

Washington State University Vancouver ILL



ILLiad TN: 556096

Borrower: NNM

Lending String: OCA,*WS2,HVC,BCN,CSO

Patron: Vrana, Caroline

Journal Title: Molecular cell

Volume: Issue:

Month/Year: 2012**Pages:** 292-302

Article Author:

Article Title: Hale, C.R. et al; Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNA

Imprint: [Cambridge, Mass.] ; Cell Press, c1997-

ILL Number: 91438801 ✓



Call #: ELSEVIER

Location:

ODYSSEY ENABLED

Charge

Maxcost: 25.00IFM

Shipping Address:

Library, ILL
Davidson College Library - ILL
209 Ridge Rd
Davidson, NC 28035-7200

Fax:

Ariel:

Email: ill@davidson.edu

Essential Features and Rational Design of CRISPR RNAs that Function with the Cas RAMP Module Complex to Cleave RNAs

Caryn R. Hale,¹ Sonali Majumdar,¹ Joshua Elmore,¹ Neil Pfister,^{1,5} Mark Compton,³ Sara Olson,⁴ Alissa M. Resch,⁴ Claiborne V.C. Glover, III,¹ Brenton R. Graveley,⁴ Rebecca M. Terns,^{1,*} and Michael P. Terns^{1,2,*}

¹Department of Biochemistry and Molecular Biology

²Department of Genetics

³Department of Poultry Science

University of Georgia, Athens, GA 30602, USA

⁴Department of Genetics and Developmental Biology, University of Connecticut Stem Cell Institute, University of Connecticut Health Center, Farmington, CT 06030-6403, USA

⁵Present address: Biological Sciences, Columbia University, New York, NY 10027, USA

*Correspondence: rterns@bmb.uga.edu (R.M.T.), mterns@bmb.uga.edu (M.P.T.)

DOI 10.1016/j.molcel.2011.10.023

SUMMARY

Small RNAs target invaders for silencing in the CRISPR-Cas pathways that protect bacteria and archaea from viruses and plasmids. The CRISPR RNAs (crRNAs) contain sequence elements acquired from invaders that guide CRISPR-associated (Cas) proteins back to the complementary invading DNA or RNA. Here, we have analyzed essential features of the crRNAs associated with the Cas RAMP module (Cmr) effector complex, which cleaves targeted RNAs. We show that Cmr crRNAs contain an 8 nucleotide 5' sequence tag (also found on crRNAs associated with other CRISPR-Cas pathways) that is critical for crRNA function and can be used to engineer crRNAs that direct cleavage of novel targets. We also present data that indicate that the Cmr complex cleaves an endogenous complementary RNA in *Pyrococcus furiosus*, providing direct in vivo evidence of RNA targeting by the CRISPR-Cas system. Our findings indicate that the CRISPR RNA-Cmr protein pathway may be exploited to cleave RNAs of interest.

INTRODUCTION

The CRISPR-Cas systems are RNA-based immune systems that protect prokaryotes from viruses, plasmids, and other invaders (see recent general reviews Deveau et al., 2010; Horvath and Barrangou, 2010; Jore et al., 2011a; Karginov and Hannon, 2010; Makarova et al., 2011; Marraffini and Sontheimer, 2010; Terns and Terns, 2011; van der Oost et al., 2009). The clustered regularly interspaced short palindromic repeat (CRISPR) loci found in prokaryotic genomes acquire short fragments of invader sequence (~30–40 base pairs in length) that are inserted between short CRISPR repeat

sequences. The CRISPR loci give rise to CRISPR RNAs (crRNAs) that each contain an invader-derived (guide) sequence and interact with CRISPR-associated or Cas proteins to form effector complexes that recognize and silence the corresponding invader (Brouns et al., 2008; Hale et al., 2008, 2009; Jore et al., 2011b; Lintner et al., 2011; Wiedenheft et al., 2011). There are multiple modules of Cas proteins (termed, for example, Cse and Csy proteins) that can function independently with crRNAs to effect CRISPR-Cas defense (Haft et al., 2005; Makarova et al., 2006, 2011; Terns and Terns, 2011); thus distinct CRISPR-Cas pathways function in various prokaryotes.

CRISPR-Cas-mediated silencing appears to occur through sequence-specific targeting of RNA or DNA. In *Streptococcus thermophilus* (Barrangou et al., 2007; Gameau et al., 2010), *Escherichia coli* (Brouns et al., 2008; Semenova et al., 2011), *Staphylococcus epidermidis* (Marraffini and Sontheimer, 2008), and *Sulfolobus solfataricus* (Gudbergdottir et al., 2011; Manica et al., 2011), there is evidence that CRISPR-Cas systems target foreign DNA. Indeed, sequence-specific cleavage of both plasmid and bacteriophage DNA was recently documented in *S. thermophilus* (Gameau et al., 2010). At the same time, the Cmr-type crRNA-Cas protein complex from *Pyrococcus furiosus* specifically cleaves complementary RNAs (Hale et al., 2009). Thus, the various CRISPR-Cas systems (formed by distinct modules of Cas proteins) may function by different mechanisms, including DNA and RNA cleavage. Interestingly, many organisms possess more than one module of Cas proteins and may resist invaders by multiple mechanisms to provide robust protection from diverse invaders.

CRISPR RNAs (previously also referred to as prokaryotic silencing [psi]RNAs) are utilized by all versions of the CRISPR-Cas system characterized to date (Brouns et al., 2008; Deltcheva et al., 2011; Gudbergdottir et al., 2011; Hale et al., 2009; Lintner et al., 2011; Manica et al., 2011; Marraffini and Sontheimer, 2008; Wiedenheft et al., 2011). The CRISPR loci are transcribed from a leader sequence region of 100–500 base pairs found at one end of each locus (Jansen et al., 2002), where promoter sequences have been identified in a few organisms

(Lillestol et al., 2009; Pui et al., 2010). Insertion of new invader-derived sequences into CRISPR loci most frequently occurs immediately downstream of the leader (Andersson and Banfield, 2008; Barrangou et al., 2007; Pourcel et al., 2005; Shah et al., 2009). The CRISPR locus transcripts are cleaved within the repeat sequence to release the individual embedded crRNAs. A series of structurally related proteins from various CRISPR-Cas systems process CRISPR RNA transcripts: Cas6 (Carte et al., 2008, 2010; Wang et al., 2011), Cse3 (Brouns et al., 2008; Gesner et al., 2011; Sashital et al., 2011), and Csy4 (Haurwitz et al., 2010; Przybilski et al., 2011). Cleavage within the repeats by these Cas endonucleases produces 1 × intermediate RNAs generally comprised of 8 nucleotides (nt) of repeat sequence (called the 5' tag), a guide sequence, and the remaining ~20–25 nt of repeat sequence at the 3' end (the 3' tag) (Brouns et al., 2008; Gesner et al., 2011; Hale et al., 2008; Haurwitz et al., 2010). These 1 × intermediates undergo various amounts of 3' end trimming (Hale et al., 2008, 2009; Jore et al., 2011b; Lintner et al., 2011; Marraffini and Sontheimer, 2008; Wiedenheft et al., 2011). However, an equivalent 8 nt 5' tag is present on mature crRNAs from a variety of bacterial and archaeal species including *E. coli* (Brouns et al., 2008), *S. epidermidis* (Marraffini and Sontheimer, 2008), *Pseudomonas aeruginosa* (Haurwitz et al., 2010), *P. furiosus* (Hale et al., 2008, 2009), *S. solfataricus* (Lintner et al., 2011; our unpublished data), *Sulfolobus tokadaei* (our unpublished data), and *Thermococcus kodakarensis* (our unpublished data). In *S. epidermidis*, the 5' tag of the crRNAs provides for discrimination between self (i.e., the complementary DNA strand in the CRISPR locus, which can base pair with the tag) and nonself (invader DNA, which does not base pair with the tag) (Marraffini and Sontheimer, 2010), but other functions of this crRNA signature sequence are not known.

