RESEARCH PAPER

RNA Biology 10:5, 738–748; May 2013; © 2013 Landes Bioscience

# Genetic determinants of PAM-dependent DNA targeting and pre-crRNA processing in *Sulfolobus islandicus*

Wenfang Peng,[1,2] Huan Li,[1] Søren Hallstrøm,[2] Nan Peng,[1] Yun Xiang Liang[1] and Qunxin She[1,2,*]

[1]State Key Laboratory of Agricultural Microbiology and College of Life Science and Technology; Huazhong Agricultural University; Wuhan, China; [2]Danish Archaea Centre; Department of Biology; University of Copenhagen; Copenhagen, Denmark

**Keywords:** CRISPR/Cas, Cascade, protospacer-adjacent motifs, crRNA biogenesis, DNA interference, *cas* mutants, *Sulfolobus islandicus*

Bacteria and Archaea encode clustered, regularly interspaced, short palindromic repeat (CRISPR) systems to confer adaptive immunity to invasive viruses and plasmids. Recent studies of CRISPR systems revealed that diverse CRISPR-associated (Cas) interference modules often coexist in different organisms but functions of *cas* genes have not been dissected in any of these systems. The crenarchaeon *Sulfolobus islandicus* encodes three distinct CRISPR interference modules, including a type IA system and two type IIIB systems: Cmr-α and Cmr-β. To study the genetic determinants of protospacer-adjacent motif (PAM)-dependent DNA targeting activity and mature CRISPR RNA (crRNA) production in this organism, mutants deleting individual genes of the type IA system or removing each of other Cas modules were constructed. Characterization of these mutants revealed that Cas7, Cas5, Cas6, Cas3′ and Cas3″ are essential for PAM-dependent DNA targeting activity, whereas Csa5, along with all other Cas modules, is dispensable for the targeting in the crenarchaeon. Cas6 is implicated as the only enzyme for pre-crRNA processing and the crRNA maturation is independent of the DNA targeting activity. Importantly, we show that Cas7 and Cas5 are essential for stabilizing the processing intermediates and mature crRNAs, respectively, and that depleting the helicase or nuclease domain of Cas3 leads to the accumulation of processing intermediates. This demonstrates that in addition to Cas6, other Cas proteins of an archaeal type IA system also contribute to crRNA processing.

## Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins constitute the adaptive and inheritable immune system in almost all archaea and more than 40% of bacteria. The antiviral nature of these systems was predicted based on bioinformatic analyses of bacterial and archaeal genomes and their genetic elements, which predicted that unique DNA sequences (spacers) of CRISPR arrays were derived from invasive genetic elements[1-3] and it was further hypothesized that the defense could occur via an RNA interference-like mechanism.[4] Subsequently, in vivo CRISPR activity has been studied in different microorganisms demonstrating that immune response by CRISPR/Cas systems exerts through three distinct stages:[5-10] (1) Pieces of DNA sequence of invading genetic elements are incorporated into a CRISPR locus at a position between the leader and the first spacer, generating new spacers; (2) Transcription of CRISPR arrays from their leaders gives precursor CRISPR RNA molecules (pre-crRNAs) that are processed to yield mature CRISPR RNAs (crRNAs) and (3) crRNAs and Cascade (CRISPR-associated complex for antiviral defense) form ribonucleoprotein complexes, exerting silencing of invasive nucleic acids.

CRISPR/Cas systems are very diverse and their reevaluation has led to the classification of three main classes, type I, II and III, which are further divided into subgroups.[7] Recently, Cas1 and Cas2, the only universally conserved Cas proteins,[4,7,11] formed a complex with Csa1 and Cas4 in *Thermoproteus tenax*, which was implicated in acquisition of new spacers, and for this reason they were classified as a new Cas module, CRISPR-associated cluster for integration of new spacers (Cascis).[12] Following the first demonstration of spacer acquisition in *Streptococcus thermophilus*,[13] insertion of new spacers between the leader sequence and the first spacer of CRISPR arrays has been shown in *Escherichia coli*[14-16] and *Sulfolobus solfataricus*.[17] Moreover, it has been shown that Cas1 and Cas2 are sufficient to yield spacer insertion in *E. coli*.[14]

Small RNA products generated from CRISPR arrays were first detected a decade ago in *Archaeoglobus fulgidus*[18] and later were found to be widespread in microorganisms. CRISPR arrays often consist of a large cluster of spacer-repeat units. Transcription of an entire CRISPR array gives pre-crRNA to be processed into effector crRNAs.[19-22] Two distinct mechanisms are known to generate

*Correspondence to: Qunxin She; Email: qunxin@bio.ku.dk
Submitted: 12/06/12; Revised: 01/25/13; Accepted: 01/26/13
http://dx.doi.org/10.4161/rna.23798

