisite stem-loop structure for cleavage [77-80].

CRISPR RNA maturation in Class 2 systems is quite different than in Class 1 systems due to the absence of Cas6 [21]. Processing of crRNAs in Type II systems requires Cas9, RNase III, and a trans-activating CRISPR RNA known as tracrRNA [81]. These tracrRNAs are abundant non-coding RNAs in Type II systems that are transcribed upstream of the leader-repeat array on the opposite strand [81]. Cas9, RNase III, and tracrRNA form a microprocessor complex due to tracrRNA possessing a complementary sequence to the repeat sequences of the pre-crRNA. Cas9 stabilizes the complex and RNase III cleaves the pre-crRNA in a site-specific manner [81,82]. The crRNA undergoes one additional processing step to remove the 5' tag region by an unknown nuclease [81]. Processing of crRNAs in Type V and Type VI systems parallels Type II systems in regards to the absence of Cas6, but differ from Type II systems due to the absence of tracrRNA [21]. In Type V systems, crRNA processing is accomplished with a single protein known as Cas12 [83]. Similar to Cas6 enzymes, Cas12 cleaves the hairpin structure of the repeat; however, the activity of Cas12 is also dependent upon metal and the length and sequence of the stem-loop [83]. Processing results in a crRNA with repeat sequence on

the 5' end of the crRNA, but not the 3' end [83,84]. Maturation in Type VI systems is accomplished with Cas13. Type IV systems are unique because they only target ssRNA via conserved Higher Eukaryotes and Prokaryotes Nucleotide Binding (HEPN) domains [85]. Although processing of the pre-CRNA is not required for defense in this system, processing does occur and produces functional crRNA species. Similar to Cas12, Cas13 recognizes the repeat sequences in a structure and sequence specific manner [85]. Cleavage by Cas 13 results in a crRNA with repeat sequence only on the 5' end of the crRNA [84,85].

Overview of Invader Silencing in CRISPR-Cas Systems

The adaptive CRISPR-Cas immune response commences with crRNA-guided and sequence-specific destruction of the invader nucleic acid (Figure 1.2). Class 1 systems contain multi-subunit effector complexes and Class 2 systems contain single-subunit complexes (Table 1.1). Class 1 effector complexes form conserved seahorse-like shapes but vary slightly between subtypes and organisms (Figure 1.7) [14,86-90]. The targeting complex of Type I systems is known as Cascade (CRISPR-associated complex for antiviral defense) [14], and the target of Cascade is double-stranded (ds)DNA [86]. Cascade is composed of five Cas proteins: Cas5, Cas6, Cas7, Cas8, and Cas11. An additional protein, Cas3, temporarily associates with Cas3 to degrade invader DNA [14,86,87]. The helical backbone of Cascade is formed by multiple copies of Cas7. After crRNA biogenesis, Cas6 remains bound to the crRNA on the 3' end and becomes part of the complex. The 5' end of the crRNA is capped by Cas5 which interacts with Cas8 to form the tail of the complex. Cas11 interacts with the Cas7 backbone forming the belly of

the complex (Figure 1.7) [86,87,91-93]. Binding of Cascade to the DNA target is triggered by recognition of the PAM sequence via Cas8. The lysine rich nature of Cascade is critical for the binding and unwinding of the dsDNA substrate [87]. Once the target DNA binds to Cascade, a conformational change occurs and is stabilized by conserved arginine residues [87]. Recognition of the PAM initiates unwinding of the DNA which results in the crRNA binding to the complementary strand and the non-complementary stand forming a R-loop structure stabilized by Cas8 and Cas11 [86,87]. Formation of the R-loop prompts binding of Cas3 to the non-complementary strand resulting in further unwinding and degradation of the non-complementary strand through the ATP- and metal-dependent helicase and HD domain DNA nuclease activity of Cas3 [93,94].

Figure 1.8 depicts the composition of Type III effector complexes known as Csm (III-A) and Cmr (III-B) complexes. Cas6 produced mature crRNAs assembly along the backbone of the complex composed of Cas7 superfamily proteins Csm3-Csm5 and Cmr1-Cmr4-Cmr6 for Csm and Cmr complexes, respectively. The crRNA is also bound by the Cas5 proteins Csm4 and Cmr3. The tail of the complexes are formed by Cas10 family proteins Csm1 and Cmr2, and the bellies are formed by Cas11 family proteins Csm2 and Cmr5 [88,90,95,96]. These complexes are very similar to Cascade with some notable differences from Cascade such as the lack of Cas6 and the lack of recruitment by a Cas3-like protein (Figure 1.7). Csm and Cmr complexes are unique from all other known CRISPR-Cas complexes due to their ability to target RNA and DNA [18,89,97-99]. As the target RNA is transcribed and released from the transcription complex, Type III complexes bind to the transcript via complementation between the crRNA and the target RNA [16,18]. Binding of the target RNA to the Csm/Cmr complex is dependent upon the

target RNA sequence within the region of complementarity and the 3' flanking region of the target RNA [68,100]. Successful binding of the target RNA to the complex results in a conformational change that triggers three highly conserved activities amongst Type III defense systems [101]—target RNA cleavage [18], DNA cleavage [98], and cyclic oligoadenylate (cOA) production [102,103]. The RNase activity occurs via conserved aspartic acid residues in Csm3 and Cmr4 [104], DNase activity is facilitated by the HD domain found in Csm1 and Cmr2 [105], and cOA species are produced via the Palm domains of Csm1 and Cmr2 [102,103]. A fourth activity of Type III defense systems involves an additional protein that is not a stable member of the complex: Csm6 (III-A) and Csx1 (III-B). Csm6 and Csx1 are highly conserved within Type III systems and contain conserved HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding) and CARF (CRISPR-associated Rossmann Fold) domains [106-108]. Csm6 and Csx1 are trans-acting ribonucleases. This activity occurs via the HEPN domain [109-111]. These enzymes contribute to defense via activation of their ribonuclease function by the Cas10 produced cOA molecules [112]. Together, these four different activates encapsulate the multi-level defense response utilized by Type III systems. Type IV systems are still being investigated and many questions remain regarding the mechanisms of interference in those systems.

Class two CRISPR-Cas systems utilize single subunit complexes to target invader nucleic acid (Table 1.1). For Type II systems that single enzyme is Cas9. The Cas9 effector complex contains Cas9, crRNA, and tracrRNA (Figure 1.9) [81]. After crRNA biogenesis, crRNA and tracrRNA remain bound to each other to form a functional guide RNA unit [81,82,113]. Along with Cas9, the targeting complex recognizes the appropriate

target DNA via PAM recognition. PAM recognition occurs via a highly conserved arginine residues within Cas9 [82,113]. Once PAM recognition occurs, the complex binds to the DNA through base-pairing between the guide RNA and the target DNA sequence. The dsDNA is then unwound resulting in R-loop formation [82]. Both strands of the dsDNA are then cleaved by different domains within Cas9 [82,113,114]. Cas9 has a bilobed architecture that undergoes conformational changes when bound to the guide RNA and target DNA [113,114]. The two lobes of Cas9 are known as the REC (α-helical recognition) and NUC (nuclease) lobes. Two highly conserved domains (HNH and RuvC) are located within the NUC lobe [82,113,114]. The DNA strand that binds to the crRNA is cleaved by the HNH and the non-complementary strand is cleaved by the RuvC domain [82,113,114].

The Cas12 effector complex of Type V systems is composed of Cas12 and crRNA (Figure 1.9). Cas12 effector complexes share functional similarities with Cas9 complexes though there are notable differences. For instance, Cas12 utilizes one distinct RNA species (crRNA) as a guide while Cas9 uses two. While the Cas9 family of proteins recognize G-rich PAMs for invader silencing, Cas12 proteins recognize T-rich PAMs [26,115]. Cas9 degrades target DNA via the RuvC and HNH domains while Cas12 only requires the RuvC domain [26,115-118]. Cas12 also degrades the dsDNA in a manner that results in staggered ends while the mechanism utilized by Cas9 leaves blunt ends on the dsDNA [26,115,117]. Similar to Cas9, Cas12 has a bilobed architecture composed of the REC and NUC lobes. The REC lobe contains two conserved REC domains and the NUC lobe contains conserved RuvC, WED (Wedge), and PI (PAM-Interacting) domains [116-118]. During interference, a central channel forms between the two lobes to

accommodate the target DNA. Similar to Cas9, PAM recognition is required for Type V systems. PAM recognition is accomplished by a groove that forms between the WED, REC1, and PI domains [116-118]. After PAM recognition, the complementary strand binds to the crRNA and the non-complementary strand forms an R-loop. Formation of the R-loop creates the appropriate positioning to facilitate target DNA cleavage by the RuvC domain [116].

Invader targeting in Type VI systems occurs using Cas13. The effector complex consists of Cas13 and crRNA and takes on a bilobed architecture [119,120]. Similar to Type II and V systems, Cas13 is composed of a REC and a NUC lobe. However, Cas13 diverges from its counterparts due to its ssRNA degradation activity rather than dsDNA cleavage activity (Figure 1.9) [85,119-122]. The REC lobe contains a conserved Helical-1 domain and the NUC lobe contains two conserved HEPN domains [119,120]. After processing the pre-crRNA to produce the functional crRNA, Cas13 remains bound to the crRNA. This results in a conformation change that is stabilized by the presence of the crRNA and activates Cas13 for target RNA cleavage [119,120]. The secondary stem-loop structure of the 5' repeat of the crRNA is essential for target RNA cleavage [120,121]. It then scans RNA transcripts for the presence of a protospacer flanking site (PFS) on either the 5' or 3' end of the target RNA [121,122]. The crRNA-induced conformational change activates target RNA cleavage due to the positioning of HEPN1 closer to HEPN2 [119]. Unlike Type III HEPN domain containing proteins Csm6 and Csx1, the ribonuclease activity of Cas13 is divalent metal-dependent [85,109,110,122]. Once activated by cognate target RNA, Cas13 complexes exhibit promiscuous ribonuclease activity and are able to then cleave non-cognate (bases do not pair with the crRNA bases) target RNA

[85,119-122]. Interestingly, Cas13 complexes from subtype VI-B was found to contain two accessory proteins Csx27 and Csx28 that either repress or enhance, respectively, Cas13 activity [122]. Figure 1.10 provides a summary of the building blocks of CRISPR-Cas systems.

Type III Systems: Effector Complex Structure

The effector complexes of Type III systems are multi-subunit structures (Figure 1.8). Csm and Cmr crRNPs form seahorse-like structures and are composed of Csm1-5 or Cmr1-6 and a mature crRNA species [88,89,95,123-125]. The assorted Cas proteins within the complexes possess structural ferredoxin-like folds containing RNA Recognition Motifs. The bodies of the complexes are composed of two intertwined filaments. The major filament is composed of multiple copies of Csm3 or Cmr4. The minor filament is composed of multiple copies of either Csm2 or Cmr5. Single copies of Csm1 and Csm4 or Cmr2 and Cmr3 form the tail of the structure, and single copies of Csm5 or Cmr1 and Cmr6 form the head of the structure. The mature crRNA species are nestled in the complexes and held in place by the major filament (Csm3-Csm4 or Cmr3-Cmr4) [90,95,96,124-126]. Type III effector complexes are large complexes. The complexes range in size from ~350 kDa in the organisms *S. thermophilus*, *Thermus thermophilus,* and *Staphylococcus epidermidis* to ~450 kDa in the organism *Saccharolobus solfataricus* [89,95,96,124,125,127,128]. The complexes are large in size due to the presence of five (for Csm) or six (for Cmr) different proteins present within the complex in addition to multiple copies being present of Csm2, Csm3, Cmr4, and Cmr5 within the complexes. The size differences between complexes is driven by the number of Cas7 proteins

present in the backbone of the particular complex which is driven by length of the crRNA [88,89]. The mature crRNAs are structurally highly conserved and very similar in sequence. The sizes of crRNAs range in size from 30 to 45 nucleotides long and maintain a conserved structure composed of two regions [18,89,127,129]. For Type III systems, mature crRNAs are composed of a 8 nt repeat-derived 5' tag region and a spacer derived region on the 3' end. Endonucleolytic cleavage by Cas6 results in a crRNA containing a 5' hydroxyl group and a 3' phosphate group [15]. Creation of the 5' tag with a 5' hydroxyl group is essential for complex formation and function [68]. The 5' tag sequences of prominent crRNA species has been found to be similar amongst organisms. For example, the tag 5'-AUUGAAAG-3' was observed in *S. solfataricus* and *P. furiosus* [123,125], 5'-ACGGAAAC-3' in *S. thermophilus* [126,130], and 5'-ACGAGAAC-3' in *L. lactis* and *S. epidermidis* [127,130]. Each component of the Csm/Cmr complex plays an important role in efficiently executing various aspects of Type III defense.

Binding of cognate target RNA to a Type III effector complex triggers three activities by the complex: RNA targeting, DNA targeting, and cOA production. Upon binding of target RNA to a Type III complex, a conformational change occurs in which the complex subunits rearrange to create a wider binding channel to accommodate the crRNA-target RNA duplex [96,101]. The wider channel formation is facilitated by Csm2/Cmr5 subunits by rotating away from the Csm3/Cmr4 subunits which exposes the crRNA while Csm1/Cmr2 subunit rotates towards the 5' end of the crRNA [96,101]. When the anti-tag of the target is capable of base-pairing with the 5' tag, Type III complexes essentially become locked into an inactive conformation that inhibits triggering of DNA targeting and cOA production activities [131]. The requirement of this conformational change for

nuclease and cOA production activities of Type III complexes but not the ribonuclease activity represents an intrinsic control mechanism for the complexes.

Type III Systems: RNA Targeting

The Csm3/Cmr4 mediated ribonuclease activity of Type III systems has been observed *in vivo* and *in vitro* [18,89,132,133]. RNA targeting in Type III systems is specific to ssRNA and is sequence-specific [88,89,124]. Sufficient base-pairing between the crRNA and target RNA is required for degradation by Csm3 or Cmr4.There is a direct correlation between the number of cleavage products generated by a given complex and the number of Csm3/Cmr4 subunits within a given complex. Multiple cleavage products are observed at 6 nt intervals when Type III complexes degrade RNA and is dependent upon the length of the crRNA [68,127]. As the length of the crRNA grows, more products will be observed due to the presence of an increased number of Csm3 or Cmr4 subunits within the complex. The 6 nt interval cleavage characteristic of Type III complexes is highly conserved [68,88-90]. This occurrence is the result of structural characteristics of the complex—specifically Csm3 and Cmr4. Both Cas7 variants form a right-hand shape including a palm, finger, and thumb [90,104]. The thumb of Csm3/Cmr4 is composed of β-hairpins that protrude from the structure; once multiple copies unite, they form a helical stack [90,104]. Upon binding of the target RNA to the crRNP, the thumbs of Csm3/Cmr4 protrude into the RNA-RNA duplex. This interaction results in the outward flipping of the base pair that interacts with the thumb. This structural interaction allows the target RNA to be cleaved and results in multiple products differing in size by 6 nts [88,134]. The ribonuclease activity of Csm3/Cmr4 is catalyzed by a highly conserved aspartic acid

residue [89,104,135]. The aspartic acid residue is critical for RNA cleavage due to the distortion of the phosphate backbone by the thumbs of Csm3/Cmr4. This results in the positioning of the scissile phosphates of the target RNA near the conserved aspartic acid [90]. Numerous *in vitro* studies across Type III systems have confirmed that the ribonuclease activity is divalent metal dependent and most commonly requires either Mg^{2+} or Mn^{2+} [18,88,89,124]. Proteins composing and stabilizing the minor filament of the effector crRNPs contribute to the ability of the complexes to bind target RNA [68,90]. Type III systems are very resilient to mutations within the region of complementarity between the crRNA and target RNA as several truncated versions of the target RNA was able to be sufficiently bound and cleaved by the Csm complex of *S. thermophilus* [89]. Furthermore, base-pairing between the 5' tag and the anti-tag region of the target RNA does not promote or inhibit RNA degradation behavior by Type III defense complexes [89,136].

Type III Systems: DNA Targeting

Type III CRISPR-Cas systems are often regarded as dual-targeting systems due to their intrinsic ability to degrade RNA and DNA. DNA cleavage is achieved via Csm1 or Cmr2. Csm1 and Cmr2 (also known as Cas10) are the signature genes of Type III systems [29,137]. The Cas10 family proteins contain five notable domains: an HD domain, two RNA recognition motif domains, a Linker domain, and a D4 domain [101]. The HD domain contains the catalytic residues for the protein's nuclease activity. Both RNA recognition motif domains contain Palm domains similar to Palm domains found in RNA polymerases, DNA polymerases, and nucleotide cyclases [105,137]. However, only

one Palm domain contains the conserved GGDD motif needed for cOA production. The Linker domain contains a zinc finger and the D4 domain is structurally important for interacting with Csm2/Cmr5 [101]. Within the HD domain, mutation of either the histidine or aspartic acid residue, or both, has inhibited nuclease activity of Cas10 proteins across several different organisms [99,100,128,138,139]. In most Type III systems, DNA cleavage occurs via the HD domain and cyclic oligoadenylate production occurs via the Palm domain [102,103]. However, exceptions to this generality have been observed. For instance, in the Type III-A system of *S. epidermidis*, sole mutation of the Palm domain *in vivo* and *in vitro* was found to halt immunity against plasmids and DNA nuclease activity [98,140]. In the heterologously expressed Type III-A systems of *L. lactis*, *S. epidermidis, and S. thermophilus*, sole mutation of the Palm domain inhibited immunity against plasmids *in vivo* while mutation of the HD domain had no effect [110]. Cas10 itself is a single-stranded DNA specific endonuclease [99,105]. Most Type III complexes appear to only degrade ssDNA or dsDNA containing a small engineered bubble via mismatched base-pairs [99,100,128]; however, evidence from the *P. furiosus* Type III-B system demonstrates the ability to either cleave the dsDNA (minus an artificial bubble) or separate strands of the dsDNA substrate to produce ssDNA regions via a PFS located on the target RNA (originally called the rPAM) [99]. Further scientific exploration of more Type III systems will be needed to determine the importance of specific PFS at the 3' end of the target RNA in various Type III systems. Unlike RNA targeting, DNA targeting has been found to be nonspecific and complementarity between the crRNA and the target DNA is not required [98-100]. Nucleotide specificity has not been rigorously explored across organisms but some thymine specificity has been observed in the organism *T.*

maritima [138]. However, all organisms currently studied have found a requirement for a divalent metal such as Ni^{2+}, Co^{2+}, Mg^{2+} or Mn^{2+} [98-100].

Another fascinating aspect of DNA targeting in Type III systems pertains to the requirement of transcription of the target DNA for nuclease activity of Cas10 to be activated [98,129,141]. Transcription of the target DNA to produce a target RNA with adequate complementarity to the crRNA is required to activate nuclease activity of the complexes, but once activated the complexes can cleave DNA of any sequence. Though the CRISPR locus itself is protected by sequence recognition via base-pairing between the 5' tag and CRISPR transcript. This characteristic has been widely observed across several organisms [98-101,126,128,138,139,142,143]; although one study conducted in *T. onnurineus* was able to reconstitute *in vitro* DNA nuclease activity by the Csm complex in the absence of target RNA [144]. The discovery of the transcription requirement has led to many subsequent discoveries about the mechanisms involved in DNA degradation. For instance, it is now known that binding rather than cleavage of the target RNA is required [98,100]. The sequence of the PFS or target adjacent region on the target RNA opposite the 5' tag region, on the crRNA, is important. DNA degradation by Type III effector complexes will typically occur with the simplest requirement of the target adjacent sequence on the target RNA is the lack of base pairing with the 5' tag [99,100,139].

Structural insights into the Type III-A complex of *S. thermophilus* revealed that the 3' region of the target RNA opposite the 5' tag on crRNA is disordered when the sequence is non-complementary to the 5' tag and forms an ordered duplex with the 5' tag when the sequence is complementary [101]. Once bound to the complex, the 8 nucleotides of the PFS of the target RNA are positioned within a positively charged

channel formed by the Linker domain and a L1 loop (within the Palm domain) of Cas10. Mutation of residues within the Linker were found to inhibit DNA nuclease and cOA production activity of the complex [101]. This structural evidence confirms that interactions between the Cas10 Linker region and PFS are crucial for DNA targeting and cOA production; these interactions are facilitated by the lack of base pairing between the 5' tag and anti-tag, at least in the *S. thermophilus* Type III-A system investigated [100].

Type III Systems: Activation of Csm6 and Csx1 via cOA Production

The third known enzymatic function of Type III complexes is production of cyclic oligoadenylate or cOA compounds. Molecules of cOA are produced by the Type III complexes via the Palm domain of Cas10 [102,103,112]. Cyclic oligoadenylate molecules consist of multiple monomers of adenosine monophosphate (AMP) assembled into a cyclic configuration. Depending on the organism of interest, the size of the rings may vary from 2 to 6 units of AMP [102,103,112,145-148]. Table 1.2 provides a summary of known cOA production for various Type III systems. Currently it is known that production of cOAs is dependent upon the Palm domain of Cas10, cognate target RNA, and the presence of ATP and Me^{2+} [102,103]. ATP is needed for the conversion and a divalent metal such as Mg^{2+} is needed for target RNA binding. A recently solved structure for the Type III-A system of *Thermococcus onnurineus* (*Ton*) has provided additional insight into the structural aspects of cOA production [147]. Using the ATP analog AMPPNP, it was found that the 3'-OH of AMPPNP forms hydrogen bonds with one of the highly conserved aspartic acid residues within the GGDD motif. The AMPPNP molecule is further stabilized by interactions between the α-phosphate of AMPPNP and divalent metals (Mn^{2+}) [147].

Previously, it was found that the cyclase activity of Cas10 is highly specific to ATP and when compared to other nucleoside triphosphates [102]. Analysis conducted in *T. onnurineus* reveals the structural basis for the preference of ATP by Cas10; conserved serine residues within the GGDD motif of Cas10 specifically recognize the adenosine residues [147].

The production of cOA molecules is well regulated by Type III systems at the target RNA level. The sequence of the PFS region of the target RNA as well as binding of the entire target RNA are critical components in the production of cyclic oligoadenylates. Creation of cOA molecules is halted by cleavage of the target RNA by Csm3 or Cmr4 [112,148]. Production of cOA molecules is a key component of the Type III defense response. Once generated by the Palm domain of Cas10, these molecules serve as allosteric activators for Csm6 and Csx1 by binding to the CARF domain of the protein and activating the distal HEPN ribonuclease active site [102,103,149,150]. Csm6 and Csx1 are rather inefficient enzymes without activation by cOA; thus, this is an important step in the defense response.

Recently, the structure of *S. islandicus* (*Sis*) Csx1 bound to cOA_4 was solved [150]. Interestingly, this homolog forms a hexamer structure. Three Csx1 dimers oligomerize via the HEPN domain to create the hexamer of about 300 kDa in size. The solved structure confirmed well established theories about the Csx1-cOA interaction such as dimerization being critical for cOA binding and the wHTH domain serving as the message transmitter between the CARF and HEPN domains [150]. Loops formed within the CARF domain are important for cOA binding. Upon cOA binding, a conformational change occurs within the loops. The loops are reshaped from an oval cleft to a cruciform orientation [150]. This

change converts the electrostatic potential in the Csx1 hexamer from neutral to polar. The binding of cOA_4 to the CARF domain of *Sis* Csx1 is stabilized by a combination of polar and hydrophobic interactions in the cruciform binding pocket [150]. Csm6 proteins do not contain a wHTH domain but they do contain a 6H domain [108]. Thus, it is possible that signals are transmitted from the CARF domain to the HEPN domain, in Csm6, via the 6H domain. More structural analysis is needed to assess the impact of this difference.

The structure of *Thermococcus onnurineus* Csm6 bound to cOA_4 was recently solved [149]. This structure is notably different from that of *S. islandicus* Csx1 because *Ton* Csm6 is capable of degrading cOA molecules while *Sis* Csx1 apparently is not. Thus, cOA molecules were only found bound to the CARF domain of *Sis* Csx1 [150]. Alternatively, cOA molecules were observed within the CARF and HEPN domains of *Ton* Csm6 [149]. In the *T. onnurineus* structure, the cOA adenosines are recognized by hydrogen bonding with tryptophan and tyrosine residues in the CARF domain. Hoogsteen edges of the adenosines bind to the tyrosine residue and the 5'-phosphates of the adenosines bind to the tryptophan residue [149]. In the HEPN domain, the adenosines of cOA_4 are recognized by residues of the HEPN motif. Specifically, the highly conserved histidine residue of the HEPN motif is positioned in a manner that facilitates the nucleophilic attack of the 2'-OH of one adenosine on the 5'-phosphate of another nearby adenosine [149]. A combination of local, rather than global, conformational changes contribute to the binding and degradation of cOA_4 by the CARF and HEPN domains.

Type III Systems: Control of the Defense Response

Studying self versus non-self recognition is a major area of exploration in the study of CRISPR-Cas immunity especially for the adaptation and invader silencing stages. Currently, it is known that non-self recognition for the defense stage of Types I, II, and V DNA-targeting CRISPR systems is executed by PAM recognition [86,113,115]. PAM-directed defense has not been observed for Type III or Type VI systems, that rely on invader RNA recognition. Alternatively, non-self recognition in Type III systems occurs through a complicated sequence specific mechanism that is still being understood. However, the simplest scenario is via lack of base pairing between certain nucleotides in the 5' tag of the crRNA and PFS of the target RNA [151]. The lack of base pairing between the 5' tag and PFS at certain positions results in crucial structural events that activate the complete defense response. Recognition of an appropriate PFS induces a conformational change that allows Type III complexes to trigger their full range of defense activities [101]. Since DNA degradation by Csm1/Cmr2 and RNA degradation by Csm6/Csx1 is presumed to be nonspecific *in vivo*, activation of those activities requires stricter control via the nature of the PFS. Lack of activation of the defense response occurs if a target RNA with an inadequate PFS binds to the Csm/Cmr crRNP. This feature has been supported in subsequent studies from various organisms [98,100,102,128,139]. A variation of this requirement was observed in *P. furiosus* with the discovery of 3 bases within the PFS of the target RNA immediately adjacent to the region of crRNA complementarity that show different abilities to support *in vivo* defense and *in vitro* DNAse activities of the complex. [99]. Since the presence and sequence of the PFS is irrelevant

for RNA targeting and RNA targeting is highly specific, the RNA targeting of the complex will be controlled by the presence or absence of transcription of the DNA target.

Dissertation Overview

The subsequent chapters of this thesis detail the scientific studies and contributions that I have made to the Type III field of CRISPR-Cas defense. Chapter 2 provides a report of the exploration efforts that I conducted using the Type III-A (Csm) CRISPR-Cas systems of *L. lactis, S. epidermidis,* and *S. thermophilus*. Chapter 3 focuses on the Type III-B system of *P. furiosus*. In that chapter, the role of Csx1 and actions employed by the Cmr targeting complex are investigated.

L. lactis is a Gram-positive bacterium that ferments lactose to lactic acid during the production of ATP. It has a wide range of applications in industry, but is predominantly an important microorganism in the food industry as it is used in the production of cheese, yogurt, and buttermilk. The lactic acid produced by *L. lactis* is also used as a moisturizer and emulsifier in the cosmetic industry and as a raw material in the pharmaceutical industry [152]. The genome of *L. lactis* does not harbor CRISPR-Cas systems, but an industrial strain was found to contain Type III-A genes on a plasmid (pKLM) that was discovered during a search for novel phage resistance mechanisms [11]. *S. epidermidis* is a gram positive bacterium that is most commonly found on human skin as a part of the epithelial microflora. This microorganism is a facultative anaerobe and is historically regarded as an innocuous commensal organism. However, emerging evaluation of *S. epidermidis* classifies it as an opportunistic pathogen due to its prevalence as the causative agent of medical device infections such as with peripheral or central

intravenous catheters [153]. *S. epidermidis* harbors widespread antibiotic resistance genes that impact public health. For instance, 75 to 90% of all hospital isolates were found to be resistant to methicillin which is the primary treatment option against staphylococcal infections [153]. The RP62A strain of *S. epidermidis* contains one Type III-A system and with two different CRISPR loci within its genome. *S. thermophilus* is a gram-positive fermentative facultative anaerobe. Similarly to *L. lactis*, it is a key organism in the food industry due to its ability to produce lactic acid and is often used in the production of cheese and yogurt. *S. thermophilus* is also widely used commercially in the probiotic industry. From an applications perspective, *S. thermophilus* has become a useful genetic tool due to its ability to be utilized as a heterologous expression host [154]. The DGCC7710 strain of *S. thermophilus* contains three CRISPR-Cas systems (I-E, II-A, and III-A) and four CRISPR loci. In chapter 2, the primary question of interest is whether the Csm6 homologs of *L. lactis*, *S.epidermidis*, and *S. thermophilus* are essential for CRISPR-Cas immunity. Analysis was conducted on recombinant Csm6 proteins to determine the enzymatic activity of the proteins. Ribonuclease activity was observed *in vitro* for 2 out of 3 of the homologs (Chapter 2). The ribonuclease activity was further characterized by determining several features of the activity including metal dependency, substrate specificity, and nucleotide specificity. All three Type III-A systems were expressed heterologously in *E. coli*. For all three systems, the ribonuclease activity of Csm6 was found to be essential and sufficient for CRISPR-Cas immunity.

P. furiosus is an obligate anerobic archaeon and a hyperthermophile due to its optimal growth temperature of 100 °C. The thermostable nature of its proteins makes it a useful source for thermostable enzymes for biotechnological applications. Due to pivotal

genetic research conducted by laboratory groups at the University of Georgia and North Carolina State University, *P. furiosus* is now a viable tool for a variety of metabolic engineering purposes [155]. The genome of *P. furiosus* contains three CRISPR-Cas systems (I-A, I-G, and III-B) and seven CRISPR loci. Chapter 3 details the investigation of the Cmr system of *P. furiosus*. I sought to determine the impact of Csx1 on CRISPR-Cas immunity for the Type III-B system of *P. furiosus*. Furthermore, analysis was conducted to reconstitute and characterize the recently discovered action of cyclic oligoadenylate production in the *P. furiosus* system. Not only was Csx1 found to be an important component of defense, contrary to previous reports, but it was found to possess self-regulation capabilities due to its ability to degrade and inactivate its allosteric activator (Chapter 3). The analysis provides additional insight into the target RNA requirements for the cOA production activity of the Cmr complex. Overall, this dissertation provides valuable insight into the functional activities and mechanistic requirements for successful CRISPR-Cas immunity for four different Type III systems.

Figure 1.1. CRISPR-Cas Components. The CRISPR-Cas system is composed of two main components: the CRISPR locus and the set genes (gray) associated with the CRISPR locus. Identical repeat sequences (black) are interspaced between unique invader derived spacer sequences (assorted colors). The leader region is AT-rich and contains the promoter for transcription of the CRISPR-locus and elements that drive spacer integration.

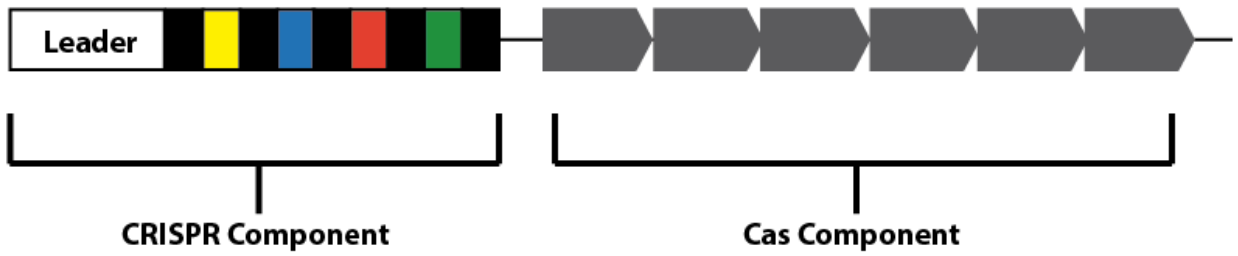


Figure 1.2. CRISPR-Cas Immunity Pathway. CRISPR-Cas immunity occurs in three main stages: adaptation, crRNA biogenesis, and invader silencing (defense). During adaptation, a fragment of invader DNA is incorporated into the CRISPR locus. The CRISPR locus is then transcribed and processed to produce mature crRNAs during crRNA biogenesis. In the final stage of the pathway, Cas proteins combine with mature crRNA to form effector crRNPs and destroy the invading nucleic acid. (Adapted from Terns & Terns, Trends in Genetics, 2014 [156])

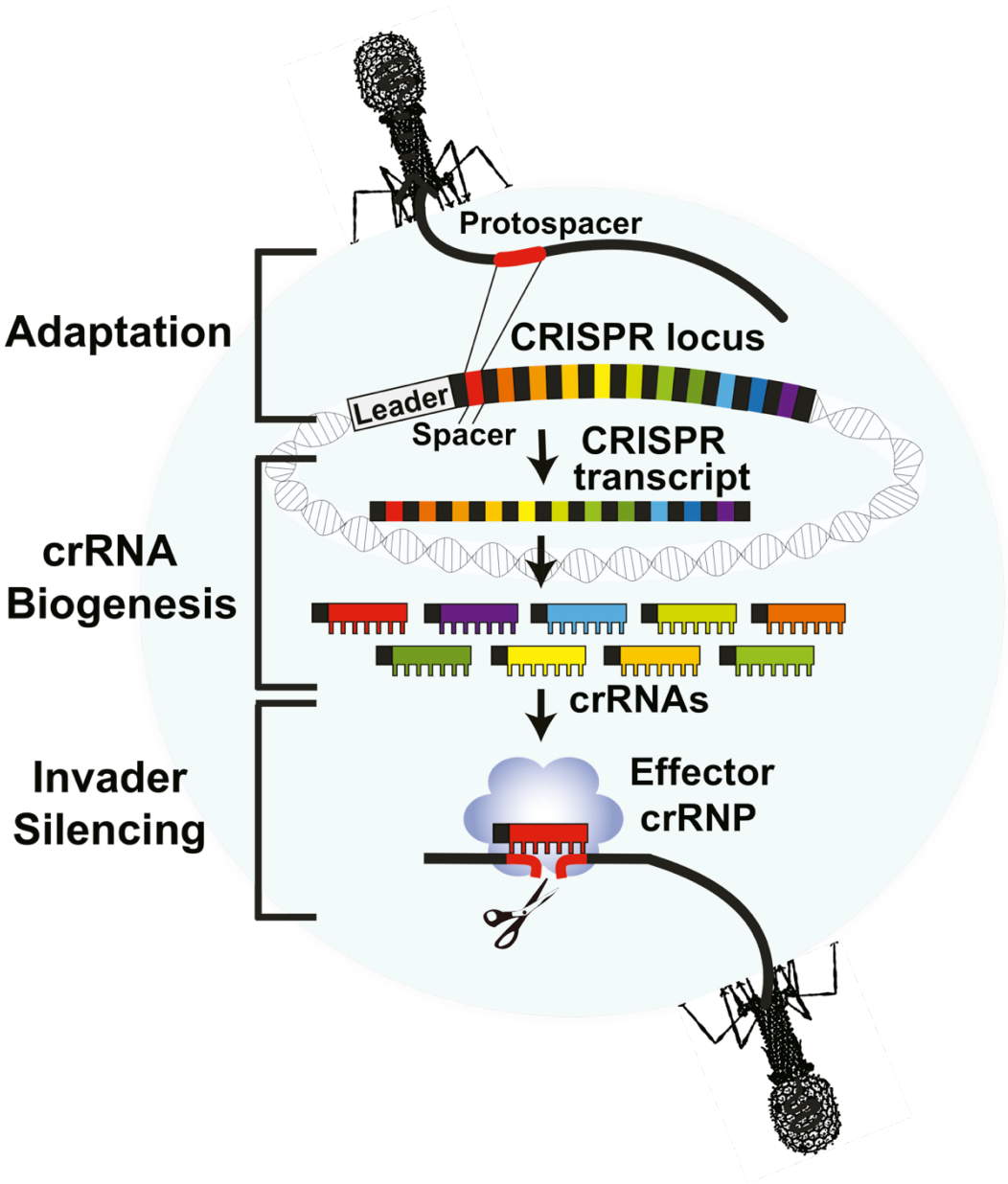

Protospacer
Adaptation
CRISPR locus
Leader
Spacer
CRISPR
transcript
crRNA
Biogenesis
crRNAs
Invader
Silencing
Effector
crRNP

Table 1.1. Classification of CRISPR-Cas Systems

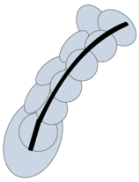
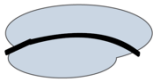
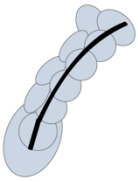
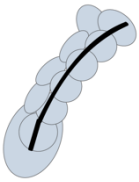
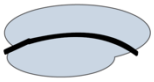
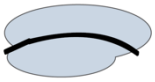
	Type I	Type II	Type III	Type IV	Type V	Type VI
Signature Protein	Cas3	Cas9	Cas10	__	Cas12	Cas13
Class	One 	Two 	One 	One 	Two 	Two 
Subtypes	I-A(Csa) I-B I-C(Csd) I-D(Csc) I-E(Cse) I-F(Csy) I-G(Cst)	II-A(Csn) II-B(Csn) II-C(Csn)	III-A(Csm) III-B(Cmr) III-C(Cmr) III-D(Csm)	IV(Csf)	V-A(Cpf) V-B V-C V-D V-E V-U1 V-U2 V-U3 V-U4 V-U5	VI-A VI-C VI-B1 VI-B2
Effector crRNP Targets	DNA	DNA	DNA RNA	__	DNA	RNA
Notable Domains	HD, RRM	RuvC, HNH	HD, Palm, RRM	__	RuvC	HEPN

Figure 1.3. Main Events of Adaptation. Successful adaptation occurs in a multi-step process. The adaptation complex (Cas1-Cas2) assembles in the presence of invader DNA. Invader DNA is processed to generated potential protospacers via degradation by host proteins and additional Cas proteins. The protospacer with the appropriate PAM is selected by the adaptation complex for integration. In the final step of the process, the spacer is inserted into the CRISPR locus at the leader-repeat end of the array.

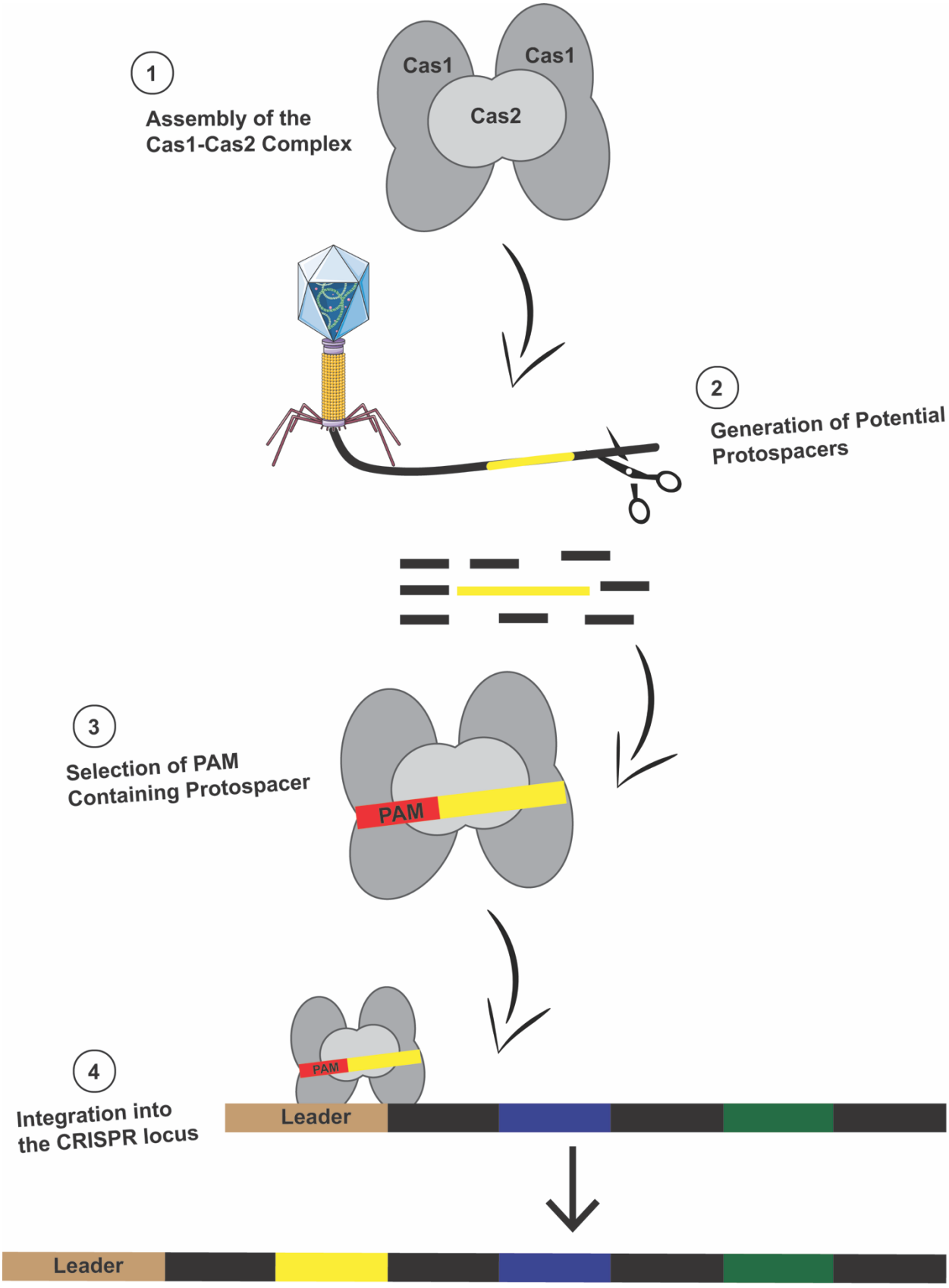

1
Assembly of the
Cas1-Cas2 Complex
Cas1
Cas1
Cas2
2
Generation of Potential
Protospacers
3
Selection of PAM
Containing Protospacer
PAM
4
Integration into
the CRISPR locus
PAM
Leader
Leader

Figure 1.4. Modes of protospacer generation for adaptation. Protospacer substrates utilized by adaptation complexes can be generated by multiple cellular processes and phenomena including stalled replication forks, activities of restriction endonucleases, and activities of RecBCD complexes. These cellular occurrences result in the production of dsDNA breaks that contain PAM sequences that are recognized by the adaptation complex and integrated into CRISPR loci. (Adapted from Jackson et al, Science, 2017 [35]).