There are three modules of Cas proteins (Cmr, Csa, and Cst) and seven CRISPR loci encoded in the genome of the hyperthermophilic archaeon *P. furiosus*. Previously, we identified and characterized a complex comprised of the Cas module RAMP or Cmr proteins (Cmr1–6) and two crRNA species in *P. furiosus* (Hale et al., 2009). The Cmr-crRNA protein complex is unique among CRISPR-Cas systems characterized to date in that it targets RNA; the complex specifically cleaves RNAs that are complementary to the crRNAs, 14 nt from the 3' end of the crRNAs (Hale et al., 2009). Activity was observed in vitro using both native and reconstituted complexes but has not been described in vivo. Here we have investigated the functionality of the crRNAs that direct cleavage by the Cmr complex. Deep sequencing of the RNAs associated with the complex as well as total crRNAs unexpectedly revealed the cleavage of a fortuitous endogenous target RNA by the Cmr complex in vivo. While multiple species of crRNAs are present in *P. furiosus*, the Cmr complex selectively includes 45 and 39 nt crRNAs regardless of the size of the invader-derived guide sequence encoded in the genome. We show that the 8 nt 5' tag sequence is not only necessary for Cmr-directed cleavage but can be used to target the complex to cleave novel RNAs, providing the potential for manipulation of the CRISPR-Cas system for experimental or pharmaceutical gene silencing.

RESULTS

Key Features of the Cmr Complex RNAs

To delineate functional features of the crRNAs that guide cleavage by the Cmr complex, we isolated active Cmr complexes using antibodies against the Cmr2 protein. Polyclonal antibodies raised against recombinant *P. furiosus* Cmr2 specifically immunoprecipitated Cmr2 (assessed by western blotting, Figure 1A), as well as the five other Cmr complex proteins (Cmr1 and Cmr3–6; assessed by mass spectrometry, S.M., C.R.H., N.P., M.C., P. Zhao, L. Wells, R.M.T., and M.P.T., unpublished data), and 39 and 45 nt species of a crRNA (assessed by northern blotting, Figure 1B). Moreover, the immunoprecipitated RNA-protein complex replicated the cleavage activity observed with both the chromatographically purified and the reconstituted Cmr complexes (Hale et al., 2009), cleaving a target RNA (complementary to one of the endogenous crRNAs) at the two sites located 14 nt upstream of the 3' ends of the complementary 39 and 45 nt crRNA species (Figure 1C).

We probed the nature of the end groups of the crRNAs immunopurified with Cmr complexes by end radiolabeling. 5' end labeling of the 39 and 45 nt RNAs was not affected by prior enzymatic treatment with phosphatase, whereas phosphatase pretreatment significantly increased 3' end labeling of the RNAs (Figure 1D), indicating that the majority of both species of crRNAs in the Cmr complex contain 5' hydroxyl and 3' phosphate end groups.

In order to further characterize the Cmr-associated crRNAs, we deep sequenced the RNAs immunopurified with the complex, as well as total *P. furiosus* small RNAs. We isolated RNAs smaller than ~70 nt by gel purification, and based on the results of end group analysis (Figure 1D), treated the RNAs with phosphatase to optimize cloning. The cDNA libraries were subject to Illumina sequencing. Figure 2A shows the numbers of small RNA sequence reads that map to each of the seven *P. furiosus* CRISPR loci from the total RNA and Cmr complex samples. A total of 2,632,773 uniquely mapped small RNA reads were obtained from the total RNA sample, of which 54.5% mapped to the CRISPR loci. In contrast, 91.2% of the 990,060 uniquely mapped reads from the Cmr complex sample mapped to the CRISPR loci.

The total RNA profiles indicate that RNAs are expressed from all seven *P. furiosus* CRISPR loci (Figure 2A, total RNA), as limited sequencing data had previously suggested (Hale et al., 2009). We analyzed the leader sequences of the seven CRISPR loci and found canonical BRE/TATA promoter sequences at a conserved position within the leaders (Figure 2B). A population of RNAs was detected that begins ~21 nt downstream of these promoter elements (and includes ~29 nt of the leader sequence at the 5' end, indicated in red in Figure 2B). (These RNAs do not represent the major crRNA species, which are described below.) These findings suggest that the CRISPR loci are transcribed by archaeal RNA polymerase via recognition of the BRE/TATA promoters within the leader sequences. The deep sequencing data also indicate that, in general, crRNAs from the leader-proximal regions of the CRISPR loci are more abundant than crRNAs encoded downstream, as had been hypothesized based on previous data (Hale et al., 2008). The molecular basis for the

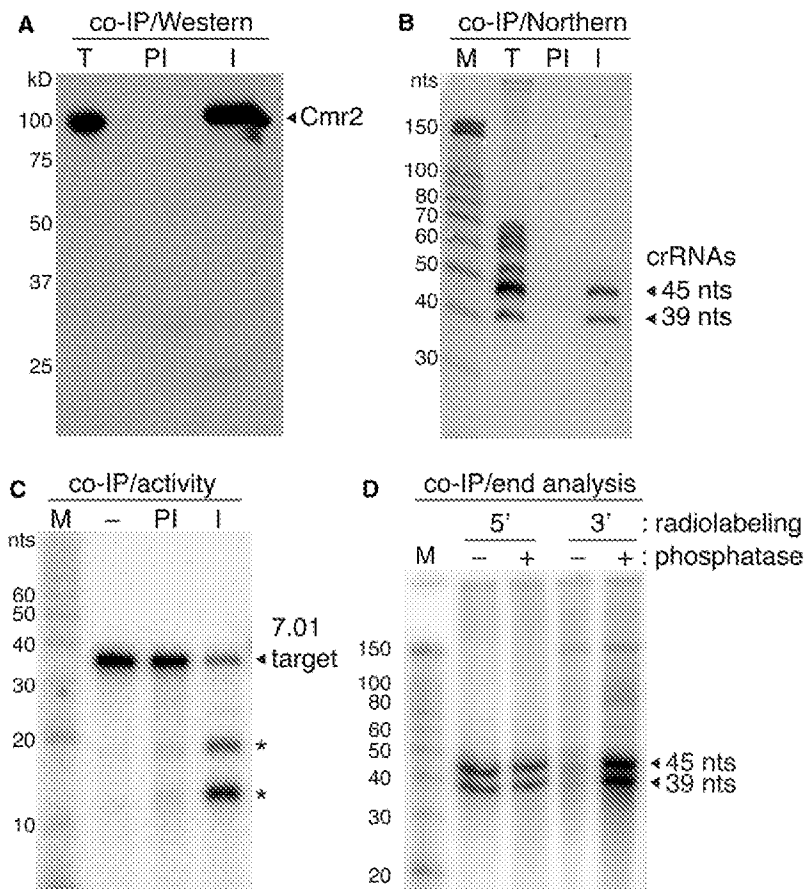


Figure 1. Cmr2 Antibodies Immunopurify Active Cmr-crRNA Complexes

(A) Cmr2 antibodies recognize and specifically immunoprecipitate ~100 kD protein. *P. furiosus* S100 extract (T) and PfCmr2 antibody immune (I) and preimmune (PI) immunoprecipitation (IP) samples were analyzed by western blotting with Cmr2 antibodies. Positions of size markers are indicated. Arrow indicates ~100 kD protein specifically recognized by immunoblotting and immunoprecipitated by Cmr2 antibodies.

(B) The Cmr2 antibodies immunopurify 45 and 39 nt crRNAs. RNAs extracted from *P. furiosus* S100 extract (T), and immune (I) and preimmune (PI) immunoprecipitation (IP) samples were analyzed by northern blotting with a probe against *P. furiosus* crRNA 7.01. The 45 and 39 nt species of 7.01 that are coimmunoprecipitated with the Cmr2 antibodies are indicated by arrows. Sizes of RNA markers (M) are shown.