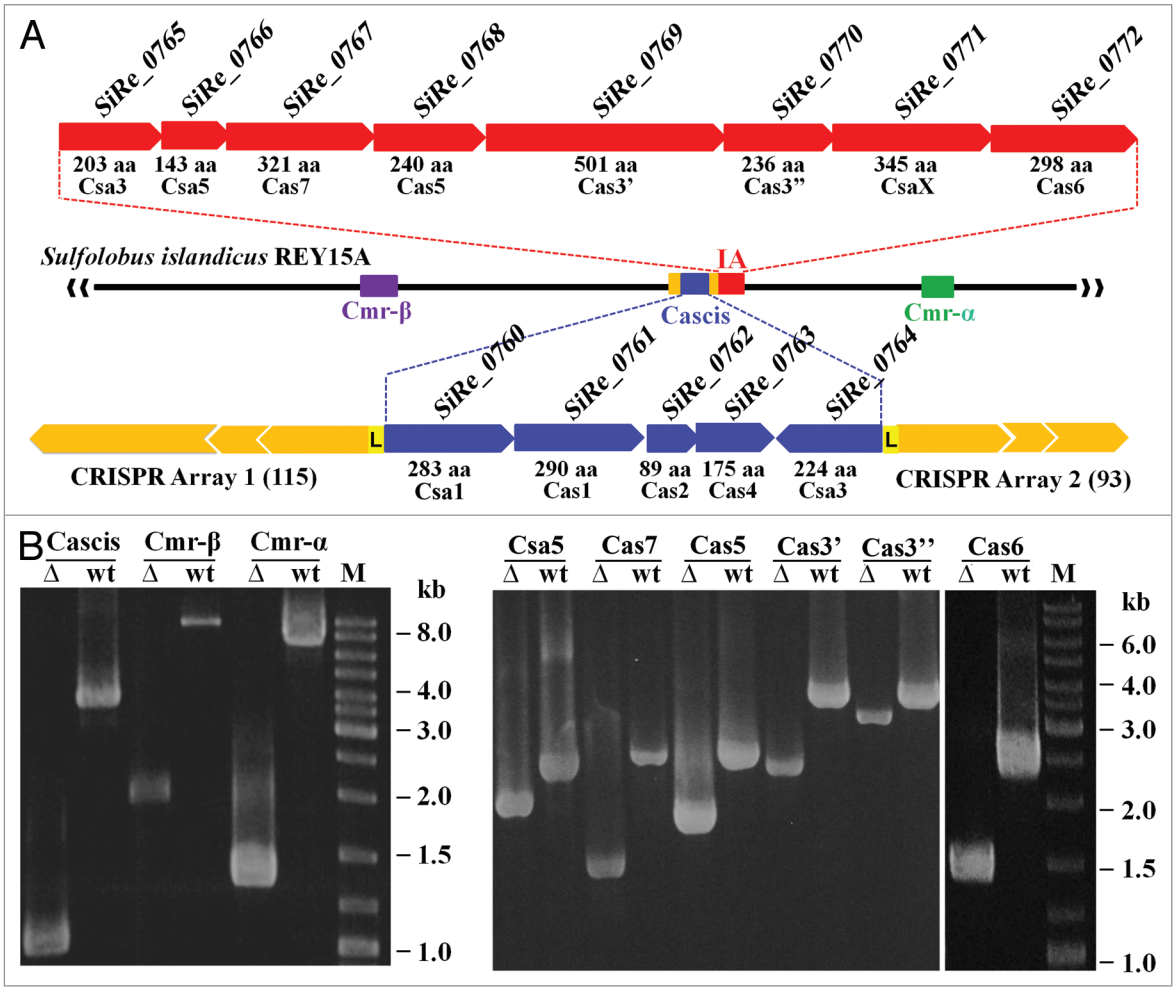


**Figure 1.** *S. islandicus* Cas modules and construction knockouts lacking a Cas module or a *cas* gene. (**A**) Schematic of CRISPR arrays and Cas modules in *S. islandicus* REY15A. (**B**) PCR verification of mutant alleles of Cas modules and *cas* genes in the mutants constructed with the MID procedure, described previously.[66] PCR was conducted with specific primers for each Cas module or *cas* gene with genomic DNAs extracted from the genetic host *S. islandicus* E233S (wt) and corresponding deletion mutant (Δ). M, DNA size marker. Cascis, CRISPR-associated cluster for integration of new spacers; IA, type IA interference module; Cmr-α and Cmr-β, two type IIIB modules containing 6 and 7 *cmr* genes, respectively.[47]

crRNAs. The first is endonuclease Cas6, a processing enzyme for type I and type III CRISPR systems,[23-27] which binds to a repeat of a pre-crRNA and cleaves it. In fact, the enzyme remains bound to the RNA after cleavage.[23,28] The second is RNase III that does the cleavage under the guidance of a trans-acting tracrRNA in type II systems.[29]