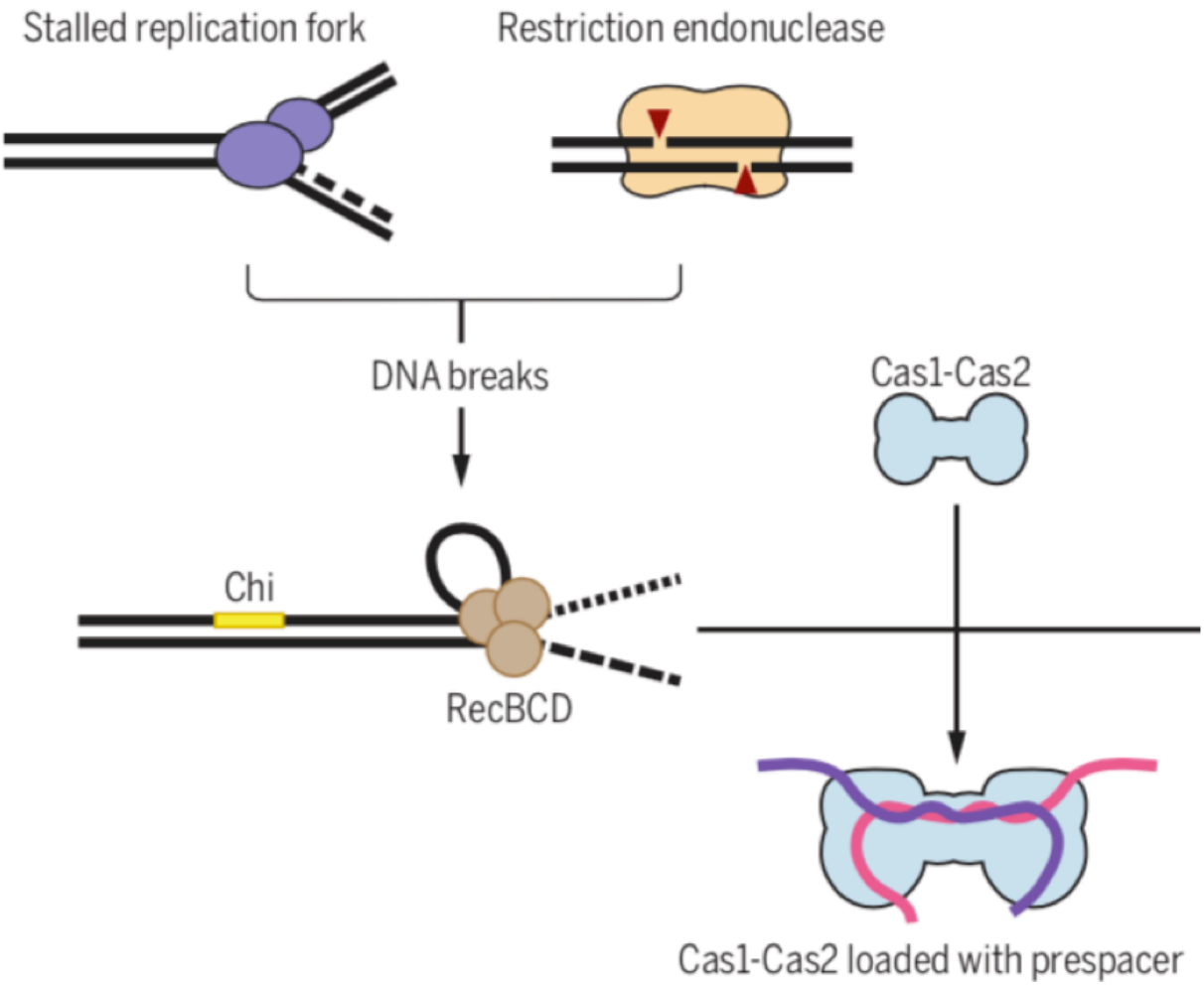

Stalled replication fork
Restriction endonuclease
DNA breaks
Cas1-Cas2
Chi
RecBCD
Cas1-Cas2 loaded with prespacer

Figure 1.5. Integration of spacers into a CRISPR locus. Newly acquired spacers are incorporated into CRISPR loci via a two-step transesterification reaction. The first nucleophilic attack occurs between the top strand of the spacer and the bottom strand of the first repeat within the array. A second attack occurs between the bottom strand of the spacer and the top strand of the repeat. Remaining gaps are filled in by DNA polymerase and DNA ligase. (Adapted from Sternberg et al, Molecular Cell, 2016 [67]).

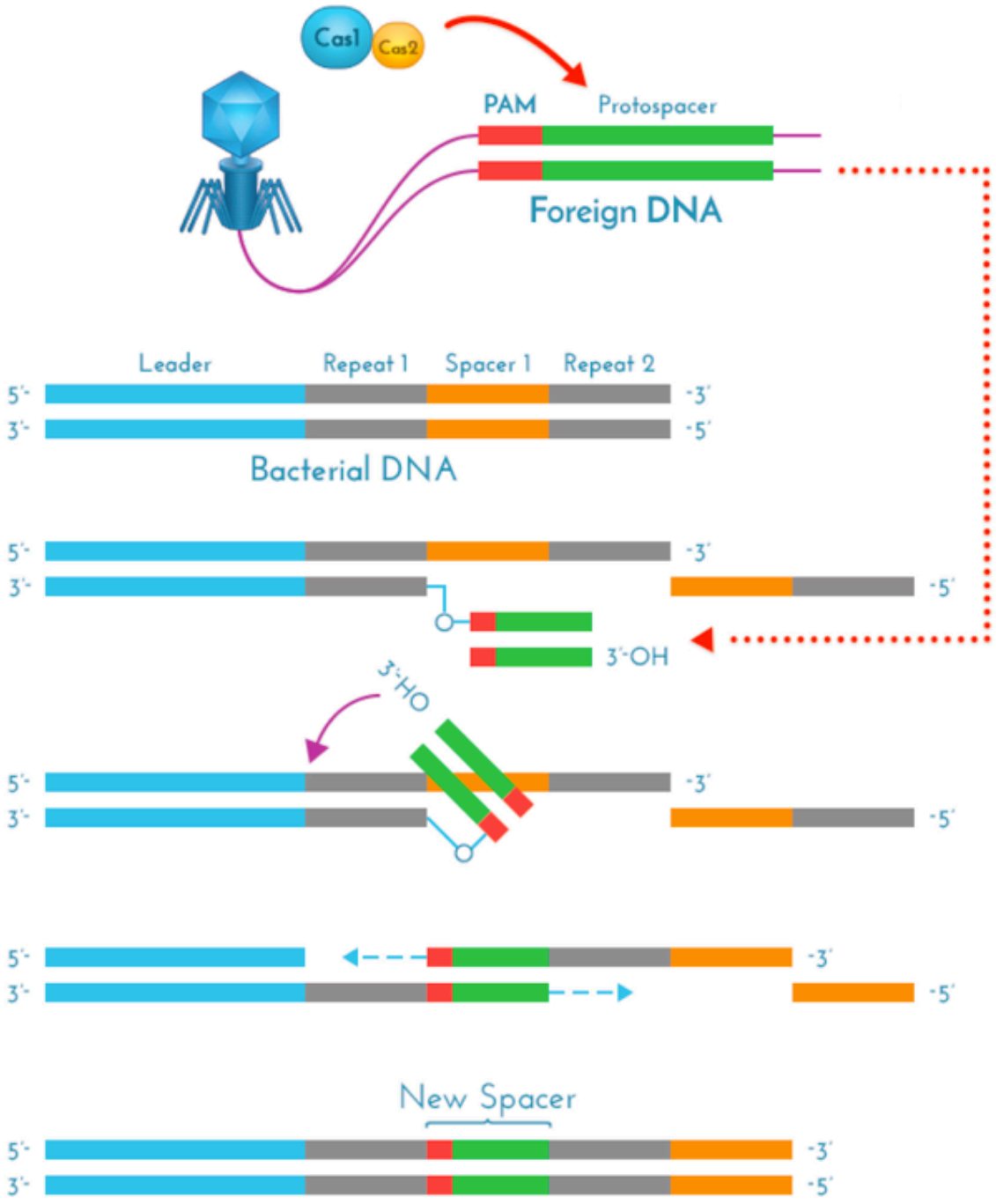

Cas1
Cas2
PAM
Protospacer
Foreign DNA
Leader
Repeat 1
Spacer 1
Repeat 2
5'-
-3'
3'-
-5'
Bacterial DNA
3'-OH
3'-HO
New Spacer

Figure 1.6. Summary of crRNA composition across CRISPR-Cas systems. Functional crRNAs are composed of repeat and spacer sequence regions**.** The structure of a functional crRNA species differs between CRISPR-Cas systems. Repeat regions are depicted in black and spacer regions are depicted in orange. All illustrations of crRNAs are shown in a 5'-3' orientation.

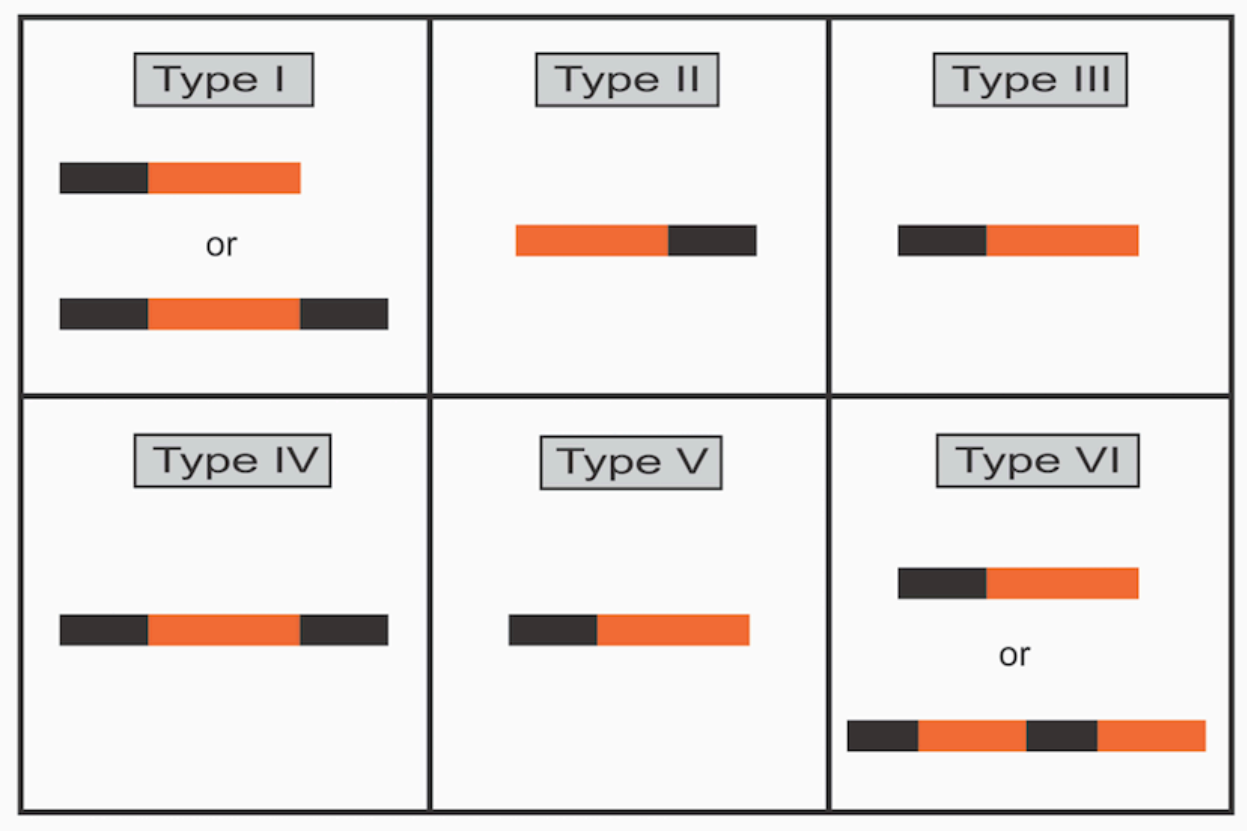


Figure 1.7. Actions involved in invader silencing for Class I systems. Invader silencing in Class I systems occurs with multi-subunit complexes**.** Successful invader silencing in Type I systems involves three prominent activities: PAM recognition, recruitment of Cas3 to Cascade, and unwinding and degradation of the DNA target by Cas3. For Type III systems, the Csm/Cmr complexes are activated upon transcription of the target RNA and base-pairing between the crRNA and target RNA. This interaction triggers four subsequent actions including degradation of the target RNA by Csm3/Cmr4, production of cyclic oligoadenylate compounds, activation of the ribonuclease activity of Csm6/Csx1, and cleavage of the target DNA by Csm1/Cmr2. (Adapted from Hille et al, Cell, 2018 [84]).

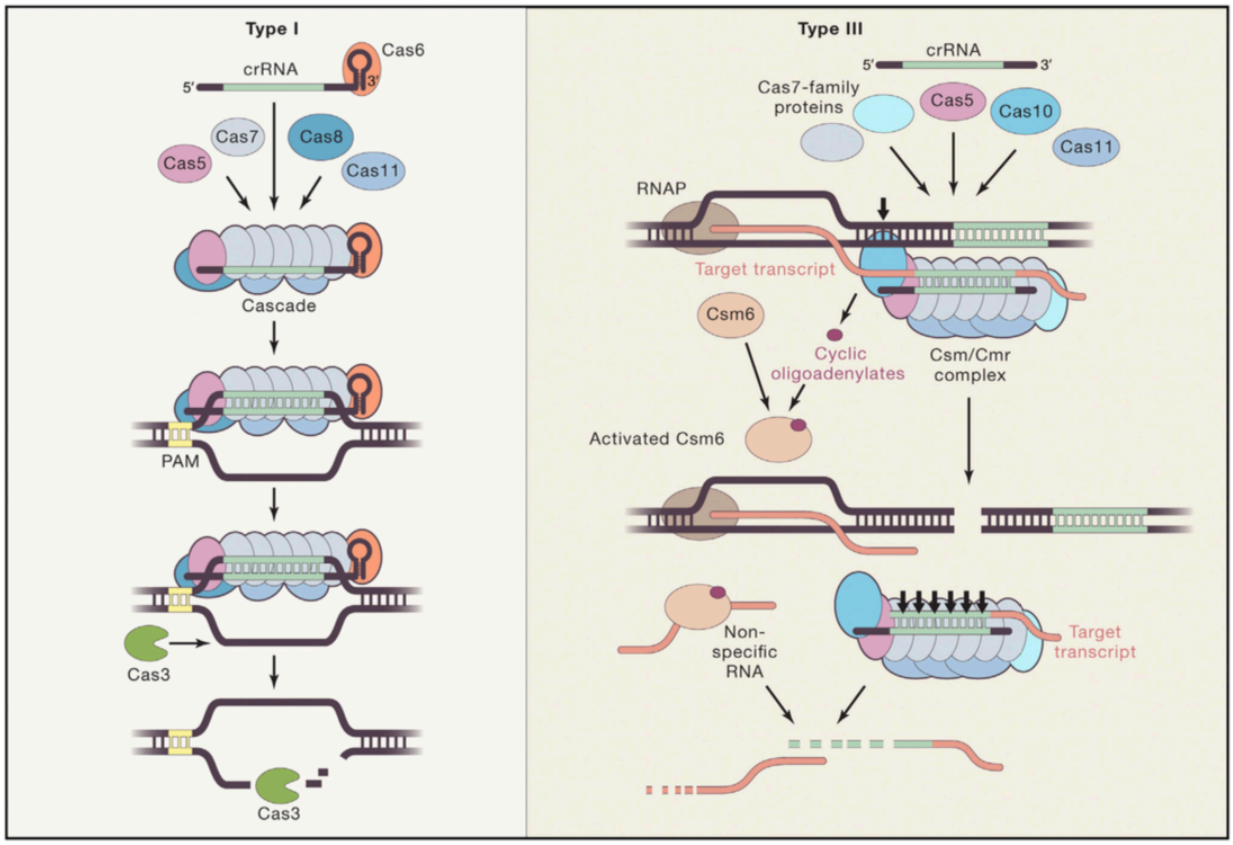


Figure 1.8. Composition of Type III effector complexes. Type III effector complexes are known as Csm or Cmr complexes. Csm complexes are composed of Csm1-5 and a mature crRNA. Cmr complexes are composed of Cmr 1-6 and a mature crRNA. Ribonuclease activity within the complexes are performed by Csm3 or Cmr4. Nuclease and cyclic oligoadenylate production activities are performed by Csm1 or Cmr2 via the HD and Palm domain, respectively.

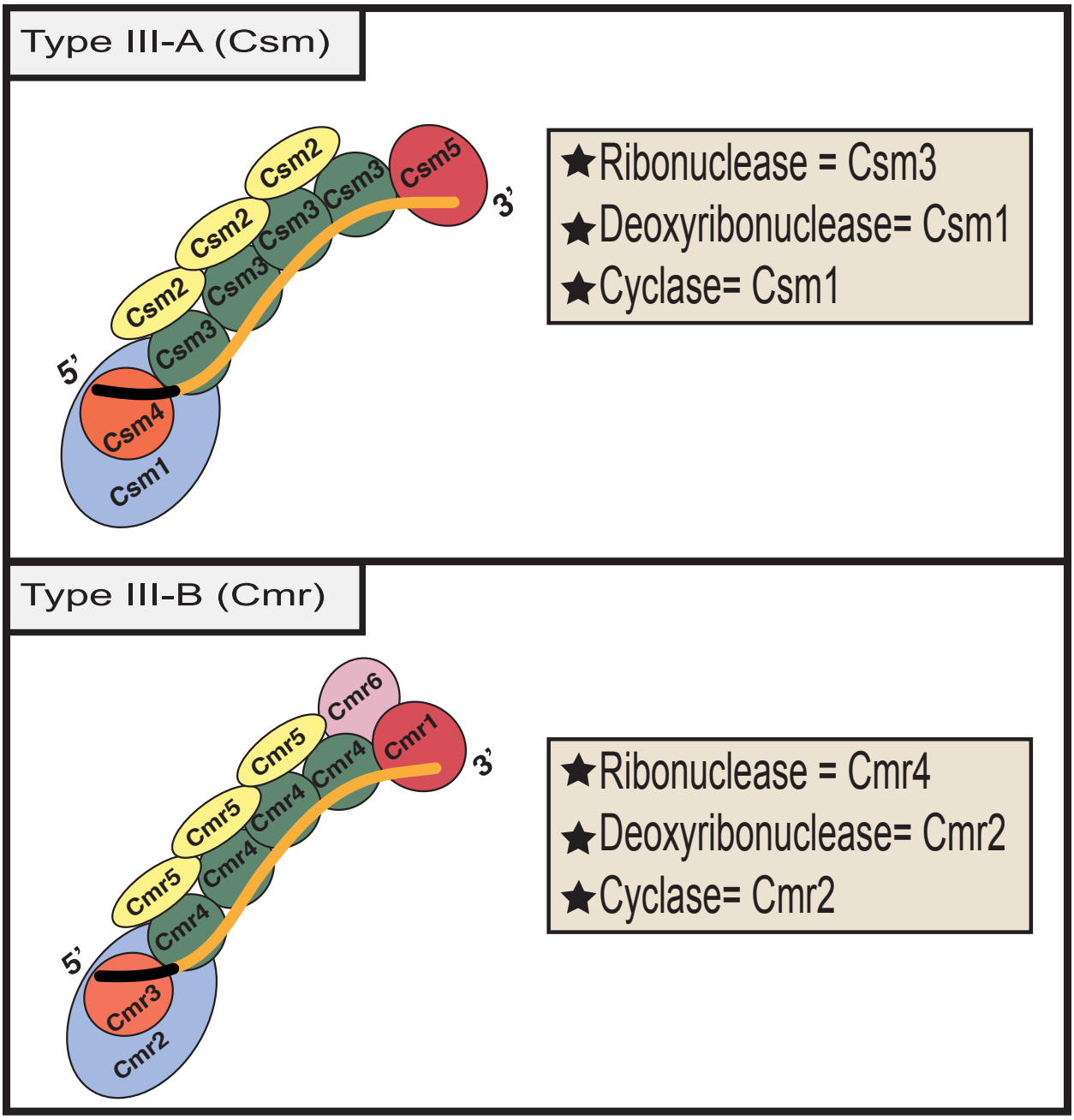

Type III-A (Csm)
Csm2
Csm3
Csm5
3'
Csm2
Csm3
Csm2
Csm3
Csm3
5'
Csm4
Csm1
★Ribonuclease = Csm3
★Deoxyribonuclease= Csm1
★Cyclase= Csm1
Type III-B (Cmr)
Cmr6
Cmr5
Cmr4
Cmr1
3'
Cmr5
Cmr4
Cmr5
Cmr4
Cmr5
Cmr4
5'
Cmr3
Cmr2
★Ribonuclease = Cmr4
★Deoxyribonuclease= Cmr2
★Cyclase= Cmr2

Figure 1.9. Actions involved in invader silencing for Class II systems. Invader silencing in Class II systems occurs via single-subunit targeting complexes. In Type II systems, Cas9 associates with tracrRNA and crRNA in its active state. Upon activation of the Cas9 targeting complex, DNA is scanned until an appropriate PAM is recognized. This leads to unwinding of the DNA and R-loop formation which triggers blunt end cleavage by Cas9. Type V systems utilize a Cas12 complex to silence invader DNA. PAM recognition by the complex results in the formation of a R-loop which prompts cleavage of the target DNA by Cas12. Invader silencing in Type VI systems occurs via the Cas13 complex. Recognition of the protospacer flanking sequence (PFS) on the target RNA activates collateral ribonuclease by Cas13. (Adapted from Hile et al, Cell, 2018 [84]).

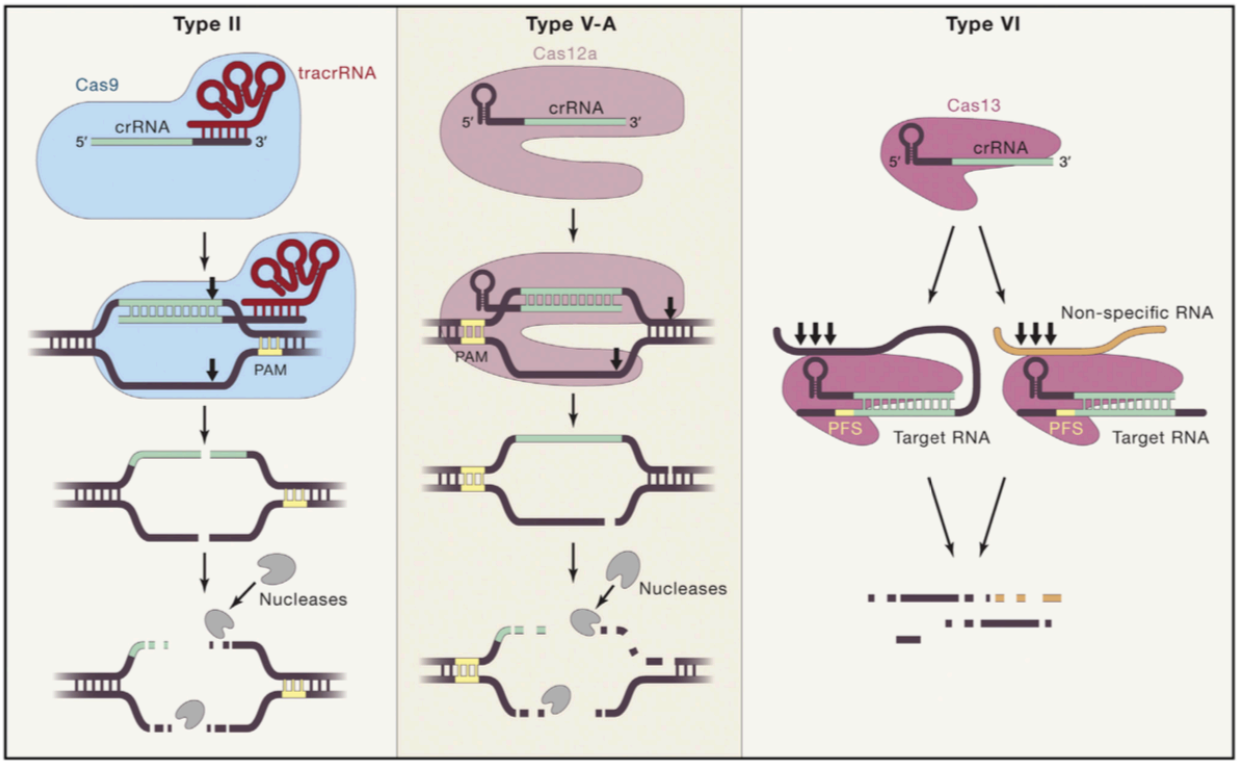


Figure 1.10. Modular organization of CRISPR-Cas systems. Summary of the building blocks of CRISPR-Cas systems. (Adapted from Koonin et al, Phil. Trans. R. Soc. B, 2019 [137]).

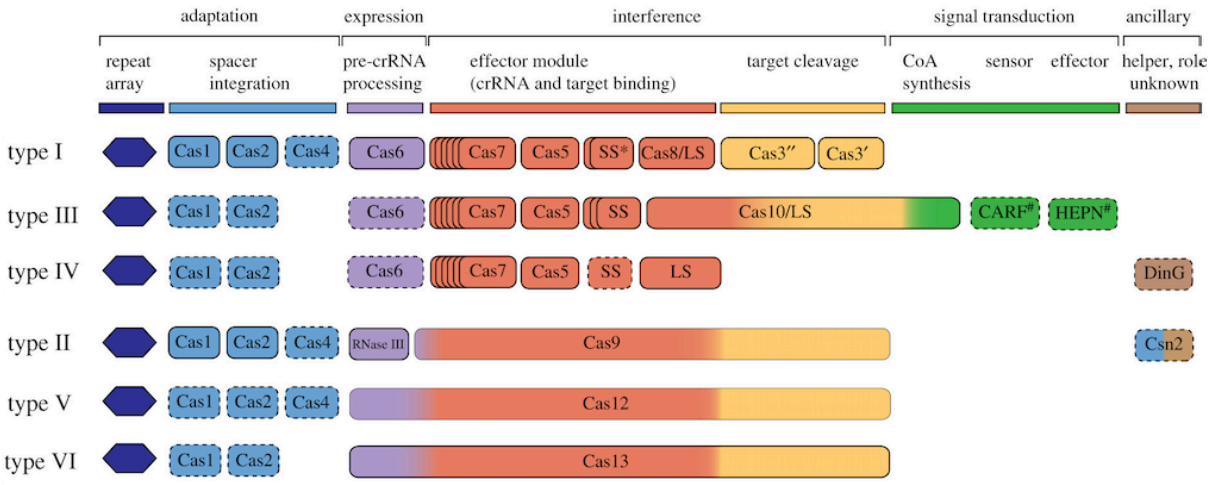


Table 1.2. Summary of cOA production for known Type III systems

Organism	cOA Species Produced	Relevant cOA Species	Reference
E. italicus	cOA_6	cOA_6	[103]
M. tuberculosis	cOA_3 cOA_4 cOA_5 cOA_6	cOA_6	[145]
P. furiosus	cOA_3 cOA_4 cOA_5 cOA_6	cOA_4	Chapter 3
S. epidermidis	cOA_3 cOA_4 cOA_5 cOA_6	unknown	[148]
S. islandicus	cOA_3 cOA_4	cOA_4	[146]
S. solfataricus	cOA_3 cOA_4 cOA_5	cOA_4	[112]
S. thermophilus	cOA_2 cOA_3 cOA_4 cOA_5 cOA_6	cOA_6	[102]
T. onnurineus	cOA_3 cOA_4	cOA_4	[147]

REFERENCES

1. Bikard, D., and Marraffini, L. A. (2012) Innate and adaptive immunity in bacteria: mechanisms of programmed genetic variation to fight bacteriophages. *Curr Opin Immunol* **24**, 15-20
2. Bernheim, A., and Sorek, R. (2018) Viruses cooperate to defeat bacteria. *Nature* **559**, 482-484
3. Chopin, M. C., Chopin, A., and Bidnenko, E. (2005) Phage abortive infection in lactococci: variations on a theme. *Curr Opin Microbiol* **8**, 473-479
4. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712
5. Stern, A., and Sorek, R. (2011) The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43-51
6. Cook, G. M., Robson, J. R., Frampton, R. A., McKenzie, J., Przybilski, R., Fineran, P. C., and Arcus, V. L. (2013) Ribonucleases in bacterial toxin-antitoxin systems. *Biochim Biophys Acta* **1829**, 523-531
7. Goeders, N., Chai, R., Chen, B., Day, A., and Salmond, G. P. (2016) Structure, Evolution, and Functions of Bacterial Type III Toxin-Antitoxin Systems. *Toxins (Basel)* **8**
8. Grissa, I., Vergnaud, G., and Pourcel, C. (2007) The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8**, 172

9. Barrangou, R., and Marraffini, L. A. (2014) CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* **54**, 234-244

10. Barrangou, R., and Oost, J. v. d. (2013) RNA-mediated Adaptive Immunity in Bacteria and Archaea. in *CRISPR-Cas Systems*, Springer Verlag, Berlin-Heidelberg. pp

11. Millen, A. M., Horvath, P., Boyaval, P., and Romero, D. A. (2012) Mobile CRISPR/Cas-mediated bacteriophage resistance in Lactococcus lactis. *PLoS One* **7**, e51663

12. van der Oost, J., Westra, E. R., Jackson, R. N., and Wiedenheft, B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**, 479-492

13. Nunez, J. K., Harrington, L. B., Kranzusch, P. J., Engelman, A. N., and Doudna, J. A. (2015) Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature* **527**, 535-538

14. Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964

15. Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev* **22**, 3489-3496

16. Hale, C., Kleppe, K., Terns, R. M., and Terns, M. P. (2008) Prokaryotic silencing (psi)RNAs in Pyrococcus furiosus. *RNA* **14**, 2572-2579

17. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355-1358

18. Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M., and Terns, M. P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945-956

19. Westra, E. R., van Erp, P. B., Kunne, T., Wong, S. P., Staals, R. H., Seegers, C. L., Bollen, S., Jore, M. M., Semenova, E., Severinov, K., de Vos, W. M., Dame, R. T., de Vries, R., Brouns, S. J., and van der Oost, J. (2012) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell* **46**, 595-605

20. Bailey, S. (2013) The Cmr complex: an RNA-guided endoribonuclease. *Biochem Soc Trans* **41**, 1464-1467

21. Koonin, E. V., Makarova, K. S., and Zhang, F. (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67-78

22. Jansen, R., Embden, J. D., Gaastra, W., and Schouls, L. M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565-1575

23. Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* **1**, e60

24. Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J.,

and Koonin, E. V. (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477

25. Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., and Koonin, E. V. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-736

26. Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., and Koonin, E. V. (2015) Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol Cell* **60**, 385-397

27. Ozcan, A., Pausch, P., Linden, A., Wulf, A., Schuhle, K., Heider, J., Urlaub, H., Heimerl, T., Bange, G., and Randau, L. (2019) Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat Microbiol* **4**, 89-96

28. Taylor, H. N., Warner, E. E., Armbrust, M. J., Crowley, V. M., Olsen, K. J., and Jackson, R. N. (2019) Structural basis of Type IV CRISPR RNA biogenesis by a Cas6 endoribonuclease. *RNA Biol* **16**, 1438-1447

29. Koonin, E. V., and Makarova, K. S. (2013) CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol* **10**, 679-686

30. Krupovic, M., Beguin, P., and Koonin, E. V. (2017) Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery. *Curr Opin Microbiol* **38**, 36-43

31. Krupovic, M., Shmakov, S., Makarova, K. S., Forterre, P., and Koonin, E. V. (2016) Recent Mobility of Casposons, Self-Synthesizing Transposons at the Origin of the CRISPR-Cas Immunity. *Genome Biol Evol* **8**, 375-386

32. Vestergaard, G., Garrett, R. A., and Shah, S. A. (2014) CRISPR adaptive immune systems of Archaea. *RNA Biol* **11**, 156-167

33. Abedon, S. T. (2011) Facilitation of CRISPR adaptation. *Bacteriophage* **1**, 179-181

34. Levy, A., Goren, M. G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520**, 505-510

35. Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., and Brouns, S. J. (2017) CRISPR-Cas: Adapting to change. *Science* **356**

36. Yosef, I., Goren, M. G., and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. *Nucleic Acids Res* **40**, 5569-5576

37. Nunez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol* **21**, 528-534

38. Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A., Phanse, S., Joachimiak, A., Koonin, E. V., Savchenko, A., Emili, A., Greenblatt, J., Edwards, A. M., and

Yakunin, A. F. (2011) A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol Microbiol* **79**, 484-502

39. He, Y., Wang, M., Liu, M., Huang, L., Liu, C., Zhang, X., Yi, H., Cheng, A., Zhu, D., Yang, Q., Wu, Y., Zhao, X., Chen, S., Jia, R., Zhang, S., Liu, Y., Yu, Y., and Zhang, L. (2018) Cas1 and Cas2 From the Type II-C CRISPR-Cas System of Riemerella anatipestifer Are Required for Spacer Acquisition. *Front Cell Infect Microbiol* **8**, 195
40. Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., and Marraffini, L. A. (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* **519**, 199-202
41. Wei, Y., Chesne, M. T., Terns, R. M., and Terns, M. P. (2015) Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in Streptococcus thermophilus. *Nucleic Acids Res* **43**, 1749-1758
42. Wright, A. V., and Doudna, J. A. (2016) Protecting genome integrity during CRISPR immune adaptation. *Nat Struct Mol Biol* **23**, 876-883
43. Wei, Y., Terns, R. M., and Terns, M. P. (2015) Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev* **29**, 356-361
44. Kieper, S. N., Almendros, C., Behler, J., McKenzie, R. E., Nobrega, F. L., Haagsma, A. C., Vink, J. N. A., Hess, W. R., and Brouns, S. J. J. (2018) Cas4 Facilitates PAM-Compatible Spacer Selection during CRISPR Adaptation. *Cell Rep* **22**, 3377-3384
45. Shiimori, M., Garrett, S. C., Chambers, D. P., Glover, C. V. C., 3rd, Graveley, B. R., and Terns, M. P. (2017) Role of free DNA ends and protospacer adjacent motifs

for CRISPR DNA uptake in Pyrococcus furiosus. *Nucleic Acids Res* **45**, 11281-11294

46. Shiimori, M., Garrett, S. C., Graveley, B. R., and Terns, M. P. (2018) Cas4 Nucleases Define the PAM, Length, and Orientation of DNA Fragments Integrated at CRISPR Loci. *Mol Cell* **70**, 814-824 e816

47. Vorontsova, D., Datsenko, K. A., Medvedeva, S., Bondy-Denomy, J., Savitskaya, E. E., Pougach, K., Logacheva, M., Wiedenheft, B., Davidson, A. R., Severinov, K., and Semenova, E. (2015) Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. *Nucleic Acids Res* **43**, 10848-10860

48. Nunez, J. K., Lee, A. S., Engelman, A., and Doudna, J. A. (2015) Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193-198

49. Dillingham, M. S., and Kowalczykowski, S. C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev* **72**, 642-671, Table of Contents

50. Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S. D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551-2561

51. Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J Bacteriol* **190**, 1401-1412

52. Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P., and Moineau, S. (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J Bacteriol* **190**, 1390-1400

53. Shah, S. A., Erdmann, S., Mojica, F. J., and Garrett, R. A. (2013) Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol* **10**, 891-899

54. Fineran, P. C., and Charpentier, E. (2012) Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology* **434**, 202-209

55. Wigley, D. B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat Rev Microbiol* **11**, 9-13

56. Fineran, P. C., Gerritzen, M. J., Suarez-Diez, M., Kunne, T., Boekhorst, J., van Hijum, S. A., Staals, R. H., and Brouns, S. J. (2014) Degenerate target sites mediate rapid primed CRISPR adaptation. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E1629-1638

57. Richter, C., Dy, R. L., McKenzie, R. E., Watson, B. N., Taylor, C., Chang, J. T., McNeil, M. B., Staals, R. H., and Fineran, P. C. (2014) Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res* **42**, 8516-8526

58. Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A., and Severinov, K. (2013) High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in E. coli. *RNA Biol* **10**, 716-725

59. Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., and Semenova, E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* **3**, 945

60. Li, M., Wang, R., Zhao, D., and Xiang, H. (2014) Adaptation of the Haloarcula hispanica CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic Acids Res* **42**, 2483-2492

61. Semenova, E., Savitskaya, E., Musharova, O., Strotskaya, A., Vorontsova, D., Datsenko, K. A., Logacheva, M. D., and Severinov, K. (2016) Highly efficient primed spacer acquisition from targets destroyed by the Escherichia coli type I-E CRISPR-Cas interfering complex. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 7626-7631

62. Nunez, J. K., Bai, L., Harrington, L. B., Hinder, T. L., and Doudna, J. A. (2016) CRISPR Immunological Memory Requires a Host Factor for Specificity. *Mol Cell* **62**, 824-833

63. Wei, Y., and Terns, M. P. (2016) CRISPR Outsourcing: Commissioning IHF for Site-Specific Integration of Foreign DNA at the CRISPR Array. *Mol Cell* **62**, 803-804

64. Yoganand, K. N., Sivathanu, R., Nimkar, S., and Anand, B. (2017) Asymmetric positioning of Cas1-2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic Acids Res* **45**, 367-381

65. Kim, J. G., Garrett, S., Wei, Y., Graveley, B. R., and Terns, M. P. (2019) CRISPR DNA elements controlling site-specific spacer integration and proper repeat length by a Type II CRISPR-Cas system. *Nucleic Acids Res*

66. Amitai, G., and Sorek, R. (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat Rev Microbiol* **14**, 67-76

67. Sternberg, S. H., Richter, H., Charpentier, E., and Qimron, U. (2016) Adaptation in CRISPR-Cas Systems. *Mol Cell* **61**, 797-808

68. Hale, C. R., Cocozaki, A., Li, H., Terns, R. M., and Terns, M. P. (2014) Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. *Genes Dev* **28**, 2432-2443

69. Charpentier, E., Richter, H., van der Oost, J., and White, M. F. (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol Rev* **39**, 428-441

70. Garside, E. L., Schellenberg, M. J., Gesner, E. M., Bonanno, J. B., Sauder, J. M., Burley, S. K., Almo, S. C., Mehta, G., and MacMillan, A. M. (2012) Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. *RNA* **18**, 2020-2028

71. Nam, K. H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M. P., and Ke, A. (2012) Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* **20**, 1574-1584

72. Richter, H., Zoephel, J., Schermuly, J., Maticzka, D., Backofen, R., and Randau, L. (2012) Characterization of CRISPR RNA processing in Clostridium thermocellum and Methanococcus maripaludis. *Nucleic Acids Res* **40**, 9887-9896

73. Carte, J., Christopher, R. T., Smith, J. T., Olson, S., Barrangou, R., Moineau, S., Glover, C. V., 3rd, Graveley, B. R., Terns, R. M., and Terns, M. P. (2014) The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in Streptococcus thermophilus. *Mol Microbiol* **93**, 98-112

74. Hatoum-Aslan, A., Maniv, I., and Marraffini, L. A. (2011) Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 21218-21222

75. Plagens, A., Tripp, V., Daume, M., Sharma, K., Klingl, A., Hrle, A., Conti, E., Urlaub, H., and Randau, L. (2014) In vitro assembly and activity of an archaeal CRISPR-Cas type I-A Cascade interference complex. *Nucleic Acids Res* **42**, 5125-5138