(C) Immunoprecipitated Cmr complexes cleave a target RNA. 5' end-labeled 7.01 target RNA (complementary to crRNA 7.01, indicated with arrow) was incubated in the absence (-) of proteins or in the presence of immune (I) and preimmune (PI) immunoprecipitation samples. Cleavage products are marked with asterisks.

(D) Coimmunopurified crRNAs possess 5' OH and 3' phosphate ends. Immunoprecipitated crRNAs were isolated and 5' or 3' labeled with (+) and without (-) prior phosphatase treatment. The 45 and 39 nt crRNA species are indicated by arrows, and RNA size markers (M) are shown.

gradient in steady-state levels of the crRNAs from each locus is currently unknown, but could reflect differences in transcription, processing, or stability. The expression gradient is noteworthy because the crRNAs that arise from the leader-proximal end of a CRISPR locus generally target the most recent invaders confronted by the system (Andersson and Banfield, 2008; Barrangou et al., 2007; Pourcel et al., 2005; Shah et al., 2009).

Deep sequencing and northern analysis indicate that a series of species of crRNAs of different lengths are present in *P. furiosus*, and that the Cmr complex specifically loads two of these crRNA species. The 5' ends of the vast majority of crRNAs in *P. furiosus*, in both the total and Cmr complex samples, are comprised of an identical 5' end tag sequence (AUUGAAAG) derived from the upstream repeat element (Figure 3A). This end is generated by Cas6 cleavage of CRISPR transcripts at this site (Carte et al., 2008, 2010; Wang et al., 2011). In total RNA, a mixture of 3' ends corresponding to crRNAs 30–70 nt in length is observed (Figure 3A), consistent with a series of RNA species detected in total RNA by northern analysis of one of the crRNAs (Figure 1B, total [T] lane). Strikingly, the 39 and 45 nt crRNA species selectively copurify with the Cmr complex (Figures 1B and 3A). The 45 nt crRNAs constitute 46% of the crRNAs associated with the Cmr complex, and the 39 nt species make up 38% (Figure 3A).

Interestingly, the deep sequence analysis indicates that the 3' ends of the Cmr complex-associated crRNAs are defined by distance from the 5' tag or absolute length of the crRNA (rather

than, for example, by the length of the guide sequence encoded in the CRISPR locus or by position relative to the 3' repeat element). Guide sequences in *P. furiosus* CRISPRs are most commonly 37 nt long, in which case the 45 nt crRNA species is comprised of an 8 nt repeat tag sequence and a full 37 nt guide sequence. (The 39 nt species simply lacks 6 nt of the guide sequence at the 3' end; see Figure 3A.) However, some guide sequences are as short as 34 nt and as long as 59 nt in *P. furiosus*, and Cmr complex crRNAs are nonetheless consistently 39 and 45 nt in length. For example, while three of the first five guide sequences in CRISPR locus 4 are 37 nt long, the second guide sequence in the locus (termed 4.02) is 35 nt; yet the 4.02 crRNAs found in the Cmr complex are 45 and 39 nt in length (Figure 3B). The 45 nt 4.02 crRNA includes 2 nt of the downstream repeat sequence. Conversely, the fifth CRISPR 4 guide sequence is 1 nt longer than standard (38 nt), and the 45 nt 4.05 crRNA lacks the last encoded nucleotide of the guide sequence (Figure 3B). Similarly, the first guide sequence in CRISPR 1 is 48 nt long (11 nt longer than is typical), and the 45 nt 1.01 crRNA found in the Cmr complex includes only 37 nt of the 48 nt guide sequence (see Figure 4B). Thus, regardless of the size of the guide sequence encoded in the genome, the crRNAs of the Cmr complex are comprised of an 8 nt 5' repeat tag and a 37 or 31 nt guide sequence.

Cleavage of a Complementary RNA In Vivo

Very few RNAs were detected that were transcribed from the opposite (nonleader) ends of the CRISPR loci in *P. furiosus*;

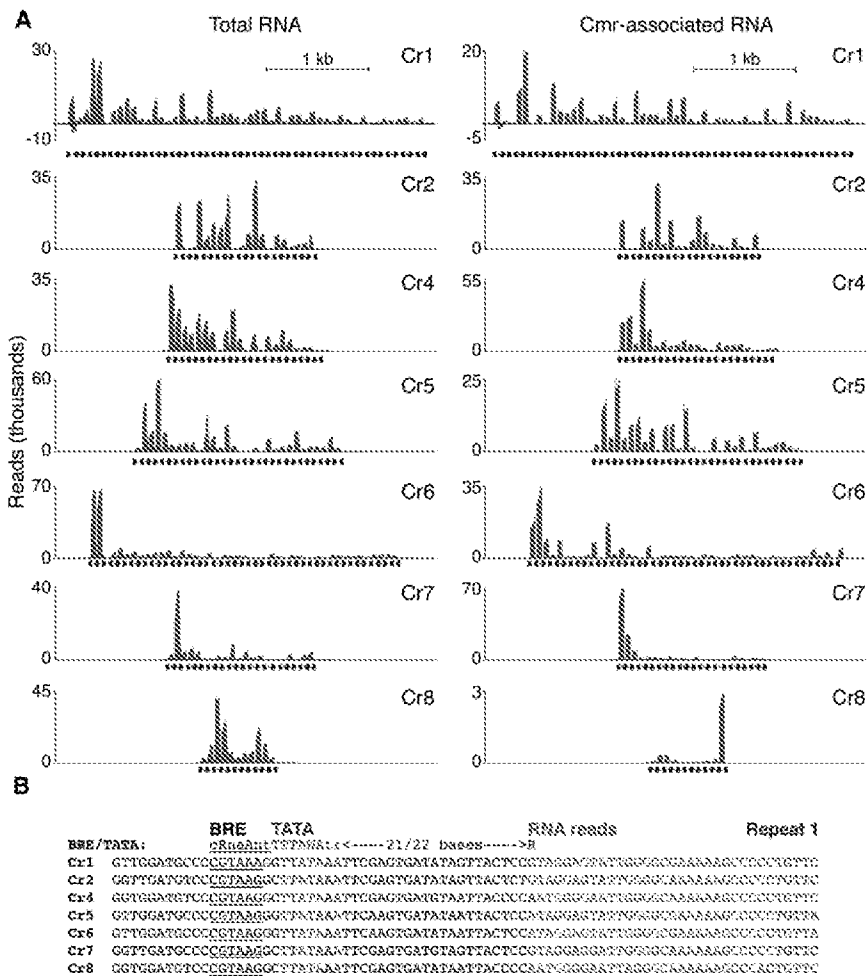


Figure 2. Deep Sequencing Profiles of crRNAs Found in Total *P. furiosus* RNA and Immunopurified Cmr Complexes

(A) Total and Cmr-associated small RNA sequencing reads that map to the seven *P. furiosus* CRISPR loci. Y axis scale bars indicate the number of reads (thousands) that contain nucleotides that map to the indicated position within the CRISPR locus. Red indicates sequence reads that map to the sense strand relative to the CRISPR leader region, and blue corresponds to reads that map to the antisense strand. The positions of CRISPR repeats are shown as black bars below the graph. An x axis scale of 1 kb is indicated. Images were generated using the UCSC archaeal genome browser (Karolchik et al., 2003).

(B) Putative promoter sequences found in the *P. furiosus* CRISPR leader regions. Alignment of the regions upstream of the first repeat of each *P. furiosus* CRISPR locus is shown. The repeats are shown in blue, the initial transcribed regions (detected by sequencing) are shown in red, the potential TATA elements are shown in pink, the potential BRE elements are underlined, and other leader sequences are shown in black. Consensus BRE/TATA elements are shown above the alignment.

however, we did detect significant antisense transcription from an apparently fortuitous promoter within CRISPR locus 1. A BRE/TATA promoter sequence is present in the antisense orientation within the second guide sequence of CRISPR locus 1, and a significant number of antisense RNAs that begin distinctly 21 nt downstream of the promoter are present in the sequence from the total RNA sample (see Figure S1, blue reads, available online). The number of antisense RNA reads is approximately one-third the number of reads observed for the sense strand crRNA reads (crRNA 1.01; Figure S1, red reads). Northern analysis reveals a series of antisense RNAs up to ~140 nt in length in total *P. furiosus* RNA (Figure 4A, 1.01 antisense, T lane).