At the final stage, it is believed that Cascade complexes containing crRNAs recruit Cas3 helicase/nuclease and exert invader silencing,[25,30-32] during which a target nucleic acid is recognized by complementarity between the protospacer and the corresponding effector crRNA. DNA targeting activity has been demonstrated for representative systems belonging to each of the main types of CRISPR systems, including the *E. coli* type IE crRNA-Cascade ribonucleoprotein complex in vitro,[33] DNA targeting activity of type II systems in *S. thermophilus* in vivo[34] and in vitro[35] and the *Staphylococcus epidermidis* type IIIA system Csm.[36] Furthermore, *P. furiosus* and *S. solfataricus* type IIIB CRISPR repeat-associated mysterious proteins (RAMP),[11] also named Cmr modules, mediate RNA targeting activity.[32,37,38]

An important point in the antiviral defense is how CRISPR systems distinguish self vs. non-self DNA. Most known CRISPR systems recognize an invasive genetic element by identifying a protospacer-adjacent motif (PAM) located either at 5'- or at 3'-flanking position to a protospacer of the genetic element.[9] PAM motifs were first recognized from genome sequence analyses,[20,39] and subsequently experimentally demonstrated by virus infection in bacteria[40,41] or by interference plasmid assay in archaea.[42-44] The other type of self vs. non-self discrimination involves repeat protection of chromosome-encoded CRISPR arrays from the DNA interference by the Csm module in *S. epidermidis*.[45]

All known *Sulfolobus* strains encode multiple CRISPR interference modules belonging to type IA or type IIIA or IIIB, whereas the number and types of modules carried in each strain vary from one to another.[46] In the recently completed genome of *S. islandicus* REY15A,[47] there are two arrays of CRISPR elements with the same repeat, a Cascis locus with five genes, a type IA CRISPR interference system and two distinct type IIIB Cmr systems (**Fig. 1**). This organism is one of the genetically tractable archaeal models[48]

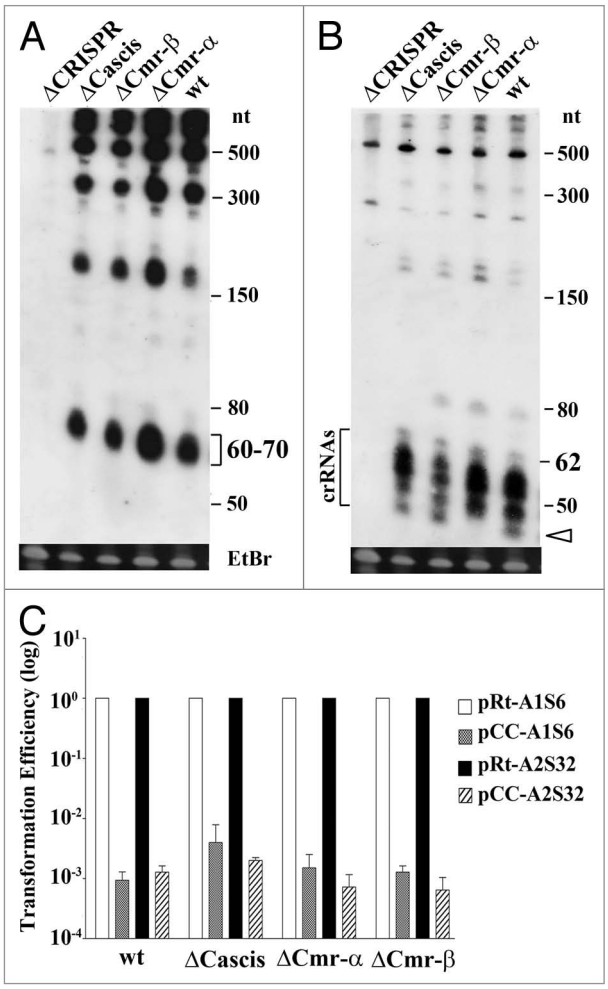


**Figure 2.** Cascis, Cmr-α and Cmr-β modules are dispensable for pre-crRNA processing and PAM motif-dependent DNA silencing. (**A and B**) Northern analyses of total RNAs prepared from the *S. islandicus* wild-type strain (wt) and knockouts of Cas modules using oligonucleotide of the repeat (**A**) or that of spacer 32 of array 2 (**B**) as a probe. Band of ca. 60–70 nt corresponds to the primary cleavage products of pre-crRNAs by Cas6. Mature crRNAs (ca. 40–60 nt) of A2S32 are also indicated. Unfilled arrowhead indicates the location of the smallest crRNA. EtBr, RNA staining with ethidium bromide as a loading control. (**C**) Plasmid interference assay of Cas module knockouts. The genetic host E233S (wt) and Cas module knockouts were used as hosts for plasmid transformation. Two artificial interference plasmids, pCC-A1S6 and pCC-A2S32, were used to test for PAM-dependent DNA interference. In the corresponding reference plasmids, the PAM motif (5'-CCA-3') was replaced with 5'-GAAAG-3' (pRt-A1S6) or 5'-ATTGAAAG-3' (pRt-R2S32). Transformation efficiencies of each interference plasmid were expressed as relative values to the efficiencies of corresponding reference plasmids, the latter of which were set to 1.0.