76. Carte, J., Pfister, N. T., Compton, M. M., Terns, R. M., and Terns, M. P. (2010) Binding and cleavage of CRISPR RNA by Cas6. *RNA* **16**, 2181-2188

77. Reeks, J., Sokolowski, R. D., Graham, S., Liu, H., Naismith, J. H., and White, M. F. (2013) Structure of a dimeric crenarchaeal Cas6 enzyme with an atypical active site for CRISPR RNA processing. *Biochem J* **452**, 223-230

78. Sefcikova, J., Roth, M., Yu, G., and Li, H. (2017) Cas6 processes tight and relaxed repeat RNA via multiple mechanisms: A hypothesis. *Bioessays* **39**

79. Shao, Y., and Li, H. (2013) Recognition and cleavage of a nonstructured CRISPR RNA by its processing endoribonuclease Cas6. *Structure* **21**, 385-393

80. Shao, Y., Richter, H., Sun, S., Sharma, K., Urlaub, H., Randau, L., and Li, H. (2016) A Non-Stem-Loop CRISPR RNA Is Processed by Dual Binding Cas6. *Structure* **24**, 547-554

81. Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., and Charpentier, E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607

82. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821

83. Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517-521

84. Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell* **172**, 1239-1259

85. East-Seletsky, A., O'Connell, M. R., Knight, S. C., Burstein, D., Cate, J. H., Tjian, R., and Doudna, J. A. (2016) Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270-273

86. Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M. R., Barendregt, A., Zhou, K., Snijders, A. P., Dickman, M. J., Doudna, J. A., Boekema, E. J., Heck,

A. J., van der Oost, J., and Brouns, S. J. (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* **18**, 529-536

87. van Erp, P. B., Jackson, R. N., Carter, J., Golden, S. M., Bailey, S., and Wiedenheft, B. (2015) Mechanism of CRISPR-RNA guided recognition of DNA targets in Escherichia coli. *Nucleic Acids Res* **43**, 8381-8391

88. Staals, R. H., Zhu, Y., Taylor, D. W., Kornfeld, J. E., Sharma, K., Barendregt, A., Koehorst, J. J., Vlot, M., Neupane, N., Varossieau, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P. J., Urlaub, H., Heck, A. J., Nogales, E., Doudna, J. A., Shinkai, A., and van der Oost, J. (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. *Mol Cell* **56**, 518-530

89. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, C., Nwokeoji, A. O., Dickman, M. J., Horvath, P., and Siksnys, V. (2014) Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. *Mol Cell* **56**, 506-517

90. Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol Cell* **58**, 418-430

91. Jackson, R. N., Golden, S. M., van Erp, P. B., Carter, J., Westra, E. R., Brouns, S. J., van der Oost, J., Terwilliger, T. C., Read, R. J., and Wiedenheft, B. (2014) Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from Escherichia coli. *Science* **345**, 1473-1479

92. Mulepati, S., Heroux, A., and Bailey, S. (2014) Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* **345**, 1479-1484

93. Hochstrasser, M. L., Taylor, D. W., Bhat, P., Guegler, C. K., Sternberg, S. H., Nogales, E., and Doudna, J. A. (2014) CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 6618-6623

94. Redding, S., Sternberg, S. H., Marshall, M., Gibb, B., Bhat, P., Guegler, C. K., Wiedenheft, B., Doudna, J. A., and Greene, E. C. (2015) Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System. *Cell* **163**, 854-865

95. Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., and White, M. F. (2013) Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol Cell* **52**, 124-134

96. Taylor, D. W., Zhu, Y., Staals, R. H., Kornfeld, J. E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J. A. (2015) Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* **348**, 581-585

97. Marraffini, L. A., and Sontheimer, E. J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843-1845

98. Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* **161**, 1164-1174

99. Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., and Terns, M. P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev* **30**, 447-459

100. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C., and Siksnys, V. (2016) Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Mol Cell* **62**, 295-306

101. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., and Wang, Y. (2019) Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. *Cell* **176**, 239-253 e216

102. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksnys, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609

103. Niewoehner, O., Garcia-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., and Jinek, M. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543-548

104. Ramia, N. F., Spilman, M., Tang, L., Shao, Y., Elmore, J., Hale, C., Cocozaki, A., Bhattacharya, N., Terns, R. M., Terns, M. P., Li, H., and Stagg, S. M. (2014) Essential structural and functional roles of the Cmr4 subunit in RNA cleavage by the Cmr CRISPR-Cas complex. *Cell Rep* **9**, 1610-1617

105. Jung, T. Y., An, Y., Park, K. H., Lee, M. H., Oh, B. H., and Woo, E. (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. *Structure* **23**, 782-790

106. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013) Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* **8**, 15

107. Grynberg, M., Erlandsen, H., and Godzik, A. (2003) HEPN: a common domain in bacterial drug resistance and human neurodegenerative proteins. *Trends Biochem Sci* **28**, 224-226

108. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front Genet* **5**, 102

109. Sheppard, N. F., Glover, C. V., 3rd, Terns, R. M., and Terns, M. P. (2016) The CRISPR-associated Csx1 protein of Pyrococcus furiosus is an adenosine-specific endoribonuclease. *RNA* **22**, 216-224

110. Foster, K., Kalter, J., Woodside, W., Terns, R. M., and Terns, M. P. (2019) The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR-Cas systems. *RNA Biol* **16**, 449-460

111. Niewoehner, O., and Jinek, M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA* **22**, 318-329

112. Rouillon, C., Athukoralage, J. S., Graham, S., Gruschow, S., and White, M. F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* **7**

113. Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569-573

114. Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., and Doudna, J. A. (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997

115. Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., and Zhang, F. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771

116. Swarts, D. C., van der Oost, J., and Jinek, M. (2017) Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol Cell* **66**, 221-233 e224

117. Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., Fedorova, I., Nakane, T., Makarova, K. S., Koonin, E. V., Ishitani, R., Zhang, F., and Nureki, O. (2016) Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell* **165**, 949-962

118. Yang, H., Gao, P., Rajashankar, K. R., and Patel, D. J. (2016) PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease. *Cell* **167**, 1814-1828 e1812

119. Liu, L., Li, X., Ma, J., Li, Z., You, L., Wang, J., Wang, M., Zhang, X., and Wang, Y. (2017) The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell* **170**, 714-726 e710

120. Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017) Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. *Cell* **168**, 121-134 e112

121. Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., and Zhang, F. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573

122. Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Zheng, K., Slaymaker, I. M., Gootenberg, J. S., Abudayyeh, O. A., Essletzbichler, P., Shmakov, S., Makarova, K. S., Koonin, E. V., and Zhang, F. (2017) Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* **65**, 618-630 e617

123. Spilman, M., Cocozaki, A., Hale, C., Shao, Y., Ramia, N., Terns, R., Terns, M., Li, H., and Stagg, S. (2013) Structure of an RNA silencing complex of the CRISPR-Cas immune system. *Mol Cell* **52**, 146-152

124. Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J., Doudna, J. A., Heck, A. J. R., Yonekura, K., van der Oost, J., and Shinkai, A. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. *Mol Cell* **52**, 135-145

125. Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Graham, S., Reimann, J., Cannone, G., Liu, H., Albers, S. V., Naismith, J. H., Spagnolo, L., and White, M. F. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol Cell* **45**, 303-313

126. Guo, M., Zhang, K., Zhu, Y., Pintilie, G. D., Guan, X., Li, S., Schmid, M. F., Ma, Z., Chiu, W., and Huang, Z. (2019) Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry. *Cell Res* **29**, 305-312

127. Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W., and Marraffini, L. A. (2013) A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J Biol Chem* **288**, 27888-27897

128. Liu, T. Y., Iavarone, A. T., and Doudna, J. A. (2017) RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. *PLoS One* **12**, e0170552

129. Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol Microbiol* **87**, 1088-1099

130. Ichikawa, H. T., Cooper, J. C., Lo, L., Potter, J., Terns, R. M., and Terns, M. P. (2017) Programmable type III-A CRISPR-Cas DNA targeting modules. *PLoS One* **12**, e0176221

131. Wang, L., Mo, C. Y., Wasserman, M. R., Rostol, J. T., Marraffini, L. A., and Liu, S. (2019) Dynamics of Cas10 Govern Discrimination between Self and Non-self in Type III CRISPR-Cas Immunity. *Mol Cell* **73**, 278-290 e274

132. Zebec, Z., Manica, A., Zhang, J., White, M. F., and Schleper, C. (2014) CRISPR-mediated targeted mRNA degradation in the archaeon Sulfolobus solfataricus. *Nucleic Acids Res* **42**, 5280-5288

133. Millen, A. M., Samson, J. E., Tremblay, D. M., Magadan, A. H., Rousseau, G. M., Moineau, S., and Romero, D. A. (2019) Lactococcus lactis type III-A CRISPR-Cas system cleaves bacteriophage RNA. *RNA Biol* **16**, 461-468

134. Benda, C., Ebert, J., Scheltema, R. A., Schiller, H. B., Baumgartner, M., Bonneau, F., Mann, M., and Conti, E. (2014) Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. *Mol Cell* **56**, 43-54

135. Zhu, X., and Ye, K. (2015) Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. *Nucleic Acids Res* **43**, 1257-1267

136. Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A. M., Glover, C. V., 3rd, Graveley, B. R., Terns, R. M., and Terns, M. P. (2012) Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol Cell* **45**, 292-302

137. Koonin, E. V., and Makarova, K. S. (2019) Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* **374**, 20180087

138. Estrella, M. A., Kuo, F. T., and Bailey, S. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Dev* **30**, 460-470

139. Han, W., Li, Y., Deng, L., Feng, M., Peng, W., Hallstrom, S., Zhang, J., Peng, N., Liang, Y. X., White, M. F., and She, Q. (2017) A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. *Nucleic Acids Res* **45**, 1983-1993

140. Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J Bacteriol* **196**, 310-317

141. Goldberg, G. W., Jiang, W., Bikard, D., and Marraffini, L. A. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* **514**, 633-637

142. Jia, N., Mo, C. Y., Wang, C., Eng, E. T., Marraffini, L. A., and Patel, D. J. (2019) Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. *Mol Cell* **73**, 264-277 e265

143. Zhang, J., Graham, S., Tello, A., Liu, H., and White, M. F. (2016) Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. *Nucleic Acids Res* **44**, 1789-1799

144. Park, K. H., An, Y., Jung, T. Y., Baek, I. Y., Noh, H., Ahn, W. C., Hebert, H., Song, J. J., Kim, J. H., Oh, B. H., and Woo, E. J. (2017) RNA activation-independent DNA targeting of the Type III CRISPR-Cas system by a Csm complex. *EMBO Rep* **18**, 826-840

145. Gruschow, S., Athukoralage, J. S., Graham, S., Hoogeboom, T., and White, M. F. (2019) Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. *Nucleic Acids Res* **47**, 9259-9270

146. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., Montoya, G., and She, Q. (2018) A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. *Nucleic Acids Res* **46**, 10319-10330

147. Jia, N., Jones, R., Sukenick, G., and Patel, D. J. (2019) Second Messenger cA4 Formation within the Composite Csm1 Palm Pocket of Type III-A CRISPR-Cas Csm Complex and Its Release Path. *Mol Cell* **75**, 933-943 e936

148. Nasef, M., Muffly, M. C., Beckman, A. B., Rowe, S. J., Walker, F. C., Hatoum-Aslan, A., and Dunkle, J. A. (2019) Regulation of cyclic oligoadenylate synthesis by the Staphylococcus epidermidis Cas10-Csm complex. *RNA* **25**, 948-962

149. Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D. J. (2019) CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. *Mol Cell* **75**, 944-956 e946

150. Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., Lopez-Mendez, B., She, Q., and Montoya, G. (2019) Structure of Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. *Nat Commun* **10**, 4302

151. Marraffini, L. A., and Sontheimer, E. J. (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568-571

152. Song, A. A., In, L. L. A., Lim, S. H. E., and Rahim, R. A. (2017) A review on Lactococcus lactis: from food to factory. *Microb Cell Fact* **16**, 55

153. Otto, M. (2009) Staphylococcus epidermidis--the 'accidental' pathogen. *Nat Rev Microbiol* **7**, 555-567

154. Lecomte, X., Gagnaire, V., Lortal, S., Dary, A., and Genay, M. (2016) Streptococcus thermophilus, an emerging and promising tool for heterologous expression: Advantages and future trends. *Food Microbiol* **53**, 2-9

155. Kengen, S. W. M. (2017) 'Pyrococcus furiosus, 30 years on'. *Microb Biotechnol* **10**, 1441-1444

156. Terns, R. M., and Terns, M. P. (2014) CRISPR-based technologies: prokaryotic defense weapons repurposed. *Trends Genet* **30**, 111-118

CHAPTER 2

THE RIBONUCLEASE ACTIVITY OF CSM6 IS REQUIRED FOR ANTI-PLASMID IMMUNITY BY TYPE III-A CRISPR-CAS SYSTEMS[1]

[1]Kawanda Foster, Joshua Kalter, Walter Woodside, Rebecca M. Terns & Michael P. Terns (2019) The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR-Cas systems, RNA Biology, 16:4, 449-460, DOI: 10.1080/15476286.2018.1493334
Reprinted here with permission of the publisher

ABSTRACT

CRISPR-Cas systems provide prokaryotes with RNA-based adaptive immunity against viruses and plasmids. A unique feature of Type III CRISPR-Cas systems is that they selectively target transcriptionally-active invader DNA, and can cleave both the expressed RNA transcripts and source DNA. The Type III-A effector crRNP (CRISPR RNA-Cas protein complex), which contains Cas proteins Csm1-5, recognizes and degrades invader RNA and DNA in a crRNA-guided, manner. Interestingly, Type III-A systems also employ Csm6, an HEPN family ribonuclease that does not stably associate with the Type III-A effector crRNP, but nevertheless contributes to defense via mechanistic details that are still being determined. Here, we investigated the mechanism of action of Csm6 in Type III-A CRISPR-Cas systems from *Lactococcus lactis*, *Staphylococcus epidermidis*, and *Streptococcus thermophilus* expressed in *Escherichia coli*. We found that *L. lactis* and *S. epidermidis* Csm6 cleave RNA specifically after purines *in vitro*, similar to the activity reported for *S. thermophilus* Csm6. Moreover, *L. lactis* Csm6 functions as a divalent metal-independent, single strand-specific endoribonuclease that depends on the conserved HEPN domain. *In vivo*, we show that deletion of *csm6* or expression of an RNase-defective form of Csm6 disrupts crRNA-dependent loss of plasmid DNA in all three systems expressed in *E. coli*. Mutations in the Csm1 palm domain, which are known to deactivate Csm6 ribonuclease activity, also prevent plasmid loss in the three systems. The results indicate that Csm6 ribonuclease activity rather than Csm1-mediated DNase activity effects anti-plasmid immunity by the three Type III-A systems investigated.

INTRODUCTION

The multiple diverse CRISPR-Cas systems function to protect bacteria and archaea from viruses, plasmids, and other mobile genetic elements [1-3]. Each system employs distinct CRISPR RNA species and Cas proteins to form effector crRNPs that engage in crRNA-guided, Cas nuclease-mediated destruction of foreign nucleic acids. Of the six major types of CRISPR-Cas systems [2,4], Types I, II, and V selectively destroy invader DNA, while Type VI is specific for invader RNA [1]. How Type IV CRISPR-Cas systems function remains to be determined. The Type III CRISPR-Cas systems are notable in that they eliminate both the RNA and DNA components of invasive mobile genetic elements [5-17].

Type III-A and the related Type III-B CRISPR-Cas effector crRNPs include five (Type III-A; Csm1-5) and six (Type III-B; Cmr1-6) protein subunits, respectively [18-24]. Both Type III systems include an additional protein – Csm6 (Type III-A) or Csx1 (Type III-B) [2] – that is not observed to be a stable component of the associated effector complex [6,7,14,17,21]. It is unclear if Csm6 and Csx1 function in *trans* or are recruited to the effector crRNPs transiently during function.

DNA targeting by Type III systems requires that the invasive DNA be actively expressed [8-12,15,17]. Interaction between the crRNA and a newly transcribed target RNA activates the DNase activity of the complex resulting in co-transcriptional destruction of the non-template strand of invader DNA as well as target RNA destruction [11]. Interestingly, Type III CRISPR-Cas system-mediated protection from viruses and plasmids relies on a DNase and two separate RNase activities [7,8,13,15,25-29].

The DNase activity of Type III CRISPR-Cas systems maps to the Csm1 (III-A) and Cmr2 (III-B) protein subunits [9-13,30-32]. The location of the Csm1 and Cmr2 DNase active site(s) is the subject of seemingly conflicting evidence: different mutational studies support a role for the conserved HD domain [9,10,12,17,31-33], the GGDD motif of the conserved palm domain [11,17,34], or both [9,10,17]. Importantly, reconsideration of the conclusions of early mutational analysis is warranted by recent findings that indicate that the palm domain of Csm1/Cmr2 functions in the synthesis of cyclic oligoadenylates [35,36], which have been shown to activate the RNase activity of Csm6 (III-A) [35,36]. The same signaling mechanism presumably also operates to activate the RNase activity of the related Csx1 (III-B) protein and recently it was determined that short linear oligoadenylates greatly stimulate Csx1 RNase activity *in vitro* [37]. The palm domain mutations may be preventing Csm6/Csx1 RNase activity rather than preventing Csm1/Cmr2 DNase activity.

One of the RNase activities of the Type III effector crRNP complexes is associated with the Csm3 protein (III-A) [5,14,38] or Cmr4 (III-B) [6,18,39-41]. Csm3 and Cmr4 subunits bind along the length of the integral crRNA and cleave bound target RNAs at regular six-nucleotide intervals [14,21,38,40,41]. The second RNase activity is provided by Csm6 (III-A) or Csx1 (III-B) proteins, which cleave a variety of RNA substrates *in vitro* [13,17,35-37,42,43]. These proteins share N-terminal CARF (Cas-associated Rossman fold) and C-terminal HEPN (Higher eukaryotes and prokaryotes nucleotide binding) domains. Evidence indicates that the catalytic activity resides in an RNase active site located within the HEPN domain [13,37,42-45]. The RNase activities of Csm6 and presumably Csx1 are stimulated by the short cyclic oligoadenylate signaling molecules

generated by the Csm1/Cmr2 subunits of the Type III effector crRNP complexes upon target RNA recognition [35-37]. The conserved N-terminal CARF domains of Csm6 and Csx1 bind the oligoadenylate ligands and allosterically activate the RNase activity [35-37,45].

The role of Csm6 in Type III-A mediated immunity is of particular interest because, while it appears to be dispensable for RNA and DNA targeting activities of the crRNP effector complex *in vitro* [5,11,12,14], it is critical for immunity to plasmids or viruses *in vivo* [7,13,15,25], at least under certain conditions [13]. Here, we have investigated Type III-A systems from three different bacteria – *Lactococcus lactis*, *Streptococcus thermophilus*, and *Staphylococcus epidermidis* – expressed as modules of component crRNAs and Csm1-6 proteins in *E. coli* [28]. All three modules confer resistance to plasmids that express transcripts containing sequences targeted by the crRNAs [28]. Our results reveal a conserved requirement for the RNase activity of Csm6 and the GGDD motif in the palm domain of Csm1 (known to activate Csm6 ribonuclease function) rather than DNase activity, in Type III-A CRISPR-Cas system-mediated immunity to plasmids.

RESULTS

L. lactis Csm6 is an HEPN-dependent ribonuclease

Previous biochemical characterization of purified *S. thermophilus* and *S. epidermidis* Csm6 proteins indicated that these proteins are single-stranded ribonucleases [13,35,43]. The RNase activity of the proteins was disrupted by mutation of the conserved histidine of the R-X_{4-6}-H RNase active site within the C-terminal HEPN domain [13,35,43]. To determine if *L. lactis* Csm6 functions as a ribonuclease, wildtype

and HEPN active site mutant (H360A) versions of the protein expressed in and purified from *E. coli* (Fig. 2.1A) were incubated with 5'-radiolabeled single- and double-stranded RNAs and DNAs, and DNA/RNA hybrid molecules (Fig. 2.1B and C). Wildtype Csm6, but not the HEPN mutant protein, selectively cleaved single-stranded RNAs. Cleavage by *L. lactis* Csm6 resulted in accumulation of distinct RNA breakdown products, suggesting possible endonucleolytic activity (Fig. 2.1B and C). The results demonstrate that *L. lactis* Csm6 is an HEPN-dependent ribonuclease that exhibits specificity for cleavage of single-stranded RNA.

L. lactis Csm6 is a divalent metal-independent endoribonuclease

Further investigation revealed that *L. lactis* Csm6 can function as an endoribonuclease. The ability of the purified protein to cleave linear and circular forms of an RNA substrate was assessed (Fig. 2.2A). Wildtype *L. lactis* Csm6 cleaved both forms of the RNA. The slightly shifted pattern of 5' radiolabeled Csm6 RNA cleavage products observed between the linear vs. circular forms of substrate RNA apparently reflects that the radiolabel is internalized in the circular RNA and indicates that RNA cleavage by Csm6 occurred at specific sites within the substrate RNA. The Csm6 HEPN mutant was unable to cleave either form. The ability of wildtype Csm6 to cleave circular RNAs shows that the protein is an endoribonuclease and does not require free RNA ends as is the case for obligate exoribonucleases.

Other HEPN family member ribonucleases have been found to cleave RNA in a divalent metal ion-independent manner [37,42-44]. We tested the hypothesis that Csm6 also cleaves RNAs without a requirement for divalent metal-ions (Fig 2.2B). We found

that RNA cleavage by Csm6 did not require addition of divalent metal ions and furthermore, the cleavage reaction was not affected by addition of the divalent metal ion chelator ethylenediaminetetraacetic acid (EDTA) even at high (10 mM) concentrations. A hallmark feature of divalent metal-independent ribonucleases is that they cleave on the 5' side of phosphodiester bonds such that a 5' hydroxyl group, and a 2'-3' cyclic phosphate or 3' phosphate groups are left on the RNA degradation products [37,42,43,46]. Using assays previously employed to map the chemical end groups of RNA cleavage products ([42] and see Methods section), we determined that digestion of RNA by *L. lactis* Csm6 resulted in products with 5' hydroxyl, and 2'-3' cyclic phosphate or 3' phosphate termini (Fig. 2.2C and D). A control RNA containing a 3' hydroxyl end group was elongated by *E. coli* poly(A) polymerase (PAP), which only adds non-templated adenosine residues to RNAs with 3' hydroxyl groups, but Csm6 RNA cleavage products were not elongated (Fig. 2.2C). Moreover, the Csm6 RNA cleavage products were resistant to cleavage by Terminator 5´-Phosphate-Dependent Exonuclease (TEX), an enzyme that degrades RNAs containing a 5' phosphate group (Fig. 2.2D).

Csm6 demonstrates a preference for cleavage at purines

Next, we investigated whether Csm6 ribonuclease activity exhibits base specificity. Previous work found that *S. thermophilus* Csm6 showed a preference for cleaving after GA and AA dinucleotides [35]. We incubated *L. lactis* Csm6 (Fig. 2.3) and *S. epidermidis* Csm6 (Fig. 2.4) with three 5' radiolabeled RNA substrates having different sequences. The RNA products were separated on sequencing gels and mapped at nucleotide resolution. To assess the specificity of any sites of Csm6 cleavage, alkaline hydrolysis

and RNase T1 ladders of each substrate RNA were used in parallel along with RNA standards differing by ten nucleotide size intervals. As specificity controls, we omitted Csm6 protein or included RNase-defective mutant forms of each Csm6 (H360A or H369A) protein. Degradation products observed in control lanes were considered non-specific and ignored for analysis. The mapping results indicate that both *L. lactis* and *S. epidermidis* Csm6 exhibit a clear preference for cleavage after purines; major cleavages were only observed after guanosine and adenosine residues (Figs. 2.3 and 2.4). *L. lactis* and *S. epidermidis* Csm6 produced nearly identical cleavage patterns for each of the tested RNA substrates, however minor differences in the particular purines recognized can be observed. Very weak cleavage activity after certain uridine residues was detected with *S. epidermidis* but not *L. lactis* Csm6. The results indicate that Csm6 proteins from three different bacterial species (*L. lactis* (Fig. 2.3), *S. epidermidis* (Fig. 2.4), and *S. thermophilus* [35]) cleave RNA preferentially after purines.

The RNase activity of Csm6 is critical for anti-plasmid immunity

L. lactis, *S. epidermidis*, and *S. thermophilus* Type III-A systems including the Csm6 protein selectively carry out anti-plasmid immunity when expressed as modules in *E. coli* [28]. Immunity was tested using two types of plasmids for each system: a Csm module plasmid and a target plasmid (Fig. 2.5A). As described previously[28], the Csm module plasmids contain all components necessary for producing Csm crRNPs and the target plasmids contain sequences with crRNA sequence homology flanked by promoter and terminator elements to confer expression of the target RNA. To determine if Csm6 is important for Type III-A immunity *in vivo*, we tested for crRNA-dependent anti-plasmid

immunity in strains lacking Csm6 or expressing RNase-defective Csm6 (Fig. 2.5). The *E. coli* strains were challenged with either a plasmid that contained a sequence that is complementary to the crRNA (target) or one that lacks crRNA complementarity (non-target). The plasmids encode ampicillin resistance. Plasmid challenge results in a high number of ampicillin resistant colonies in strains lacking a functional III-A CRISPR-Cas system (Fig. 2.5A). Strains with an active Type III-A CRISPR-Cas system yield a low number of ampicillin resistant colonies at the plated dilutions (Fig. 2.5A). The wildtype *L. lactis*, *S. epidermidis*, and *S. thermophilus* Type III-A systems successfully conferred anti-plasmid immunity against plasmids containing the crRNA target sequences, but not those that lacked target sequences, as was also previously demonstrated [28] (Fig. 2.5B-D, wildtype Csm6, target and non-target plasmids).

For all three Type III-A systems tested, anti-plasmid immunity was lost when the Csm6 protein was not expressed or if an RNase-defective HEPN mutant (H to A) protein was substituted (Fig. 2.5B-D). Strains in which the Csm6 protein was omitted or mutated typically resulted in 4-5 orders of magnitude more colonies than wildtype strains containing Csm6 when plasmids with crRNA targets were assayed. In the absence of csm6 (ΔCsm6), there appears to be a slight (one order of magnitude or less) reduction in colony number when comparing the data of plasmids with crRNA targets vs. those lacking targets. Wildtype and HEPN mutant *L. lactis* Csm6 proteins were recovered at comparable levels from the strains (Fig. 2.S1), suggesting that the phenotype of the Csm6 HEPN mutant was due to disruption of ribonuclease activity rather than simply to loss of stability. Furthermore, biochemical analysis of *T. thermophilus* Csm6 showed that the same HEPN point mutation did not hinder Csm6 homodimerization further indicating that

Csm6 structure is not obviously disrupted by the mutation [43]. To determine whether or not Csm6 was functioning as an integral component of the Type III-A effector crRNPs, the effector complex was isolated and evaluated for the presence of any associated Csm6 protein (Fig. 2.S1). Affinity purification of the 6x histidine-tagged Csm3 subunit confirmed that Csm1-5 but not Csm6 are stable components of the *L. lactis* effector crRNP complexes (Fig. 2.S1), in agreement with findings with other Type III-A systems [7,14]. Our results show that anti-plasmid activity of the Type III-A systems from *L. lactis*, *S. epidermidis,* and *S. thermophilus* critically depends upon the Csm6 protein. Furthermore, the evidence indicates that anti-plasmid immunity depends on the HEPN domain and the ability of Csm6 to catalyze RNA cleavage.

The palm domain of Csm1 is needed for anti-plasmid immunity

The Csm1 proteins of Type III-A systems cleave DNA *in vitro* [11,12,30-33]. It is not clear whether the DNase active site resides in the HD domain and/or the palm domain [7,11,12,32,33] or whether the DNase activity is vital for CRISPR-Cas immunity. Moreover, mutation of the palm domain of Csm1 has recently been shown to disrupt Csm6 ribonuclease activity by blocking cyclic oligoadenylate production [35,36], complicating the interpretation of previous studies where reduction of Csm6 activity may have contributed to phenotypes being examined [7,11].

To assess whether the two conserved domains (HD and palm) of the Csm1 subunit are important for anti-plasmid immunity in *L. lactis*, *S. epidermidis*, and *S. thermophilus* Csm systems including Csm6 *in vivo*, we examined the activity of *E. coli* strains expressing Csm1 HD domain (HD to AA) or palm domain (GGDD to GGAA) mutants (Fig.

2.6). For all three Type III-A systems, mutations in the palm domain significantly reduced the anti-plasmid immunity compared to strains with wildtype Csm1. In contrast, anti-plasmid immunity was not observably affected by mutations in the HD domain. The results reveal that the GGDD motif of the palm domain but not the HD domain of Csm1 is critical for plasmid resistance in all three Type III-A systems investigated (Fig. 2.6).

DISCUSSION

Type III-A CRISPR-Cas systems are selectively activated when a crRNA target sequence on a virus or plasmid is transcribed [7,11,12,15,28,32]. Interestingly, Type III-A systems include three distinct nucleases that can function to target the destruction of both RNA and DNA molecules of the invaders. Target RNA binding by the Type III-A effector crRNP complex (consisting of Csm1-5 and crRNA) is the central trigger that prompts the DNase (Csm1) and ribonuclease (Csm3 and Csm6) activities [11-13,35,36]. It is not yet clear how these three nucleases are deployed to eliminate invaders or the hierarchical importance of each nuclease for CRISPR-Cas immunity. In this study, we focused on dissecting the role of the Csm6 protein in anti-plasmid immunity using three different Type III-A systems comprised of components from *L. lactis*, *S. epidermidis*, and *S. thermophilus*. Although Csm6 is apparently not stably associated with the Type III-A crRNP effector complex (Fig. 2.S1 [7,14], we found that it is vital for anti-plasmid immunity in each of the three bacterial systems tested (Fig. 2.5). Our biochemical studies confirmed that Csm6 proteins function as divalent metal-independent endoribonucleases that employ the HEPN RNase active site to cleave single-stranded RNA molecules (Fig. 2.1 and 2.2). *In vivo* analysis clearly established that Csm6 RNase activity is important for

anti-plasmid immunity; the ability of the Type III-A systems to execute anti-plasmid immunity was nearly eliminated by mutations that disrupt RNase activity of Csm6 (Fig. 2.5). In addition, we show that mutations in the Csm1 palm domain that have been shown to inhibit activation of Csm6 by cyclic oligoadenylate signaling [35,36] also disrupted plasmid defense (Fig. 2.6).

Csm6 enzymes cleave RNAs after purines

Our detailed biochemical characterization of the *L. lactis* Csm6 protein showed that this protein exhibits all of the hallmark features of other previously characterized HEPN superfamily RNases [44]. Indeed, Csm6 and related Csx1 (Type III-B) proteins from various bacterial and archaeal species including *L. lactis* (Figs. 2.1-3), *S. epidermidis* (Fig. 2.4 and [13,43]), *S. thermophilus* [35], *T. thermophilus* [43], *P. furiosus* [42], *P. horikoshii [*43] and *S. islandicus [*37] all appear to function as endoribonucleases that specifically cleave single-stranded RNA using an HEPN RNase active site. Like other divalent metal-independent ribonucleases, Csm6 and Csx1 cleave on the 5' side of the phosphodiester bond to produce RNA degradation products with 5' hydroxyl and a 2'-3' cyclic phosphate or 3' phosphate group end groups (Fig. 2.2B and [17,36,42].

Our detailed mapping revealed that both *L. lactis* and *S. epidermidis* Csm6 RNases selectively cleave after purines (Fig. 2.3 and 2.4). This specificity for cleavage after certain guanine and adenosine residues was also reported for the *S. thermophilus* Csm6 protein [35]. No base specificity was reported in an earlier study that characterized RNA cleavage by *S. epidermidis* Csm6 [13,36], however, it is clear that discrete RNA cleavage products accumulated in the previous study, and our results with the same enzyme predict that the

observed Csm6 RNA cleavage products would map to purines if high resolution mapping studies were performed ([13,36] and Fig. 2.4). Lack of base specificity was also reported for RNA cleavage by *T. thermophilus* Csm6 [43], but was based only on testing of homo-polymer RNA substrates and not mixed-sequence RNA substrates [43]. In contrast, the Csx1 protein of the *P. furiosus* Type III-B system cleaves RNAs with a strict specificity for adenosine residues [42]. It is not clear whether the observed purine specificity of Csm6 or Csx1 proteins has any physiological significance other than to provide an effective means to cleave a wide range of possible viral, plasmid or host RNAs encountered during invader defense. Our findings that Csm6 proteins cleave after certain purines but not others in single-stranded RNA substrates (Fig. 2.3 and 2.4) and do not cleave double-stranded RNA (Fig. 2.1), suggests that the enzyme may cleave at exposed single-stranded regions within RNA secondary structures.

Role of Csm6 RNase activity in anti-plasmid immunity

The importance of Csm6 and the Type III-B-associated HEPN ribonuclease Csx1 in immunity against plasmids and viruses appears to vary with a number of biological conditions. For example, in *S. epidermidis* it appears that Csm6 is not required for phage immunity if the phage target sequence is expressed early in the phage infection cycle [13]. However, Csm6 is critical if the target DNA is expressed later in the cycle or if the target sequence includes mutations expected to weaken the immunity afforded by the DNase (Csm1) and RNase (Csm3) activities of the Type III-A effector crRNPs [13]. Deletion of *S. epidermidis csm6* results in disruption of immunity against conjugative plasmids [7], but not transformed plasmids [13]. In contrast, deletion of *Staphylococcus*

aureus csm6 partially reduces immunity against transformed plasmids [47]. Similarly, for unclear reasons, anti-plasmid immunity by Type III-B systems requires Csx1 in *S. sulfolobus* [8] but not *P. furious* [9]. Some of these findings may reflect a conditional dependence of Csm6 (and Csx1) in defense wherein these RNases function as auxiliary systems to the Type III crRNP effector complexes and are essential when the effector crRNP complexes fail to efficiently eliminate the invader DNA [13]. For three Type III-A systems tested with transformed plasmids, we found that Csm6 was essential for anti-plasmid immunity and that a functional HEPN RNase domain was integral for this function of Csm6 (Fig. 2.5).

Importantly, our finding that deletion or mutation of Csm6 severely disrupts anti-plasmid immunity (Fig. 2.5) indicates that DNA destruction by the Type III-A effector crRNP is not sufficient for efficient anti-plasmid immunity in our assay. However, the mechanism by which Csm6 affects immunity against plasmids will require further investigation. The plasmid transcript that is targeted in our system (an artificial transcript engineered to contained a crRNA target sequence [28] is not expected to be essential for plasmid maintenance. However, there is evidence that Csm6 cleavage activity is not strictly confined to the region of crRNA interaction *in vivo,* but spreads several kilobases in either direction from the crRNA interaction site [13]. Thus, it is possible that Csm6-mediated RNA cleavage spreads to encompass mRNAs essential for plasmid replication resulting in plasmid loss, or to host mRNAs resulting in death of bacterial cells containing a plasmid. A similar mechanism has been proposed to explain *E. coli* growth defects observed during anti-plasmid immunity by the exclusively RNA targeting Cas13a (Type VI) CRISPR-Cas system[48].

Key function of the palm domain of Csm1 in anti-plasmid immunity

There is considerable evidence that the Csm1 (Type III-A) and Cmr2 (Type III-B) proteins (members of the Cas10 superfamily) function both as target RNA-activated DNases [7,9-13,15,17,32,36] and cyclic oligoadenylate synthases (that activate the RNase activities of associated Csm6 and Csx1 proteins) [35-37]. We found that anti-plasmid immunity by three bacterial Type III-A systems requires an intact Csm1 palm motif, which is required for the cyclic oligoadenylate synthesis and possibly also for DNase activity [7,11,30] (Fig. 2.6). Mutation of the HD domain, which has also been implicated in DNase activity in some systems [31-33,35], did not disrupt anti-plasmid activity (Fig. 6).

It is unclear whether the disruption in anti-plasmid immunity observed with mutation of the Csm1 palm domain in our study is the result of disruption of DNase or oligoadenylate polymerase activity, or both. Similar palm domain mutations have been shown to be critical for DNase activity of *S. epidermidis* Csm1 [11,30], but not *S. thermophilus* Csm1 [12], *in vitro*. This suggests that the importance of the Csm1 palm domain observed in our work (Fig. 2.6) is via the activation of Csm6 RNase activity, at least in the case of the *S. thermophilus* system. The lack of effective anti-plasmid immunity in the absence of Csm6 indicates that DNA destruction by the Type III-A effector crRNP is not sufficient for anti-plasmid immunity by any of the 3 systems investigated (Fig. 2.5). A straightforward explanation for the impact of the Csm1 palm domain mutations in our work (Fig. 2.6) is that oligoadenylate signaling and Csm6 RNase activation (rather than DNA cleavage) is required for anti-plasmid immunity.

Comprehensive genome sequence analyses reveal that Type III CRISPR-Cas systems represent one of the most common, as well as highly diverse and variable types of CRISPR-Cas systems [2]. In addition to the Type III-A and III-B systems, there are also uncharacterized Type III-C and III-D systems that are predicted to encode equivalents to Csm1 or Cmr2 subunits with catalytically inactive HD or palm domains, respectively [2]. The diversity in Type III components predicts natural variation in the roles of the protein homologs in protection against plasmid and viral invaders. Detailed characterization of members of each of the four subtypes of Type III systems is needed to tease out both common features as well as the predicted mechanistic differences.

MATERIALS AND METHODS

Cloning and Mutagenesis

The construction of the wildtype *L. lactis*, *S. epidermidis*, and *S. thermophilus* Csm module plasmids (encoding Csm1-6, Cas6 and CRISPR array; pACYC), *L. lactis csm6* gene deletion plasmid, wildtype Csm6-only plasmids (pT7), and the crRNA target and non-target plasmids (pTrcHis), were reported [28]. The subsequent *csm6* or *csm1* gene deletions or mutations of the components of each Csm module plasmid as well as plasmids encoding just Csm6 were generated using either Quikchange site directed mutagenesis (Stratagene), splicing by overlap extension PCR, or inverse PCR, as described next. Quikchange site directed mutagenesis was performed on *csm6* genes of the Csm module plasmids (*S. epidermidis* H369A and *S. thermophilus* H347A) or Csm6-only plasmids (*L. lactis* H360A and *S. epidermidis* H369A) to create the desired HEPN domain mutation. Inverse PCR was used to delete the *csm6* genes from the *S. epidermidis* and S. *thermophilus* systems. Splicing overlap extension PCR was used to make constructs with the following wildtype or mutant *csm6* and *csm1* genes: *L. lactis* wildtype and HEPN mutant (H360A) (6x histidine-tagged or untagged) *csm6*, *L. lactis* HD (H13A/D14A) and palm domain (D576A/D577A) *csm1* mutations, and *S. thermophilus* HD (H15A/D16A) and palm domain (D575A/D576A) *csm1* mutations. To generate Csm1 HD domain (H14A/D15A) and palm domain (D586A/D587A) mutants in the *S. epidermidis* module, the wildtype *csm1* gene was first subcloned into Blunt II TOPO vector with the Zero Blunt II TOPO kit (Thermofisher) prior to inverse PCR mutagenesis and resubcloning (using BamHI and NdeI restriction enzymes) to replace the wildtype *csm1* gene in the Csm module plasmid with the mutant *csm1* genes. All DNA oligonucleotides were from

Eurofins Genomics and sequences used for cloning can be found in Table 2.S2 of the supplemental materials. Successful mutagenesis for all plasmids was confirmed via DNA sequencing (Eurofins Genomics).

Affinity purification of histidine-tagged Csm proteins

N-terminal 6x-histidine tagged Csm6 protein was expressed using pT7 transformed *E. coli* BL21-Star strain (DE3, Thermo Scientific) grown at 37°C with 50 μg/ml of kanamycin to an OD_{600} of 0.6. Protein expression was induced at room temperature overnight using Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM for *L. lactis* Csm6 containing cells or 0.3 mM for *S. epidermidis* Csm6 containing cells. Cells were resuspended in lysis buffer (40 mM Tris-HCl (pH 7.5), 200 mM NaCl, 10 mM Imidazole) and lysed via sonication. Insoluble material was removed by centrifugation at 14,000 rpm for 30 minutes and filtered through a 0.8 μm syringe filter (Corning Incorporated) prior to purification of Csm6 proteins using gravity affinity chromatography and Ni-NTA resin (Thermo Scientific). The lysate was incubated with the resin (pre-equilibrated with lysis buffer) for 1 hr at 4°C then washed with the lysis buffer, wash buffer A (40 mM Tris-HCl (pH 7.5), 200 mM NaCl, 20 mM Imidazole), and wash buffer B (40 mM Tris-HCl (pH 7.5), 200 mM NaCl, 30 mM Imidazole). Csm6 was eluted using four elution buffers containing increasing amounts of imidazole (75, 150, 250, 500 mM). Eluted *S. epidermidis* Csm6 was dialyzed (40 mM Tris-HCl (pH 7.5), 200 mM NaCl) and purified a second time using gravity affinity chromatography in the same manner. Wildtype and mutant Csm6 proteins were purified under identical conditions. For purification of either the 6x histidine-tagged, *L. lactis* Csm6 or Csm3 proteins expressed from the *L. lactis* Csm

module (Fig. 2.S1), proteins were affinity purified using a batch method that involved incubating the soluble lysate with Ni-NTA resin for 1 hour at 4°C prior to washing and elution with imidazole as described above. The purity of proteins was assessed using (10 % and 12.5% polyacrylamide) SDS-PAGE and Coomassie blue staining.