The antisense CRISPR transcripts found in *P. furiosus* would be predicted to be recognized and targeted by Cmr complexes containing the crRNA encoded on the opposite strand of CRISPR locus 1 (crRNA 1.01). Previous in vitro studies established that the Cmr complex cleaves complementary RNAs 14 nt upstream of the 3' ends of the 39 and 45 nt crRNAs (Hale et al., 2008). Accordingly, in this case, the 1.01 crRNAs would be predicted to generate 5' products of 45 and 39 nt (as well as 3' products of various lengths) upon cleavage of the antisense

transcripts. And indeed, distinct antisense RNAs of 45 and 39 nt were observed in northern analysis of the total RNA (with a probe to the 5' region of the antisense transcript; Figure 4A, 1.01 antisense, T lane), suggesting that the Cmr complex targets and cleaves the endogenous CRISPR1 antisense transcript in *P. furiosus*.

We also found in both RNA sequencing and northern analysis that the 45 and 39 nt putative 5' cleavage products are specifically coimmunopurified with the Cmr complex (Figure 4A, 1.01 antisense, immune [I] versus preimmune [PI] lanes; and Figure 4B, blue reads, versus Figure S1, blue reads), suggesting that the products remain associated with the complex to some extent following cleavage. The 45 and 39 nt antisense RNAs do not possess the 5' repeat tag sequence, which we hypothesize is important for recognition and function of crRNAs with the Cmr complex. However, because of the coincidence in the size of these RNAs with Cmr complex crRNA species, we tested the possibility that the antisense RNAs are functional guide RNAs. We found that the immunopurified Cmr complexes do not cleave a target RNA complementary to the antisense RNAs, indicating that the antisense RNAs do not guide cleavage by the Cmr complex (Figure 4C, antisense RNA target). At the same time, the immunopurified complexes do cleave an RNA corresponding to the ~140 nt antisense RNA transcript, generating 5' products of the same sizes observed in vivo (45 and 39 nt; Figure 4C, crRNA target). Together, the findings indicate that the Cmr complex recognizes and cleaves endogenous complementary RNAs in vivo.

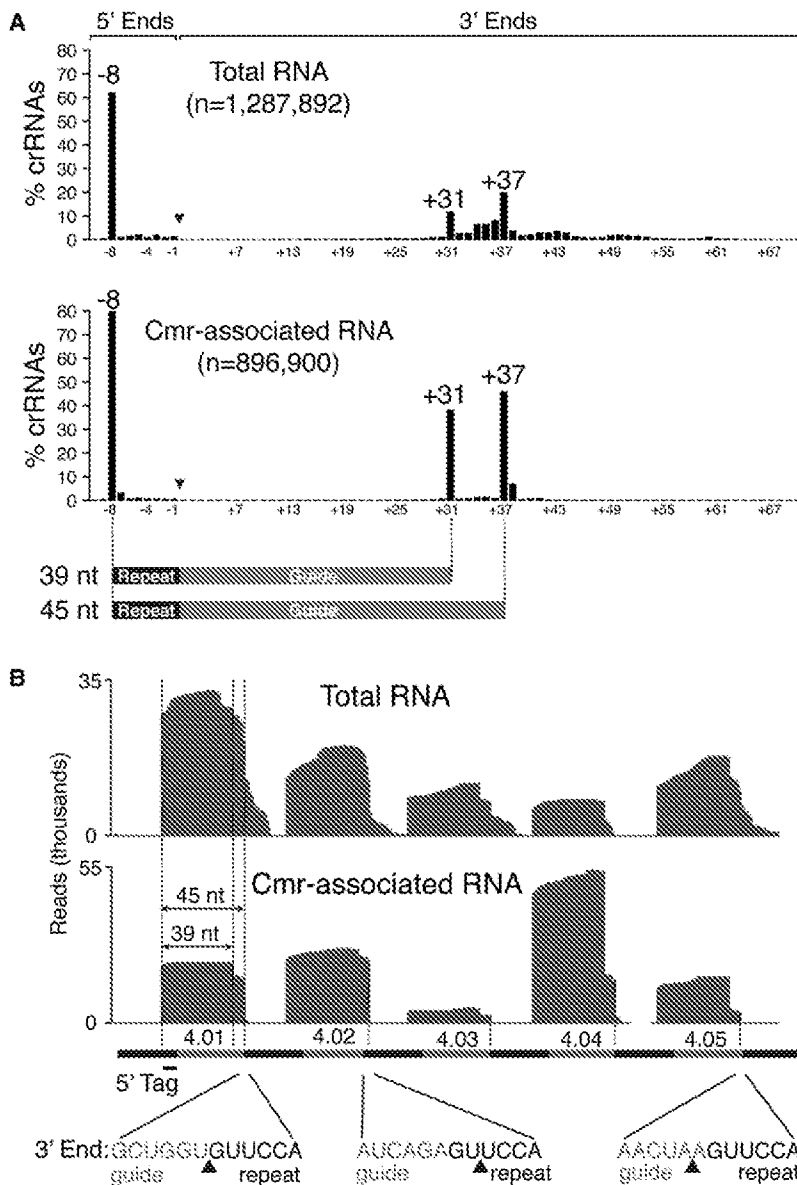


Figure 3. Cmr Complex crRNAs Contain a Conserved 8 nt 5' Sequence Tag and Are of Defined 39 and 45 nt Lengths

(A) The 5' and 3' ends of Cmr-associated crRNAs are defined relative to the upstream repeat element. Graphs show the percentage of crRNAs sequenced in total *P. furiosus* RNA (upper panel) and immunopurified Cmr complexes (lower) with ends that map at indicated positions relative to the 5' repeat-guide junction (arrowhead). The Cmr-associated crRNA species are illustrated below with 5' repeat tag (black) and guide sequence (green) indicated. RNAs from CRISPR8 were excluded from this analysis due to a repeat sequence polymorphism (see text). Read numbers analyzed (n) are indicated.

(B) The length of the Cmr complex crRNAs is independent of the length of the genome-encoded guide sequence. Deep sequence profiles of crRNAs 4.01–4.05 in total *P. furiosus* RNA (upper panel) and immunopurified Cmr complexes (lower) are shown relative to repeat (black) and guide (green) sequence elements (see also Figure 2A). Dashed lines mark the 3' ends of the 45 nt crRNA species. (The 5' end and 3' end of the 39 nt species are also indicated for crRNA 4.01.) The sequence at the 3' guide-repeat junction is shown below for crRNAs 4.01, 4.02, and 4.05. Guide sequences are shown in green, and repeat sequences are shown in black. Arrowheads indicate the locations of the 3' ends of the 45 nt crRNA species.

from locus 8 were significantly underrepresented in the Cmr complex (Figure 2A and Figure S2). The reduced presence of CRISPR locus 8 RNAs in the Cmr complex correlates with a difference in the 5' repeat tag found on these RNAs. Most of the crRNAs from locus 8 possess a 7 nt AUUGAAG (rather than 8 nt AUUGAAAG) repeat tag due to a variation in the repeat sequences within this CRISPR locus (Figure 5C). The exception is crRNA 8.11 (the last crRNA encoded by locus 8), which retains the canonical 8 nt repeat tag sequence, and this crRNA is selectively enriched in the Cmr complex relative to the other crRNAs from CRISPR locus 8 (Figures 2A and 5C). Together, these results

indicate that the CRISPR repeat tag is critical for the formation of functional Cmr complexes by crRNAs.