with which a novel gene knockout method has been established[49] and plasmid interference assays were developed for testing interference activities.[42,50] This has enabled us, for the first time, to conduct genetic analyses of *cas* gene functions in a microbe carrying complex CRISPR/Cas systems. Deletion mutants lacking Cascis, Cmr-α or Cmr-β module and those deleting each of the genes encoding components of the putative IA-Cascade complex were constructed and analyzed pre-crRNA processing and CC PAM motif-mediated targeting. We show that type IA module is responsible for the PAM-dependent interference. Evidence has been provided to demonstrate that other Cas proteins also play a role in modulating pre-crRNA processing by Cas6.

## Results

**Cascis, Cmr-α and Cmr-β modules are dispensable for pre-crRNA processing and CC PAM motif DNA targeting.** Previously, it was shown that CRISPR arrays were transcribed and processed into short RNA molecules in different *Sulfolobus* species[20,44,51] and that *Sulfolobus* CRISPR systems conferred DNA targeting activity.[42] However, as *Sulfolobus* contain multiple Cas interference modules, functions of each Cas module and individual *cas* genes in these molecular processes have not been studied. Here, we systematically studied possible functions of four Cas modules in *S. islandicus* REY15A, i.e., a type IA interference module, a Cas cluster for integration of new spacers (Cascis) and two type IIIB modules Cmr-α and Cmr-β modules (**Fig. 1A**). Mutant strains lacking each entire module of Cascis, Cmr-α or Cmr-β were constructed, from which total RNAs were prepared and subsequently analyzed by northern hybridization as described in Materials and Methods.

Results of the northern hybridization of RNAs prepared from the wild-type strain were very similar to those of the northern analyses of the mutants (**Fig. 2A and B**). The radiolabeled repeat oligonucleotide probe hybridized to RNAs of 60–70 nt in size. These RNAs are comparable in size with the mature crRNAs produced by an *S. solfataricus* Cas6, the primary product of pre-crRNAs was processed by Cas6, yielding spacers carrying 8 nt 5'-handle and 16 nt 3'-handle of the repeat.[25] A few processing intermediates were also hybridized by the repeat probe showing the sizes equivalent to the RNAs carrying 3, 5 and 7 spacers (**Fig. 2A**). When analyzed with an oligonucleotide of array 2 spacer 32 as probe, an RNA of ca. 62 nt as well as RNAs < 60 nt were identified (**Fig. 2B**). The former could represent a primary processing product of Cas6 cleavage and the latter should be resulted from further processing/maturation of the primary product. Again, the probe identified the same RNA bands in ΔCascis, ΔCmr-α and ΔCmr-β, comparing with the wild-type strain except for the smallest crRNA species (ca. 40 or 41 nt) that is absent from ΔCmr-α (**Fig. 2B**), a result consistent with the functional analysis of Cmr-α in this crenarchaeon.[50] Furthermore, none of these bands was present in the hybridization of the total RNAs prepared from ΔCRISPR, a strain lacking both arrays of CRISPR, indicating that the hybridization was specific for both probes. Taken together, this indicated that whereas ΔCmr-α is responsible for the maturation of the smallest crRNA, none of these Cas modules play an important role in pre-crRNA processing in this archaeon.

These strains were then tested for DNA targeting mediated by CC PAM motif using two different interference plasmids, pCC-A2S32 and pCC-A1S6. These plasmids carry the artificial protospacer sequence of array 2 spacer 32 (A2S32) or that of array 1 spacer 6 (A1S6) where each spacer is preceded by the PAM motif (5'-CCA-3') of *Sulfolobus*.[20] In the corresponding