Preparation of RNA and DNA substrates

Synthetic RNAs were purchased from Integrated DNA Technologies and DNAs from Eurofins Genomics. The RNA and DNA sequences are given in Table 2.S1. The RNA and DNA were 5' end labeled as previously described [42]. Double-stranded oligonucleotides and circularized RNAs were generated as previously described [42].

Csm6 nuclease cleavage assay

To determine the nuclease activity of *L. lactis* Csm6 proteins against ssRNA, a 20 μl reaction consisting of 500 nM of Csm6, ~0.5-1.5 nM of radiolabeled RNA, 1x assay buffer (20 mM Tris-HCl (pH 7.5), 100 mM NaCl) was incubated at 30°C for 1 hour. In order to test substrate specificity for *L. lactis* Csm6, ssRNA was used in conjunction with dsRNA, ssDNA, dsDNA, and a dsDNA-RNA hybrid. The substrates were created as previously described [42]. Substrate specificity reactions were performed in a 20 μl reaction consisting of 500 nM of Csm6, ~0.5-1.5 nM of radiolabeled RNA or DNA, 10 mM EDTA, and 1x assay buffer and incubated at 30°C for 1 hour. Reactions were stopped by adding Gel Loading Buffer II (Life Technologies) and visualized by using denaturing 7M urea 15% polyacrylamide gels followed by autoradiography.

RNA cleavage product end group analysis

5’ and 3’ chemical end group analysis of Csm6 RNA cleavage products was conducted by using a combination of linear and circular substrates created as described above. First, wildtype *L. lactis* Csm6 was incubated with linear and circular RNA in standard cleavage assay conditions described above. PCI extraction and ethanol precipitation was then performed on the reactions and the resulting RNA was normalized based on cpm. Analysis of the 5’ end group was conducted by incubating linear and circular RNAs with 1 U Terminator 5’-Phosphate-Dependent Exonuclease (TEX) (Epicentre), 2,500 cpm of RNA, and 1x terminator reaction buffer A (Epicentre). Reactions were incubated at 30°C for 1 hour. Analysis of the 3’ end group was conducted by treating linear RNA with 5 U *E. coli* poly (A) polymerase (NEB), 15,000 cpm of RNA, 1 mM ATP, and 1x PAP reaction buffer A (NEB). Reactions were incubated at 37°C for 15 minutes. Enzyme activity was halted by adding Gel Loading Buffer II (Life Technologies). The RNAs were visualized after electrophoresis on denaturing 7 M urea 15% polyacrylamide gels for TEX reactions and denaturing 8.3 M urea 20% polyacrylamide gels for PAP reactions followed by autoradiography.

Mapping Csm6 RNA cleavage sites

RNA1, RNA5, and RNA7 substrates (sequences given in Table 2.S21) were incubated with Csm6 under reaction conditions supporting cleavage as described above with the following differences: 250 nM *S. epidermidis* Csm6 was used with 15 mM EDTA, 150 nM *L. lactis* Csm6 and 5 nM EDTA was used for RNA1 and RNA5 while 250 nM *L. lactis* Csm6 and 15 mM EDTA was used for RNA7. All reactions were incubated at 30°C

and samples were removed at four time points over the course of 1 hour. Reactions were stopped by a combination of Gel Loading Buffer II (Life Technologies) and placing the reactions on ice. Three ladders were used for mapping purposes: a Decade Marker (Life Technologies), a partial alkaline hydrolysis ladder (Ambion), and an RNase T1 ladder (Ambion) as described [42]. Reactions and ladders were separated on denaturing 15% polyacrylamide sequencing gels and the RNAs were visualized by autoradiography.

Plasmid interference assay

Plasmid interference assays were performed similarly to previously described experiments [28]. Electrocompetent *E. coli* BL21-Ai (Thermo Scientific) strains containing a Csm module (pACYC; chloramphenicol resistance) were transformed with 100 ng of either target or non-target plasmid (pTrcHis vector; ampicillin resistance). Electroporation was performed using a Gene Pulser II Electroporation System (Bio-Rad); 950 μl of SOC medium was immediately added to the 50 μl of the electrocompetent host cells and shaken at 200 rpm for 90 mins at 37°C. A series of six sequential 1:10 dilutions were made for each transformation and 5 μl of each dilution was spotted onto LB agar plates containing 34 μg/ml of chloramphenicol, 50 μg/ml of ampicillin, and 0.2% arabinose. Arabinose was omitted from the LB agar plates for the *L. lactis* and *S. epidermidis* systems in experiments shown in figure 6 since these systems were previously determined to induce plasmid resistance in the absence of arabinose induction of Csm module components [28]. Plates were incubated overnight at 37°C and images of the colonies were captured using white light in Gel Doc XR (Bio-Rad). Assays were done in triplicate.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed

Acknowledgments

We thank members of the Terns laboratory and Claiborne Glover for helpful discussions.

Funding

This work was supported by National Institutes of Health grants R35GM118160 (M.P.T.), R01GM54682 (M.P.T. and R.M.T.), and 1F31GM125365 (K.F.).

Supplemental Material

Supplemental material is available online

Figure 2.1. *L. lactis* Csm6 is a single-strand-specific ribonuclease that employs an HEPN active site. (A) The upper diagram highlights the conserved domains of Csm6 which include an N-terminal CARF domain and a C-terminal HEPN domain containing the R-X_{4-6}-H RNase active site motif. The lower panel shows purified wildtype and HEPN active site mutant (H360A) *L. lactis* 6x-His tagged Csm6 proteins following SDS-PAGE and Coomassie staining. Protein size standards (M) in kDa. (B) Wildtype and HEPN mutant *L. lactis* Csm6 were tested for nuclease activity using ^{32}P-labeled RNA (RNA1) compared to no protein (-) control. Reactions were incubated for 60 mins. and the RNA was resolved by denaturing polyacrylamide gel electrophoresis. The arrow indicates full length RNA and the bracket indicates Csm6-specific cleavage products. RNA size standards (M) in nucleotides. Results shown are representative of seven repeats of this experiment. (C) The ability of wildtype and HEPN Mutant *L. lactis* Csm6 to degrade a range of single- and double-stranded RNA and DNA substrates was tested (RNA2, RNA2+RNA3, DNA1, DNA1+DNA2, DNA1+RNA2). See Table S2.1 for sequences of the RNA and DNA substrates. Results shown are representative of two repeats of this experiment.

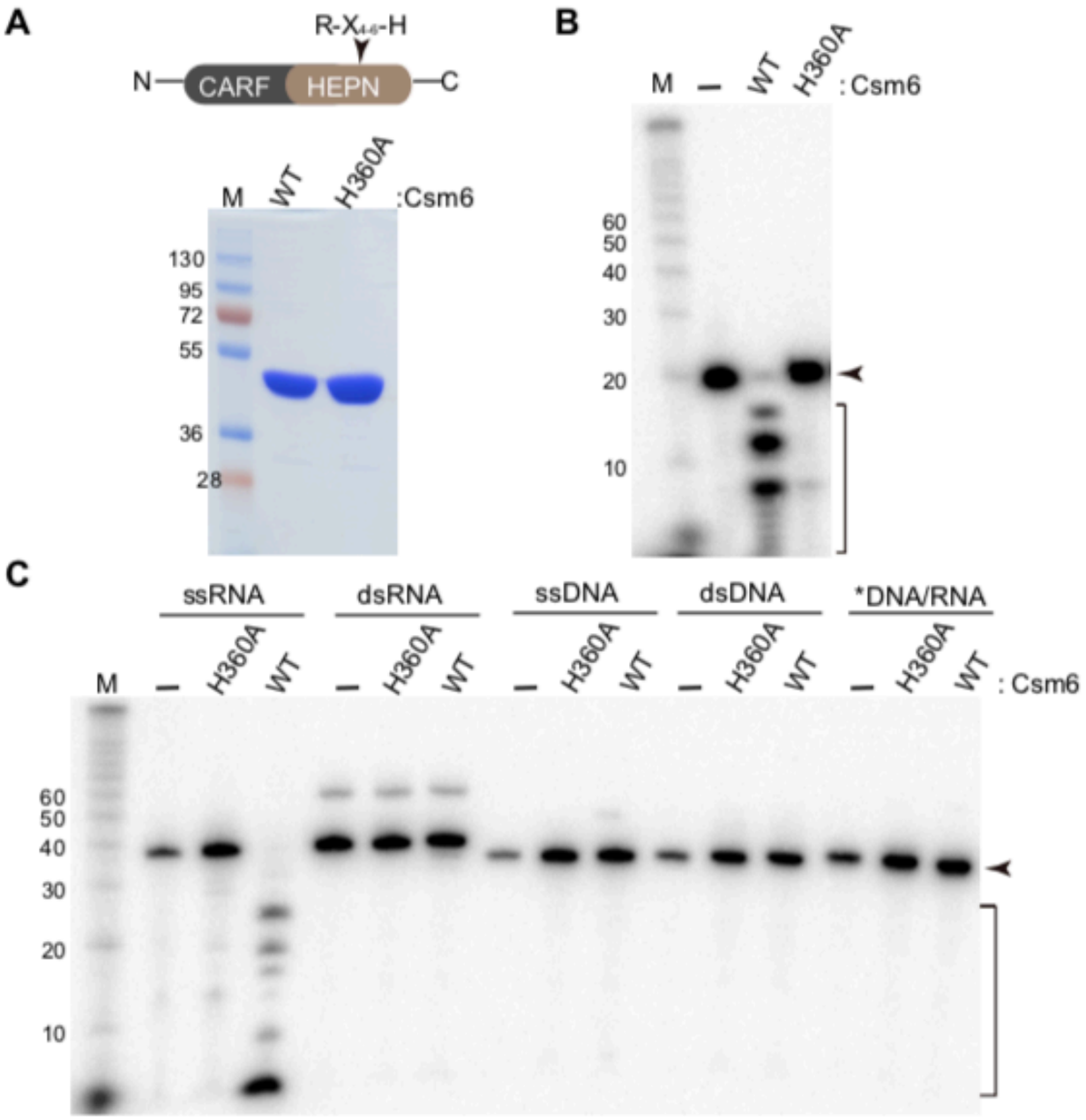

A
R-X4-6-H
N
CARF
HEPN
C
M
WT
H360A
:Csm6
130
95
72
55
36
28
B
M
—
WT
H360A
:Csm6
60
50
40
30
20
10
C
ssRNA
dsRNA
ssDNA
dsDNA
*DNA/RNA
M
—
H360A
WT
:Csm6
60
50
40
30
20
10

Figure 2.2. *L. lactis* Csm6 is a divalent metal-independent endoribonuclease and cleaves on the 5' side of the phosphodiester bond. (A) Endoribonuclease activity of *L. lactis* Csm6 was tested by incubating wildtype (WT), HEPN mutant (H360A) or no (-) Csm6 protein with linear or circular forms of a 5' radiolabeled RNA (RNA4; Table S2.1) as described in Figure 2.1. Results shown are representative of three repeats of this experiment. (B) Divalent metal-dependency of Csm6 RNA cleavage activity was tested by incubating wildtype *L. lactis* Csm6 with radiolabeled RNA (RNA5; Table S2.1) in the absence (-) or presence of increasing amounts of EDTA (0.1, 0.5, 1, 5, and 10 mM) and compared to a no Csm6 protein (-) control. Results shown are representative of two repeats of this experiment. (C) The enzyme poly(A) polymerase (PAP) was utilized to determine the 3' terminal chemical end group of Csm6-specific RNA degradation products. The RNA (RNA6; Table S2.1) was incubated in the presence or absence (-) of Csm6 followed by treatment with or without (-) PAP. Results shown are representative of four repeats of this experiment. (D) The enzyme Terminator 5'-Phosphate-Dependent Exonuclease (TEX) was utilized to determine the 5' terminal chemical end groups of Csm6 RNA degradation products. The RNA (Circularized or linear RNA6; Table S2.1) was incubated in the presence of absence (-) of Csm6 followed by treatment with or without (-) TEX. In panels A-D, the RNAs were resolved by denaturing polyacrylamide gel electrophoresis and full-length RNA is indicated with an arrow, while Csm6-specific RNA cleavage products are indicated with a bracket. In panel C, RNAs elongated by poly(A) polymerase are indicated with stars while in panel D, a star indicates a TEX-specific, minor RNA breakdown product. Results shown are representative of three repeats of this experiment.

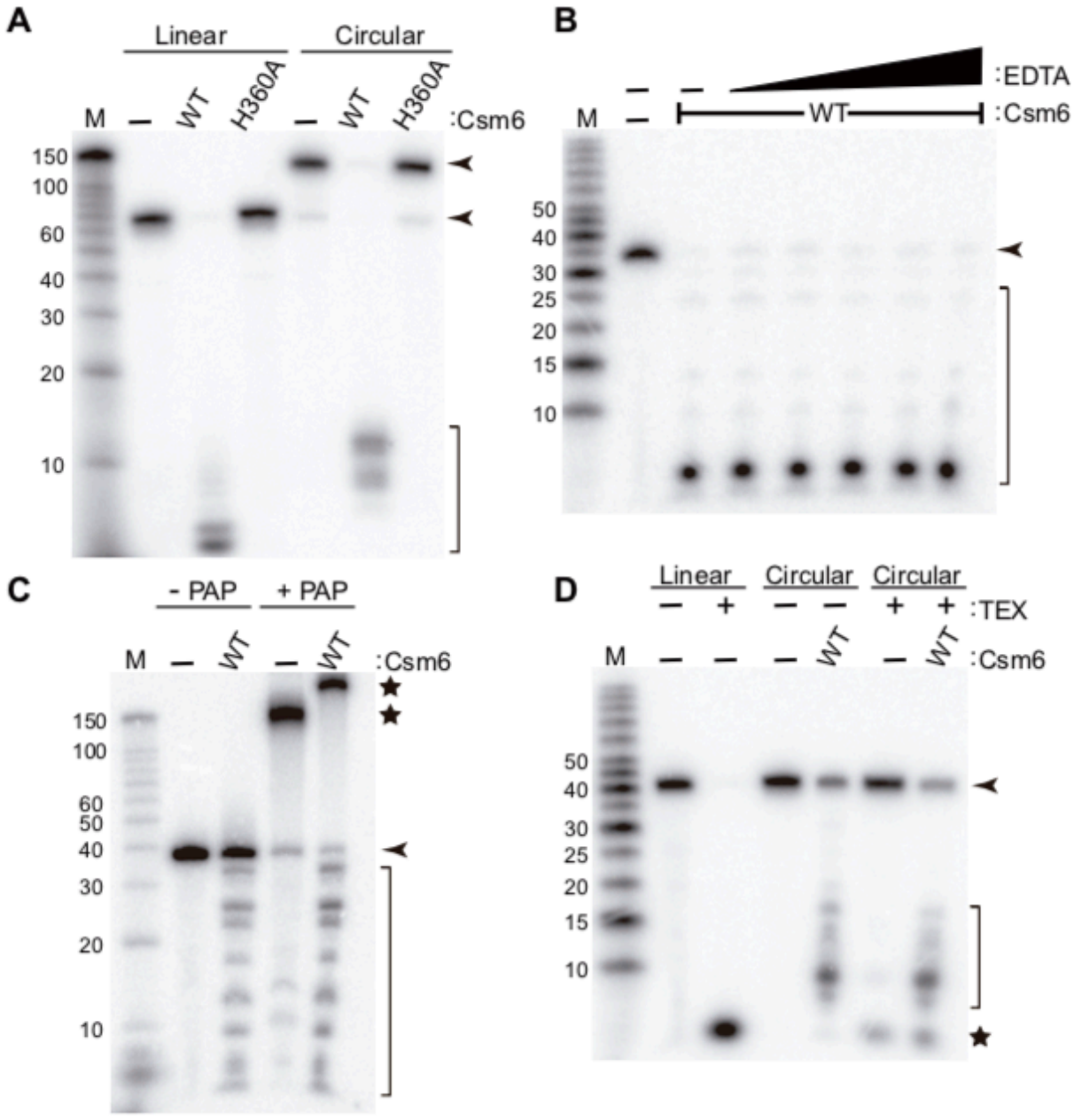
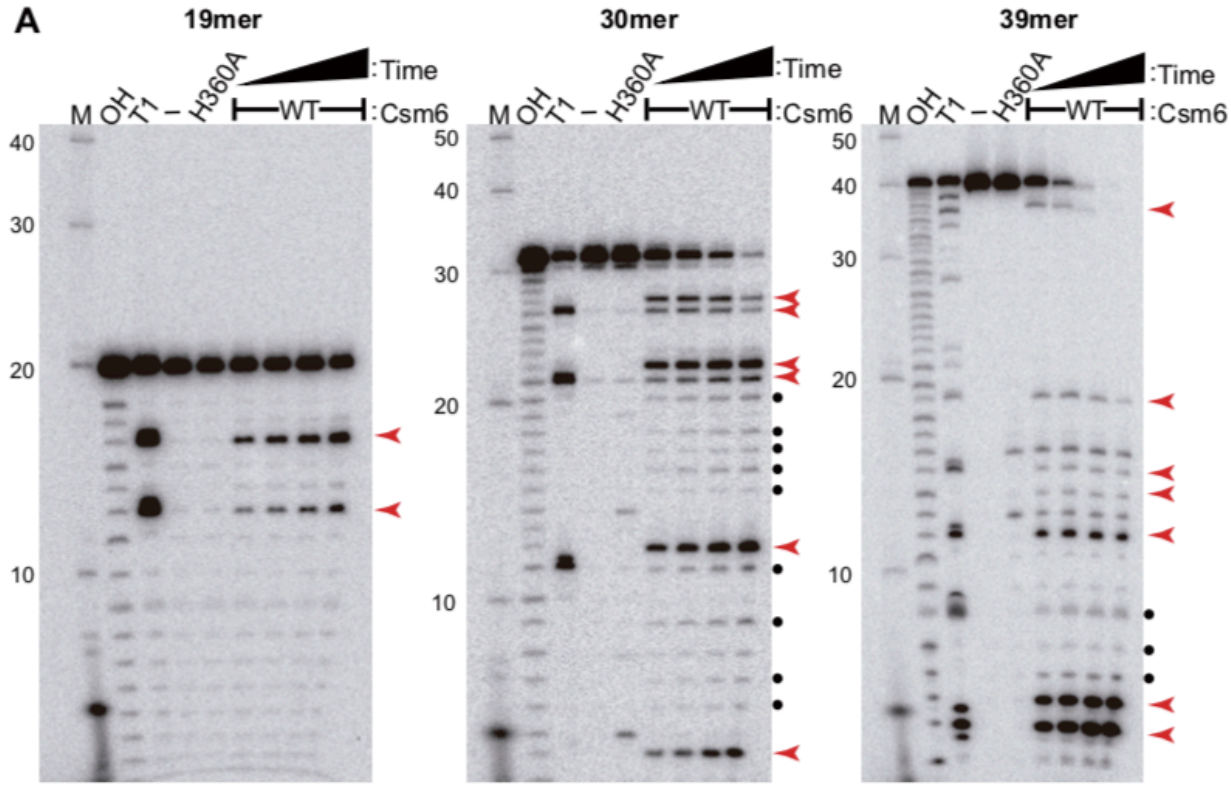

A
Linear
Circular
M
–
WT
H360A
–
WT
H360A
:Csm6
150
100
60
40
30
20
10
B
:EDTA
M
–
–
WT
:Csm6
50
40
30
25
20
15
10
C
- PAP
+ PAP
M
–
WT
–
WT
:Csm6
150
100
60
50
40
30
20
10
D
Linear
Circular
Circular
–
+
–
–
+
+
:TEX
M
–
–
–
WT
–
WT
:Csm6
50
40
30
25
20
15
10

Figure 2.3. *L. lactis* Csm6 cleaves RNA with a preference for adenosines and guanosines. (A) Three different RNA sequences (19mer (RNA1), 30mer (RNA5), and 39mer (RNA7); sequences given in panel B and Table S2.1) were incubated with no protein (-) or wildtype or H360A mutant *L. lactis* Csm6 for 5, 15, 30, and 60 minutes (a 60-minute incubation time was chosen for mutant Csm6 and no protein control). The RNAs were resolved by denaturing sequencing gel electrophoresis alongside 5' radiolabeled RNA markers (M), RNase T1 ladders (T1) and alkaline hydrolysis ladders (OH). Vertical bars show cleavage sites. Nucleotides at major cleavage sites are shown in red bold and with red arrows in panel A. Nucleotides at minor cleavage sites are shown in black bold and with black dots in panel A. Results shown are representative of multiple repeats of this experiment. RNA1 was tested six times and RNA7 was tested two times.

A

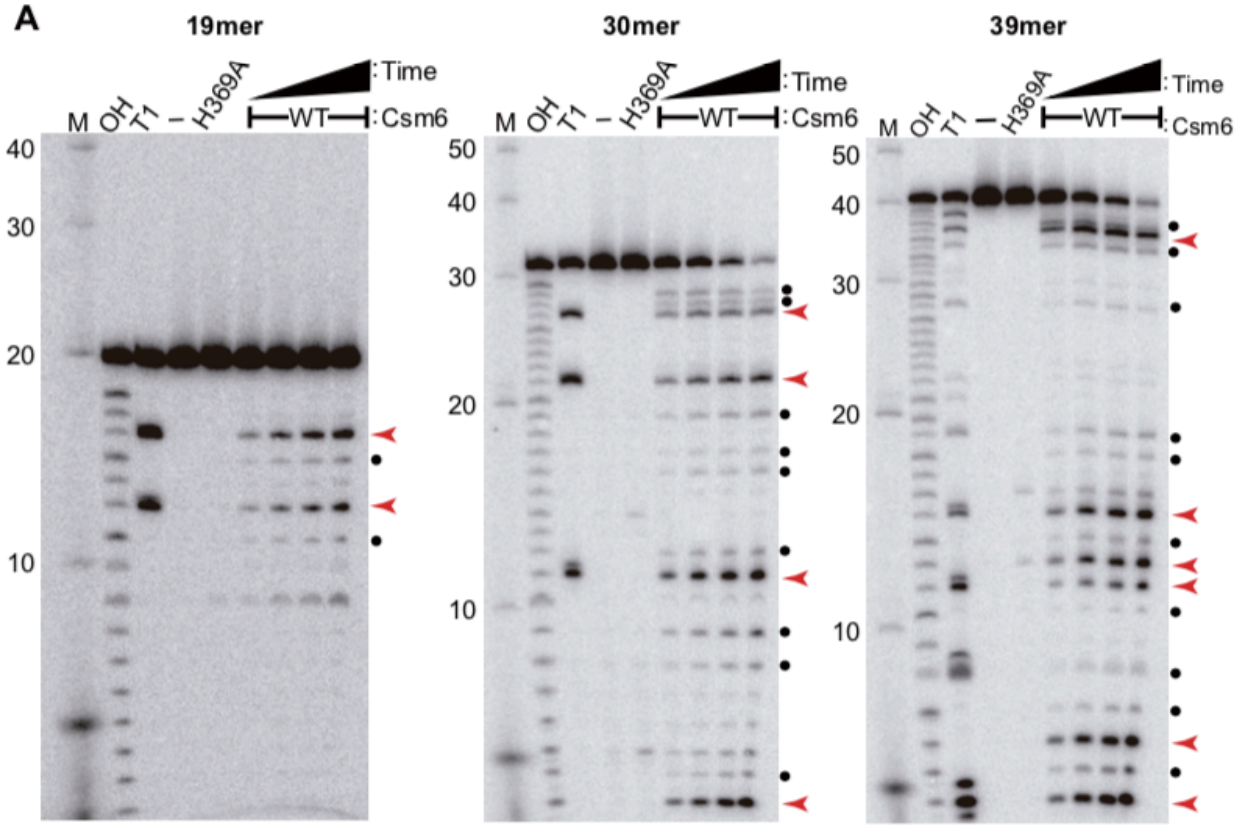


B

	Lactococcus lactis Csm6
Substrate	**Sequence (5'-3')**
19mer	CCCCCCCCCCAUG AUG AUG
30mer	GUUA CA A UA AG A CCA A A A UA G A AUUG A AAG
39mer	AUUG A A A G UUG UA G UAUG CGGUCCUUGCGGCUGAGA GCA

Figure 2.4. *S. epidermidis* Csm6 cleaves RNA with a preference for adenosines and guanosines. Three different RNA sequences (19mer (RNA1), 30mer (RNA5), and 39mer (RNA7); sequences given in panel B and Table S2.1) were incubated with no protein (-) or wildtype or H369A mutant *L. lactis* Csm6 for 5, 15, 30, and 60 minutes (a 60-minute incubation time was chosen for mutant Csm6 and no protein control). The RNAs were resolved by denaturing sequencing gel electrophoresis alongside 5' radiolabeled RNA markers (M), RNase T1 ladders (T1) and alkaline hydrolysis ladders (OH). Red arrows indicate major and black dots indicate minor Csm6-specific RNA cleavage products. (B) Cleavage products were mapped back to each RNA. Vertical bars show cleavage sites. Nucleotides at major cleavage sites are shown in red bold and with red arrows in panel A. Nucleotides at minor cleavage sites are shown in black bold and with black dots in panel A. Results shown are representative of multiple repeats of this experiment. RNA1 and RNA5 were tested three times and RNA7 was tested one time.

A



B

	Staphylococcus epidermidis Csm6
Substrate	**Sequence (5'-3')**
19mer	CCCCCCCCCCA**U** **G** A**U** **G** AUG
30mer	GU**U** **A** CAA**U** **A** A**G** **A** CCA**A** **A** A**U** A**G** AAUU**G** **A** **A** AG
39mer	AUU**G** **A** **A** **A** **G** U**U** **G** **U** **A** **G** UA**U** **G** CGGUCCUU**G** CGGCU**G** A**G** **A** GCA

Figure 2.5. Csm6 RNase activity is required to support Type III-A anti-plasmid immunity. (A) *In vivo* plasmid interference assay. *E. coli* strains harboring plasmids (orange; chloramphenicol selectable) expressing Csm crRNPs from *L. lactis*, *S. epidermidis*, or *S. thermophilus* were transformed with target and non-target plasmids (blue; ampicillin selectable) as previously described [28]. The target plasmid contains a sequence (yellow) that is complementary to expressed crRNA while the non-target plasmid has a sequence (green) that lacks crRNA homology. Serial ten-fold dilutions of transformed cells are spotted onto plates containing chloramphenicol and ampicillin and CRISPR-Cas mediated plasmid loss (defense) is indicated by a reduction in colonies. (B-D) Plasmid interference assays were performed with Csm modules containing or lacking (ΔCsm6) wildtype Csm6 or an RNase-defective form of Csm6 (HEPN mutant) in *L. lactis* (B), *S. epidermidis* (C), and S. *thermophilus* (D) Type III-A systems. Results shown are representative of three different experiments per system.

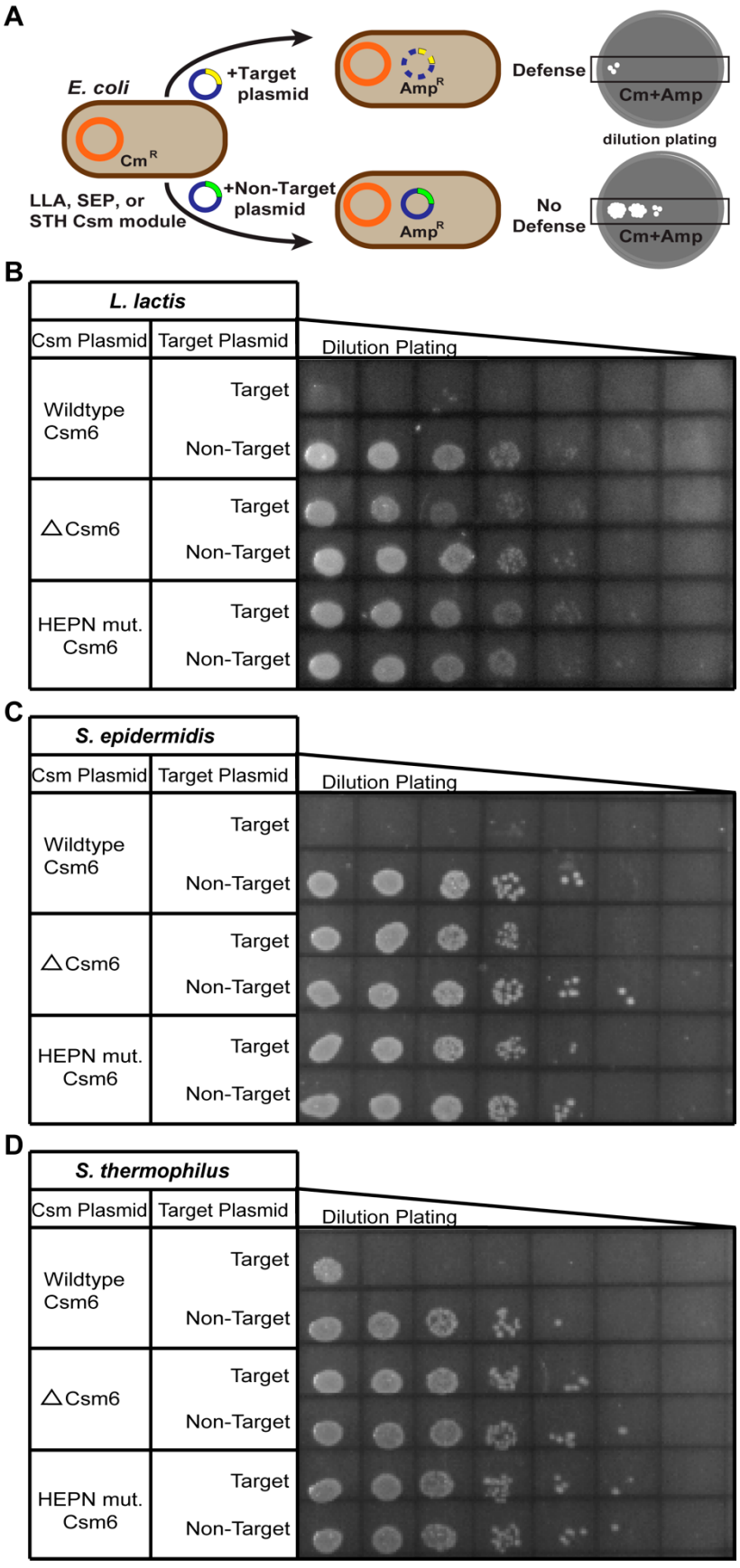

A
E. coli
Cm^R
+Target plasmid
Amp^R
Defense
Cm+Amp
dilution plating
LLA, SEP, or STH Csm module
+Non-Target plasmid
Amp^R
No Defense
Cm+Amp
B
L. lactis
Csm Plasmid
Target Plasmid
Dilution Plating
Wildtype Csm6
Target
Non-Target
△Csm6
Target
Non-Target
HEPN mut. Csm6
Target
Non-Target
C
S. epidermidis
Csm Plasmid
Target Plasmid
Dilution Plating
Wildtype Csm6
Target
Non-Target
△Csm6
Target
Non-Target
HEPN mut. Csm6
Target
Non-Target
D
S. thermophilus
Csm Plasmid
Target Plasmid
Dilution Plating
Wildtype Csm6
Target
Non-Target
△Csm6
Target
Non-Target
HEPN mut. Csm6
Target
Non-Target

Figure 2.6. Mutation of the palm domain of Csm1 prevents anti-plasmid immunity. Plasmid interference assays were carried out as described in Fig. 2.5 except that the effects of HD and palm domain mutations of the Csm1 protein were compared to that of wildtype Csm1. The effects of Csm1 mutations on plasmid defense were tested for *L. lactis* (A), *S. epidermidis* (B) and *S. thermophilus* (C) Type III-A systems. Results shown are representative of three different experiments per system.

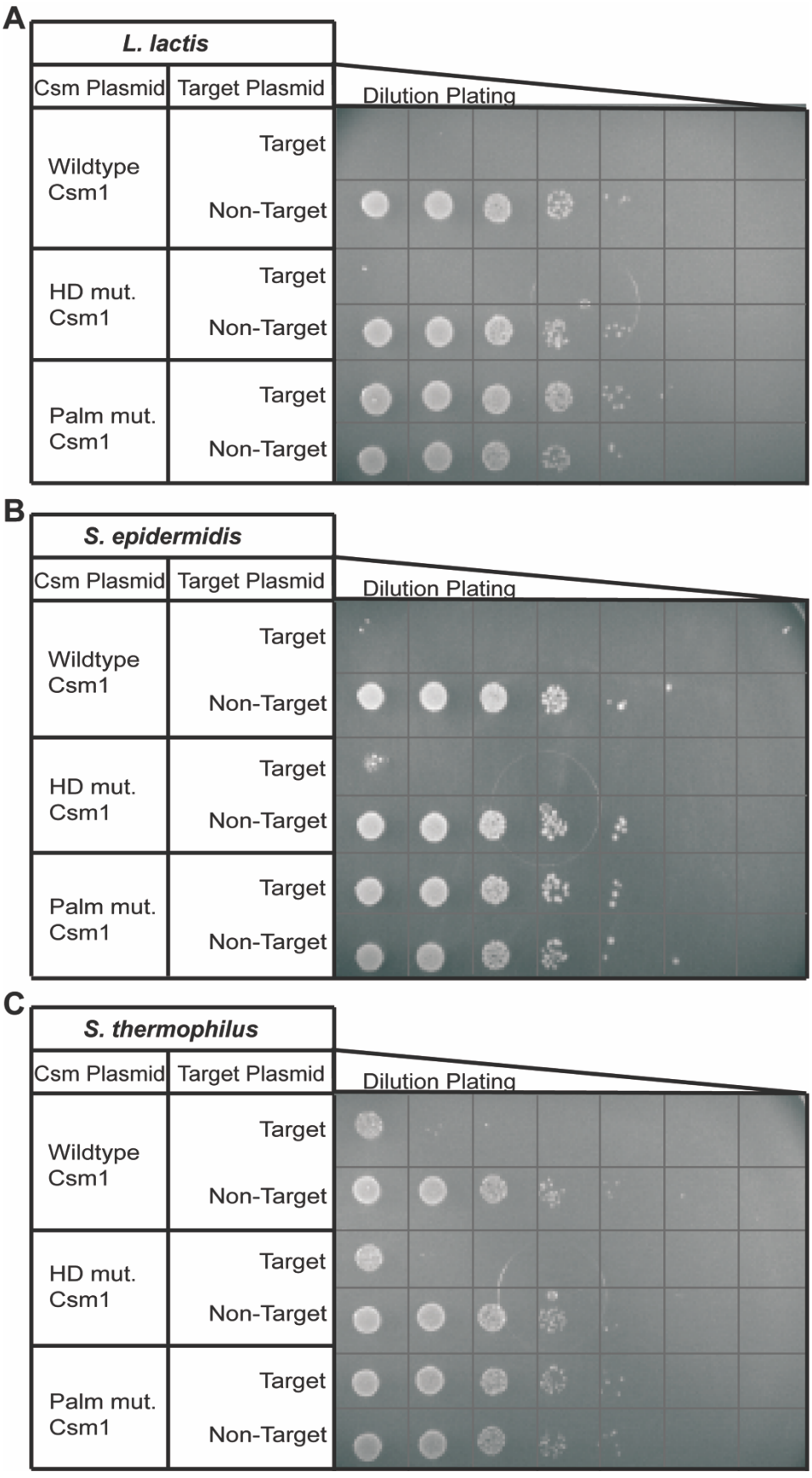

A
L. lactis
Csm Plasmid
Target Plasmid
Dilution Plating
Wildtype Csm1
HD mut. Csm1
Palm mut. Csm1
Target
Non-Target
B
S. epidermidis
C
S. thermophilus

Table 2.S1: RNA and DNA substrates used in this study

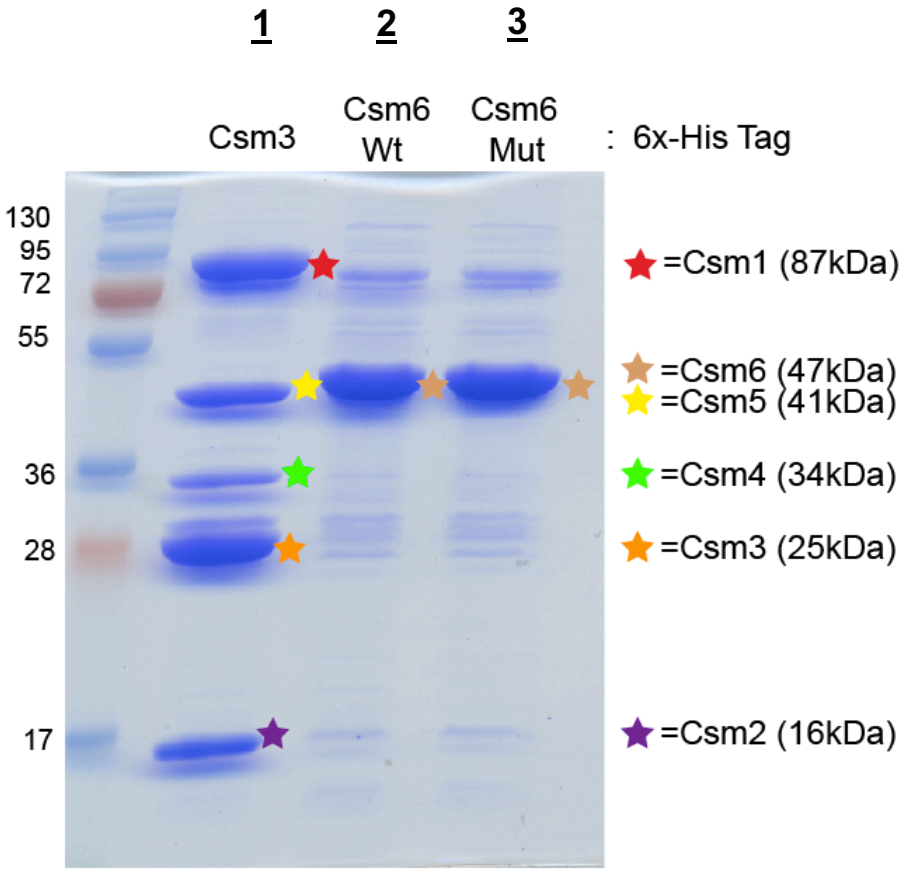
Name	Associated Figure(s)	Sequence (5'-3')
RNA1	1,3,4	CCCCCCCCCCAUGAUGAUG
RNA2	1	UUGUAGUAUGCGGUCCUUGCGGCUGAGAGCACU UCAG
RNA3	1	CUGAAGUGCUCUCAGCCGCAAGGACCGCAUACUA CAA
DNA1	1	CTGAAGTGCTCTCAGCCGCAAGGACCGCATACTAC AA
DNA2	1	TTGTAGTATGCGGTCCTTGCGGCTGAGAGCACTTC AG
RNA4	2	AUUGAAAGUUGUAGUAUGCGGUCCUUGCGGCUG AGAGCACUUCAGUCGUUAUCUCUUACGAAGUCUU
RNA5	2-4	GUUACAAUAAGACCAAAAUAGAAUUGAAAG
RNA6	3	AUGAGUAUUCAACAUUUCCGUGUCGCCCUUAUUC CCU
RNA7	3 & 4	AUUGAAAGUUGUAGUAUGCGGUCCUUGCGGCUG AGAGCA

Table 2.S2: Oligos Used for mutagenesis

Oligo	Alteration	Sequence (5'-3')
LLACsm6H360A-FP (2771)	Csm6 H360A	AGGACCGCAATAAAGTTGCAGCAAGTCT GCAGGCATTTGAC
LLACsm6H360A-RP (2772)	Csm6 H360A	GTCAAATGCCTGCAGACTTGCTGCAACTT TATTGCGGTCCT
SEPCsm6H369A-FP (2767)	Csm6 H369A	CCTGCGTAATAGCATTGCCGCAAACCTG GATACCCTGAAC
SEPCsm6H369A-RP (2768)	Csm6 H369A	GTTCAGGGTATCCAGGTTTGCGGCAATG CTATTACGCAGG
AIOCsm6_Tag/Mut#1 (3672)	LLA H360A &/or tag	ACTGGAAGAACTGAAATAACCTAGGCGC TTCAACGGAACG
AIOCsm6_Tag/Mut#2 (3673)	LLA H360A &/or tag	AACTGCGCTAATCAGGATTTTCATACCGC TGCTTCCATGGTGGTGATGATGATGCATT ATTTACCTCCTTTATCCATG
AIOCsm6_Tag/Mut#3 (3674)	LLA H360A &/or tag	ATGAAAATCCTGATTAGCGCAGTT
AIOCsm6_Tag/Mut#4 (3675)	LLA H360A &/or tag	TGCAACTTTATTGCGGTCCT
AIOCsm6_Tag/Mut#5 (3676)	LLA H360A &/or tag	ACCGCAATAAAGTTGCAGCAAGTCTGCA GGCATTTGAC
AIOCsm6_Tag/Mut#6 (3677)	LLA H360A &/or tag	GTTTTCCATGAGCAAACTGAAACGTTTTC ATCGCTCTGGAGTGA
Csm6_Del_InvPCR #1 (4419)	SEP & STH Csm6 del	TGCCCTACTTGGTCTCCGAGAC
Csm6_Del_InvPCR #2 (4420)	SEP & STH Csm6 del	CTCGAGAGGTTACAGCCTGCATAATG
STHCsm6H347A-FP (2769)	Csm6 H347A	CAGCCGTAATAAAGTTGCCGCAAGCCTG AGTCCGCTG
STHCsm6H347A-RP (2770)	Csm6 H347A	CAGCGGACTCAGGCTTGCGGCAACTTTA TTACGGCTG
WW_LLAAIO_mut2R (3186)	LLA H13A/D14A	GTGCCACGATAGATAATTTTGCCAATAGC AGCCAGCAGGCTACCACAAACC
WW_LLAAIO_mut3F (3187)	LLA H13A/D14A	ATTGGCAAAATTATCTATCGTGGCAC

WW_LLAAIO_mut4 R (3188)	LLA & D576A/D577A	CGCCACCGGCATAAATAAC
WW_LLAAIO_mut5 F (3189)	LLA D576A/D577A	GTTATTTATGCCGGTGGCGCCGCCCTGT TTATGATTGGTGCATGGC
P1_revision (3197)	LLA H13A/D14A & D576A/D577A	AAGCAACCTCGAGGCTGTGGTCTA
P6_revision (3198)	LLA H13A/D14A	AAAAGCCATATGTTAGCGTTCACG
SthSOE_P1 for (3892)	STH H15A/D16A & D575A/D576A	GGGAGAGGATCCATAAAGGAGG
SthSOE_P2 rev (3893)	STH H15A/D16A	TGCACGCTGAATAACTTTACCAATGGCTG CCAGCAGTGCACCATAGAACAG
SthSOE_P3 for (3894)	STH H15A/D16A	ATTGGTAAAGTTATTCAGCGTGCA
SthSOE_P4 rev (3895)	STH D575A/D576A	ACCACCGGCATAGATAATGGA
SthSOE_P5 for (3896)	STH D575A/D576A	TCCATTATCTATGCCGGTGGTGCAGCAG TTTTTGCAATTGGTAGCTGGC
SthSOE_P6 rev (3897)	STH H15A/D16A & D575A/D576A	CGCAGCCATATGTTAATCTTTACG
Sep csm1F (5043)	SEP H14A/D15A & D586A/D587A	CCTCTAGTGGCTGGCTAAG
Sep csm1R (5044)	SEP H14A/D15A & D586A/D587A	CGCAGCCATATGTTAGTCG
Sep ggdd F (5045)	SEP D586A/D587A	GCAGCACTGTTTCTGATTGGTGCATGG
Sep ggdd R (5046)	SEP D586A/D587A	TCCGCCACTATAAATTGCGGTAA
Sep hd F (5047)	SEP H14A/D15A	GCAGCAATTGGCAAAATTATCTATCGTAG CG
Sep hd R (5048)	SEP H14A/D15A	CAGCAGGCTGCCATACAT

Figure 2.S1: Purification of tagged Csm proteins. Purification of Csm proteins using co-expression plasmids that contain a 6x-His tag on either Csm3 or Csm6 following SDS-Page and Coomassie staining. Plasmids are expressed in *E. coli* BL21-AI cells and protein purification was performed as detailed on page 86 except 1L of LB broth and 0.2% arabinose were used for expression and induction, respectively. The first lane shows that only Csm1-5 co-purify when Csm3 is His-tagged. Lanes 2 and 3 show that no additional proteins co-purify with Csm6 when either Csm6 wildtype or mutant is His-tagged.