For CRISPR-Cas systems that target DNA, the 5' repeat tag on the crRNAs provides a key function in distinguishing self from nonself. In organisms with these systems, it is important for crRNAs to avoid targeting the complementary DNA sequence found in the host CRISPR locus itself, and evidence indicates that in these systems potential targets with complementarity to the tag sequence (found in the CRISPR locus but not in invader targets) are not silenced (Marraffini and Sontheimer, 2010). We tested the effect of target complementarity to the tag sequence on RNA cleavage by the Cmr complex, and observed no significant reduction in activity (Figure 5D). The antisense transcript that is cleaved in vivo by the Cmr complex also is complementary to the tag of crRNA 1.01 (Figure 4). Thus, RNA targeting by the

The CRISPR Repeat Tag Sequence Is Required for Functional Cmr Complex crRNAs

Its conservation suggests that the 5' CRISPR repeat tag is an important element of functional Cmr complex crRNAs. We investigated the importance of the tag by testing several tag sequence mutants for the ability to guide cleavage by reconstituted Cmr complexes. Removal of the repeat tag sequence from crRNA 7.01 eliminates cleavage of a complementary target RNA (Figure 5A). Substitution of the entire tag sequence to its complement (AUUGAAAG → UAACUUUC) or of the first two nucleotides of the tag sequence (AUUGAAAG → GAUGAAAG) also prevents function of the crRNA, as does the addition of a G at the 5' end of the tag (Figure 5B).

Analysis of CRISPR locus 8 provided further insight into the basis for the importance of the 5' repeat tag in vivo. crRNAs

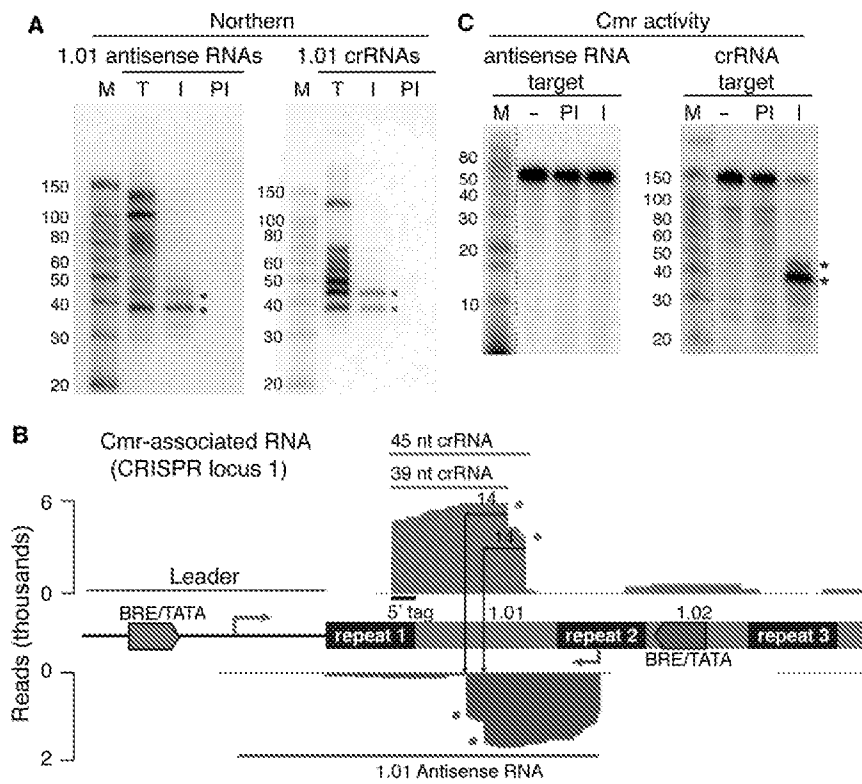


Figure 4. Antisense RNA Found in *P. furiosus* Is Cleaved by the Cmr-crRNA Complex

(A) Antisense RNA from the region of crRNA 1.01 is detected by northern analysis, and two species are associated with the Cmr complex. Northern analysis of total *P. furiosus* RNA (T) and RNAs isolated from immune (I) or preimmune (PI) Cmr2 immunopurifications was probed for antisense (left panel) or sense 1.01 crRNAs (right panel). Colored dots indicate the primary species of antisense RNAs (blue dots) and crRNAs (red dots) specifically associated with the Cmr complex (and correspond with dots in panel B). Sizes of radiolabeled RNA marker (M) are noted.

(B) The two antisense RNAs associated with the Cmr complex map downstream of a potential fortuitous promoter and correspond to the expected products of crRNA 1.01-guided cleavage. Deep sequencing profiles of Cmr-associated crRNAs in the region encoding crRNA 1.01 are shown. Red indicates RNAs transcribed from the leader region, and blue corresponds to reads from the opposite strand. The positions of CRISPR repeats (black), guide sequences (green), and BRE/TATA promoters (red and blue) are indicated. A blue line represents the apparent full-length antisense RNA observed by deep sequencing total RNA (see Figure S1). Cleavage of the antisense RNA by the Cmr complex 14 nt from the 3' ends of the 45 and 39 nt species of crRNA 1.01 (red dots) would produce antisense RNA products of the sizes observed by deep sequencing and northern analysis (blue dots).

(C) Immunopurified Cmr complexes cleave the antisense RNA. Radiolabeled RNA targets for the antisense RNAs (left panel) and crRNAs (right panel) were incubated in the absence (-) of proteins or in the presence of immune (I) or preimmune (PI) immunopurified preparations, and the products were analyzed following denaturing gel electrophoresis. (The target for the crRNAs is 140 nt in length and corresponds to the full-length antisense RNA; see C.) The cleavage products obtained by incubation of the crRNA target with the immunopurified Cmr complex (asterisks) are approximately the same sizes as the endogenous antisense RNA species that are immunopurified with the Cmr complexes (A, left panel). Noncontiguous lanes from the same gel are indicated by dashed lines.

Cmr complex does not require a lack of complementarity between the target and the crRNA repeat tag.

crRNAs Can Be Engineered to Direct the Cmr Complex to Cleave Novel Target RNAs

Our findings indicate that a 5' repeat tag is necessary for a crRNA to function to guide cleavage by the Cmr complex. To determine whether the Cmr complex can be directed to cleave novel target RNAs by engineered crRNAs, we designed crRNAs with the 5' AUUGAAAG tag sequence and guide sequences complementary to a random target sequence (X') or to a sequence near the 5' end of the β -lactamase (*bla*) mRNA. Cmr complexes reconstituted with the crRNA designed to cleave the random target sequence specifically cleaved the X' target RNA but not a *bla* target sequence (Figure 6A). Similarly, the crRNA designed against the 5' region of the *bla* mRNA directed cleavage of the *bla* target sequence but not the X' target RNA (Figure 6A). The cleavage product sizes are consistent with target cleavage 14 nt from the 5' ends of the crRNAs. The 5' *bla* crRNA and a crRNA designed against an internal region of the *bla* mRNA were also tested for the ability to direct cleavage in the context of the full-length *bla* mRNA. The 5' radiolabeled *bla* mRNA (~860 nt) was cleaved by Cmr complexes reconstituted with both *bla*

crRNAs, yielding products of the expected sizes (~17 nt for the 5' *bla* crRNA and ~236 nt for the internal *bla* crRNA) (Figure 6B). These results demonstrate that the Cmr complex can be directed to cleave novel target RNAs at a desired site by crRNAs designed with an appropriate repeat tag sequence and a guide region complementary to the target cleavage site.