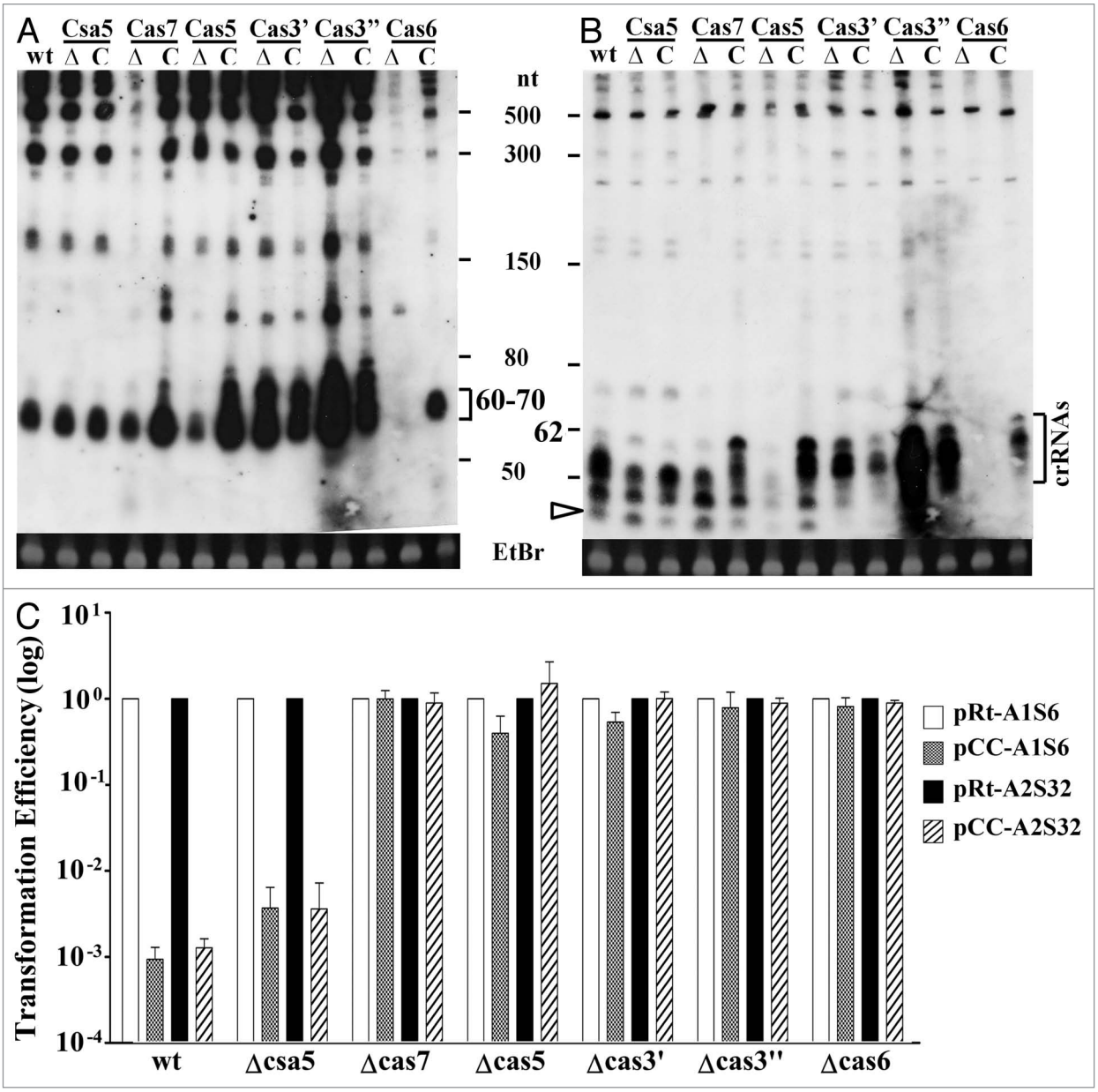


**Figure 3.** Functions of Cas proteins of type IA interference module in pre-crRNA processing and PAM-dependent DNA silencing. (**A and B**) Northern analyses of total RNAs prepared from the *S. islandicus* wild-type strain (wt), knockouts (Δ) of each *cas* gene of type IA interference module and their complementation strains (**C**) using oligonucleotide of the repeat (**A**) or that of spacer 32 of array 2 (**B**) as a probe. Band of ca. 60–70 nt corresponds to the primary cleavage products of pre-crRNAs by Cas6. Mature crRNAs (ca. 40–60 nt) of A2S32 are also indicated. Unfilled arrowhead indicates the location of the smallest crRNA. EtBr, RNA staining with ethidium bromide as a loading control. (**C**) Plasmid interference assay of knockout mutants using pCC-A2S32 and pCC-A1S6 as interference plasmids and pRt-A2S32 and pRt-A1S6 as reference plasmids. Transformation efficiencies of each interference plasmid were expressed as relative values to the efficiencies of corresponding reference plasmids, the latter of which were set to 1.0.

reference plasmids, each PAM motif was replaced with 5 or 8 nt of the repeat sequence, yielding a structure resembling the genetic organization of CRISPR arrays in the chromosome. As shown in **Figure 2C**, when each mutant was employed as host for transformation, hundreds to thousands fold lower rates of transformation were yielded with pCC-A2S32 and pCC-A1S6 comparing with their reference plasmids, indicating none of these modules is essential for the CC motif-mediated DNA targeting activity.