REFERENCES

1. Jackson, R. N., van Erp, P. B., Sternberg, S. H., and Wiedenheft, B. (2017) Conformational regulation of CRISPR-associated nucleases. *Curr Opin Microbiol* **37**, 110-119
2. Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., and Koonin, E. V. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-736
3. Jackson, R. N., and Wiedenheft, B. (2015) A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. *Mol Cell* **58**, 722-728
4. Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., and Koonin, E. V. (2015) Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol Cell* **60**, 385-397
5. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, C., Nwokeoji, A. O., Dickman, M. J., Horvath, P., and Siksnys, V. (2014) Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. *Mol Cell* **56**, 506-517

6. Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M., and Terns, M. P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945-956

7. Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J Bacteriol* **196**, 310-317

8. Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol Microbiol* **87**, 1088-1099

9. Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., and Terns, M. P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev* **30**, 447-459

10. Estrella, M. A., Kuo, F. T., and Bailey, S. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Dev* **30**, 460-470

11. Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* **161**, 1164-1174

12. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C., and Siksnys, V. (2016) Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Mol Cell* **62**, 295-306

13. Jiang, W., Samai, P., and Marraffini, L. A. (2016) Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* **164**, 710-721

14. Staals, R. H., Zhu, Y., Taylor, D. W., Kornfeld, J. E., Sharma, K., Barendregt, A., Koehorst, J. J., Vlot, M., Neupane, N., Varossieau, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P. J., Urlaub, H., Heck, A. J., Nogales, E., Doudna, J. A., Shinkai, A., and van der Oost, J. (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. *Mol Cell* **56**, 518-530

15. Goldberg, G. W., Jiang, W., Bikard, D., and Marraffini, L. A. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature*

16. Peng, W., Feng, M., Feng, X., Liang, Y. X., and She, Q. (2015) An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res* **43**, 406-417

17. Han, W., Li, Y., Deng, L., Feng, M., Peng, W., Hallstrom, S., Zhang, J., Peng, N., Liang, Y. X., White, M. F., and She, Q. (2016) A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. *Nucleic Acids Res*

18. Benda, C., Ebert, J., Scheltema, R. A., Schiller, H. B., Baumgartner, M., Bonneau, F., Mann, M., and Conti, E. (2014) Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. *Mol Cell* **56**, 43-54

19. Taylor, D. W., Zhu, Y., Staals, R. H., Kornfeld, J. E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J. A. (2015) Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science*

20. Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol Cell* **58**, 418-430

21. Staals, R. H., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J., Doudna, J. A., Heck, A. J., Yonekura, K., van der Oost, J., and Shinkai, A. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. *Mol Cell* **52**, 135-145

22. Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., and White, M. F. (2013) Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol Cell* **52**, 124-134

23. Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Graham, S., Reimann, J., Cannone, G., Liu, H., Albers, S. V., Naismith, J. H., Spagnolo, L., and White, M. F. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol Cell* **45**, 303-313

24. Spilman, M., Cocozaki, A., Hale, C., Shao, Y., Ramia, N., Terns, R., Terns, M., Li, H., and Stagg, S. (2013) Structure of an RNA silencing complex of the CRISPR-Cas immune system. *Mol Cell* **52**, 146-152

25. Millen, A. M., Horvath, P., Boyaval, P., and Romero, D. A. (2012) Mobile CRISPR/Cas-mediated bacteriophage resistance in Lactococcus lactis. *PLoS One* **7**, e51663

26. Marraffini, L. A., and Sontheimer, E. J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843-1845

27. Marraffini, L. A., and Sontheimer, E. J. (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568-571

28. Ichikawa, H. T., Cooper, J. C., Lo, L., Potter, J., Terns, R. M., and Terns, M. P. (2017) Programmable type III-A CRISPR-Cas DNA targeting modules. *PLoS One* **12**, e0176221

29. Cao, L., Gao, C. H., Zhu, J., Zhao, L., Wu, Q., Li, M., and Sun, B. (2016) Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of Staphylococcus aureus. *Int J Med Microbiol* **306**, 686-696

30. Ramia, N. F., Tang, L., Cocozaki, A. I., and Li, H. (2014) Staphylococcus epidermidis Csm1 is a 3'-5' exonuclease. *Nucleic Acids Res* **42**, 1129-1138

31. Jung, T. Y., An, Y., Park, K. H., Lee, M. H., Oh, B. H., and Woo, E. (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. *Structure* **23**, 782-790

32. Liu, T. Y., Iavarone, A. T., and Doudna, J. A. (2017) RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. *PLoS One* **12**, e0170552

33. Park, K. H., An, Y., Jung, T. Y., Baek, I. Y., Noh, H., Ahn, W. C., Hebert, H., Song, J. J., Kim, J. H., Oh, B. H., and Woo, E. J. (2017) RNA activation-independent DNA targeting of the Type III CRISPR-Cas system by a Csm complex. *EMBO Rep* **18**, 826-840

34. Silas, S., Lucas-Elio, P., Jackson, S. A., Aroca-Crevillen, A., Hansen, L. L., Fineran, P. C., Fire, A. Z., and Sanchez-Amat, A. (2017) Type III CRISPR-Cas

systems can provide redundancy to counteract viral escape from type I systems. *Elife* **6**

35. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksnys, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609

36. Niewoehner, O., Garcia-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., and Jinek, M. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature*

37. Han, W., Pan, S., Lopez-Mendez, B., Montoya, G., and She, Q. (2017) Allosteric regulation of Csx1, a type IIIB-associated CARF domain ribonuclease by RNAs carrying a tetraadenylate tail. *Nucleic Acids Res* **45**, 10740-10750

38. Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W., and Marraffini, L. A. (2013) A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J Biol Chem* **288**, 27888-27897

39. Zhu, X., and Ye, K. (2015) Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. *Nucleic Acids Res* **43**, 1257-1267

40. Hale, C. R., Cocozaki, A., Li, H., Terns, R. M., and Terns, M. P. (2014) Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. *Genes Dev* **28**, 2432-2443

41. Ramia, N. F., Spilman, M., Tang, L., Shao, Y., Elmore, J., Hale, C., Cocozaki, A., Bhattacharya, N., Terns, R. M., Terns, M. P., Li, H., and Stagg, S. M. (2014) Essential structural and functional roles of the Cmr4 subunit in RNA cleavage by the Cmr CRISPR-Cas complex. *Cell Rep* **9**, 1610-1617

42. Sheppard, N. F., Glover, C. V., 3rd, Terns, R. M., and Terns, M. P. (2016) The CRISPR-associated Csx1 protein of Pyrococcus furiosus is an adenosine-specific endoribonuclease. *RNA* **22**, 216-224

43. Niewoehner, O., and Jinek, M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA* **22**, 318-329

44. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013) Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* **8**, 15

45. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front Genet* **5**, 102

46. Yang, W. (2011) Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys* **44**, 1-93

47. Guan, J., Wang, W., and Sun, B. (2017) Chromosomal Targeting by the Type III-A CRISPR-Cas System Can Reshape Genomes in Staphylococcus aureus. *mSphere* **2**

48. Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., and Zhang, F. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573

CHAPTER 3

REGULATION OF THE RNA AND DNA NUCLEASE ACTIVITIES REQUIRED FOR PYROCOCCUS FURIOSUS TYPE III-B CRISPR-CAS IMMUNITY[1]

[1]Kawanda Foster, Sabine Grüschow, Scott Bailey, Malcolm F. White, Michael P. Terns. To be submitted to Nucleic Acids Research

ABSTRACT

Type III CRISPR-Cas prokaryotic immune systems provide anti-viral and anti-plasmid immunity via a dual mechanism of RNA and DNA destruction. Upon target RNA association, Type III crRNP effector complexes become activated to cleave both target DNA (via Cas10) and target RNA (via Cas7). Moreover, trans-acting endoribonucleases, Csx1 or Csm6, support the immune response elicited by some Type III complexes by destroying both invader and host RNAs. Here, we characterize how the RNAse and DNAse activities associated with Type III-B immunity in *Pyrococcus furiosus* (*Pfu*) are regulated by target RNA features and second messenger signaling events. *In vivo* mutational analyses reveal that both the DNAse activity of Cas10 and the RNAse activity of Csx1 can each direct effective anti-plasmid immunity. Biochemical analyses confirmed that the Cas10 Palm domain converts ATP into cyclic oligoadenylate (cOA) compounds that significantly enhance the ribonuclease activity of *Pfu* Csx1. Furthermore, the adenosine-specific endoribonuclease, *Pfu* Csx1, is capable of degrading cOA molecules to provide an auto-inhibitory off-switch of Csx1 activation. Activation of both the DNase and cOA generation activities require target RNA binding but are stimulated by overlapping but distinct target RNA 3' protospacer flanking sequences. Our results highlight the complex regulatory mechanisms controlling Type III-B immunity.

INTRODUCTION

CRISPR-Cas is a heritable immune system present in prokaryotes to protect against threats such as viruses and mobile genetic elements. The immune response is implemented in three major steps: adaptation, crRNA biogenesis, and defense [1,2].

During adaptation, a fragment invader DNA is incorporated into the CRISPR locus. During crRNA biogenesis, the CRISPR locus is transcribed and processed to produce mature crRNAs. During the final stage of defense, mature crRNAs interact with Cas proteins to form effector complexes. These effector complexes then complete the immune response by destroying the DNA and/or RNA of the invader [1,2]. CRISPR-Cas systems are classified into two main classes, each containing three major system types [3-5]. Class I CRISPR systems contain Types I, III, and IV; class II CRISPR systems contain Types II, V, and VI. Class I systems are characterized by containing multi-subunit effector complexes while class II systems utilize effector complexes that contain a single-subunit effector complex [3,4]. Invader DNA is targeted in Types I and II [6,7], RNA is targeted in Type VI systems [8], and DNA and RNA is targeted in Type III systems [9]. The targeting mechanisms of the Type IV system is still largely unknown [4,10,11].

Type III effector complexes are composed of several Cas proteins oriented around a Cas7 subunit backbone [12-19]. Type III systems are composed of four subtypes including two abundant major subtypes, III-A (Csm) and III-B (Cmr), and two relatively minor subtypes (III-C and III-D) [2]. Prior investigations into Type III systems have uncovered several activities of these systems including RNA and DNA degradation [9,17,20-26]. III-A effector crRNP complexes contain Csm 1-5 Cas proteins and II-B effector complexes contain Cmr 1-6 [15,17]. Type III systems specifically recognize target RNA protospacers and subsequently destroy DNA and RNA in a transcription-dependent manner [25,27,28]. The RNA targeting capabilities of Csm and Cmr crRNP complexes occur via a highly-conserved aspartic acid residue in the ribonuclease subunits (Csm3 or Cmr4) of Type III systems [16,17,26,29]. The DNA targeting activity requires a highly-

conserved HD nuclease domain in the Cas10 subunit of the complex (also known as Csm1 or Cmr2) [20,23,25,30]. The DNA degradation activity of Type III systems is nonspecific but requires activation by binding a specific target RNA that is complementary to the mature crRNA species present within the effector complex [9,22,24,31]. The intrinsic control mechanisms of Type III systems was further elucidated when studies found that that Type III systems produce a cyclic oligoadenylate (cOA) second messenger compound via the highly-conserved GGDD motif with the Palm domain of Cas10 [32,33]. Like DNA targeting, binding of the target RNA to the Type III effector complex is essential for production of cOA molecules [32,34,35]. The prevalence of this activity in Type III systems has been observed in several other Type III systems [19,34-38].

An additional layer of diversity within Type III systems pertains to the presence of the Csm6/Csx1 family of proteins that have been found to exhibit endoribonuclease activity [39-42]. Csm6 proteins are typically associated with Csm CRISPR systems and Csx1 proteins are typically associated with Cmr CRISPR systems [4]. Csm6 and Csx1 proteins share two highly conserved domains: the HEPN (Higher Eukaryotes and Prokaryotes Nucleotide binding) domain and the CARF (CRISPR Associated Rossman Fold) domain [43,44]. The ribonuclease activity of both Csx1 and Csm6 is facilitated by the HEPN motif [39-42]. Recently, the CARF domain was shown to facilitate activation of the ribonuclease activity of Csm6 and Csx1 proteins via binding of cOA molecules to the CARF domain of these two proteins [32-38].

Pyrococcus furiosus (*Pfu*) is a hyperthermophilic archaeon that contains three distinct functional CRISPR-Cas systems: I-A (Csa), I-G (Cst), and III-B (Cmr) [45,46] (Supplemental Figure 3.S1). Type I-A and I-G effector crRNP complexes are active

against DNA targets while the III-B Cmr crRNP complex targets the destruction of both DNA and RNA targets provided the target is transcribed [20,21,29,47,48]. Consistent with other Type III systems, the *Pfu* III-B system specifically cleaves bound target RNA in 6 nt intervals and nonspecifically cleaves target DNA in a transcription-dependent manner [20,29]. Furthermore, in addition to interaction with the target RNA protospacer through crRNA guide base-pairing, the DNA targeting activity of the *Pfu* Type III-B crRNP additionally requires the presence of a short three nucleotide sequence element located immediately 3' of the protospacers sequence within the activating target RNA. Previously formally referred to as the rPAM [RNA protospacer-adjacent motif] but now more generally indicated as the RNA PFS [protospacer flanking sequence], a term used in the field for type III and type VI RNA-targeting CRISPR systems [8,20,49]). Conflicting evidence in the Type III field exists in regards to whether Csm6 and Csx1 are crucial for Type III CRISPR-Cas defense [20,27,39,40,50]. Previous investigation into *Pfu* Csx1 found that deleting the *csx1* gene did not prevent anti-plasmid interference *in vivo* but a double mutation within the HD and Palm domains of *cmr2* abrogated immunity. The HD domain was confirmed to contribute to DNase activity but the contribution of the Palm domain remained unknown [20]. Other Type III systems have found Csm6/Csx1 family proteins to be important for defense under cellular conditions of low transcription [51]. Thus, we investigated the *Pfu* Cmr system further to determine if *Pfu* Csx1 is a key component of the Cmr defense response. We demonstrate that *Pfu* Csx1 is a contributing and sufficient component of Type III-B defense. Furthermore, we show that the *Pfu* Cmr complex utilizes the GGDD motif of the Palm domain of Cmr2 to convert ATP into cOA molecules that activate the ribonuclease activity of Csx1 and that Csx1 is also capable of

auto- regulating RNAse activity through specific recognition and cleavage of the cOA signaling molecules . The effects of altering the three nucleotide PFS elements adjacent to the RNA protospacer revealed that only specific subsets of purine-rich sequences trigger activation of the DNAse or cOA generation capabilities of the Cmr2 protein and show that the two important activities for immunity are differentially controlled by distinct molecular mechanisms.

MATERIALS AND METHODS

P. furiosus strains and growth conditions

All *P. furiosus* strains utilized in this study are listed in Supplemental Table 3.S1. Strains were grown at 90°C under strict anaerobic conditions using defined medium. Defined media for cultures and plates were prepared as previously detailed [52]. Cultures were grown in either 5 or 20 mL volumes and inoculated with either 1% inoculum or a single isolated colony. Cultures were incubated for 16-24 hours and plates were incubated for 64 hours. Uracil (20 μM) and/or 5-fluoroorotic acid (5-FOA, 2.75 mM) were supplemented in the media for selective conditions.

Newly generated *P. furiosus* strains were produced using homologous recombination of transformed SOE-PCR (splicing by overlap extension polymerase chain reaction) constructs as previously reported [20]. Complementation strains were generated by adding a modified wildtype *cmr2* or *csx1* gene back onto the *P. furiosus* genome that contained restriction sites that enabled detection of the introduced genes from wildtype genes and did not effect the coding potential (Supplemental Figure 3.S4).

A BclI-HF (5'-TGATAA-3' → 5'-TGATCA-3') restriction site was introduced into *cmr2* and a NruI-HF (5'-TCGGGA-3' → 5'-TCGCGA-3') restriction site was introduced into *csx1*. Primers used to make the complementation strains can be found in Supplemental Table 3.S3. All newly generated strains underwent at least three rounds of strain purification using minus uracil selective media. Successful strain generation was confirmed via PCR gene amplification and DNA sequencing (Eurofins Genomics).

Recombinant protein expression and purification

The genes encoding *P. furiosus* Cmr1-6 and Csx1 proteins were amplified via PCR and cloned into modified versions of pET24-D (Cmr4, Cmr5, Csx1), pET101-D (Cmr1-1), and pET200-D(Cmr2, Cmr3, Cmr6) as previously detailed [29,42]. All constructs contain a 6x-histidine tag on either the N-terminal (Cmr2-6, Csx1) or C-terminal (Cmr1-1) of the corresponding gene. Recombinant protein expressions were performed in *E. coli* BL21-RIPL cells (DE3,Novagen). Expression cultures for wildtype and mutant proteins were grown in 1L (Cmr1-1, Cmr4, Cmr5, Csx1), 2L (Cmr2, Cmr3), or 4L (Cmr6) batches at 37°C. Luria broth (Cmr1-1, Cmr2 – Cmr5, Csx1; RPI) or Terrific broth (Cmr6;RPI) medium was used for cultures and supplemented with either 50 µg/mL kanamycin sulfate (Cmr2 – Cmr6, Csx1) or 100 µg/mL ampicillin (Cmr1-1) for selection. Cultures were grown to an OD_{600} of 0.7 at 37°C then induced with 0.5 mM iso-propyl-β-D-thiogalactopyranoside (IPTG) at 24°C overnight. Cells were pelleted then resuspended in lysis buffer (40 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM Imidazole) containing one protease inhibitor tablet (Roche) and lysed via sonication. Cell lysates underwent a thermal precipitation by incubating in a 75°C bead bath for 20 minutes. Insoluble material was removed by

centrifugation at 14,000 rpm for 20 minutes at 4°C and filtered through a 0.8 μM syringe filter (Corning Incorporated). Proteins were purified using gravity affinity chromatography and either Ni-NTA resin (Cmr2, Cmr3, Cmr4, Cmr5, Cmr6, Csx1; Thermo Scientific) or Talon Cobalt resin (Cmr1-1; Clontech). Cell lysates were rotated with pre-rinsed and equilibrated resin for 1 hour at 4°C. The proteins were then washed with the lysis buffer and wash buffer (40 mM Tris-HCl (pH 7.5), 500 mM NaCl, 20 mM Imidazole). The proteins were then eluted using four different elution buffers containing increasing amounts of imidazole (50, 100, 200, 500 mM). Wildtype and HEPN mutant Csx1 proteins were dialyzed and underwent a second round of gravity affinity chromatography. Buffer exchange was performed using Slide-A-Lyzer Dialysis Cassettes (Thermo Scientific) in elution buffer with 0 mM Imidazole. Protein concentrations were assessed using Qubit protein concentration assays (Invitrogen) and purity was assessed via SDS-Page and Coomassie blue staining analysis.

Recombinant Csx1 mutagenesis

The wildtype gene encoding Csx1 (PF1127) was cloned into a modified pET24D vector. The Csx1 HEPN mutant (Csx1-HEPN$_m$) contains a H436A mutation and was created as previously described [42]. CARF domain mutants were created using the wildtype vector and either inverse PCR or Quikchange site directed mutagenesis (Stratagene). Mutagenesis primers are provided in Supplemental Table 3.S3. Amino acid residues 121-127 were deleted from Csx1 using inverse PCR to create a mutant form of the protein predicted to not be able to organize a functional CARF motif (Csx1-CARF$_m$). Quikchange was used to create site-specific mutations within the cOA binding pocket of

the CARF domain of Csx1. Two cOA binding mutants were created: Csx1-INAA (I169A,N170A) and Csx1-INQQ (I169Q,N170Q). Successful mutagenesis for all mutant plasmids was confirmed via DNA sequencing (Eurofins Genomics).

Purification of Csx1-HEPN$_m$ was performed as described above. Purification of wildtype, Csx1-CARF$_m$, Csx1-INAA, and Csx1-INQQ proteins was performed using a batch method. The batch method involved affinity purifying the proteins by incubating the soluble lysate with Ni-NTA resin for 1 hour at 4°C then performing subsequent washes and elutions in a 15 mL centrifuge tube and spinning at 4,000 rpm for 2 mins in between each step. Protein concentrations were assessed using Qubit assays and purity was assessed via SDS-Page and Coomassie blue staining analysis.

Preparation of ssRNA and dsDNA substrates

Synthetic RNAs (7.01 crRNA and 7.01 Target RNA) were purchased from Integrated DNA Technologies and DNAs from Eurofins Genomics. The RNA and DNA sequences can be found in Supplemental Table 3.S3. Target RNA was 5'-end labeled using γ-^{32}P-ATP, gel purified, eluted, extracted, and precipitated as previously described [42]. 7.01 crRNA was gel purified, eluted, extracted, and precipitated prior to using in assays. Double-stranded (ds)DNA substrate used in Figure 3.2 was made by first gel purifying the single-stranded oligos (DNA Target 1 & 2) on 7M urea denaturing 15% polyacrylamide gels. The single-stranded (ss)DNAs were then eluted, extracted, and precipitated. DNA Target 2 was then 5'-end labeled, gel purified, eluted, extracted, and precipitated as previously described [42]. Radiolabeled DNA Target 2 was then annealed to unlabeled DNA Target 1. DNA Target 2 was mixed with a 1.3x molar excess of DNA

Target 1 in annealing buffer (100 mM Tris (pH 8), 1 M NaCl, 10 mM EDTA) and incubated for 5 minutes at 95°C, followed by temperatures decreasing by 1° each minute, down to 22°C. The dsDNA product was then gel purified on a non-denaturing 10% polyacrylamide gel alongside a radiolabeled ssDNA (DNA target 2) control to confirm annealing. The product was then eluted in DNA elution buffer (40 mM Tris-HCl (pH 7.5), 500 mM NaCl) then extracted and precipitated using phenol/chloroform/isoamyl alcohol (PCI) and EtOH.

In vitro transcribed target RNAs were created to activate the DNase, RNase, and cOA production activities of the *Pfu* Cmr crRNP. The target RNAs were produced as previously reported [20]. A DNA template with a T7 phage promoter sequence was generated by incubating a target plasmid listed in Supplemental Table 3.S2 with IVT primers listed in Supplemental Table 3.S3. The PCR products were then gel extracted using the Zymoclean Gel Recovery Kit (Zymo Research) and MEGAshortscript T7 kit from Invitrogen was used to produce the target RNAs which were subsequently were gel purified from denaturing gels, eluted, extracted, and precipitated prior to adding to the assays. Target RNA concentrations were determined using the Qubit RNA BR Assay Kit (Invitrogen) and quality was assessed using 7M urea denaturing 15% polyacrylamide gels and ethidium bromide staining (Supplemental Figure 3.S2).

Plasmid silencing assay

The plasmid silencing assays were performed similarly to a previously described process [20]. Liquid cultures of *P. furiosus* were allowed to reach mid-to-late log phase of growth. 100 μL of liquid culture was transformed with 1 ng of either Target (pJE65) or No Target (pJE47) control plasmid. The transformations were incubated for 15-45 minutes at

room temperature prior to plating. Each transformation mixture was split between two plates and spread onto solid defined media lacking uracil. The plates were incubated at 90°C in an anaerobic chamber. Plates were observed for colony growth and counted after 64 hours of incubation. Results shown represent two replicates.

In vitro Cmr activity assays

Nuclease activity assays were performed similarly to methods previously described [20,29]. Recombinant Cmr proteins were purified individually as described above and first assembled with the 7.01 crRNA to form the crRNP. For RNase assays, the Cmr crRNP was assembled by preincubating 500 nM of each Cmr protein (50 nM of Cmr2) with RNA assay buffer (20 mM Tris-HCl (pH 7.5), 250 mM NaCl, 1.5 mM $MgCl_2$), and 12.5 nM of 7.01 crRNA for 25 minutes at 70°C. After the preincubation, 0.5–1.5 nM of radiolabeled 7.01 Target RNA was added to the reaction and incubated for one hour at 70°C. 1 unit of Proteinase K (NEB) was then added to each reaction and incubated for 30 minutes at 37°C prior to gel electrophoresis. For DNase assays, the Cmr crRNP was assembled by preincubating 500 nM of each Cmr protein (50 nM of Cmr2) with DNA assay buffer (20 mM Tris-HCl (pH 7.5), 250 mM NaCl, 1.5 mM $MgCl_2$, 200 μM $NiCl_2$), and 50 nM of 7.01 crRNA for 25 minutes at 70°C. After the preincubation, 100 nM of 7.01 Target RNA and 1 nM of radiolabeled dsDNA was added to the reaction and incubated for one hour at 70°C. Unless otherwise indicated, DNase assays were completed with target RNA containing a 5'-GGG-3' PFS sequence. One unit of Proteinase K was then added to each reaction and incubated for 30 minutes at 37°C prior to gel electrophoresis.

Cyclic oligoadenylate production assays were performed by assembling the Cmr crRNP as described above with 500 nM of each Cmr protein (50 nM of Cmr2), RNA assay buffer, and 50 nM of 7.01 crRNA. After assembly, 0.5 mM of ATP (NEB), 5 nM of α-^{32}P-ATP (3000 Ci/mmol; Perkin Elmer), and 100 nM of target RNA was added to the reaction and incubated for one hour at 70°C. Unless otherwise indicated, cOA production assays were completed with target RNA containing a 5'-GGG-3' PFS sequence. All Reactions were stopped by adding Gel Loading Buffer II (Life Technologies) and visualized by using 7M urea denaturing 15% polyacrylamide gels followed by autoradiography. Cyclic oligoadenylate reactions were also ran on 8M denaturing 20% polyacrylamide sequencing gels. Decade Markers (Life Technologies) and partial alkaline hydrolysis ladders (Ambion) were generated as previously described [39,42].

In vitro Csx1 activation assays

Activation with the native Cmr produced cOA

Csx1 activity assays were performed as previously reported with modifications [42]. Ribonuclease activity of Csx1 was assessed by incubating 500 nM of Csx1 with 0.5–1.5 nM of radiolabeled target RNA in assay buffer (20 mM Tris-HCl (pH 7.5), 200 mM NaCl) 1 hour at 70°C. In order to assess activation of Csx1, 10 or 20 nM of Csx1 was incubated with radiolabeled target RNA and assay buffer in the presence of unlabeled cOA for 1 hour 70°C. Unlabeled native cOA was generated as described above by omitting α-^{32}P-ATP. The unlabeled cOA was then extracted similarly to published methods [53]. Five reaction volumes of phenol/chloroform/isoamyl alcohol (PCI, 125:24:1 at pH 4.5; Ambion) was added to the reaction and vortexed for 30 seconds. The mixture

was then centrifuged for 45 mins at 20,000 rpm at 4°C then the aqueous layer incubated with five reaction volumes of chloroform (Fisher Scientific) vortexed and centrifuged. The aqueous layer was extracted, aliquoted, and stored at -80°C in single use aliquots. All reactions were stopped by adding Gel Loading Buffer II and visualized by using 7M urea denaturing 15% polyacrylamide gels followed by autoradiography.

Activation with synthetic cOA_4 and cOA_6

Csx1 ribonuclease activation assays with cOA_4 and cOA_6 species was performed as described above except synthetic cOAs from BIOLOG Life Science Institute were used. Several concentrations of cOA_4 and cOA_6 were used ranging from 0.0185 pM to 1.85×10^4 pM per reaction as indicated in the figure legend.

Activation with Csx1-treated cOA

Csx1 activation with cOA or $A_4>P^{¥}$ was performed by first generating and extracting unlabeled cOA as described above. Cmr crRNP produced cOA was then incubated with 600 nM of wildtype Csx1 or a mock (no protein) control for 30 mins at 70°C. The PCI-Chloroform extracted products (as described above) were added at various dilutions indicated in the figure legend and incubated with 20 nM of Csx1, 5'-end labeled target RNA, and assay buffer for 1 hour at 70°C.

Mass spectrometry

Unlabeled cOA production assays were performed as described above except the reactions were incubated for 2 hours. Liquid chromatography high resolution mass

spectrometry (LC-HRMS) analysis was performed on a Thermo Scientific Velos Pro instrument equipped with HESI source and Dionex UltiMate 3000 chromatography system as previously described [36].

RESULTS

Pfu Csx1 is a key component of the Type III-B anti-plasmid immunity defense response

The *P. furiosus* (*Pfu*) Type III-B crRNP effector complex is composed of Cmr1-6 and a crRNA (Figure 3.1A). Previous *in vitro* studies with the Cmr complex demonstrated that the HD domain of Cmr2 exhibits nuclease activity while the GGDD motif of the Palm domain had no impact on nuclease activity [20]. Individual mutations of either the HD (H13A/D14A) domain or Palm (D673A/D674A) domain of Cmr2, or a single deletion of the *csx1* gene did not interfere with plasmid silencing but double mutations in the HD and Palm prevented immunity [20]. To determine if *Pfu* Csx1 is a critical component of the *Pfu* Cmr defense response, we generated double mutants in the context of a HD or Palm mutation. Eleven strains were generated to determine if *Pfu* Csx1 is required for plasmid silencing (Figure 3.1). After growing the *P. furiosus* strains, cells were transformed with either a target plasmid or a no target plasmid and plated on selective media (Figure 3.1B). The transformation plasmid confers uracil prototrophy on *P.furiosus* thus successful uptake of the plasmid (no CRISPR-Cas defense) will result in an abundance of colony formation while failed uptake of the plasmid (CRISPR-Cas defense) will result in little to no colony formation. As previously observed, Cmr defense was active and functioning in

the null, Cmr only, Cmr2-HD_m, Cmr2-$Palm_m$, and $Csx1_\Delta$ strains but not in the Cmr2-HD_m/$Palm_m$ double mutant strain (Figure 3.1C). Further analysis with additional strains demonstrated that CRISPR-Cas defense was lost when *Pfu* Csx1 was deleted or mutated in the presence of an Cmr2-HD mutation but not a Cmr2-Palm mutation (Figure 3.1C). Defense was interrupted for strains in which Csx1 was deleted or mutated within the HEPN (H436A) motif. Cis-complementation was performed to determine if CRISPR-Cas immunity could be rescued by restoring either wildtype *cmr2* or *csx1* in the Cmr2-HD_m/$Csx1_\Delta$ double mutant strain. For both complementation strains (Cmr2-HD_m/$Csx1_c$ and $Cmr2_c$/$Csx1_\Delta$) plasmid silencing activity was restored. The results reveal that the ribonuclease activity of Csx1 is an active contributor to plasmid silencing in the III-B system and is sufficient to implement a successful defense response.

Production of the cyclic oligoadenylate second messenger in the III-B Cmr system of *P. furiosus* is Palm domain and target RNA dependent

We tested the Type III-B system of *P.furiosus* for production of cyclic oligoadenylate compounds. To assay cOA production, we first confirmed that the Type III-B system was functioning properly by assaying for ribonuclease and nuclease activities as previously observed [20,29]. Five different functional mutants of the Cmr crRNP complex including wildtype, Cmr2-HD_m, Cmr2-$Palm_m$, Cmr4-D26N, and Cmr2-HD_m/$Palm_m$ were incubated with either 5'-end radiolabeled ssRNA or dsDNA (Figure 3.2A). As expected, RNAse activity was observed for all complexes except Cmr4-D26N and DNAse activity was only observed for wildtype, Cmr2-$Palm_m$, and Cmr4-D26N complexes (Figure 3.2B-C). To test for cOA production, the same five complexes were

incubated with α-^{32}P-ATP and analyzed using denaturing gel electrophoresis. Conversion of the α-^{32}P-ATP was only observed for complexes with an intact Palm domain. Additionally, presence of the target RNA was required for cOA production (Figure 3.2D). These findings indicate that the *Pfu* Type III system possesses the highly conserved activity of cOA production which is dependent upon interactions between the target RNA and the Palm domain.

Pfu Csx1 ribonuclease activity is activated by cyclic oligoadenylate second messenger compounds produced by the Cmr complex

Next we sought to determine if the cyclic oligoadenylate compound produced by *Pfu* Cmr is an activator of *Pfu* Csx1. Other Type III systems have shown that oligoadenylate compounds can be used to activate Csx1 and Csm6 [32,33,35,37,54]. Using a lower than usual concentration of wildtype *Pfu* Csx1 (20 nM), we assayed for ribonuclease activity by first incubating each Cmr crRNP complex with unlabeled ATP then deproteinizing the reactions with a phenol-chloroform extraction. The deproteinized reaction product solutions were then incubated with *Pfu* Csx1 and radiolabeled ssRNA. Cleavage of the ssRNA by Csx1 was only observed for reactions that contained product of reactions in which the Palm domain was present (Figure 3.2E). The results indicate that cOA production by the Cmr crRNP complex is functionally utilized to activate the RNA degradation capabilities of *Pfu* Csx1.

Residues within the CARF domain of *Pfu* Csx1 facilitate activation of the ribonuclease activity of Csx1 using cyclic oligoadenylate compounds

Pfu Csx1 contains two conserved domains: the HEPN and CARF domains (Figure 3.3A) [43,44,55]. The ribonuclease activity of Csx1 is executed by the HEPN domain and the CARF domain is the site of activation by cOA molecules [12,37,42,54]. To investigate if the CARF domain of *Pfu* Csx1 supports its RNase activity we started by determining the location of the *Pfu* Csx1 CARF domain using protein sequence and structural analysis (Figure 3.3A). The led us to hypothesize that two main putative regions contribute to the cOA compound successfully activating *Pfu* Csx1: the CARF motif folding region and the cOA binding region. To address this question, we tested 5 different recombinant Csx1 proteins including 3 CARF domain mutants (Figure 3.3B). *Pfu* Csx1 CARF domain mutants: $CARF_m$, INAA, and INQQ. Csx1-$CARF_m$ contains a deletion of residues 121-127, Csx1-INAA contains a double mutation of residues 169 and 170 to alanine (I169A,N170A), and Csx1-INQQ contains a double mutation of residues 169 and 170 to glutamine (I169Q,N169Q). The impact of these three mutations was tested *in vivo* and *in vitro* (Figure 3.3). To explore whether the CARF mutations would affect the ability of Csx1 RNase activity to be activated, we completed *in vitro* assays (Figure 3.3C-E). We first compared the activity of a low concentration of Csx1 (10nM) versus a high concentration of Csx1 (500nM) with cOA produced from wildtype *Pfu* Cmr complexes. When comparing the five different Csx1 proteins, only wildtype *Pfu* Csx1 was able to be activated by the cOA molecules at the low concentration (Figure 3.3C-E). These findings reveal relevant residues of the CARF domain in *Pfu* Csx1; residues 121-127 and 169-170 of *Pfu* Csx1 are utilized for activation of the CARF domain by the *Pfu* Cmr produced cOA molecules.

To determine if the CARF mutations impacted plasmid silencing, we incorporated CARF mutations into *P. furiosus*. Seven different strains were compared using the same approach as in Figure 1. We found that, when paired with Cmr2-HD_m, Csx1-$CARF_m$ and Csx1-INQQ both led to a loss in CRISPR defense (Figure 3.3F). These findings reveal important residues located with the CARF domain that assist in activation of *Pfu* Csx1.

Pfu Csx1 is activated by cOA_4 species produced by *P. furiosus* Cmr complexes

Type III systems have been found to produce multiple sizes of cOA molecules. Various cyclic oligoadenylate compounds have been observed that range from two to six nucleotides in size [32-37,56]. We sought to determine the sizes of cOA species produced by the *Pfu* Cmr complex. Using LC-HRMS analysis, cyclic oligoadenylate species produced from wildtype and Cmr4-D26N *Pfu* Cmr complexes were analyzed. The extracted ion chromatograms reveal that cOA_3 and cOA_4 are the most abundant species produced by *Pfu* Cmr complexes with no major differences between wildtype and mutant complexes (Figure 3.4A and Supplemental Figure 3.S5). The data reveal that cOA_4 is the most abundant species and is approximately twice as abundant as cOA_3; there were traces of cOA_5 and cOA_6 which made up than 5% of the cOAs. Next we explored the identity of the relevant *Pfu* Csx1 cOA activator. Prior studies have found that Type III complexes can produce multiple sized cOA species but not all can activate Csm6 or Csx1 [32,35,36]. Currently, only cOA_4 or cOA_6 have been found to be the relevant activator for Csm6 and Csx1 proteins. Additionally, given the homodimerization of *Pfu* Csx1 we hypothesized that cOA_3 would not be a biologically relevant activator. However since multiple different species are produced by the *Pfu* Cmr complexes, it is unknown which

cOA molecules are capable of activating *Pfu* Csx1. To test activation of Csx1 based on size, we completed Csx1 activation assays using synthetic cyclic nucleotides—c-tetraAMP and c-hexaAMP (Figure 3.4B). The dimerization of Csx1 and the commercial availability of pure cOA_4 and cOA_6 molecules led us to only test these two versions of the activator. Varying concentrations of synthetic cOA_4 and cOA_6 were incubated with *Pfu* Csx1 in the presence of 5'-end labeled ssRNA. Restoration of ribonuclease activity at levels comparable to the native *Pfu* Cmr activator was observed using cOA_4 but not cOA_6. A low level of cleavage activity was observed for reactions containing synthetic cOA_6. The results indicate that cOA_4 is the major metabolite produced by *Pfu* Cmr and the relevant ribonuclease activator for *Pfu* Csx1.