DISCUSSION

The RNA Targeting Branch of the CRISPR-Cas Immune System

There is tremendous diversity in the CRISPR-Cas immune pathways that protect prokaryotes from invaders (recently reviewed in Deveau et al., 2010; Garrett et al., 2011; Horvath and Barrangou, 2010; Jore et al., 2011a; Karginov and Hannon, 2010; Makarova et al., 2011; Marraffini and Sontheimer, 2010; Terns and Terns, 2011; van der Oost et al., 2009). The different CRISPR-Cas systems are populated by distinct modules of proteins, and the mechanisms of key steps in the pathways, including crRNA biogenesis and invader silencing, can be fundamentally different. In this work, we investigated the Cmr effector complex, which is unique among systems characterized to date in that it targets RNAs rather than DNAs (Hale et al., 2009). This

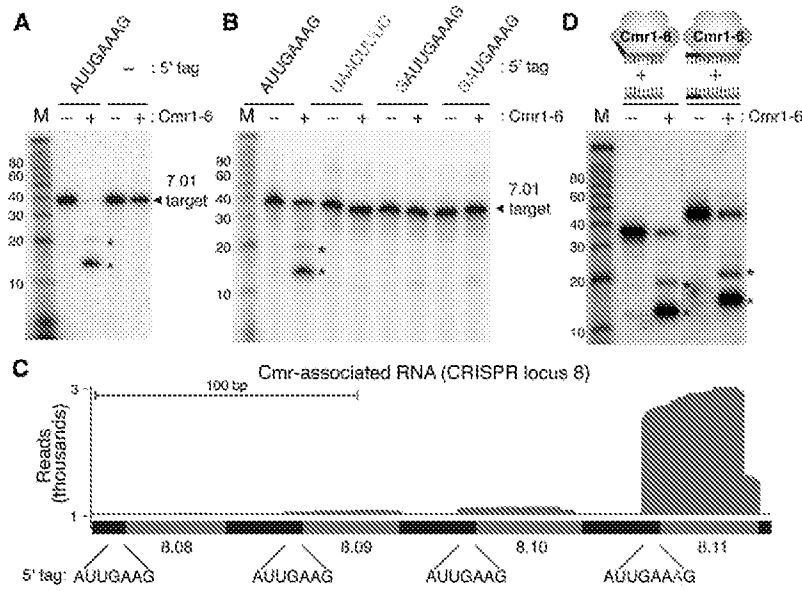


Figure 5. The 5' Repeat Tag Is Required for Formation of Functional Cmr Complexes

(A) The 5' repeat tag is required for Cmr complex function. Reconstitutions were performed with recombinant Cmr1-6 proteins and the 45 nt wild-type 7.01 crRNA (AUUGAAAG tag) or crRNA 7.01 lacking the tag (-), and activity was tested against the radiolabeled 7.01 target RNA. RNA marker sizes are indicated. The cleavage product observed with the wild-type crRNA is indicated with an asterisk.

(B) Mutations to the tag sequence prevent function. Reconstitutions were performed with the 45 nt wild-type 7.01 crRNA (AUUGAAAG tag) or crRNA 7.01 with the indicated mutations in the tag sequence (mutated or inserted nucleotides are highlighted in red) and tested in cleavage assays as described in (A). The cleavage product observed with the wild-type crRNA is indicated with an asterisk.

(C) crRNAs with variant tag sequences are underrepresented in the Cmr complex. Deep sequencing profiles of Cmr-associated crRNAs in the region encoding crRNAs 8.08–8.11 are shown (see also Figure 2A). The positions of CRISPR repeats (black) and guide sequences (green) are indicated. The 5' tag sequence is shown below for crRNAs 8.08–8.11. An additional nucleotide present in the wild-type tag of crRNA 8.11 is highlighted in red.

(D) Complementarity to the tag does not prevent function. Cmr complexes reconstituted with wild-type crRNA 7.01 were tested for the ability to cleave the 7.01 target RNA or a target with complementarity to the tag sequence (as illustrated). Cleavage products are indicated with asterisks. Noncontiguous lanes from the same gel are indicated with a dashed line.

RNA-targeting branch of the CRISPR-Cas immune system appears to play an important role in the war against viruses and other mobile genetic elements, given that ~70% of the archaea and 30% of the bacteria with CRISPR-Cas systems include the Cmr module (Shah and Garrett, 2011). Our results provide direct evidence that the CRISPR-Cmr system cleaves complementary RNAs in vivo (Figure 4). Moreover, we have identified the key structural features of the crRNAs associated with active Cmr complexes (Figures 1–3 and 5) and engineered crRNAs capable of directing Cmr complexes to efficiently cleave selected RNA targets at predictable sites (Figure 6).

Anatomy of the Cmr Complex Guide RNAs

The crRNAs associated with the Cmr complex are of two sizes, 39 and 45 nt in length, both with an 8 nt CRISPR repeat sequence at the 5' end (5' tag) followed by either 31 or 39 nt of invader guide sequence (Figures 1 and 3, and Hale et al., 2009). Unlike other crRNAs found in *P. furiosus*, the Cmr crRNAs contain little or no CRISPR repeat sequence at the 3' end (Figure 3). Our studies revealed that the crRNAs present in Cmr complexes are essentially restricted to an absolute length of 39 or 45 nt, which may be defined by a mechanism that measures from the 5' end of the crRNA or from the 5' tag sequence, but does not depend on the position of the guide sequence–3' repeat sequence boundary (Figure 3). Both the 39 and 45 nt crRNAs have 5' hydroxyl and 3' phosphate end groups (Figure 1D). Overall, the 39 and 45 nt forms are found in roughly equal proportion in the purified Cmr complexes (Figure 3A); however, the proportion of the two forms for an individual crRNA can vary significantly (see crRNAs 4.01 and 4.05, Figure 3B), and the basis for this variation is unclear. It is not yet known whether individual Cmr complexes contain one or more crRNAs; however, functional Cmr complexes can

be reconstituted with either the 39 or the 45 nt crRNA species, indicating that both forms are not simultaneously required for cleavage activity and hinting that each Cmr complex may harbor a single crRNA.

While the 5' ends of the Cmr complex crRNAs are presumably generated by cleavage of CRISPR transcripts by the site-specific ribonuclease Cas6 (Carte et al., 2008, 2010; Wang et al., 2011), we do not yet know how the 3' ends of mature crRNAs are generated or how the 39 and 45 nt crRNAs are selectively incorporated into Cmr complexes. This exquisite selectivity may simply be because the Cmr complex proteins preferentially bind crRNAs of these lengths. Alternatively, the proteins may associate with longer crRNA species (e.g., the 1× intermediate generated by Cas6 cleavage that includes an additional 22 nt of the repeat at the 3' end), and the 3' ends may be generated following formation of the complexes. The nuclease(s) responsible for crRNA 3' end trimming could be an intrinsic subunit of the Cmr complex or an extrinsic enzyme that may or may not be specific for crRNA processing. The precisely measured 3' ends of the Cmr complex crRNAs could result, for example, from a combination of nonspecific nuclease activity and specific RNA protection by Cmr proteins. The 45 nt crRNA and the other crRNA size forms found in *P. furiosus* (which contain longer 3' ends; see Figures 1 and 3A) are components of other RNP complexes currently under investigation (Hale et al., 2008, and our unpublished data). The *P. furiosus* genome encodes 21 Cas proteins in addition to the six Cmr proteins, including Csa and Cst subtype proteins that are also expected to associate with crRNAs to form additional effector complexes. Future work will elucidate the mechanisms by which distinct crRNA species are sorted among CRISPR-Cas complexes within a cell.