**Functions of Cas6, Cas7, Cas5 and Csa5, components of the aCASCADE complex in *S. islandicus*.** Recently, an archaeal Cascade complex was isolated from *S. solfataricus*.[25] It contained Cas7, Cas5, Cas6 and Csa5 and was loaded with crRNAs of 60–70 nt, which correspond to primary cleavage products by Cas6. Each of these Cascade proteins has a homolog in *S. islandicus* REY15A showing 82–93% amino acid sequence identity. Since both *S. solfataricus* and *S. islandicus* exhibited PAM-dependent DNA targeting activity, we were interested in studying

possible functions of these Cas proteins in this process. In-frame deletion mutants were constructed for all four *cas* genes, giving Δcas7, Δcas5, Δcsa5 and Δcas6. Their complementation plasmids were constructed with a highly efficient expression vector pSeSD[52] and transformed into corresponding mutants to yield complementation strains. Both mutants and complementation strains were used for northern analysis of pre-crRNA processing and for the interference plasmid assay.

First, we examined pre-crRNA processing in Δcas6 and its complementation strain. As shown in **Figure 3**, depleting Cas6 completely inactivated pre-crRNA processing since neither intermediates nor mature crRNAs were present in the mutant strain. These bands appeared again in the complementation strain where the Cas6 protein was expressed from the expression plasmid pSe-cas6. These results indicated that *cas6* is essential for crRNA production in *S. islandicus* and that it is the only enzyme responsible for pre-crRNA processing in this organism.

Next, we analyzed processing intermediates and crRNAs in Δcsa5, Δcas5 and Δcas7 from which different results were obtained. Northern analysis of Δcsa5 and its complementation strain showed the same hybridization pattern and RNA band intensities that were essentially identical to the results of the wild-type strain. Hybridization signals of processing intermediates were absent or greatly weakened in Δcas7, while crRNAs were maintained at very similar levels. The exactly opposite results were obtained for Δcas5 where mature crRNAs were hardly detectable but normal levels of processing intermediates accumulated. Furthermore, genetic complementation of mutant deficiency by episomic expression of each corresponding gene from a plasmid restored the hybridization pattern of the wild-type strain (**Fig. 3**). These results indicated that Cas7 is essential for stabilizing RNA intermediates of pre-crRNA processing and it could function in the absence of Csa5 or Cas5, and that Cas5 is essential for maintaining mature crRNAs whereas Csa5 does not appear to have a role in pre-crRNA processing and maturation.

These mutants were also tested for PAM motif-mediated DNA targeting. As shown in **Figure 3C**, whereas Δcas7 and Δcas5 lost the capability of conferring DNA interference to pCC-A2S32 and pCC-A1S6, depleting Csa5 did not compromise the inference activity.

Taken together, we have demonstrated that only three of the four putative Cascade proteins, i.e., Cas6, Cas7 and Cas5 are essential for the PAM-mediated DNA interference. All these Cas proteins function in pre-crRNA processing in which Cas6 is the primary processing enzyme in vivo whereas Cas7 and Cas5 stabilize processing intermediates and crRNAs, respectively. However, Csa5 does not appear to have a role in either process, suggesting that Csa5 may not represent an integral part of the Cascade that mediates PAM-directed DNA silencing.

**Cas3 helicase/nuclease functions both in processing and targeting.** Cas3 is required for all known type I CRISPR systems for interference. In some CRISPR systems such as those in *Sulfolobus* species, Cas3 is divided into two parts: Cas3' containing a helicase domain and Cas3" carrying a nuclease domain. To reveal their role in DNA interference, Δcas3' and Δcas3" were investigated for silencing of pCC-A2S32 and pCC-A1S6 interference plasmids. As shown in **Figure 3C**, both mutants showed high transformation efficiencies with all tested plasmids, including both interference and control plasmids, indicating that the interference activity was inactivated in the mutants.

When pre-crRNA processing was studied by northern analysis, both mutants showed increased levels of processing intermediates compared with the wild-type strain and a much elevated level of mature crRNA was observed for Δcas3" (**Fig. 3**). Interestingly, while genetic complementation of the mutant deficiency by episomic expression of *cas3'* or *cas3"* reversed the strong accumulation of the intermediates, a few new RNA species (60 < crRNAs < 80 nt), which first appeared in Δcas3' and Δcas3", persisted in their complementing strains. These results suggest that the helicase and nuclease activities of Cas3 function in resolving the processing intermediates and that the nuclease domain could play a role in crRNA degradation.

## Discussion

In this work, we have studied functions of different interference modules in pre-crRNA processing and PAM-mediated DNA targeting in *S. islandicus* REY15A. Mutants lacking each module were constructed and analyzed for mature crRNA and their processing intermediates and for DNA interference. This has revealed that the type IA interference module mediates PAM-dependent DNA targeting and it functions independently from other Cas/Cmr modules that co-exist in *S. islandicus* REY15A. We also show that depleting Cas6 inactivates pre-crRNA processing and that inactivation of Cas7, Cas5, Cas3' or Cas3" impairs processing of pre-crRNAs. This unravels an important difference between in vivo and in vitro results since biochemical characterization of a *S. solfataricus* Cas6 protein, which is 93% identical to this *S. islandicus* Cas6 in amino acid sequence, demonstrated that Cas6 was capable of generating mature crRNAs by itself.[25]