Degradation of cOAs by *Pfu* Csx1 inactivate the molecules as activators of *Pfu* Csx1

The ribonuclease activity of Csm6 and Csx1 proteins is robust and non-specific to CRISPR transcripts [39,41,42,51]. It is not well understood how the ribonuclease activity of Csm6 and Csx1 is regulated to prevent complete destruction of vital cellular RNAs once activated by cOAs. Regulation of the nonspecific and robust ribonuclease activity of Csm6 and Csx1 proteins has remained a major question in the study of Type III systems. Recently, a new class of distinct CARF-domain containing proteins discovered in *Sulfolobus solfataricus* have exhibited cOA nuclease activity that halts Csx1 activity [57]. Interestingly, *S. solfataricus* Csx1 was not able to degrade cOA molecules [57]. Previous characterization of *Pfu* Csx1 revealed an adenosine specificity for endoribonuclease activity [42]. That knowledge combined with the lack of a CARF domain-containing ring

nuclease homolog in *P. furiosus* lead us to hypothesize that *Pfu* Csx1 serves as its own ring nuclease. We explored this hypothesis by first testing whether incubation of Csx1 with native cOAs would result in a phenotype, indicating cOA cleavage, detectable by denaturing gel electrophoresis. We incubated wildtype and four mutant versions of Csx1 with radiolabeled cOAs (Figure 3.5A). Detectable shifts in the migration of the *Pfu* cOAs were observed for reactions containing Csx1, Csx1-CARF$_m$, Csx1-INAA, and Csx1-INQQ (Figure 3.5A). The most prominent shifts occurred with reactions containing Csx1 and Csx1-CARF proteins and no shifts were observed for the reactions containing Csx1-HEPN$_m$. Next we tested whether degradation of the cOAs would result in the inability of the oligoadenylates to serve as activators of *Pfu* Csx1 ribonuclease activity. In order to explore this question we treated cOA$_4$ with and without Csx1. Samples were then extracted and tested for activated ribonuclease activity (Figure 3.5B). Activation of *Pfu* Csx1 was observed for the majority of the reactions containing cOA while reactions containing A2>P$^{¥}$ did not result in activation. The results indicate an additional capability of Pfu Csx1 to degrade cOA molecules in an autoregulation feedback loop.

Activation of Cmr2-mediated nuclease activity and production of cOAs in the *Pfu* Cmr system is dependent upon distinct protospacer flanking sequences

Prior investigation into the defense mechanisms of the *P. furiosus* Cmr system revealed that DNA cleavage activation is not only target RNA dependent but also PFS (previously referred to as rPAM) dependent (Figure 3.6A) [20]. The PFS requirement was observed *in vivo* for defense assays as well as *in vitro* for DNA cleavage assays. Successful CRISPR-Cas defense or ssDNA cleavage required a PFS with either a NGN,

NNG, or NAA sequence [20]. Given this newly observed activity for *Pfu* Cmr, we questioned whether the first 3 positions of the PFS would be required for cOA production. In order to determine the impact of the PFS on cOA production, we generated 23 different *in vitro* RNAs with varying sequences within the first three positions of the PFS region. DNA cleavage and cOA production was tested in parallel for all 23 target RNAs. As expected, only the 17 target RNAs containing a NGN, NNG, or NAA protospacer flanking sequence activated DNA cleavage via the HD domain of Cmr2 (Figure 3.6B). However, only 7 of those 17 target RNAs activated cOA production (Figure 3.6C). The target RNAs eliciting strong cOA production activity contain a 5'-NGR-3' sequence and target RNAs eliciting weak cOA production contain a 5'-NRR-3'. A summary of all rPAM dependent activities are listed in Table 3.1. The results indicate that the PFS specificity requirement persists for cOA activation in the *Pfu* Cmr system but slightly differs than the requirements needed for activating the nuclease activity of Cmr2.

DISCUSSION

Previous and current investigations of the Type III system of *P. furiosus* has provided valuable insight into the mechanisms involved in Cmr defense [20,29,42]. Consistent with other Type III systems, *P. furiosus* Cmr employs a multi-pronged approach to successfully execute CRISPR-Cas immunity (Figure 3.7). Transcription-dependent activation of the *Pfu* Cmr complex results in four main enzymatic activities: ssRNA nuclease via the conserved aspartic acid residue Cmr4, ssDNA nuclease activity via the HD domain of Cmr2, cyclase activity via the GGDD motif of Cmr2, and activation of ssRNA nuclease activity via the HEPN motif of Csx1. Base-pairing between the crRNA

and target RNA results in multiple cleavage products that differ in size by 6 nucleotides (Figure 3.2B). Recognition of the appropriate protospacer flanking sequence triggers ssDNA nuclease and cOA generation activities by Cmr2 (Figure 3.2C-D). Production of cOA via the Palm domain leads to activation of the ribonuclease activity of Csx1 using residues located with the CARF domain of *Pfu* Csx1 (Figure 3.3). Activation of the ribonuclease activity of *Pfu* Csx1 can either lead to two modes for Csx1: viral/plasmid immunity or dormancy/cell death. In the cell death outcome, the nonspecific ribonuclease activity is unregulated and destroys any ssRNA including cellular RNA. However, during immunity mode, *Pfu* Csx1 autoregulates its own ribonuclease activity by degrading its own activator (Figure 3.7). In this study we demonstrate the autoregulating capability of *Pfu* Csx1 and further develop the role of the PFS in *Pfu* Cmr defense.

The role of *Pfu* Csx1 in Type III-B defense

Previous investigation into the role of *Pfu* Csx1 in anti-plasmid immunity failed to uncover its contributions to defense due to the limited number of *Pfu* mutant strains tested [20]. Upon further analysis, we found that Csx1 is an active contributor to Cmr defense (Figure 3.1C). When evaluating three of the main enzymatic activities of the Cmr defense machinery, we found that two are sufficient in executing defense: ssDNA cleavage by Cmr2-HD and ssRNA cleavage by Csx1. Interestingly, the cyclase activity of Cmr2-Palm was found to contribute to Cmr defense via the activation of Csx1 rather than DNA cleavage. Given that the Palm domain of Cmr2 was also found to not contribute to DNA cleavage *in vivo*, it appears that the sole function of the Palm domain of *Pfu* Cmr2 is to produce cyclic oligoadenylate compounds that activate the ribonuclease activity of Csx1.

These findings uncover a novel dual execution method for Type III systems *in vivo* in which the Cmr2 and Csx1 members of this system make equal contributions to immunity. Given that DNA cleavage by Cmr2 and RNA cleavage by Csx1 are alone sufficient to implement a successful defense response in *P. furiosus*, both components must be inhibited or eliminated for invaders to evade the immune system. These results differ from previously studied Type III systems that either observed a dominant role for Csm6 [39] or Csm1 [51]. In each case, Csm1 or Csm6 either did not contribute or made auxiliary contributions to defense. Here we demonstrate the requirement for *Pfu* Csx1; solidifying its prominent role as an essential component in Type III defense

P. furiosus Cmr employs multiple mechanisms for defense

The intrinsic ribonuclease activity of Type III effector complexes was originally found using the *Pfu* Cmr complex [21,29]. Early analysis conducted with *Pfu* Csx1 provided novel biochemical data confirming the adenosine-specific endoribonuclease activity of Csx1 [42]. Continued work with the *Pfu* Cmr complex revealed the existence of a unique PFS sequence [20]. CRISPR-Cas immunity, *in vivo,* as well as DNA targeting activities were found to be transcription and PFS dependent [20]. In this study, we contribute to the understanding of *Pfu* Cmr and Type III systems by reconstituting a new Type III behavior—cOA production (Figure 3.2). Generation of cOA was found to be Palm dependent which is consistent with previous works [32,33,35]. The *in vitro* results provide additional support to the results observed in Figure 3.1 that indicate a direct connection between the cyclase activities of the Palm domain and the ribonuclease activities of Csx1.

Identification of important residues within the *Pfu* Csx1 CARF domain

Upon observing Palm domain dependent activation of *Pfu* Csx1, we aimed to determine if residues with the CARF domain supported the activation of Csx1 and identify the relevant residues for activation. Unlike the HEPN motif, the critical residues for cOA binding within the CARF domain varies significantly amongst Csx1 and Csx6 homologs [33,41,54,56,58]. Thus we conducted mutational analysis to provide important insight into the functional CARF residues of *Pfu* Csx1. Our analysis indicated two putative regions within the CARF that impact *Pfu* Csx1 activation. Based on our analysis, it is likely that these residues contribute to CARF domain folding (residues 121-127) and cOA binding (residues 169 and 170). Mutations generated in both regions prevented activation of Csx1 *in vitro* (Figure 3.3C-D) as well as *in vivo* (Figure 3.3F). As expected, mutation of both CARF regions in the context of an HD domain mutation resulted in loss of plasmid immunity. These findings indicate that activation of *Pfu* Csx1 is needed to make Csx1 a suitable defense component.

Identification of the *Pfu* Cmr second messenger activator

Different cOA production profiles have been observed amongst Type III systems [32,35,36]. In order to assess the cOA profile for *P. furiosus* Cmr we utilized liquid chromatography high resolution mass spectrometry. Cyclic oligoadenylate compounds ranging in size from 3 to 6 nucleotides were detected (Figure 3.4A). The most abundant species were cOA_3 and cOA_4. Given the homodimerization of *Pfu* Csx1, we hypothesized that cOA_4 would be the relevant activator for Csx1 [55]. The use of synthetically produced cOA_4 and cOA_6 compounds confirmed that cOA_4 is a relevant and the prominent activator

for Csx1 (Figure 3.4B). Turning production of cOA molecules off has been shown to be connected to target RNA cleavage by Csm3/Cmr4 [34,36]. We evaluated the cOA profiles for wildtype and Cmr4-D26N mutant *Pfu* Cmr complexes and saw no detectable difference between the two complexes. However, only one time-point was analyzed for this study so it is likely that a difference between both complexes would be observed upon further analysis of additional timepoints.

Autoregulation of *Pfu* Csx1 ribonuclease activity

Given the robust nature of cOA production in the *Pfu* Cmr system even in the presence of the Cmr4-D26N mutation, we theorized that the degradation of cOA molecules by a ring nuclease would be a prominent feature of cOA regulation for *P. furiosus*. Currently, examples exist of Csm6 proteins degrading cOA compounds while evidence of Csx1 proteins degrading cOAs has not been observed [56-58]. In fact, recent structural and biochemical analysis of *Sulfolobus islandicus* Csx1 found that cOA_4 molecules bind to the CARF domain but *S. islandicus* Csx1 does not degrade the cOA_4 molecules [12]. We explored the ring nuclease functions of *Pfu* Csx1 and found that the enzyme exhibited the behavior (Figure 3.5). Upon comparing five different versions of *Pfu* Csx1 (wildtype, Csx1-$CARF_m$, Csx1-$HEPN_m$, Csx1-INAA, and Csx1-INQQ) we observed strong degradation of cOAs for wildtype protein after only 5 minutes (Figure 3.5A). After 30 minutes, we observed no degradation phenotype for the HEPN mutant and weakened degradation for the CARF mutants (Figure 3.5B). Collectively the results indicate a direct requirement of the HEPN motif and an indirect role of the CARF domain for the ring nuclease activity of *Pfu* Csx1. It is possible that the CARF mutations impact an important structural change

needed for cOA binding and thus results in weakening but not complete inhibition of the activity. This would explain why the residues that we theorize are required for cOA binding (I169N170) had a greater negative impact on the ring nuclease activity than the region we suspect to be important for folding of the CARF domain (residues 121-127).

The identity of the *Pfu* Csx1 cOA cleavage reaction is a remaining question of this analysis. Based on analysis conducted in similar homologs two linear ring nuclease products have been observed: A_2>P and A_4>P [56,57]. The ring nuclease products were found to contain a 2'3' cyclic phosphate denoted by '>P'. Reactions between ring nucleases and cOAs have detected A_4>P initially as an intermediate product and A_2>P as the final product over time. Given the time length utilized in this analysis, it is feasible that the identity of the product is A_4>P. Alternatively, it is possible that A_2>P is the product observed (Figure 3.5A). Activation with the linearized cleavage product failed to activate *Pfu* Csx1 ribonuclease activity (Figure 3.5B). It is possible that either a change in structure (conversion to A_4>P) or structure and size (A_4>P) inactivates the oligoadenylate as a suitable activator for *Pfu* Csx1. Previously analysis conducted in *S. islandicus* Csx1 demonstrated that a linear ligand consisting of four adenosines successful activated the ribonuclease activity of *S. islandicus* Csx1 but the same linear ligand with only two adenosines failed to activate the protein [54]. Thus, there is evidence for both possibilities and further analysis will be needed to determine the cleavage products for *Pfu* Csx1. Overall, the results indicate a self-regulation ability of *Pfu* Csx1 via degradation of cOA molecules. Given the nonspecific nature of *Pfu* Csx1, it is likely that this feature serves as a built in off switch to control the activity of Csx1 once activated in the defense response.

Contributions of the PFS to CRISPR-Cas immunity

The discovery of an PFS sequence for the *Pfu* Cmr system was a novel and insightful occurrence [20]. Here we aimed to utilize that knowledge and the discovery of the cOA production activity to determine if PFS selection played a role in the production of cOA compounds. Comparison of 23 different protospacer flanking sequences revealed three different outcomes depending on the sequence of the PFS (Figure 3.6). Either no CRISPR defense was triggered, or defense with only DNAse activity was triggered, or defense with DNAse and cOA generation activities were triggered. All PFSs that triggered CRISPR defense *in vivo* also resulted in DNAse activity *in vitro*; however, only a subset of those PFSs resulted in cOA production *in vitro* (Table 3.1). The purine specificity within position +2 and +3 of the PFS in addition to the difference in activation of DNAse and cOA synthesis reveals an additional layer of regulation for the Type III-B system of *P. furiosus*. The results are consistent with recent structural analysis in *S. thermophilus* that revealed RNA-protein interactions between Csm1 and the PFS that dictated whether DNAse only, cOA synthesis only, or both activities were triggered [19]. In that study, positions +2 and +3 were found to be critical for important structural features and triggering both activities. In *Pfu*, the difference in requirements represents a novel dual-layered method of Csx1 activation regulation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

This work was supported by The National Institutes of Health [R35GM118160 to M.P.T. and 1F31GM125365 to K.F.].

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENT

We thank members of the Terns and White laboratories for helpful discussions and guidance. We also express gratitude to Janet Westpheling for providing critical guidance on generating the complementation strains.

Figure 3.1. Csx1 is required for anti-plasmid immunity in the *P. furiosus (Pfu)* III-B CRISPR-Cas system. (A) Composition of the *Pfu* Cmr crRNP complex. The Cmr effector complex is composed of Cmr1 – 6 and a mature crRNA (black & orange) containing a 5' tag (black) 8 nucleotides in length.. Cyclic oligoadenylate compounds (orange) produced by the Palm domain of Cmr2 bind to the CARF domain of Csx1. (B) Plasmid Silencing Assay Setup. Liquid *Pfu* cultures are transformed with either target or no target plasmid. The target plasmid contains a transcribed target region complementary to the 7.01 crRNA of *Pfu*. Both plasmids contain the *pyrF* gene to facilitate growth in the absence of uracil. (C) Colonies produced by transforming 11 different *Pfu* strains. Strains were transformed with the target plasmid (blue) or no target plasmid (gray). Colonies obtained per μg of plasmid DNA are plotted with error bars indicating the standard deviation for two replicates.

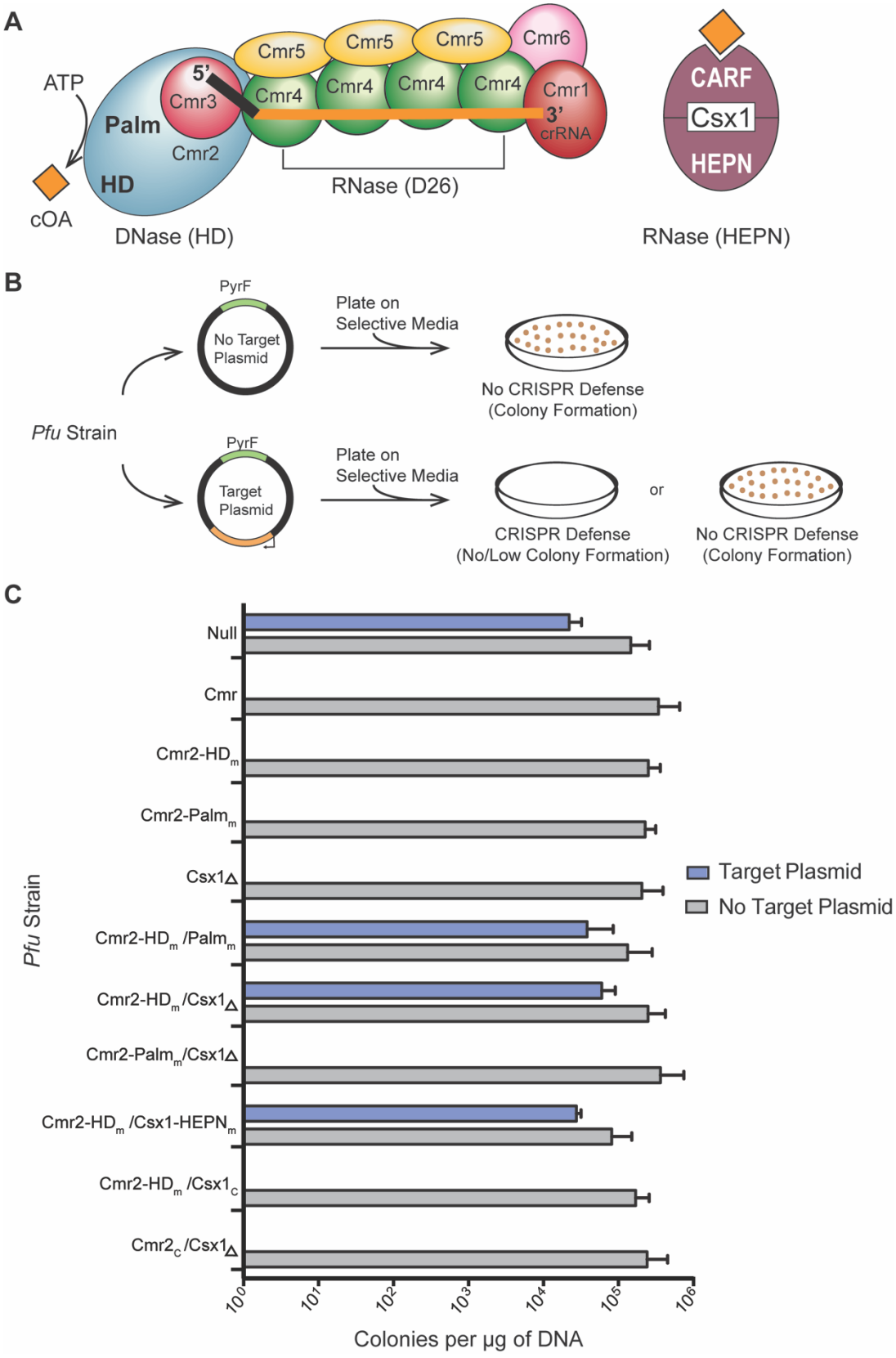

A
ATP
cOA
Palm
HD
Cmr2
5'
Cmr3
Cmr5
Cmr5
Cmr5
Cmr6
Cmr4
Cmr4
Cmr4
Cmr4
Cmr1
3'
crRNA
RNase (D26)
DNase (HD)
CARF
Csx1
HEPN
RNase (HEPN)
B
PyrF
No Target Plasmid
Plate on Selective Media
No CRISPR Defense (Colony Formation)
Pfu Strain
PyrF
Target Plasmid
Plate on Selective Media
or
CRISPR Defense (No/Low Colony Formation)
No CRISPR Defense (Colony Formation)
C
Null
Cmr
Cmr2-HDm
Cmr2-Palmm
Csx1Δ
Cmr2-HDm /Palmm
Cmr2-HDm/Csx1Δ
Cmr2-Palmm/Csx1Δ
Cmr2-HDm /Csx1-HEPNm
Cmr2-HDm /Csx1C
Cmr2C/Csx1Δ
Pfu Strain
Target Plasmid
No Target Plasmid
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Colonies per µg of DNA

Figure 3.2. The ribonuclease activity of Csx1 is activated by cOA species produced by the *Pfu* effector complex in a Palm domain dependent mechanism. (A) Purification of *Pfu* complex proteins. HD and Palm mutations are located within Cmr2. D26N mutation is located within Cmr4. (B) RNAse activity of each *Pfu* Cmr complex. Radiolabeled 7.01 Target RNA was incubated with Cmr complexes and reaction products were visualized by urea-PAGE. Radiolabeled RNA size standards (M) in nucleotides were used. The black arrow indicates the full length substrate and the black asterisks indicate cleavage products. (C) DNAse activity of each *Pfu* Cmr complex. Each complex was incubated with dsDNA (label located on DNA Target 2) and tested for nuclease activity. Reactions were visualized as described in part B. (D) cOA production activity of the *Pfu* Cmr crRNP. The five Cmr complexes were tested for cOA production in presence (+) and absence (-) of 7.01 Target RNA. Reaction products were run on a denaturing gel. A radiolabeled alkaline hydrolysis ladder (M2) and size standard (M) were added to the urea-PAGE analysis.. (E) Activation of wildtype Csx1 by *Pfu* cOAs. Csx1 was tested for ribonuclease activity under activating (20 nM Csx1) conditions. Reactions were visualized as described in part B.

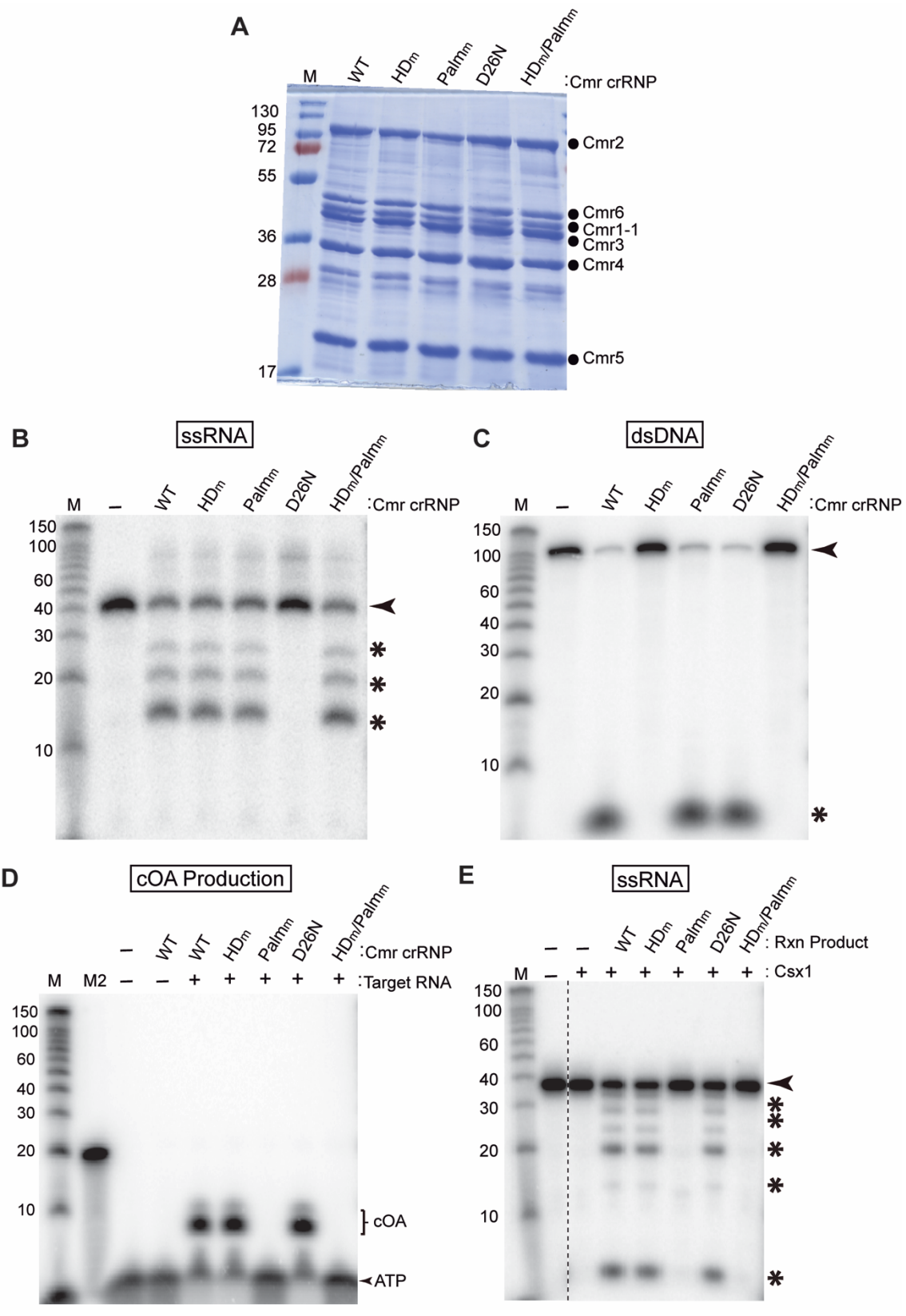

A
M
WT
HDm
Palmm
D26N
HDm/Palmm
:Cmr crRNP
130
95
72
55
36
28
17
Cmr2
Cmr6
Cmr1-1
Cmr3
Cmr4
Cmr5
B
ssRNA
M
–
WT
HDm
Palmm
D26N
HDm/Palmm
:Cmr crRNP
150
100
60
40
30
20
10
C
dsDNA
D
cOA Production
M
M2
:Target RNA
cOA
ATP
E
ssRNA
: Rxn Product
: Csx1

Figure 3.3. Activation of Csx1 is CARF domain dependent. (A) Ribbon structure of *Pfu* Csx1 (PDB: 4EOG) indicating the CARF and HEPN domains. (B) SDS-page analysis of *Pfu* Csx1 protein purifications. WT_{column} refers to protein purified using a column method and WT_{batch} refers to protein purified using a batch method. (C-E) *In vitro* activation of Csx1 mutants. Wildtype Csx1 along with four Csx1 mutants ($CARF_m$, INAA, INQQ, $HEPN_m$) was tested for ribonuclease activity under non-activating (500 nM Csx1) and activating (10 nM Csx1) conditions. A radiolabeled alkaline RNA size marker (M) was added to the urea-PAGE analysis. The black arrow indicates the full length RNA substrate and the black asterisks indicate cleavage products. (F) *In vivo* plasmid silencing assay results for CARF domain mutants.

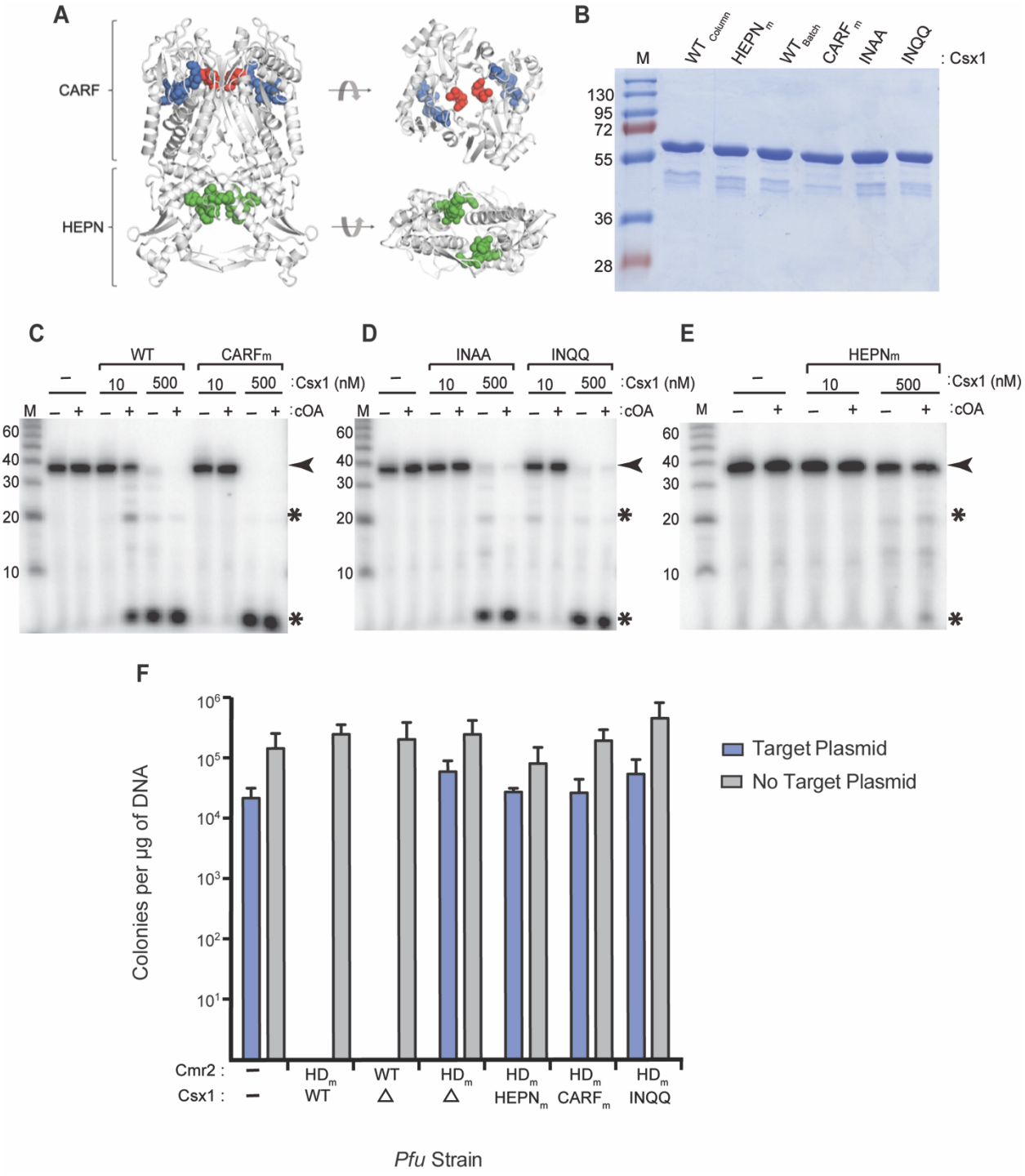

A
CARF
HEPN
B
M
WT Column
HEPNm
WT Batch
CARFm
INAA
INQQ
: Csx1
130
95
72
55
36
28
C
WT
CARFm
–
10
500
10
500
:Csx1 (nM)
M – + – + – + – + – + :cOA
60
40
30
20
10
D
INAA
INQQ
–
10
500
10
500
:Csx1 (nM)
M – + – + – + – + – + :cOA
60
40
30
20
10
E
HEPNm
–
10
500
:Csx1 (nM)
M – + – + – + :cOA
60
40
30
20
10
F
Colonies per µg of DNA
10^6
10^5
10^4
10^3
10^2
10^1
Target Plasmid
No Target Plasmid
Cmr2 : – HDm WT HDm HDm HDm HDm
Csx1 : – WT Δ Δ HEPNm CARFm INQQ
Pfu Strain

Figure 3.4. cOA_4 is the relevant activator for *Pfu* Csx1. (A) UV chromatogram (258 nm) and MS chromatogram (extracted ion chromatogram for *m/z* 494.6, 659.1, 823.6, 988.2) for cOA samples from *Pfu* Cmr wildtype complexes. (B) *Pfu* Csx1 activation assays with synthetic cOA. Wildtype Csx1(20 nM) was incubated with increasing concentrations of synthetic cOA_4 and cOA_6. For each activator, 1.85, 18.5, 185, and 1.85 x 10^4 pM concentrations were used. A radiolabeled alkaline RNA size marker (M) was added to the urea-PAGE analysis. The black arrow indicates the full length RNA substrate and the black asterisks indicate cleavage products.

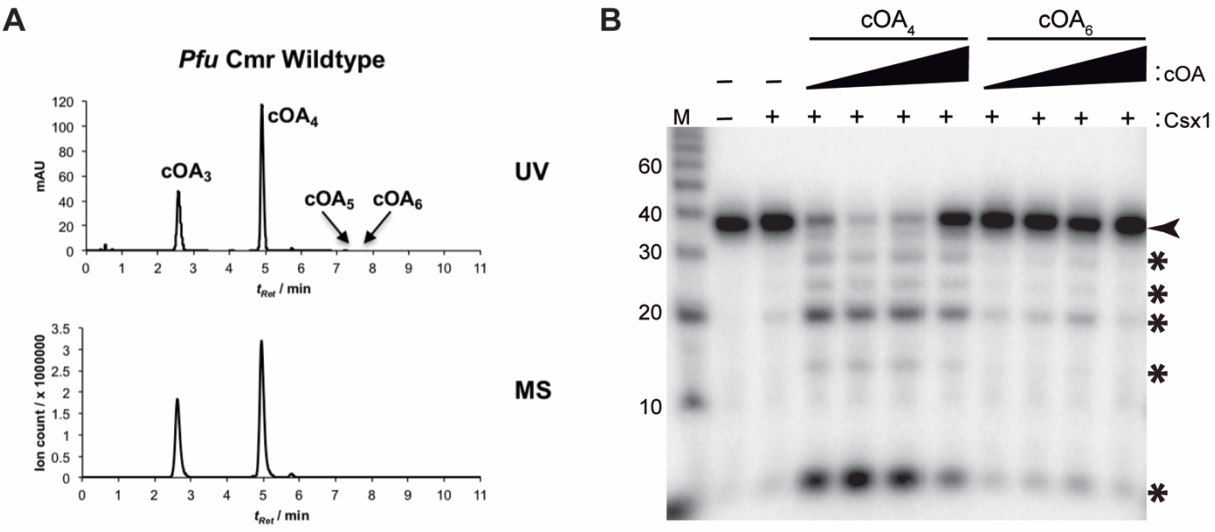


Figure 3.5. Degradation of cOAs by *Pfu* Csx1 inactivates the activators. (A) Sequencing gel reactions of Csx1 proteins incubated with *Pfu* cOAs. Five different *Pfu* Csx1 proteins were incubated with radiolabeled cOAs for two timepoints-5 and 20 mins. Urea-page analysis on sequencing gels was performed. Radiolabeled RNA size marker (M) and alkaline hydrolysis ladder (M) and were included in the analysis. (B) Effect of Csx1 treatment on cOA activation activity. Unlabeled *Pfu* cOAs were treated with 600 nM of wildtype Csx1 or no protein for 30 minutes. The plus extraction set of reactions were tested for activation of wildtype Csx1 by incubating 20 nM of wildtype Csx1 with decreasing amounts of A_2>P¥ or cOA_4— undiluted, 2^{-1}, 10^{-1}, 10^{-2}, and 10^{-4} dilutions. ¥ indicates a putative designation for this reaction product. Native cOA without the 30 minute incubation or extraction was utilized as a control (N).

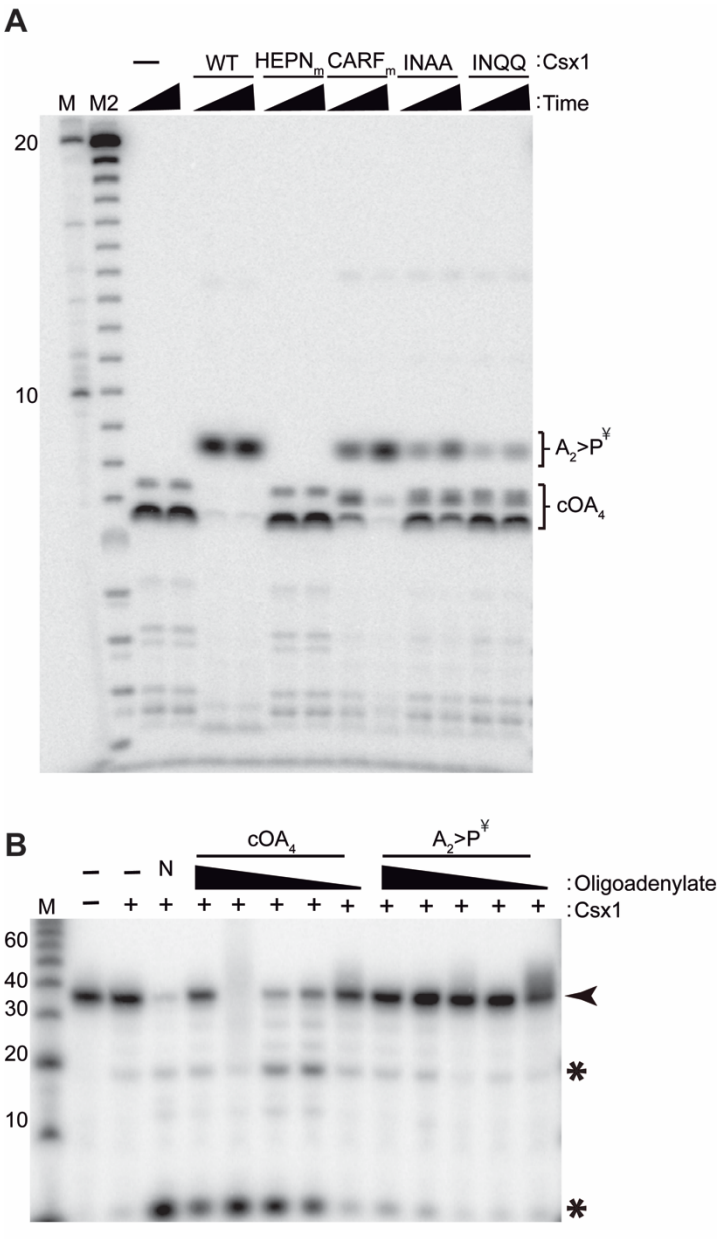

A
—
WT
HEPN$_m$
CARF$_m$
INAA
INQQ
:Csx1
M
M2
:Time
20
10
A$_2$>P$^{¥}$
cOA$_4$
B
cOA$_4$
A$_2$>P$^{¥}$
N
:Oligoadenylate
:Csx1
60
40
30
20
10

Figure 3.6. *Pfu* Cmr DNase and cOA production activities are rPAM dependent. (A) Composition of the *Pfu* Cmr crRNP. The Cmr effector complex is composed of Cmr1 – 6 and a mature crRNA (black & orange) containing a 5' tag (black). The *Pfu* Cmr complex is activated by binding to target RNA (purple) that is complementary to the crRNA. The target RNA contains a PFS (red) within it that is located adjacent to the target sequence. (B) Target RNA dependent activation of DNase activity by the *Pfu* complex. Cmr crRNP complexes were incubated with radiolabeled dsDNA in the presence of 23 different target RNAs. Each target RNA contains a different sequence in the PFS region of the target RNA. Urea-page analysis on sequencing gels was performed. Radiolabeled RNA size markers (M) were included in the analysis. The black arrow indicates the full length RNA substrate and the black asterisks indicate cleavage products. (C) Target RNA dependent activation of cOA production activity by the *Pfu* complex using the same target RNAs mentioned in part B. Products were analyzed as described in part B. An alkaline hydrolysis ladder (M2) was included in the urea-PAGE analysis.