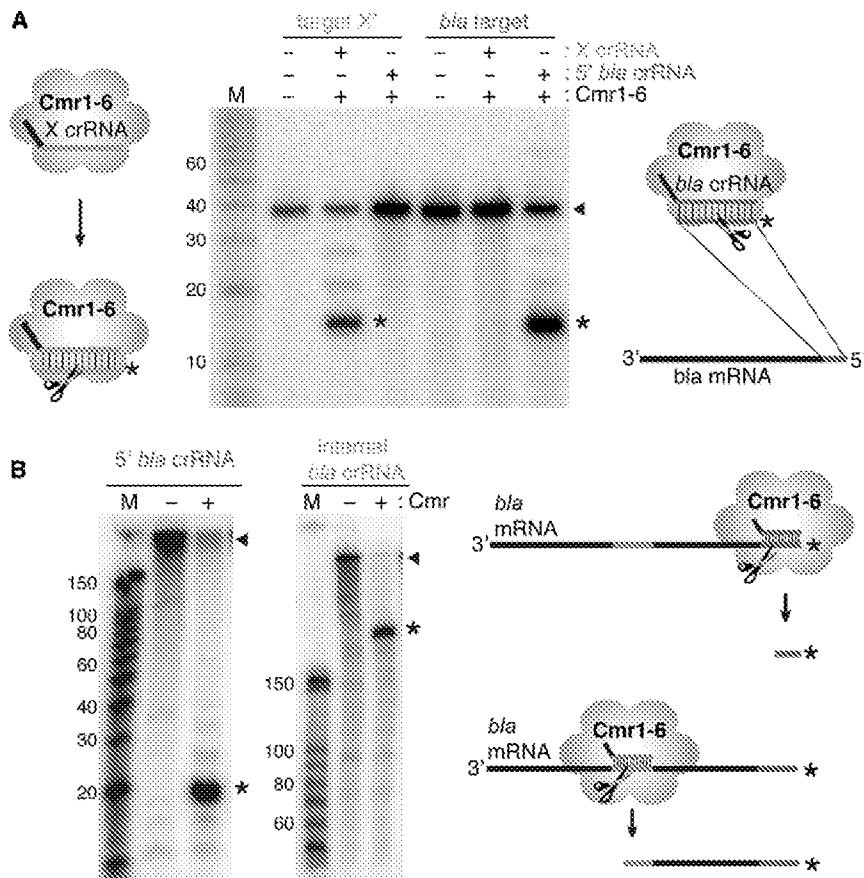


Figure 6. Engineered crRNAs Direct the Cmr Complex to Cleave Intended Target RNAs

(A) Specific cleavage of novel target RNAs. Cmr complexes were reconstituted with crRNAs designed to target a random sequence (X crRNA, X' target, orange) or a fragment from the 5' end of the β -lactamase (*bla*) mRNA (5' *bla* crRNA, *bla* target, red) and tested for the ability to cleave both target RNAs. Cleavage products are indicated with asterisks. Diagrams illustrate engineered crRNAs with 5' repeat tag (black) and guide sequence elements (orange or red) interacting with and cleaving the intended 5' radiolabeled target.

(B) Directed cleavage of predicted sites in β -lactamase mRNA. Engineered crRNAs target two locations on the full-length *bla* mRNA. The full-length *bla* mRNA was subject to cleavage by Cmr complexes reconstituted with crRNAs targeting a region near the 5' end of the mRNA (5' *bla* crRNA, red) or a sequence approximately 200 nt from the 5' end of the mRNA (internal *bla* crRNA, blue). Cleavage products are indicated by asterisks. Diagrams illustrate engineered crRNAs interacting with and cleaving the *bla* mRNA.

Directing Cmr Complexes to Cleave RNAs of Choice

In this work, we have demonstrated that crRNAs can be designed to target the Cmr complex to cleave novel RNA substrates (Figure 6). crRNAs comprised of a 37 nt complementary guide sequence

The 5' Tag Is Critical for Cmr Complex crRNA Function

The results presented here indicate that the 5' tag sequence is necessary for crRNA function in Cmr complexes (Figure 5). We found that the sequence, and perhaps size, of the 5' tag is important for Cmr complex crRNA function; even subtle mutations are detrimental to Cmr activity. We hypothesize that the 5' tag is important as a binding site for one or more Cmr proteins. Consistent with this idea, we found that crRNAs from CRISPR locus 8 with variant 5' repeat tags are preferentially excluded from Cmr complexes in vivo (Figure 5C). As discussed above, the basis for exclusion of longer crRNA species, which possess the 5' tag, from Cmr complexes in *P. furiosus* is not yet clear. The presence of a 5' tag sequence is a conserved feature of crRNAs in many CRISPR-Cas systems (Brouns et al., 2008; Hale et al., 2008, 2009; Haurwitz et al., 2010; Lintner et al., 2011; Marraffini and Sontheimer, 2008). A correlation exists between specific CRISPR-Cas modules and the CRISPR repeat sequences found together in various organisms, suggesting that the different systems may recognize characteristic 5' tags. However, in *P. furiosus* the crRNAs encoded by the seven CRISPR loci generally share the same 5' tag sequence, which evidently functions with the Csa and Cst as well as Cmr modules. We hypothesize that the 5' tag is also important in effector complex function in the other systems where it is found. Our results reveal the functional importance of the 5' tag in Cmr effector complex function.

and 5' tag were used to direct specific cleavage of RNAs including the β -lactamase mRNA in vitro. Cleavage product sizes indicate that the target was predictably cut 14 nt downstream of the 3' end of the engineered crRNA. These findings will enable efforts to exploit the site-specific cleavage function of Cmr-crRNA complexes for in vivo gene knockdown experiments. Our studies have defined the minimal components for Cmr complex function (Cmr1-6 and a crRNA; Figure 6 and Hale et al., 2009), making it conceivable to express functional Cmr complexes in eukaryotes as well as additional prokaryotes to provide a method of RNA silencing. Our results indicate that Cmr-crRNA complexes have the potential to target destruction of mRNAs encoding any gene of interest and to become important research tools with industrial and biomedical applications.

RNA Targeting versus DNA Targeting CRISPR-Based Gene Silencing Systems

Our findings demonstrate that there are significant differences in the modes of target nucleic acid recognition for CRISPR-based gene silencing systems that target RNA and DNA. For example, while DNA targeting by at least some CRISPR-Cas systems requires a short protospacer-adjacent motif (PAM) in the invading DNA target (Deveau et al., 2008; Garneau et al., 2010; Gudbergsson et al., 2011; Mojica et al., 2009; Semenova et al., 2011), our work suggests that Cmr complexes function independently of PAMs. Cmr-crRNA complexes can cleave target RNAs that

completely lack flanking sequences or contain a variety of flanking sequences (that do not match predicted PAMs; Figures 1, 4, and 6 and Hale et al., 2009). For DNA-targeting systems, the requirement for a PAM sequence in the invader avoids damage to the host genome at the site of the corresponding guide sequence in the CRISPR locus (where PAMs are not found in the flanking repeat sequences). The DNA-targeting system of *Staphylococcus epidermidis* further demonstrates an explicit requirement for mismatches between the 5' tag of the crRNA and the target DNA strand, which also prevents targeting the host CRISPR locus (Marraffini and Sontheimer, 2010). In contrast, for the RNA-targeting Cmr complex (where host genome damage is not a risk), target RNAs containing sequences with perfect homology to the 5' tag of the crRNA are efficiently cleaved (Figures 4 and 5D). Evidence indicates that there are also differences in the interactions with the guide region of the crRNA. A specific 7–8 nt region of the crRNA located immediately downstream of the 5' tag (dubbed the “seed sequence”) was found to be important for initial high-affinity recognition of targets for certain DNA-targeting CRISPR-Cas systems (Semenova et al., 2011; Wiedenheft et al., 2011). Base-pairing between the 3' end of the crRNA and target DNA was dramatically less important. The opposite trend is observed for RNA targeting, where interaction with the 3' end of the crRNA is critical for function and target RNAs that lack the “seed sequence” region are efficiently cleaved (Hale et al., 2009). The molecular bases of these differences in CRISPR-Cas RNA and DNA targeting pathways remain to be explored. The lack of restrictions in the flanking sequence content in RNA targeting provides for more flexibility in substrate interaction.

EXPERIMENTAL PROCEDURES

Antibody Preparation and Coimmunoprecipitation Reactions

IgY antibodies were raised against recombinant *P. furiosus* Cmr2 as described (Carte et al., 2010). Preimmune and immune Cmr2 antibodies were coupled either to anti-IgY-conjugated agarose beads (Gallus Immunotech) for northern analysis, end analysis, or small RNA sequencing, or to CarboLink beads (Pierce) for western analysis.