A Cascade complex (IA-Cascade) isolated from *S. solfataricus* was reported consisting of at least four components, Cas7, Cas5, Csa5 and Cas6.[25] Characterization of other type I Cascade complexes, including the IE-Cascade of *E. coli* and the IF Cascade of *Pseudomonas aeruginosa*, revealed five components (Cse1, Cse2, Cas7, Cas5 and Cas6e) for type IE[19,30] and four components Csy1, Csy2, Csy3 and Cas6f (Csy4) for type IF.[53] Importantly, all Cascade components and Cas3 are shown to be essential for the DNA-targeting activity in IE and IF systems.[19,30,54] Here, we show that Csa5, one of the Cascade components, is dispensable for the PAM-directed DNA targeting in *S. islandicus*. It is important to point out though that a functional *S. solfataricus* Cascade complex remains to be demonstrated since the isolated one did not show any cleavage activity.[25]

Interestingly, although non-essential to DNA interference, Csa5 is encoded in all known archaea carrying a type IA CRISPR system, including all known *Sulfolobus* species,[47,55-58] *Methanocaldococcus jannaschii*,[59] *Thermoproteus tenax*[12] and four *Pyrobaculum* species,[60] whereas Csa5 is absent from all known type IB CRISPR systems, including *Haloferax volcanii*,[43]

*Clostridium thermocellum* and *Methanococcus maripaludis*.[27] Currently, the role of Csa5 in type 1A CRISPR system remains to be demonstrated.

In the *E. coli* type IE system, cleavage of pre-crRNAs by Cas6e yields mature crRNAs that bind to IE-Cascade to form ribonucleoprotein complexes.[33] It has been shown that Cse1 component of the IE-Cascade recognizes the PAM motif of a protospacer of invasive DNAs[61] and complementarity between the crRNA and the protospacer then guides the cleavage of foreign DNA by Cas3.[31] Currently, the counterpart of Cse1 remains to be identified in the *Sulfolobus* IA-Cascade complex if this complex functions in analogy to the *E. coli* IE-Cascade.

Examination of crRNA processing in *S. islandicus* by northern hybridization showed that the processing was inactivated by Cas6 depletion and the inactivation only occurred in Δcas6 (**Fig. 3**). This is consistent with the results that *P. furiosus* and *S. solfataricus* Cas6 proteins are capable of generating crRNA in vitro.[25,62] Together, these studies demonstrate that Cas6 is both necessary and sufficient for pre-crRNA processing in these archaea.

Furthermore, our study of pre-crRNA processing in the *cas* gene knockouts of *S. islandicus* has revealed three interesting features: (1) mature crRNAs are stabilized in Δcas7, (2) processing intermediates are stabilized in Δcas5 and (3) processing intermediates are accumulated in Δcas3' and Δcas3". Interestingly, Cas7 forms protein filaments in *S. solfataricus*.[25] Putting all these facts together, possible mechanistic details are revealed for pre-crRNA processing in *S. islandicus.* (1) Immediately after transcription, pre-crRNAs are bound by Cas7 protein filaments, protecting the large RNA molecules from degradation; (2) Cas6 proteins bind to repeat sequences of pre-crRNAs and make a cut at each repeat, (3) as both Cas7 and Cas6 proteins remain bound to the processing products, interaction between Cas3' and Cas3" and the complex will facilitate the remodeling of new smaller complexes and (4) single spacer unit complexes can only be stabilized by recruiting Cas5, which either constitute the Cascade or a basis for generating Cascade complexes. Possibly, Cas5 may also modulate Cas3 activities that could lead to crRNAs degradation.

The proposed strategy is apparently not used in the type I systems of *E. coli* and *P. aeruginosa* because relative band intensities of intermediates and crRNAs did not change upon depletion of any component of the Cascade complexes in these organisms.[19,54] Neither has it been reported that the *E. coli* Cas7 forms protein filaments. In fact, these type I systems are further diverged, whereas only Cas6 is essential in pre-crRNA processing in *S. islandicus*, two protein factors are reported to be essential for in vivo pre-crRNA processing in the bacterial systems, namely Cas5 and Cas6 in *E. coli*[19,63] and Cas6f (Csy4) and Csy2 in *P. aeruginosa*.[54]

Then, the question is why it is relevant for *S. islandicus* to employ such a strategy in crRNA processing. This archaeon contains three distinct interference modules that are presumed to employ different types of mature crRNA for their interference. The type IA system most likely utilizes 60–70 nt crRNAs, which are components of the *S. solfataricus* Cascade complex.[25] The Cmr-β module most likely uses 30–50 nt crRNAs, which are associated with the *S. solfataricus* type IIIB Cmr reported by Zhang et al.,[38] because the two Cmr systems share high-sequence similarity and exhibit the same gene synteny. Although the Cmr-α complex has not been isolated thus far, here we show that depleting the complex eliminates the smallest crRNA, which is ca. 40–41 nt crRNA for A2S32. More recently, interference activity has been demonstrated for the *S. islandicus* Cmr-α, which confers a protospacer transcript-dependent DNA interference.[50] Further investigation of the distinct Cas interference modules in *S. islandicus* will yield important insights into regulation of pre-crRNA processing and maturation as well as mechanisms of coordinating different interference activities that co-exist intracellularly.