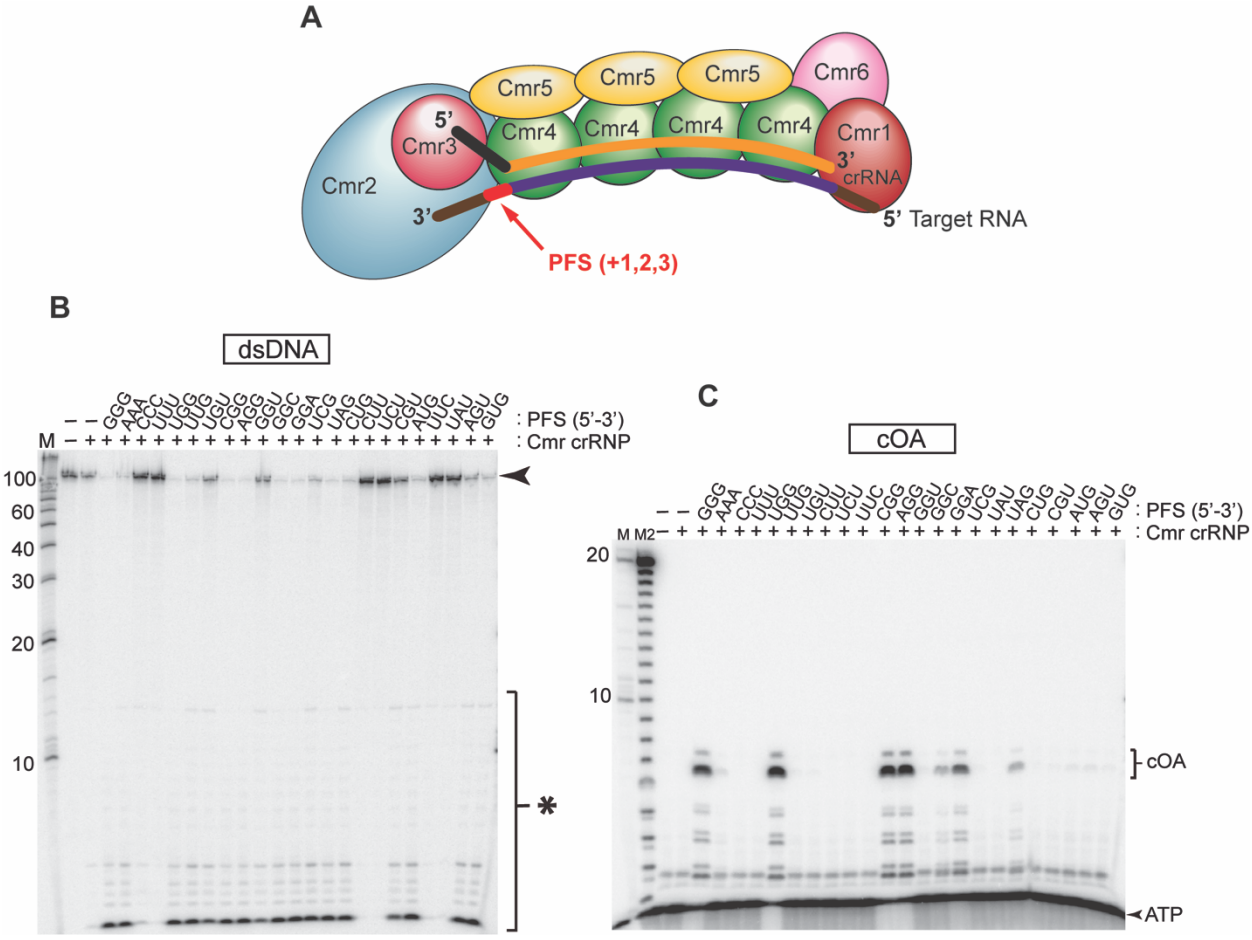
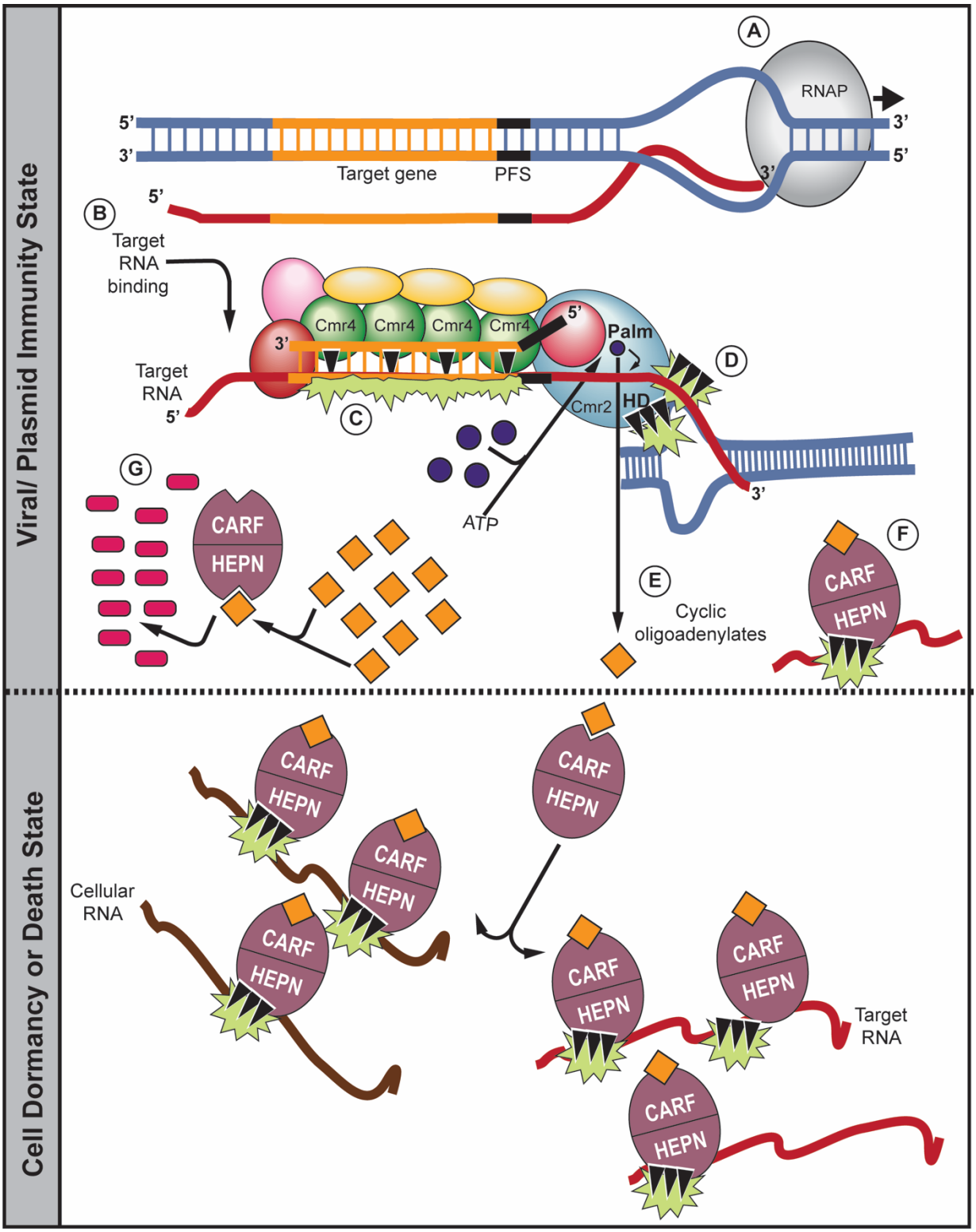

A
Cmr5
Cmr5
Cmr5
Cmr6
5'
Cmr3
Cmr4
Cmr4
Cmr4
Cmr4
Cmr1
3'
crRNA
Cmr2
3'
5' Target RNA
PFS (+1,2,3)
B
dsDNA
: PFS (5'-3')
: Cmr crRNP
M
100
60
40
30
20
10
*
C
cOA
M M2
20
10
cOA
ATP

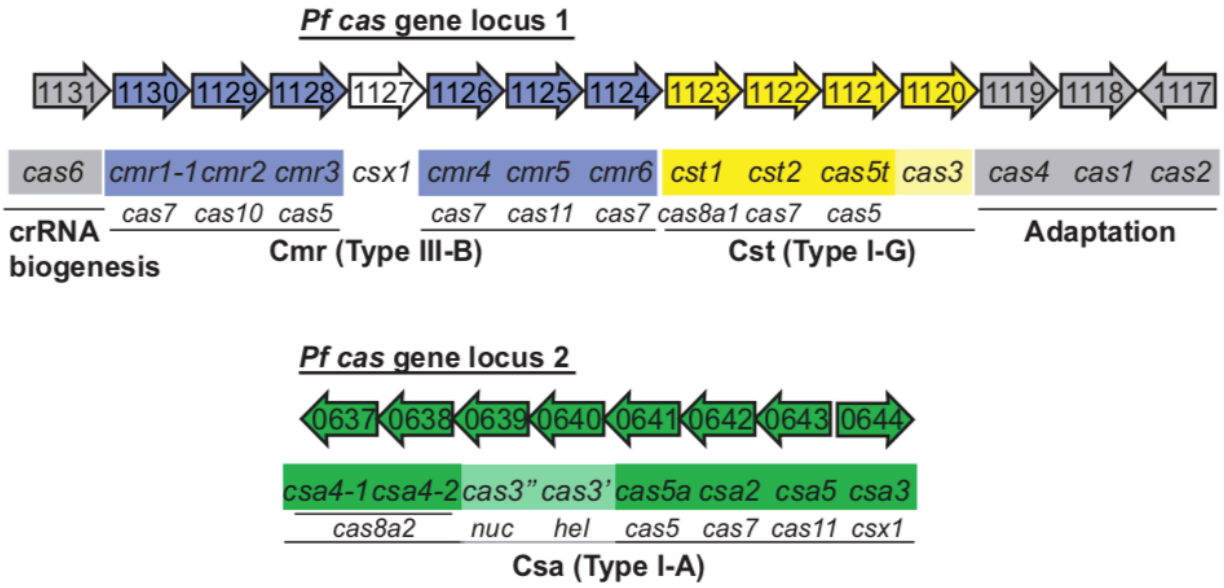
Table 3.1. Summary of rPAM sequence requirements for Pfu Cmr activities. Recap of PFS dependent CRISPR defense, DNase activity, and cOA production activity results observed for target RNAs used in this study and previously reported [20]. + indicates the activity was observed, - denotes the activity was not observed, and +/- indicates the activity was weakly observed.

Target RNA PFS (5'-3')		CRISPR Defense *(in vivo)*	DNase *(in vitro)*	cOA *(in vitro)*
NNN	GGG	+	+	+
	AAA	+	+	–
	CCC	–	–	–
	UUU	–	–	–
NGG	GGG	+	+	+
	AGG	+	+	+
	CGG	+	+	+
	UGG	+	+	+
GGN	GGG	+	+	+
	GGA	+	+	+
	GGC	+	+	+/–
	GGU	+	+	–
UNG	UGG	+	+	+
	UAG	+	+	+/–
	UCG	+	+	–
	UUG	+	+	–
NUG	GUG	+	+	–
	AUG	+	+	–
	CUG	+	+	–
	UUG	+	+	–
NGU	GGU	+	+	–
	AGU	+	+	–
	CGU	+	+	–
	UGU	+	+	–
UNU	UGU	+	+	–
	UAU	–	–	–
	UCU	–	–	–
	UUU	–	–	–
YUY	CUU	–	–	–
	UUC	–	–	–

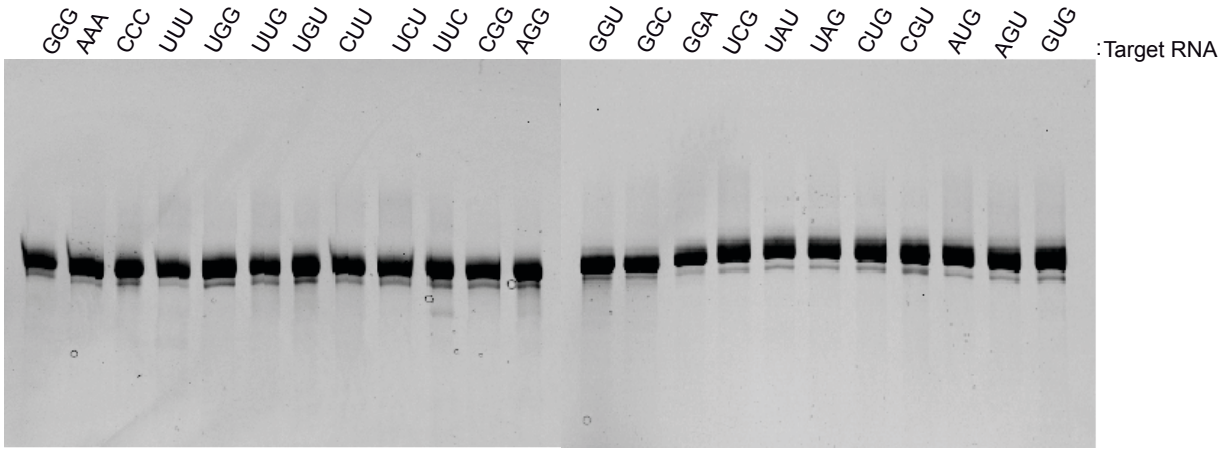
Figure 3.7. Model for activation of defense activities by the *Pfu* Cmr CRISPR-Cas machinery. Invader silencing in the Type III-B system of P. furiosus begins with the transcription of the target DNA region (A). Production of target RNA leads to base pairing between the Cmr crRNP and target RNA (B). Binding of the target RNA and recognition of the PFS results in conformational changes that trigger three actions by *Pfu* Cmr: RNA cleavage by Cmr4 (C), DNA cleavage by the HD domain of Cmr2 (D), and ATP dependent production of cyclic oligoadenylates by the Palm domain of Cmr2 (E). Cyclic oligoadenylate compounds are released from the *Pfu* Cmr complex and activate *Pfu* Csx1 ribonuclease activity by binding to the CARF domain (F). Under conditions of controlled immunity, the nonspecific RNase activity of *Pfu* Csx1 is auto-regulated by conversion of cOA_4 to the inactive product A_2>P¥ (G). Under conditions of cell dormancy or death, *Pfu* Csx1 nonspecifically degrades cellular and target RNA.


Viral/ Plasmid Immunity State
A
RNAP
5'
3'
Target gene
PFS
3'
5'
5'
B
Target RNA binding
Cmr4
Cmr4
Cmr4
Cmr4
3'
5'
Palm
D
Target RNA
5'
C
Cmr2
HD
3'
ATP
G
CARF
HEPN
E
Cyclic oligoadenylates
F
CARF
HEPN
Cell Dormancy or Death State
CARF
HEPN
CARF
HEPN
Cellular RNA
CARF
HEPN
CARF
HEPN
CARF
HEPN
CARF
HEPN
Target RNA
CARF
HEPN

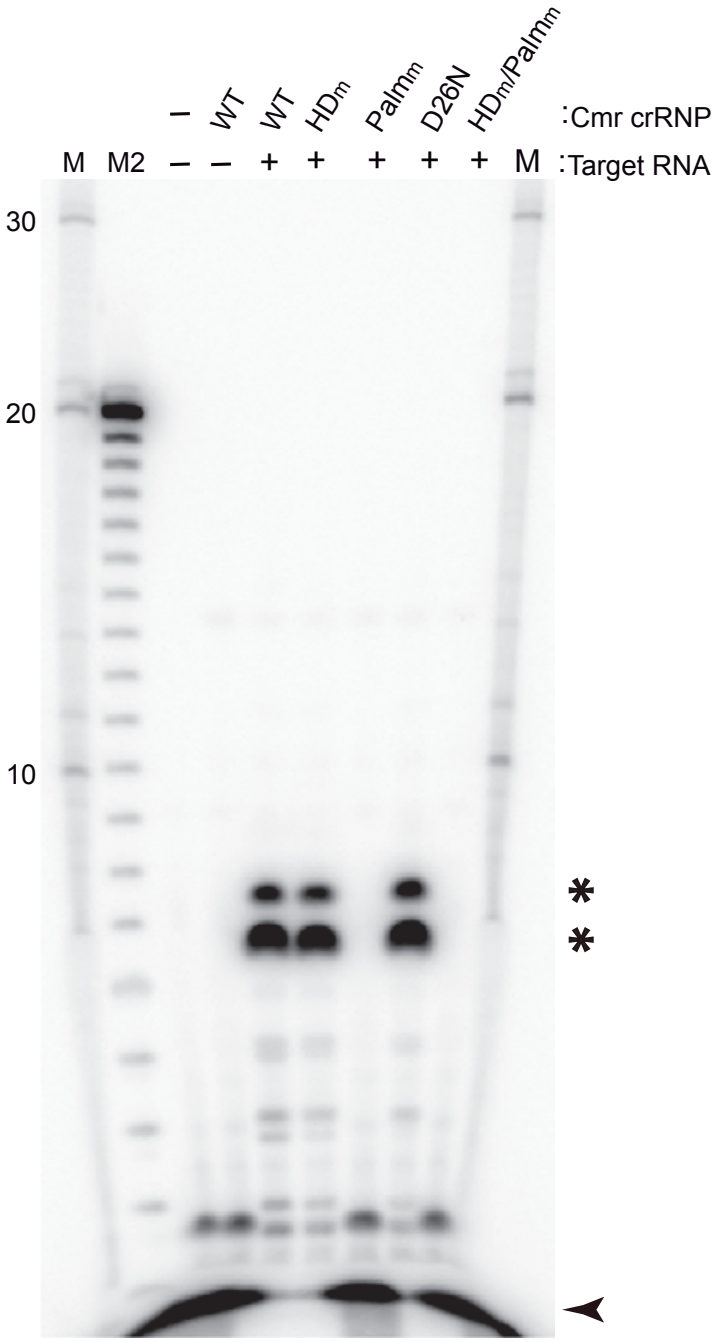
Supplemental Figure 3.S1. *Pyrococcus furiosus cas* gene locus organization. The genome organization and annotations of the predicted *cas* genes were adapted from the NCBI database (http://www.ncbi.nlm.nih.gov). *P. furiosus* contains three CRISPR-different Cas systems: Type III-B Cmr (blue), Type I-G Cst (yellow), and Type I-A Csa (green). Specific *cas* gene superfamily designations are indicated below corresponding *csa*, *cst*, and *cmr* genes. Adaptation (gray) and crRNA biogenesis (gray) genes are shared amongst all three systems. Csx1 is not a stable member of the Cmr crRNP but the *csx1* gene (white) is often found in association with Type III-B systems and is encoded in between the *cmr* genes in Pfu.



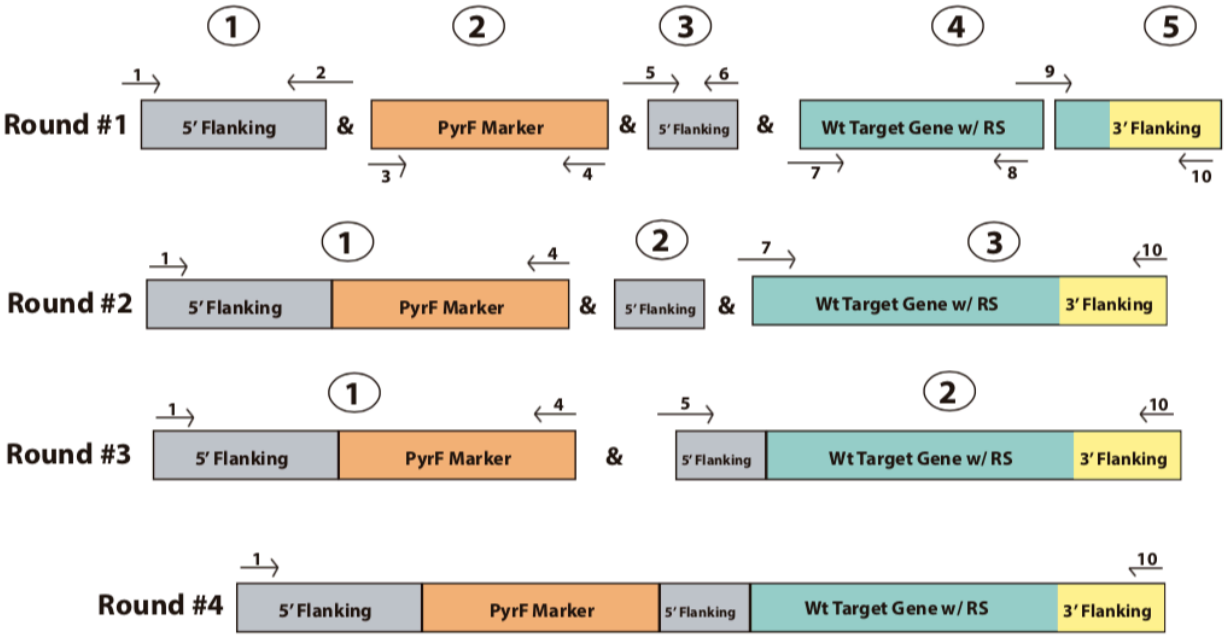
Supplemental Figure 3.S2. *In vitro* assay target RNAs. Target RNAs utilized for *in vitro* assays were created using *in vitro* transcription (Invitrogen). Each target RNA contains a different sequence in the PFS region (5'-3') of the target RNA. 100ng of each RNA was ran on 15% denaturing PAGE gels and stained with EtBr.



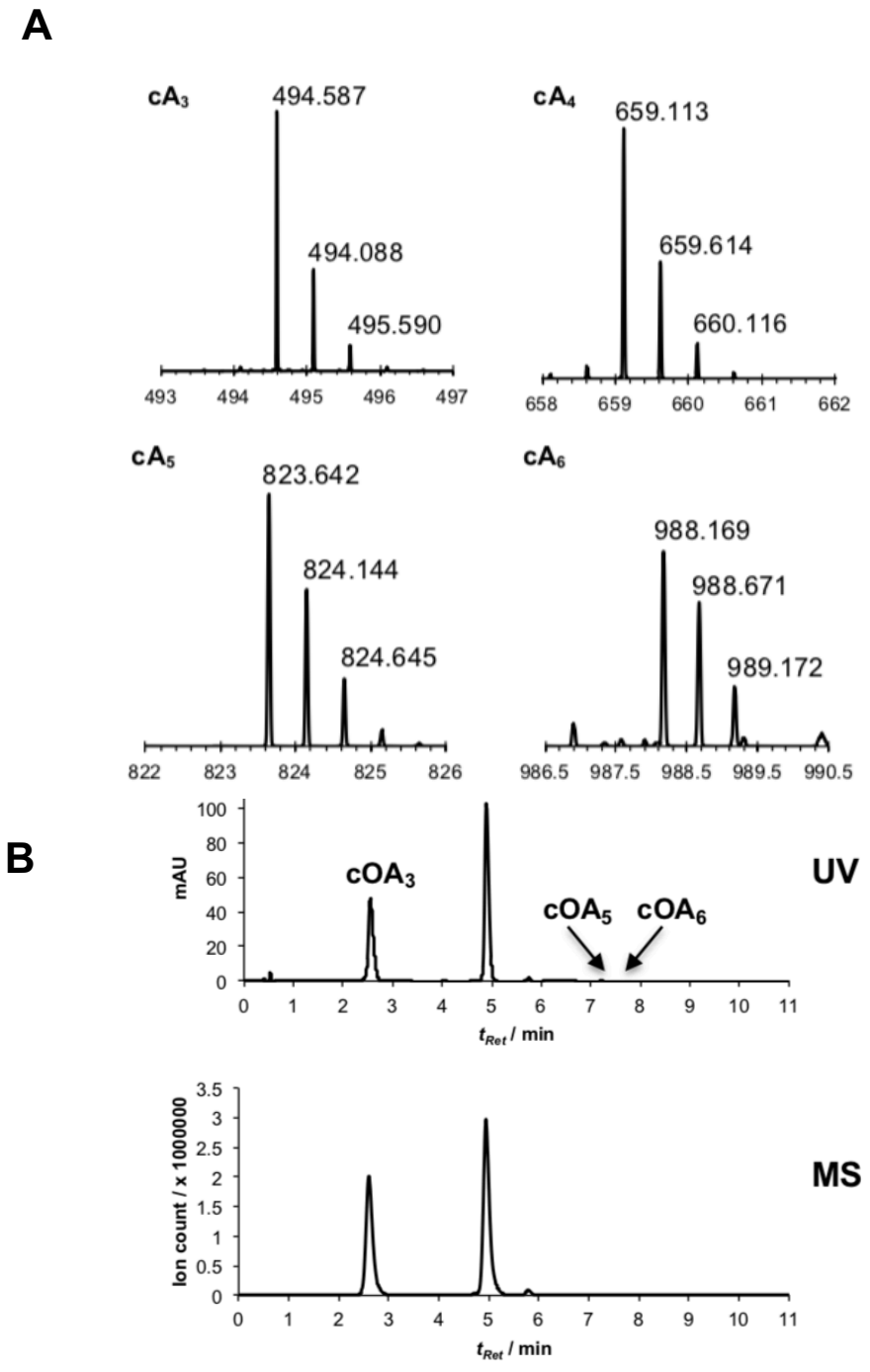
Supplemental Figure 3.S3. cOA production activity of the *Pfu* Cmr crRNP. Five different Cmr complexes were tested for cOA production by incubating the complexes in the presence (+) and absence (-) of 7.01 Target RNA. Reaction products are shown on a sequencing gel. ATP is indicated by an arrow and cOA species are indicated with asterisks.



Supplemental Figure 3.S4. Modified SOE-PCR product creation approach for *P. furiosus* complementation strains. Two complementation strains were generated for this study ($Cmr2_c/Csx1_{\Delta}$ and $HD_m/Csx1_c$). The strains were generated using a modified version of the process detailed previously [20]. Construction of the SOE-PCR transformation product was altered to complement the wildtype genes back onto the genome with a wobble-base mutation coding for a new restriction enzyme site (RS). Construction of the SOE-PCR product required four rounds of PCR and 10 primers. Round #1 generates 5 different products, round #2 generates 3 different products, round #3 generates 2 different products, and the fourth round generates the final PCR product. The final product is then transformed into *P. furiosus.*



Supplemental Figure 3.S5. *Pfu* Cmr cOA mass spectrometry analysis. (A) Mass spectra for each cOA produced by *Pfu* Cmr (B) UV chromatogram (258 nm) and MS chromatogram (extracted ion chromatogram for *m/z* 494.6, 659.1, 823.6, 988.2) for cOA samples from *Pfu* Cmr4 D26N mutant complexes.



Supplemental Table 3.S1: *Pyrococcus furiosus* Strains

Strain	Relevant Characteristics	Source
Null (TPF20)	Δ*pyr*F ,Δ*trp*AB, ΔCmr (Δ*PF1124-PF1130*), ΔCst (Δ*PF1121-1123*) , & ΔCsa (Δ*PF0637-0644*)	[20]
Cmr (TPF15)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*) & ΔCsa(Δ*PF0637-0644*)	[20]
HD_m (TPF63)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), & *cmr2*-H13A/D14A	[20]
$Palm_m$ (TPF27)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), & *cmr2*-D673A/D674A	[20]
$Csx1_\Delta$ (TPF24)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), & Δ*csx1*(ΔPF1127)	[20]
HD_m / $Palm_m$ (TPF37)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), & *cmr2*-H13A/D14A, D673AD674A	[20]
$Palm_m$ / $Csx1_\Delta$ (TPF119)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-D673A/D674A, & Δ*csx1*(ΔPF1127)	This study

HD_m / $Csx1_{\Delta}$ (TPF118)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & Δ*csx1*(ΔPF1127)	This study
HD_m / $Csx1_{HEPN}$ (TPF128)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & *csx1*-H436A	This study
HD_m / $Csx1_{CARF}$ (TPF136)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & *csx1*-Δ121-127	This study
HD_m / $Csx1_{INA}$ (TPF137)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & *csx1*-I169A/I170A	This study
HD_m / $Csx1_{INQ}$ (TPF138)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & *csx1*-I169Q/I170Q	This study
HD_m / $Csx1_C$ (TPF131)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), *cmr2*-H13A/D14A, & +*csx1*(+PF1127)	This study
$Cmr2_c$ / $Csx1_{\Delta}$ (TPF132)	Δ*pyr*F ,Δ*trp*AB ,ΔCst (Δ*PF1121-1123*), ΔCsa(Δ*PF0637-0644*), +*cmr2 (+PF1129)*, & Δ*csx1*(ΔPF1127)	This study

Supplemental Table 3.S2: Plasmids

Plasmid	Relevant Characteristics	Purpose	Source
PyrF-SOE (pJFW18)	AprR general cloning vector with *E.coli* OriT, *Pfu* Pgdh-pyrF cassette, and *Pfu* OriC for replication in *P.furiosus*	*Pfu* strain construction	[59]
No target (pJE47)	pJFW18 derivative; Tk-csg promoter/Tk-chiA terminator expression cassette	Plasmid silencing assays (Figures 3.1, 3.3)	[20]
Target (pJE65)	pJE47 derivative; 7.01 spacer, GGG flank, target strand transcribed	Plasmid silencing assays (Figures 3.1, 3.3) IVT Target RNA Production (Figures 3.2, 3.4; Table 3.1)	[20]
Target (pJE186)	pJE47 derivative; 7.01 spacer, TTT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE187)	pJE47 derivative; 7.01 spacer, TTC flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]

Target (pJE189)	pJE47 derivative; 7.01 spacer, TTG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE190)	pJE47 derivative; 7.01 spacer, TCT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE193)	pJE47 derivative; 7.01 spacer, TCG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE194)	pJE47 derivative; 7.01 spacer, TAT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE197)	pJE47 derivative; 7.01 spacer, TAG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE198)	pJE47 derivative; 7.01 spacer, TGT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE201)	pJE47 derivative; 7.01 spacer, TGG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]

Target (pJE202)	pJE47 derivative; 7.01 spacer, CTT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE205)	pJE47 derivative; 7.01 spacer, CTG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE207)	pJE47 derivative; 7.01 spacer, CCC flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	
Target (pJE214)	pJE47 derivative; 7.01 spacer, CGT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE217)	pJE47 derivative; 7.01 spacer, CGG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE221)	pJE47 derivative; 7.01 spacer, ATG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE228)	pJE47 derivative; 7.01 spacer, AAA flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]

Target (pJE230)	pJE47 derivative; 7.01 spacer, AGT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE233)	pJE47 derivative; 7.01 spacer, AGG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE237)	pJE47 derivative; 7.01 spacer, GTG flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE246)	pJE47 derivative; 7.01 spacer, GGT flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE247)	pJE47 derivative; 7.01 spacer, GGC flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]
Target (pJE248)	pJE47 derivative; 7.01 spacer, GGA flank, target strand transcribed	IVT Target RNA Production (Figure 3.6; Table 3.1)	[20]

Supplemental Table 3.S3: Oligos

***In vitro* Assay Oligos**

Oligo	Sequence (5'-3')	Purpose
7.01 crRNA	AUUGAAAGUUGUAGUAUGCGGUCCUU GCGGCUGAGAGCACUUCAG	crRNA used for assembling Cmr crRNPs
7.01 Target RNA	CUGAAGUGCUCUCAGCCGCAAGGACC GCAUACUACAAGGG	RNA substrate used for Cmr crRNP and Csx1 cleavage assays
DNA Target 1	AGCGACAGAGTTTAGGTATGAATGATT AGATAGAGGTAGTGAAGTGGTGTGAG ATGGAAGGAGAGGATAGTAGAAGGGT AGTTAGTATTAGATAGAGGTAAGTGAT CCCAGTACGTCGTAGTA	DNA substrate used for Cmr crRNP cleavage assays (annealed to DNA target 2)
DNA Target 2	TACTACGACGTACTGGGATCACTTACC TCTATCTAATACTAACTACCCTTCTACT ATCCTCTCCTTCCATCTCACACCACTTC ACTACCTCTATCTAATCATTCATACCTA AACTCTGTCGCT	DNA substrate used for Cmr crRNP cleavage assays (annealed to DNA target 1)

SOE-PCR Construct Primers

Primer	Sequence (5'-3')	Purpose
Pgdh_PyrF_F	GATTGAAAATGGAGTGAGCTGAG	Used with pJFW18 to amplify the PyrF marker for all constructs
Pgdh_PyrF_R	TTATCTTGAGCTCCATTCTTTCACC	Used with pJFW18 to amplify the PyrF marker for all constructs
ΔCsx1_1	GGCAGAATTTACCCCCTTCC	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains
ΔCsx1_2	CTCAGCTCACTCCATTTTCAATCTC ATTCCCATATCCCTCCTAAAGC	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains
ΔCsx1_5	GGTGAAAGAATGGAGCTCAAGATAA TCCCACAATAGGGAAAGTTGG	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains
ΔCsx1_6	TCATTCCCATATCCCTCCTAAAGC	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains
ΔCsx1_7	GCTTTAGGAGGGATATGGGAATGA CTGCAAATCTCGCTTATGAAG	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains

ΔCsx1_8	CCTTTGCCCTGGGAGTTACA	Used to add a Csx1 deletion into existing $Palm_m$ or HD_m strains
Csx1_H436A_1	TAAGAAACTTGCTTGGATTGTTG	Used to add a Csx1 H436A mutation into the existing HD_m strain
Csx1_H436A_2	CTCAGCTCACTCCATTTTCAATCTG CTATAAAGTTACGAACAACATTTG	Used to add a Csx1 H436A mutation into the existing HD_m strain
Csx1_H436A_5	GGTGAAAGAATGGAGCTCAAGATAA TTTAGAGAGGCAAGTAGAGTG	Used to add a Csx1 H436A mutation into the existing HD_m strain
Csx1_H436A_6	TGCTATAAAGTTACGAACAACATTT G	Used to add a Csx1 H436A mutation into the existing HD_m strain
Csx1_H436A_7	CAAATGTTGTTCGTAACTTTATAGCA GCATCTGGATTTGAGTATAACATTG TC	Used to add a Csx1 H436A mutation into the existing HD_m strain
Csx1_H436A_8	AGTTCCTGTCTCAGCATTAATT	Used to add a Csx1 H436A mutation into the existing HD_m strain

Csx1_CARF_1	CAGGATGCGGAATCAAAGATA	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARFdel_2	CTCAGCTCACTCCATTTTCAATCAATTGTAATTCTTTCATCTTTTAACATATTC	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARF_5	GGTGAAAGAATGGAGCTCAAGATAAAGAAATACCTTGCCAATTCTAG	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARFdel_6	AATTGTAATTCTTTCATCTTTTAACATATTC	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARFdel_7	GAATATGTTAAAAGATGAAAGAATTACAATTGTCTTTGGCAACATTACAGTAG	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARF_8	AAGACCCTTCTTTATTTTCTGTTC	Used to delete residues 121-127 from Csx1 in the existing HD_m strain
Csx1_CARF_IN_1	GAAAGCCCTCTTCTATTTCTC	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain

Csx1_CARF_IN_2	CTCAGCTCACTCCATTTTCAATCCC CATGAGTTAGGTCTAAGT	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain
Csx1_CARF_IN_5	GGTGAAAGAATGGAGCTCAAGATAA TTCGGAAGTTGTGGAAGATAC	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain
Csx1_CARF_IN_6	CCCATGAGTTAGGTCTAAGT	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain
Csx1_CARF_INA_7	ACTTAGACCTAACTCATGGGGCTGC TTTCATGCCCACCTTTACTTAC	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain
Csx1_CARF_INQ_7	ACTTAGACCTAACTCATGGGCAGCA GTTCATGCCCACCTTTACTTAC	Used to create an I169A/I170A or I169Q/I170Q mutation in

		Csx1 in the existing HD_m strain
Csx1_CARF_IN_8	TCCAGAATGTGCTATAAAGTTACG	Used to create an I169A/I170A or I169Q/I170Q mutation in Csx1 in the existing HD_m strain
Csx1_NruI_1	AGGCCTCAAGAATTGTGTTG	Used to complement wildtype Csx1 back into the HD_m / $Csx1_{\Delta}$ strain
Csx1_NruI_2	CTCAGCTCACTCCATTTTCAATCCA TATCCCTCCTAAAGCAAG	Used to complement wildtype Csx1 back into the HD_m / $Csx1_{\Delta}$ strain
Csx1_NruI_5	GGTGAAAGAATGGAGCTCAAGATAA ATTTCTCAACTCCCACAATAGG	Used to complement wildtype Csx1 back into the HD_m / $Csx1_{\Delta}$ strain
Csx1_NruI_6	CATATCCCTCCTAAAGCAAG	Used to complement wildtype Csx1 back into the HD_m / $Csx1_{\Delta}$ strain
Csx1_NruI_7	CTTGCTTTAGGAGGGATATGGGAAT GAGAGTTTTGGTAACTAC	Used to complement wildtype Csx1 back into the HD_m / $Csx1_{\Delta}$ strain

Csx1_NruI_8	CGAATATTTGGATAAAAGCTGGCA	Used to complement wildtype Csx1 back into the HD_m / $Csx1_\Delta$ strain
Csx1_NruI_9	TGCCAGCTTTTATCCAAATATTCG**C** GACGTAGAAGATTACCTTAACAAAA	Used to complement wildtype Csx1 back into the HD_m / $Csx1_\Delta$ strain
Csx1_NruI_10	TGATTAGGGCCATATCATTATCC	Used to complement wildtype Csx1 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_1	CACAAAGGAAAGAAGCAGAGT	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_2	CTCAGCTCACTCCATTTTCAATCTCT CTTTGATGTTAACCACTCCAA	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_5	GGTGAAAGAATGGAGCTCAAGATAA GGAGGAGAATTAAGAAGTATCAAAG	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_6	TCTCTTTGATGTTAACCACTCCAA	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain

Cmr2_BclI_7	TTGGAGTGGTTAACATCAAAGAGAA ACTTTTTGTATACCTTCATGATCC	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_8	ATCAAACACATGGGACATAATGG	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_9	CCATTATGTCCCATGTGTTTGAT**C**AA AAGGTATTATCCAGTGTGGATTAG	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain
Cmr2_BclI_10	GTTCTTGTACTATTTCTCCAACTTTC	Used to complement wildtype Cmr2 back into the HD_m / $Csx1_\Delta$ strain

***P. furiosus* Strain Screening Oligos**

Primer	**Sequence (5'-3')**	**Mutation**
ΔCsx1_seq_For	GTGTTGGAGTGGGTGAGGAG	Csx1 deletion
ΔCsx1_seq_Rev	TCTGGAGATATTTGCCGTTAATC	Csx1 deletion
Csx1_H436A_Seq_F	GGTTGCCAGTTCAGAACAAT	Csx1-H436A
Csx1_H436A_Seq_R	AGGGCCATATCATTATCCTTTG	Csx1-H436A
Csx1_CARF_SeqF	AGCTTTCTAGAGAGAAGAAAGTA	Csx1 Δ121-127
Csx1_CARF_SeqR	AAATATTCTTGTAAGCTTTGCAGC	Csx1 Δ121-127

Csx1_CARF_IN_SeqF	TCATCATATGAGTTCTTAACGTTAG	Csx1- I169A/I170A or Csx1- I169Q/I170Q
Csx1_CARF_IN_SeqR	TTCATAAGCGAGATTTGCAGCT	Csx1- I169A/I170A or Csx1- I169Q/I170Q
Csx1_CARF_IN_SeqI	GTCTTTGGCAACATTACAGTAG	Csx1- I169A/I170A or Csx1- I169Q/I170Q
Csx1_NruI_SeqF	ATGCAGGAAGTGAAAGTGTG	Csx1 complementation
Csx1_NruI_SeqInt	TTCACAAGAAATAAGGGAGGAC	Csx1 complementation
Csx1_NruI_SeqR	GTCACCTCTTTAAGAGTAACTTTG	Csx1 complementation
Cmr2_BclI_SeqF	ATGGTGGTTCAGAGCTTTG	Cmr2 complementation
Cmr2_BclI_SeqInt	GGAGAGAGAAAAGTTACAGAAG	Cmr2 complementation
Cmr2_BclI_SeqR	TGGCTTCTTTTCATGCATATCC	Cmr2 complementation

Csx1 *in vitro* Mutagenesis Oligos

Primer	**Sequence (5'-3')**	**Mutation**
Csx1_CARFdel_F	GTCTTTGGCAACATTACAGTAG AG	Csx1 ∆121-127
Csx1_CARFdel_R	AATTGTAATTCTTTCATCTTTTAA CATATTCTC	Csx1 ∆121-127
Csx1_ 169A170A _F	GTTTACTTAGACCTAACTCATG GGGCTGCTTTCATGCCCACCTT TACTTAC	Csx1- I169A/I170A

Csx1_ 169A170A _R	GTAAGTAAAGGTGGGCATGAAA GCAGCCCCATGAGTTAGGTCTA AGT	Csx1- I169A/I170A
Csx1_ 169Q170Q _F:	GTTTACTTAGACCTAACTCATG GGCAGCAGTTCATGCCCACCTT TACTTAC	Csx1- I169Q/I170Q
Csx1_ 169Q170Q _R:	GTAAGTAAAGGTGGGCATGAAC TGCTGCCCATGAGTTAGGTCTA AGT	Csx1- I169Q/I170Q
Csx1_ CARFSeq _Fwd	GCTGATACAATGGCCAACT	Csx1 ∆121-127 or Csx1- I169A/I170A or Csx1- I169Q/I170Q
Csx1_ CARFSeq _Rev	GAATTGTGTCCTCCCTTATTTC	Csx1 ∆121-127 or Csx1- I169A/I170A or Csx1- I169Q/I170Q

IVT Template PCR Primers

Primer	**Sequence (5'-3')**
pJE47_IVT_T7_F	TAATACGACTCACTATAGGGAGACAACACTTAGTAGGGGCTA
pJE47_IVT_R	GCTTCCTTAGCTGTTTCTCCA

Supplemental Table S4: Cmr Gene Sequences

Gene	Sequence (5'-3')
cmr2	GTGGTTAACATCAAAGAGAAACTTTTTGTATACCTTCATGATC CACCAGACAAGGCTCTAAAAATTGAAAATCATGAGGAAAGGT CAAAAAAGATATTAAGTTCTGGCAATATCCAGTACTCGAGAA CGGACAAAGTTAAACAAGCAGATGCACTTTCTTCTAAGACTC AGAGATTTATAATTCGAACAAAGGAAAATAAAGAGCCAGTAA TAGATTTTTTGGGTAGATCTTCAGGAAAGTACTTCCATGTTG GATATCCTGTTTTTATACACCCCATATCCACAGAAATTAAGAG GTATGAAACACTTGAAAAGTACATAGACCTTGGCAGGAGTAA TAGAGGGGAAAGATTTGTTAACGAGTTTTTGGAAAGGGTTTC AAAGCTTGAAGGCGATGTTCTCAAAGAGGTCTTTGAAGATGC TAGTAACAAATTTAAAGGAGAAGAGAGTAAACAGTGGGCCTA CATCTGGCAGTTTTATCCCGTAAAACTCAAAGAAGGAGTCAA GGAATTTGCCAAGTCAGAGTTAAAACTTAAAGAGGAAGAAGC AGAAAAGTTTGCAGAGGAATTTGTTAACCTCCCAGCTGATAC AAGATTTCCAGATCATGCAATTTGGACCCATTTAGACTTAACT TCCGCATTATCCGTTAAGGATCCCACTTTGCTCAGGATCAAA ATAGTTCCAGTTCAACCTTTTATTGCCAATTCAAGAAAGCAGT TAGATCTCTGGGCCTCCAGTCATCTCCTTTCAATGCTTATGT ATAAAGCTTTAGAGGTGATAGTGGACAAGTTCGGGCCAGAA CATGTAATCTATCCATCTCTAAGGGATCAACCCTTCTTCTTGA