CoIP/Western Analysis

Western blotting of coIP material was performed by standard procedures.

CoIP/Northern Analysis

Trizol LS reagent (Invitrogen) was used to isolate RNAs from the coIP samples and *P. furiosus* extract (total RNA). Northern analysis was performed as described previously (Hale et al., 2008) with modifications (see the Supplemental Experimental Procedures). Probe sequences can be found in Table S1.

CoIP/Activity Assays

CoIP beads were washed into 40 mM HEPES (pH 7.0), 500 mM KCl and added to a 20 μ l cleavage reaction containing radiolabeled target RNA (5,000 cpm). Reactions were incubated for 3 hr at 70°C with mixing. RNAs were isolated by Proteinase K treatment, phenol extraction, and ethanol precipitation; separated on 7M urea 15% polyacrylamide gels; and visualized by phosphor imaging.

CoIP/End Analysis

Trizol LS was used to isolate small RNAs from coIP samples. The RNAs were treated with thermostable shrimp alkaline phosphatase (Promega). Treated and nontreated RNAs were subject to either (1) 5' radiolabeling with

γ -³²P-ATP (MP Biomedicals) via polynucleotide kinase (Ambion), or (2) 3' radiolabeling via T4 ssRNA Ligase 1 (NEB) with cytidine-3',5'-bisphosphate [5'-³²P] (3,000 ci/mmol, MP Biomedicals). Both labeling reactions were carried out for 1 hr at 37°C, and the RNAs were isolated, separated by denaturing gel electrophoresis, and visualized by phosphor imaging.

Small RNA Library Preparation

For preparation of total and Cmr-associated small RNA libraries, Trizol was used to extract RNAs from *P. furiosus* cells or from a coIP reaction, respectively. The total small RNA library was prepared with 50 μ g of total RNA. Briefly, the RNAs were phosphatase treated to remove 3' phosphates, followed by 3' ligation to an adaptor. These products were treated with T4 polynucleotide kinase before 5' ligation to an additional adaptor. The 5' and 3' ligated products were subject to reverse transcription, RNase H digestion, and PCR.

Illumina Sequencing and Analysis

cDNA libraries were visualized by electrophoresis, quantitated by Agilent Bioanalyzer and NanoDrop, diluted to 2 pM, and subject to 76 cycles of sequencing on the Illumina Genome Analyzer IIx. Sequence reads were trimmed of the 3' linker, and reads 18–76 nt in length were aligned to the *P. furiosus* genome using Bowtie (Langmead et al., 2009). The 5' and 3' ends of reads mapping to the CRISPR loci were calculated using a custom PERL script available upon request.

Recombinant Cmr Protein Production and Activity Assays

Expression and purification of recombinant Cmr proteins were performed as described (Hale et al., 2009) with modifications (see the Supplemental Experimental Procedures). RNA cleavage assays were performed as described (Hale et al., 2009). Briefly, all six recombinant Cmr proteins (500 nM each) were preincubated with unlabeled crRNA (0.05 pmol) prior to addition of 5' radiolabeled target RNAs (5,000 cpm). Reactions were incubated for 1–3 hr at 70°C and isolated RNAs analyzed by denaturing gel electrophoresis and phosphor imaging.

RNA Preparation

Sequences of RNA/DNA oligos used for RNA preparation are found in Table S1, and detailed information about RNA preparation is found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the complete high-throughput data set analyzed in this work is GSE33256.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures, one table, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.molcel.2011.10.023.

ACKNOWLEDGMENTS

We thank the University of Connecticut Health Center Translational Genomics Core Facility for use of the Illumina Genome Analyzer IIx. This work was supported by National Institutes of Health (NIH) grants RO1GM54682 (including ARRA funds) to M.P.T. and R.M.T. and RO1GM062516 to B.R.G. This work was funded in part by an award from the Technology Commercialization Office of the University of Georgia Research Foundation, Inc.

Received: August 4, 2011

Revised: October 7, 2011

Accepted: October 27, 2011

Published online: January 5, 2012

REFERENCES

- Andersson, A.F., and Banfield, J.F. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**, 1047–1050.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Deveau, H., Garneau, J.E., and Moineau, S. (2010). CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* **64**, 475–493.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71.
- Garrett, R.A., Shah, S.A., Vestergaard, G., Deng, L., Gudbergdottir, S., Kenchappa, C.S., Erdmann, S., and She, Q. (2011). CRISPR-based immune systems of the Sulfolobales: complexity and diversity. *Biochem. Soc. Trans.* **39**, 51–57.
- Gesner, E.M., Schellenberg, M.J., Garside, E.L., George, M.M., and Macmillan, A.M. (2011). Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **18**, 688–692.
- Gudbergdottir, S., Deng, L., Chen, Z., Jensen, J.V., Jensen, L.R., She, Q., and Garrett, R.A. (2011). Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* **79**, 35–49.
- 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. 10.1371/journal.pcbi.0010060.
- Hale, C., Kleppe, K., Terns, R.M., and Terns, M.P. (2008). Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **14**, 2572–2579.
- 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.
- 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.
- Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170.
- Jansen, R., Embden, J.D., Gastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565–1575.
- Jore, M.M., Brouns, S.J., and van der Oost, J. (2011a). RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* 10.1101/cshperspect.a003657.
- 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., et al. (2011b). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **18**, 529–536.
- Karginov, F.V., and Hannon, G.J. (2010). The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* **37**, 7–19.
- Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., et al. (2003). The UCSC Genome Browser Database. *Nucleic Acids Res.* **31**, 51–54.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.
- Lillestøl, R.K., Shah, S.A., Brugger, K., Redder, P., Phan, H., Christiansen, J., and Garrett, R.A. (2009). CRISPR families of the crenarchaeal genus Sulfolobus: bidirectional transcription and dynamic properties. *Mol. Microbiol.* **72**, 259–272.
- Lintner, N.G., Kerou, M., Brumfield, S.K., Graham, S., Liu, H., Naismith, J.H., Sdano, M., Peng, N., She, Q., Copie, V., et al. (2011). Structural and functional characterization of an archaeal CASCADE complex for CRISPR-mediated viral defense. *J. Biol. Chem.* **286**, 21643–21656.
- Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and Koonin, E.V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**, 7.
- 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., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477.
- Manica, A., Zebec, Z., Teichmann, D., and Schleper, C. (2011). In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol. Microbiol.* **80**, 481–491.
- Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845.
- Marraffini, L.A., and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568–571.
- Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740.
- Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663.
- Przybilski, R., Richter, C., Gristwood, T., Clulow, J.S., Vercoe, R.B., and Fineran, P.C. (2011). Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol.* **8**, 517–528.
- Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **75**, 1495–1512.
- Sashital, D.G., Jinek, M., and Doudna, J.A. (2011). An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* **18**, 680–687.
- Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* **108**, 10098–10103.
- Shah, S.A., and Garrett, R.A. (2011). CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems. *Res. Microbiol.* **162**, 27–38.
- Shah, S.A., Hansen, N.R., and Garrett, R.A. (2009). Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles

and implications for their inhibitory mechanism. *Biochem. Soc. Trans.* *37*, 23–28.

Terns, M.P., and Terns, R.M. (2011). CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* *14*, 321–327.

van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M., and Brouns, S.J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* *34*, 401–407.

Wang, R., Preamplume, G., Terns, M.P., Terns, R.M., and Li, H. (2011). Interaction of the Cas6 ribonuclease with CRISPR RNAs: recognition and cleavage. *Structure* *19*, 257–264.

Wiedenheft, B., van Duijn, E., Bultema, J., Waghmare, S., Zhou, K., Barendregt, A., Westphal, W., Heck, A., Boekema, E., Dickman, M., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* *108*, 10092–10097.