	AGTTCTACCTGGGGGAAAACATAGGTGATGAAATCTTAGTTG
	CAAACTTGCCTAACAAAGCGCTTGCAATAGTCTCAGGAAAGG
	AGGCTGAAAAGATTGAAGAAGAAATCAAGAAAAGAATTAGGG
	ATTTCCTACTCCAACTGTACAGAGAAGCTGTTGATTGGGCAG
	TTGAAAATGGAGTAGTAAAAGTGGATAGAAGTGAAAAGGATA
	GCATGCTCAAGGAAGCATATCTTAAAATTGTGAGGGAGTACT
	TCACCGTCTCGATAACCTGGGTATCTCTTTCCGAAAAGGAGG
	ATATCTATCAAGTAACAGAGAACGCGGGTCTCTCGGATGAAG
	ATGTTAAGAAGTGGCTAAAGTTTGCAGAAAAGAAAGAAAATA
	GTAGAGTTCTCGAGAGGATTGCAATATACCCACTTTTGGTAA
	AGATATTGGATAGCCTGGGAGAGAGAAAAGTTACAGAAGAA
	AGGTTCGAAAAAAGCGAACAACTCAAAGGATGGAAGTGCCA
	CGTTTGTGGTGAGAATCTTGCAATTTTTGGAGACATGTACGA
	TCACGATAATCTTAAGAGTTTGTGGCTTGATGAGGAACCATT
	ATGTCCCATGTGTTTGATAAAAAGGTATTATCCAGTGTGGATT
	AGGAGTAAAACTGGACAGAAAATAAGGTTTGAGTCGGTGGT
	AGATGTTGCACTTCTGTACAAGAACTGGAGGAAGATATTTGA
	CGAGAAGTATGGAAAAGACCTAGTCTCAAAGGCTAGGGAAG
	TTAGTGAAGACTTCGTAAAGGACAATATGCTAGTAGATTCGG
	ATCTATACTATTCTTCAACCTGGGAATCTGGACTTTCTAAAAA
	GCTCAAAAATAAGAAAGAGATTGATGAGGAAAAAGTTAAGGA
	AGTTGTTGACTTCTTAAATGCGGCTTATAAAGAAATCGGTAAT
	CCACCAAAGTACTATGCTATTCTAGTTATGGATGGCGACGAT

	ATGGGGAAAGTTATTTCAGGAGAGGTGCTTGGAGAAATATCA ACTAGAATTCATCCAAATATTAGGGATTACGTTGAAATTCCAG AAGCAAAATATTACTCCACCCCGCAGGTTCACGTGGCTATAA GCCAAGCATTGGCTAACTTTTCGATAAGGGAAGTTAGATCCG TAGTTAAAGACGAGGGATTGCTAATATACGCTGGAGGGGAT GATGTCCTAGCAATTTTGCCAGTCGACAAAGCTTTAGAAGTT GCATATAAGATAAGGAAAGAATTTGGCAAGAGCTTTGAAAAT GGTTCTCTTCTCCCAGGTTGGAAGTTGAGTGCTGGAATTTTG ATAGTCCATTATAAGCATCCATTGTATGACGCCCTAGAAAAG GCAAGAGATCTTCTCAATAATAAAGCAAAAAACGTTCCAGGA AAAGATACACTAGCTATAGGCCTACTTAAGAGGAGTGGTTCC TACTATATCTCCCTAGTGGGATGGGAATTAATTAGGGTCTTC TACAACTCAGAGCTGAGGAAAAAGCTATTGGAAGAGAAAGG TGGAGTGGGAAAGAGGTTCATTTATCATGTGCTCAGAGAAGT TGATACTTGGCCAAAAGTTGGAATAGACGAGATGCTTAAGTT TGAGGTGATTAGACATATCAGGGGAAGGAACAAAGAAGAAA CTAAAGAGCTCAGAGAAAAGATCTATGGAGAAATAAAGGATC TTCTTGAGCATGTAAGAGGGAACAATGAAGTTGAAAAAGTTA GAGGCTTATTCACATTTCTAAAAATAATCACGGACGCGGAGG TGTTTCCATGA
csx1	ATGGGAATGAGAGTTTTGGTAACTACATGGGGTAATCCCTTC CAGTGGGAACCAATAACATATGAATACAGAGGAATCAAAGTT AAAAGCAGAAATACCTTGCCAATTCTAGTCAAGACTCTTGAG

	CCAGAGAGGATTCTAATCCTTGTGGCTGATACAATGGCCAAC TACTATGATTCAGGAAAAAATAAGCCAGAAATAGAAGAAAAA TCGTTTTCGTCTTATTCGGAAGTTGTGGAAGATACAAAAGAA AGGATACTATGGCACATAAAAGAGGAGGTCATTGAAGAACTC CGTGAGGAAGATCCTGAGCTTGCTAAGAAAATTGAGAATATG TTAAAAGATGAAAGAATTACAATTGAAGTTCTTCCCGGCGTT GGAGTCTTTGGCAACATTACAGTAGAGGGAGAAATGCTTGA CTTCTATTATTATGCCACATACAAGTTGGCCGAATGGTTGCC AGTTCAGAACAATTTAGAGGTTTACTTAGACCTAACTCATGG GATAAATTTCATGCCCACCTTTACTTACAGAGCCCTAAGAAA CTTGCTTGGATTGTTGGCCTACTTGTACAATGTAAAGTTTGA GATAGTTAATTCAGAACCTTATCCCCTGGGGGTTTCACAAGA AATAAGGGAGGACACAATTCTCCATATTAGGGAAATTGGAGA GGGAGTAGTTCGTCCTAGACCACAGTATTCTCCAGTAGAAG GAAAGCTTTACTGGAATGCATTTATAAGCTCTGTAGCCAATG GCTTCCCGTTAGTCTTTGCCAGCTTTTATCCAAATATTCGGG ACGTAGAAGATTACCTTAACAAAAAGCTTGAGGAATTCCTGG TGGGAATTGAGGTTGGGGAGAGAGAAGATGGAAAACCTTAT GTTAAAAGAGAGAAAGCTCTTGACAGGAGCTTTAAGAATGCT TCTAAGCTCTACTATGCTTTAAGAGTGTTCAATACAAAATTCC AAAACTATCCAAAAAAAGAAGTTCCTATTGAAGAAATAATGGA GATATCAAAGATATTCGAGTCTCTTCCCAGGATTGGAATTATT TTAGAGAGGCAAGTAGAGTGGCTAAGAAATTTAGTATATGGA

	AGATTATGGTATGAAAATGGAGAACAGAAAATAAAGAAGGGT CTTTTAGAGATTATCAAGGATAAGAAGGATAAAAGGAAAGAG GCCGAAGCTCTTAAAAAAGGGAAGACAATATCTTTAGCCGAA GCTGCAAAGCTTACAAGAATATTTTCTCCGAGTGGAGAAAGA ATAGAGACAATAGAATCTCCAAATGTTGTTCGTAACTTTATAG CACATTCTGGATTTGAGTATAACATTGTCTATGTGAAATATGA TAGACTAAGTGATAGGCTGTACTTTTTCTATAAGGATAAAGAA AAAGCTGCAAATCTCGCTTATGAAGCCCTTTTATATAGGGGT GAAAAAGAATGA

REFERENCES

1. van der Oost, J., Westra, E. R., Jackson, R. N., and Wiedenheft, B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**, 479-492
2. Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell* **172**, 1239-1259
3. Koonin, E. V., and Makarova, K. S. (2019) Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* **374**, 20180087
4. Koonin, E. V., Makarova, K. S., and Zhang, F. (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67-78
5. Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., and Koonin, E. V. (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477
6. Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M. R., Barendregt, A., Zhou, K., Snijders, A. P., Dickman, M. J., Doudna, J. A., Boekema, E. J., Heck, A. J., van der Oost, J., and Brouns, S. J. (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* **18**, 529-536
7. Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., and Doudna, J. A. (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997

8. Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., and Zhang, F. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573

9. Jia, N., Mo, C. Y., Wang, C., Eng, E. T., Marraffini, L. A., and Patel, D. J. (2019) Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. *Mol Cell* **73**, 264-277 e265

10. Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., and Koonin, E. V. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-736

11. Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2018) Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J* **1**, 325-336

12. Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., Lopez-Mendez, B., She, Q., and Montoya, G. (2019) Structure of Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. *Nat Commun* **10**, 4302

13. Mulepati, S., Heroux, A., and Bailey, S. (2014) Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* **345**, 1479-1484

14. Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol Cell* **58**, 418-430

15. Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., and White, M. F. (2013) Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol Cell* **52**, 124-134

16. Staals, R. H., Zhu, Y., Taylor, D. W., Kornfeld, J. E., Sharma, K., Barendregt, A., Koehorst, J. J., Vlot, M., Neupane, N., Varossieau, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P. J., Urlaub, H., Heck, A. J., Nogales, E., Doudna, J. A., Shinkai, A., and van der Oost, J. (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. *Mol Cell* **56**, 518-530

17. Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J., Doudna, J. A., Heck, A. J. R., Yonekura, K., van der Oost, J., and Shinkai, A. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. *Mol Cell* **52**, 135-145

18. Taylor, D. W., Zhu, Y., Staals, R. H., Kornfeld, J. E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J. A. (2015) Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* **348**, 581-585

19. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., and Wang, Y. (2019) Structure Studies of the CRISPR-Csm

Complex Reveal Mechanism of Co-transcriptional Interference. *Cell* **176**, 239-253 e216

20. Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., and Terns, M. P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev* **30**, 447-459
21. Hale, C. R., Cocozaki, A., Li, H., Terns, R. M., and Terns, M. P. (2014) Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. *Genes Dev* **28**, 2432-2443
22. Han, W., Li, Y., Deng, L., Feng, M., Peng, W., Hallstrom, S., Zhang, J., Peng, N., Liang, Y. X., White, M. F., and She, Q. (2017) A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. *Nucleic Acids Res* **45**, 1983-1993
23. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C., and Siksnys, V. (2016) Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Mol Cell* **62**, 295-306
24. Liu, T. Y., Iavarone, A. T., and Doudna, J. A. (2017) RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. *PLoS One* **12**, e0170552
25. Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* **161**, 1164-1174
26. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, C., Nwokeoji, A. O., Dickman, M. J., Horvath, P., and Siksnys, V. (2014) Programmable RNA shredding

by the type III-A CRISPR-Cas system of Streptococcus thermophilus. *Mol Cell* **56**, 506-517

27. Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol Microbiol* **87**, 1088-1099

28. Goldberg, G. W., Jiang, W., Bikard, D., and Marraffini, L. A. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* **514**, 633-637

29. Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M., and Terns, M. P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945-956

30. Estrella, M. A., Kuo, F. T., and Bailey, S. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Dev* **30**, 460-470

31. Guo, M., Zhang, K., Zhu, Y., Pintilie, G. D., Guan, X., Li, S., Schmid, M. F., Ma, Z., Chiu, W., and Huang, Z. (2019) Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry. *Cell Res* **29**, 305-312

32. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksnys, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609

33. Niewoehner, O., Garcia-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., and Jinek, M. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543-548

34. Nasef, M., Muffly, M. C., Beckman, A. B., Rowe, S. J., Walker, F. C., Hatoum-Aslan, A., and Dunkle, J. A. (2019) Regulation of cyclic oligoadenylate synthesis by the Staphylococcus epidermidis Cas10-Csm complex. *RNA* **25**, 948-962

35. Rouillon, C., Athukoralage, J. S., Graham, S., Gruschow, S., and White, M. F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* **7**

36. Gruschow, S., Athukoralage, J. S., Graham, S., Hoogeboom, T., and White, M. F. (2019) Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. *Nucleic Acids Res* **47**, 9259-9270

37. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., Montoya, G., and She, Q. (2018) A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. *Nucleic Acids Res* **46**, 10319-10330

38. Jia, N., Jones, R., Sukenick, G., and Patel, D. J. (2019) Second Messenger cA4 Formation within the Composite Csm1 Palm Pocket of Type III-A CRISPR-Cas Csm Complex and Its Release Path. *Mol Cell* **75**, 933-943 e936

39. Foster, K., Kalter, J., Woodside, W., Terns, R. M., and Terns, M. P. (2019) The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR-Cas systems. *RNA Biol* **16**, 449-460

40. Jiang, W., Samai, P., and Marraffini, L. A. (2016) Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* **164**, 710-721

41. Niewoehner, O., and Jinek, M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA* **22**, 318-329

42. Sheppard, N. F., Glover, C. V., 3rd, Terns, R. M., and Terns, M. P. (2016) The CRISPR-associated Csx1 protein of Pyrococcus furiosus is an adenosine-specific endoribonuclease. *RNA* **22**, 216-224

43. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013) Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* **8**, 15

44. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front Genet* **5**, 102

45. Majumdar, S., Zhao, P., Pfister, N. T., Compton, M., Olson, S., Glover, C. V., 3rd, Wells, L., Graveley, B. R., Terns, R. M., and Terns, M. P. (2015) Three CRISPR-Cas immune effector complexes coexist in Pyrococcus furiosus. *RNA* **21**, 1147-1158

46. Terns, R. M., and Terns, M. P. (2013) The RNA- and DNA-targeting CRISPR-Cas immune systems of Pyrococcus furiosus. *Biochem Soc Trans* **41**, 1416-1421

47. Majumdar, S., Ligon, M., Skinner, W. C., Terns, R. M., and Terns, M. P. (2017) Target DNA recognition and cleavage by a reconstituted Type I-G CRISPR-Cas immune effector complex. *Extremophiles* **21**, 95-107

48. Majumdar, S., and Terns, M. P. (2019) CRISPR RNA-guided DNA cleavage by reconstituted Type I-A immune effector complexes. *Extremophiles* **23**, 19-33

49. Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Zheng, K., Slaymaker, I. M., Gootenberg, J. S., Abudayyeh, O. A., Essletzbichler, P., Shmakov, S., Makarova, K. S., Koonin, E. V., and Zhang, F. (2017) Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* **65**, 618-630 e617

50. Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J Bacteriol* **196**, 310-317

51. Rostol, J. T., and Marraffini, L. A. (2019) Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat Microbiol* **4**, 656-662

52. Lipscomb, G. L., Stirrett, K., Schut, G. J., Yang, F., Jenney, F. E., Jr., Scott, R. A., Adams, M. W., and Westpheling, J. (2011) Natural competence in the hyperthermophilic archaeon Pyrococcus furiosus facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. *Appl Environ Microbiol* **77**, 2232-2238

53. Rouillon, C., Athukoralage, J. S., Graham, S., Gruschow, S., and White, M. F. (2019) Investigation of the cyclic oligoadenylate signaling pathway of type III CRISPR systems. *Methods Enzymol* **616**, 191-218

54. Han, W., Pan, S., Lopez-Mendez, B., Montoya, G., and She, Q. (2017) Allosteric regulation of Csx1, a type IIIB-associated CARF domain ribonuclease by RNAs carrying a tetraadenylate tail. *Nucleic Acids Res* **45**, 10740-10750

55. Kim, Y. K., Kim, Y. G., and Oh, B. H. (2013) Crystal structure and nucleic acid-binding activity of the CRISPR-associated protein Csx1 of Pyrococcus furiosus. *Proteins* **81**, 261-270

56. Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D. J. (2019) CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. *Mol Cell* **75**, 944-956 e946

57. Athukoralage, J. S., Rouillon, C., Graham, S., Gruschow, S., and White, M. F. (2018) Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277-280

58. Athukoralage, J. S., Graham, S., Gruschow, S., Rouillon, C., and White, M. F. (2019) A Type III CRISPR Ancillary Ribonuclease Degrades Its Cyclic Oligoadenylate Activator. *J Mol Biol* **431**, 2894-2899

59. Farkas, J., Chung, D., DeBarry, M., Adams, M. W., and Westpheling, J. (2011) Defining components of the chromosomal origin of replication of the hyperthermophilic archaeon Pyrococcus furiosus needed for construction of a stable replicating shuttle vector. *Appl Environ Microbiol* **77**, 6343-6349

CHAPTER 4

DISCUSSION

The wealth of knowledge pertaining to the mechanisms involved in Type III defense has grown and advanced rapidly over the past decade. From initially thinking that Type III complexes were single activity destroyers to now uncovering their multiple activities, these systems have truly shown themselves to be more complicated and multi-faceted compared to their counterparts. The diversity amongst CRISPR-Cas systems is a recurring theme in the study of this biological process. This fact is becoming more and more relevant especially in regards to their classification. As the understanding of CRISPR-Cas biology grows, the systems become trickier to classify due to differences observed at the main type level, subtype level, and organism level. Consistent independent functional studies are essential to sorting out these differences and advancing the field. The work described in this thesis provides much needed information about four different Type III systems particularly in regards to understanding the activities of Csm6 and Csx1 and the overall contribution of these proteins to Type III CRISPR-Cas immunity.

The Nonspecific Ribonuclease Activity of Csm6 and Csx1

At the inception of this dissertation, there was virtually no published biochemical data about the Csm6/Csx1 family of proteins. Bioinformatical analysis revealed two conserved domains for this group of proteins. It was predicted that this family of proteins

possessed ribonuclease activity via the HEPN motif (R-X_{4-6}-H) located within the HEPN domain with the potential to be allosterically regulated by the CARF domain [1,2]. *In vivo* data from two organisms (*Streptococcus thermophilus* and *Sulfolobus islandicus*) revealed that Csm6 and Csx1 were required for anti-plasmid immunity in both systems, but many questions remained regarding the functional activity of the proteins and the connection between Csm6/Csx1 and the effector complexes [3,4]. Eventually, additional studies were published that confirmed the ribonuclease activity of Csm6 and Csx1 [5-10]. The findings in this dissertation have provided additional insight into the *in vitro* ribonuclease activity of Csm6 for two additional Csm6 proteins from *Lactococcus lactis (Lla)* and *Staphylococcus epidermidis* (*Sep*) (Chapter 2). *L. lactis* Csm6 was found to be a single-stranded RNA specific nuclease and the highly conserved histidine residue of the HEPN motif was found to eliminate the ribonuclease activity of both *Lla* and *Sep* Csm6. Both proteins were found to cleave ssRNA nonspecifically at the substrate level but with specificity at the nucleotide level [8]. Several different RNA substrates were tested and both proteins were able to degrade all RNAs without bias very similar to the Csx1 protein from *Pyrococcus furiosus* (*Pfu*) [5]. These findings provided much needed information about these mysterious proteins. A consistent lack of specificity for a specific CRISPR related RNA target across several different Csm6/Csx1 proteins in independent studies solidified the fact that these proteins can degrade CRISPR related and non-CRISPR related targets. This characteristic of Csm6 and Csx1 proteins has many implications for the cell; predominantly whether the nonspecific nature occurs within the cell and mechanisms of control of these nonspecific RNases. Currently, independent studies conducted in *S. epidermidis* have found that the nonspecific degradation activity

of Csm6 found *in vitro* also occurs *in vivo* [7,11]. Csm6 has been found to degrade host, plasmid, and phage transcripts in vivo [7,11]. Furthermore, the activity appears to be dependent upon levels of transcription for the organism *S. epidermidis* [11]. Collectively, these findings about the conserved ribonuclease activity of Csm6 and Csx1 proteins demonstrate that they are an integral component of CRISPR-Cas defense rather than simply auxiliary proteins as once thought. The nonspecific nature of the cleavage activity of Csm6 and Csx1 proteins also makes them a powerful defense tool for the prokaryotes harboring them. Phages are constantly evolving to evade bacterial defense mechanisms and CRISPR-Cas is no different. The ability to target RNA nonspecifically provides these systems with a defense mechanism that is harder to evade. Moving forward, more independent studies for additional homologs will be needed to determine the full scope of *in vivo* activity of this family of proteins.

Although Csm6 and Csx1 proteins are broadly nonspecific at the substrate level, specificity at the nucleotide level has been observed. *Pfu* Csx1 was found to only degrade RNA substrates with adenosine residues and Csm6 from *S. thermophilus* (*Sth*) were found to recognize GA or AA dinucleotides [5,9]. Mapping analysis presented in this thesis (Chapter 2) demonstrate that *Lla* and *Sep* Csm6 proteins cleave ssRNA at A or G nucleotides [8]. Alternatively, analysis conducted with *Thermus thermophilus* Csm6 observed no nucleotide specificity [6]. Thus, the presence of nucleotide specificity appears to differ between homologs but carries a strong preference for purines. In addition to its nucleotide specificity, *Lla* Csm6 was found to be a divalent metal-independent endoribonuclease that degrades ssRNA on the 5' side of phosphodiester bonds resulting in a 5' hydroxyl group and a 2'-3' cyclic phosphate or 3' phosphate group

[8]. These characteristics are consistent amongst Csm6 and Csx1 homologs [5,6]. The consistency of these aspects of the activity of the enzymes would be very useful for using them for application purposes such as a control for understanding the behaviors for other CRISPR and non-CRISPR related enzymes and as a tool for biotech purposes. For example, Csm6 was used in nucleic acids diagnostics using the platform specific high-sensitivity enzymatic reporter unlocking (SHERLOCK). In this application, Csm6 was utilized to amplify signals from Cas13 cleavage [12].

The Role of Csm6 and Csx1 in Type III CRISPR-Cas Immunity

During production of this thesis, a major question for Type III systems pertained to the role of the ribonuclease activity of Csm6 and Csx1 in the overall process of CRISPR-Cas immunity. The question remains a prominent topic in the field due to conflicting answers to this question at the organism level [3,4,7,8,12]. Conflicting results were obtained for the requirement of Csm6 in defense in *S. epidermidis* [4,7,10], Csx1 was found to be important for defense in *S. islandicus* [3] but not in *P. furiosus* [12]. In Chapter 2, three separate Csm6 homologs were found to be required *in vivo* for CRISPR-Cas defense [8]. A mutation within the HEPN motif or complete deletion of the gene for all three homologs resulted in loss of CRISPR-Cas immunity. Furthermore, inhibition of the ribonuclease activity was sufficient for preventing defense in all three systems and the DNA degradation activity of Cas10 was not required. This initial set of *in vivo* results were limited but conveyed a possible conditional role for Csm6 and Csx1 in the overall defense process. The results also suggested varying levels of impact of Csm6/Csx1 homologs between organisms. Additional mutational analysis of more than one Cas protein of the

Type III-B system conducted in *Pfu* (Chapter 3) confirmed the role of *Pfu* Csx1 in Type III-B immunity. Though simply deleting *Pfu csx1* had no impact on anti-plasmid immunity [12], the deletion of *csx1* or the interruption of the ribonuclease activity of Csx1 in conjunction with mutation of the HD domain of Cas10 impacted immunity. Double mutations of the Palm domain of Cas10 and Csx1 did not impact immunity indicating the functional connectedness of the two entities (Chapter 3). These findings indicate a multi-pronged approach for defense in *P. furiosus* and a key role for Csx1 in conferring immunity in wild type *Pfu* cells. Furthermore, the DNA degradation activity of Cas10 and the RNA degradation activity of Csx1 are both independently sufficient to enact defense. This differs from other studies in which Csm6 was found to be required but not sufficient for a complete defense response [10,11]. A role for *Sep* Csm6 was found when considering levels of transcription and levels of DNA copies of the CRISPR target [11] Due to the logistical limitations of controlling transcription in *P. furiosus*, the connection between levels of transcription and *Pfu* Csx1 activity has yet to be explored. Additional investigation of this in *P. furiosus* will be needed to further understand the role of these proteins in the context of transcription.

Additional clarity about the role of Csm6/Csx1 was provided when it was discovered that the Cas10 subunit of the Type III effector complexes produces a second messenger molecule that binds to the CARF domain of Csm6 to activate ssRNA activity catalyzed by the proteins' HEPN ribonuclease domain [9,10]. Subsequent studies have confirmed the widespread presence and conserved nature of this activity amongst Type III systems [13-18]. Data presented in Chapter 3 of this thesis confirms that cyclic oligoadenylate (cOA) production occurs in the *P. furiosus* Type III-B system. Production

of cOA molecules in *Pfu* is dependent upon on the Cas10 Palm domain (and GGDD motif) and activates Csx1 ribonuclease activity via residues located within the CARF domain. The activity was confirmed *in vivo* and *in vitro* (Chapter 3). The discovery of this feature of Type III systems has been a breakthrough for Type III systems for multiple reasons. Firstly, it has solidified the role of Csm6/Csx1 in defense. Second, it has provided a functional connection between the Type III effector complexes and Csm6/Csx1. Csm6 and Csx1 proteins are not stable members of their respective complexes; thus, the connection between the two entities has always remained mysterious until now. Additionally, discovery of the cOA molecule has provided much needed insight into methods of control of the nonspecific activity of Csm6 and Csx1 proteins. Currently, control of the ribonuclease activity of Csm6/Csx1 occurs at the RNA level and protein level. Production of cOA molecules has been found to inversely correlate with ribonuclease activity of Csm3 or Cmr4 [16,17]. As degradation of the target RNA by Csm3/Cmr4 increases, levels of cOA decline and activation of Csm6/Csx1 subsides. The temporal nature of the activation provides a built in “on and off” switch for Csm6 and Csx1. Regulation has also been found at the protein level. In *Sulfolobus solfataricus*, a distinct ring nuclease was found to cleave cOA molecules and thus prevent activation of *S. solfataricus* Csx1 [19]. Two Csm6 homologs have been shown to degrade cOA molecules themselves [20,21]. In Chapter 3, *Pfu* Csx1 itself was found to alter the cOA species. Modification of the cOA molecules by *Pfu* Csx1 resulted in lack of activation of the ribonuclease activity of Csx1 by those modified cOA species. Moving forward, additional investigation into the specific biochemical modification of cOA by *Pfu* Csx1 is needed. Not all Csx1 proteins possess the ability to self-modify the cOA molecules like *Pfu* Csx1. For

example, the Csx1 protein of *S. solfataricus* was not able to modify its own cOA [19]. This variation provides a new and exciting line of investigation for Type III systems. Once more studies are published, it will be interesting to uncover how conserved this activity is and whether or not there are evolutionary factors or conditional triggers that impact the behavior.

Future studies of the defense pathway of Type III systems should explore the question of recruitment of Csm6 and Csx1 to their respective complexes. Csm6 and Csx1 are not stable members of Type III complexes [22,23]. Thus, it is currently believed that the second messenger cOA molecules diffuse through the cell and bind to Csm6 or Csx1 without the proteins ever physically interacting with the complex. No studies have explored the possibility of a temporary physical connection between Csm6/Csx1 and the Type III complex. It is possible that target RNA binding triggers a temporary physical interaction between Csm6/Csx1 and the complex that has not yet been discovered. Evidence from Cascade complexes in Type I systems sets a precedent for this form of mechanism with Cas3 [24-26]. In Type I systems, Cas3 is a conditional member of the Cascade complex. Recruitment of Cas3 to the Cascade is triggered by R-loop formation after PAM recognition. It is possible that a similar mechanism exists in Type III systems. The discovery of this occurrence would represent a major breakthrough in the study of Type III systems.

The Role of the Target RNA PFS in Type III CRISPR-Cas Immunity

An important feature of the Type III-B system of *P. furiosus* is the existence of a three nucleotide PFS in the target RNA located 3' to the region of crRNA complementarity

and its impact on immunity in this organism. The results reported in this dissertation solidify the importance and robustness of PFS recognition for this system as previously discussed [12]. Interestingly, no evidence for the presence of a PFS requirement in other Type III systems has been reported though systematic analyses to reveal this feature have not yet been performed in most cases [27]. It is possible that more, but not all, Type III systems rely on PFS regions for defense as a result of evolutionary events or organismal differences. A specific implication for the use of PFS recognition in the *Pfu* Cmr system, in addition to controlling activation of the Cmr complex, may pertain to the form of DNA that *Pfu* Cmr is able to interact with. It is well established that Cmr2 is a ssDNA-specific deoxyribonuclease when tested in isolation from the other members of the complex [12,28,29]. Except in the *Pfu* Cmr system, all currently studied Type III systems have only observed *in vitro* nuclease activity with the entire Csm/Cmr complex when exposing the complex to ssDNA or dsDNA with an engineered artificial bubble [30-32]. However, when challenged with a fragment of dsDNA with or without an artificial bubble, *Pfu* Cmr was able to act on both types of DNA. This activity is PFS-dependent [12]. In Type I systems, PAM recognition is a critical step in unwinding of the dsDNA in preparation for cleavage by Cas3 [24,26]. A similar mechanism with PFS may be occurring in Type III systems, but on the RNA rather than DNA level, that has yet to be uncovered. Moving forward, investigation of this significant difference between *Pfu* Cmr and other Type III systems would benefit from novel structural analysis of the *Pfu* Cmr complex bound to target RNAs with and without PFS elements as well as additional biochemical analysis. Recent structural studies have provided evidence showing the interaction between the 5' tag of the crRNA and the PFS of the target RNA is critical for

stabilizing conformational changes that activate DNase and cyclase activities [18,33]. Thus, it is possible that unknown structural changes may be occurring within the *Pfu* Cmr complex upon PFS detection.

Divergence of Type III Systems

Although Type III systems share many commonalities in regards to function and mechanistic features, notable differences have been observed between Type III-A (Csm) and Type III-AB Cmr systems as well as between CRISPR-Cas containing organisms. This is likely due to the complexity of the CRISPR-Cas systems themselves and the organisms in which they inhabit. For example, differences in nucleotide specificity have been observed for Csm6 and Csx1 homologs [5,6,9,34]. In Chapter 2 of this dissertation, a preference for purines for the *L. lactis* and *S. epidermidis* Csm6 homologs was reported. However, similar analysis for *T. thermophilus* Csm6 resulted in a lack of evidence for nucleotide specificity. Such a difference does not dramatically affect the main functions and features of Csm6 and Csx1 homologs but would have significant implications if these proteins were to be repurposed in research or industry as tools for programmable enzymatic activities. The structures of Csm6 and Csx1 are highly conserved owing to the presence of the HEPN and CARF domains but there are notable differences among the homologs [2,6,35]. Csm6 proteins contain a 6H domain in between the HEPN and CARF domains while Csx1 contain a HTH domain in that region. The HTH domain of Csx1 has been shown to transmit the activation signal from the CARF to HEPN domain [36]. It is likely that the 6H domain region serves a similar function in Csm6 but that question has yet to be formally investigated and the impact of this difference elucidated.

Another prominent example pertains to whether Csm6 or Csx1 is important and/or sufficient *in vivo* for CRISPR-Cas immunity. In Chapter 2 of this thesis, it was reported that only Csm6 and the Palm domain of Csm1 were required for immunity [8]. The Palm domain is most likely contributing to the ribonuclease activity and not the DNA nuclease activity in those systems. In fact, even after robust and comprehensive testing *in vitro*, DNA degradation activity was not observed for the *L. lactis* Csm complex (unpublished data). Additional independent analysis into the *L. lactis* system of Csm also found that it degraded RNA *in vivo* but failed to degrade DNA *in vivo* [34]. Furthermore, a recently published study using the Csm system of *Mycobacterium tuberculosis* resulted in the same outcome that was observed in Chapter 2—only the RNA cleavage by Csm6 was needed for defense [13]. If a subset of Csm systems only target RNA as the primary defense response then it is possible that the plasmid loss witnessed for these *in vivo* studies arises from the failure of the antibiotic resistance gene to be transcribed and thus the cells are unable to confer resistance under selective conditions. Also, Type III-C and III-D systems exist and these system have Cas10 features that are naturally devoid of HD and Palm domains [38]. It is possible that these emerging subtypes perform similarly *in vivo* as the Csm systems of *L. lactis* and *M. tuberculosis*.

As the wealth of knowledge about Type III systems grows, more divergent features and nuances are uncovered. For instance, binding of the target RNA to the Type III complex is an essential step in the mechanism of action of these proteins. However, investigation into the Csm system of *Thermococcus onnurineus* was found to degrade ssDNA *in vitro* in the absence of target RNA [37]. Additionally, the composition of the mature crRNAs in Type III systems is highly conserved and required for functional

complexes [38-40]. However, recent analysis of the Type III-A system of Mycobacterium tuberculosis observed functional complexes with non-mature crRNA species that were not processed on the 3' end of the RNA. It appears the more Type III systems are investigated, the more the extent of diversity amongst the systems is revealed.

Recognition of DNA Targets in Type III Systems

Another interesting aspect of Type III systems that has yet to be deeply explored is the interaction between the target DNA and Type III complex. Currently no structures have been solved for Csm or Cmr bound to target DNA. Since Csm1 and Cmr2 are the nucleases associated with each complex, it is believed that the DNA directly and solely interacts with those subunits. This hypothesis is further supported by the fact that target RNA binding is required to activate Type III complexes and cleavage of the target RNA inactivates Type III complexes. Under these conditions it is unlikely that the majority of complex would bind to both nucleic acids. However, no published biochemical or structural evidence exists to confirm these hypotheses. A long-standing theory amongst Type III researchers is that DNA cleavage is coupled to transcription of the target DNA [12,31,32]. However, recent *in vitro* evidence from the Type III system of *S. epidermidis* shows that DNA cleavage is not exclusively coupled to the transcription process and thus may have multiple routes of being cleaved by Csm/Cmr complexes [41]. Perhaps there are unknown non-Cas elements that facilitate the recognition and positioning of the target DNA. More exploration into this area is needed.

Concluding Remarks

The data provide clear and supporting evidence for the necessity of HEPN family proteins in Type III defense. In the transcription-dependent Type III systems, the immune response remains dormant until cognate target RNA is recognized by the effector complex. Target RNA is the key to unlocking these multi-faceted systems. Outstanding unknowns remain surrounding target RNAs such as structural considerations, modes of regulation, specificity, tolerance, and the potential for collateral activity. This work has contributed to uncovering this interaction by characterizing three different HEPN family homologs, reconstituting cyclic oligoadenylate, and discovering new target RNA requirements using a combination of *in vivo* and *in vitro* approaches. The evidence presented in this dissertation provides critical knowledge about Type III systems and will be of great use for future applications of Type III systems.

REFERENCES

1. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013) Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* **8**, 15
2. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front Genet* **5**, 102
3. Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol Microbiol* **87**, 1088-1099
4. Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J Bacteriol* **196**, 310-317
5. Sheppard, N. F., Glover, C. V., 3rd, Terns, R. M., and Terns, M. P. (2016) The CRISPR-associated Csx1 protein of Pyrococcus furiosus is an adenosine-specific endoribonuclease. *RNA* **22**, 216-224
6. Niewoehner, O., and Jinek, M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA* **22**, 318-329
7. Jiang, W., Samai, P., and Marraffini, L. A. (2016) Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* **164**, 710-721

8. Foster, K., Kalter, J., Woodside, W., Terns, R. M., and Terns, M. P. (2019) The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR-Cas systems. *RNA Biol* **16**, 449-460

9. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksnys, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609

10. Niewoehner, O., Garcia-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., and Jinek, M. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543-548

11. Rostol, J. T., and Marraffini, L. A. (2019) Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat Microbiol* **4**, 656-662

12. Gootenberg, J. S., Abudayyeh, O. O., Kellner, M. J., Joung, J., Collins, J. J., and Zhang, F. (2018) Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **360**, 439-444

13. Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., and Terns, M. P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev* **30**, 447-459

14. Gruschow, S., Athukoralage, J. S., Graham, S., Hoogeboom, T., and White, M. F. (2019) Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. *Nucleic Acids Res* **47**, 9259-9270

15. Guo, T., Zheng, F., Zeng, Z., Yang, Y., Li, Q., She, Q., and Han, W. (2019) Cmr3 regulates the suppression on cyclic oligoadenylate synthesis by tag complementarity in a Type III-B CRISPR-Cas system. *RNA Biol* **16**, 1513-1520

16. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., Montoya, G., and She, Q. (2018) A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. *Nucleic Acids Res* **46**, 10319-10330

17. Nasef, M., Muffly, M. C., Beckman, A. B., Rowe, S. J., Walker, F. C., Hatoum-Aslan, A., and Dunkle, J. A. (2019) Regulation of cyclic oligoadenylate synthesis by the Staphylococcus epidermidis Cas10-Csm complex. *RNA* **25**, 948-962

18. Rouillon, C., Athukoralage, J. S., Graham, S., Gruschow, S., and White, M. F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* **7**

19. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., and Wang, Y. (2019) Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. *Cell* **176**, 239-253 e216

20. Athukoralage, J. S., Rouillon, C., Graham, S., Gruschow, S., and White, M. F. (2018) Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277-280

21. Athukoralage, J. S., Graham, S., Gruschow, S., Rouillon, C., and White, M. F. (2019) A Type III CRISPR Ancillary Ribonuclease Degrades Its Cyclic Oligoadenylate Activator. *J Mol Biol* **431**, 2894-2899

22. Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D. J. (2019) CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. *Mol Cell* **75**, 944-956 e946

23. Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., and White, M. F. (2013) Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol Cell* **52**, 124-134

24. Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J., Doudna, J. A., Heck, A. J. R., Yonekura, K., van der Oost, J., and Shinkai, A. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. *Mol Cell* **52**, 135-145

25. Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M. R., Barendregt, A., Zhou, K., Snijders, A. P., Dickman, M. J., Doudna, J. A., Boekema, E. J., Heck, A. J., van der Oost, J., and Brouns, S. J. (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* **18**, 529-536

26. Westra, E. R., van Erp, P. B., Kunne, T., Wong, S. P., Staals, R. H., Seegers, C. L., Bollen, S., Jore, M. M., Semenova, E., Severinov, K., de Vos, W. M., Dame, R. T., de Vries, R., Brouns, S. J., and van der Oost, J. (2012) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell* **46**, 595-605

27. Xiao, Y., Luo, M., Dolan, A. E., Liao, M., and Ke, A. (2018) Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science* **361**

28. Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O., and Marraffini, L. A. (2017) Broad Targeting Specificity during Bacterial Type III CRISPR-Cas Immunity Constrains Viral Escape. *Cell Host Microbe* **22**, 343-353 e343

29. Jung, T. Y., An, Y., Park, K. H., Lee, M. H., Oh, B. H., and Woo, E. (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. *Structure* **23**, 782-790

30. Ramia, N. F., Tang, L., Cocozaki, A. I., and Li, H. (2014) Staphylococcus epidermidis Csm1 is a 3'-5' exonuclease. *Nucleic Acids Res* **42**, 1129-1138

31. Estrella, M. A., Kuo, F. T., and Bailey, S. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Dev* **30**, 460-470

32. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C., and Siksnys, V. (2016) Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Mol Cell* **62**, 295-306

33. Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* **161**, 1164-1174

34. Jia, N., Mo, C. Y., Wang, C., Eng, E. T., Marraffini, L. A., and Patel, D. J. (2019) Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. *Mol Cell* **73**, 264-277 e265

35. Millen, A. M., Samson, J. E., Tremblay, D. M., Magadan, A. H., Rousseau, G. M., Moineau, S., and Romero, D. A. (2019) Lactococcus lactis type III-A CRISPR-Cas system cleaves bacteriophage RNA. *RNA Biol* **16**, 461-468

36. Kim, Y. K., Kim, Y. G., and Oh, B. H. (2013) Crystal structure and nucleic acid-binding activity of the CRISPR-associated protein Csx1 of Pyrococcus furiosus. *Proteins* **81**, 261-270

37. Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., Lopez-Mendez, B., She, Q., and Montoya, G. (2019) Structure of Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. *Nat Commun* **10**, 4302

38. Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2018) Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J* **1**, 325-336

39. Park, K. H., An, Y., Jung, T. Y., Baek, I. Y., Noh, H., Ahn, W. C., Hebert, H., Song, J. J., Kim, J. H., Oh, B. H., and Woo, E. J. (2017) RNA activation-independent DNA targeting of the Type III CRISPR-Cas system by a Csm complex. *EMBO Rep* **18**, 826-840

40. Carte, J., Christopher, R. T., Smith, J. T., Olson, S., Barrangou, R., Moineau, S., Glover, C. V., 3rd, Graveley, B. R., Terns, R. M., and Terns, M. P. (2014) The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in Streptococcus thermophilus. *Mol Microbiol* **93**, 98-112

41. Carte, J., Pfister, N. T., Compton, M. M., Terns, R. M., and Terns, M. P. (2010) Binding and cleavage of CRISPR RNA by Cas6. *RNA* **16**, 2181-2188

42. Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev* **22**, 3489-3496

43. Liu, T. Y., Liu, J. J., Aditham, A. J., Nogales, E., and Doudna, J. A. (2019) Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble. *Nat Commun* **10**, 3001