## Materials and Methods

***Sulfolobus* strains, growth conditions and transformation.** All experiments were performed with the model organism of *Sulfolobus islandicus* REY15A originally isolated from an enrichment culture obtained from samples collected from hotsprings in Iceland.[64] Its genetic host E233S is double deletion mutant carrying a large deletion in *pyrEF* genes and a complete deletion of *lacS* gene.[65]

All *Sulfolobus* strains used in this study are listed in **Table 1**. The strains that carry *pyrEF* genes either on the chromosome or on a plasmid were grown at 76°C in the SCV (0.2% sucrose, 0.2% casamino acids plus 1% vitamin solution) selective medium, whereas the medium was supplemented with 20 μg/ml uracil (SCVU) when growing *pyrEF*-deficient strains. *Sulfolobus*-competent cells were prepared as described previously[65] and plasmids were introduced into the cells by electroporation and transformed cells were allowed to form colonies on SCV plates.

**Construction of deletion mutants of *S. islandicus*.** A novel gene knockout method named marker insertion and unmarked target gene deletion (MID) developed by Zhang et al.[66] was employed to construct *cas* gene knockouts. In MID manipulation, each plasmid vector contains three homologous sequence arms, i.e., left, right and target gene arms. These homologous arms were carefully designed such that the insertion of a marker cassette would yield a functional target gene at the first step of recombination. Primers for amplification of these arms for each Cas module or *cas* gene by polymerase chain reaction are listed in Table 3. Ligation of these arms to the *E. coli* vector pUC19 gave pMID-Cascis, -Cmr-α, -Cmr-β, -cas7, -cas5, -cas6, -csa5, -cas3' and -cas3". These knockout plasmids were linearized with Pvu II before transformation. Transformants were obtained on SCV plates, whereas mutant cells were recovered by counter selection of *pyrEF* with 5'-fluoroorotic acids (5'-FOA). After three rounds of single colony purification, the mutants were analyzed for their mutant alleles by PCR with specific primers for each target gene or gene region and their mutant genotypes were verified (**Fig. 1B**).

Complementation plasmids were constructed with pSeSD, a highly efficient expression vector recently reported for *Sulfolobus*.[52] The expression is under the control of $P_{araS\text{-}SD}$, a synthetic promoter that confers up to one-third of the fully induced level of expression in a sucrose medium.[52] Each *cas* gene was amplified by PCR from *S. islandicus* total DNA with a Pfu DNA polymerase using specific primers listed in **Table 3**. After cleavage with the enzymes indicated in each PCR primer, the DNA fragments were

**Table 1.** Strains used in this study

| Strain | Genotype and features | Source |
|---|---|---|
| *S. islandicus* E233S | Δ*pyrEF*Δ*lacS* | Deng et al.[65] |
| ΔCRISPR | CRISPR arrays are not detectable by PCR | Gudbergsdottir et al.[42] |
| ΔCascis | All genes in Cascis module, including *csa3*, *csa1*, *cas1*, *cas2* and *cas4* are deleted | This work |
| ΔCmr-α | Deletion of the type IIIB Cmr locus with 6 genes | This work |
| ΔCmr-β | Deletion of the type IIIB Cmr locus with 7 genes | This work |
| Δcsa5 | In-frame deletion of *csa5* | This work |
| Δcas7 | In-frame deletion of *cas7* | This work |
| Δcas5 | In-frame deletion of *cas5* | This work |
| Δcas3′ | In-frame deletion of *cas3′* | This work |
| Δcas3″ | In-frame deletion of *cas3″* | This work |
| Δcas6 | Carrying deletion of the entire *cas6* gene plus 277 bp downstream flanking sequence | This work |

**Table 2.** Plasmids used in this study

| Plasmid | Features | Source |
|---|---|---|
| pSeSD | A *Sulfolobus-E. coli* shuttle vector carrying an expression cassette controlled under a synthetic strong promoter $P_{araS\text{-}SD}$ | Peng et al.[52] |
| pCC-A1S6 | An artificial interference plasmid; pSeSD carrying the interference module of the sequence of spacer 6 of array 1 preceded by the 5'-CCA-3' motif | This work |
| pRt-A1S6 | Reference plasmid of pCC-A1S6; the PAM motif (5'-CCA-3') was replaced with the last 5 nt of the repeat sequence (5'-GAA AG-3') | This work |
| pCC-A2S32 | An artificial interference plasmid; pSeSD carrying the interference module of the sequence of spacer 32 of array 2 preceded by the 5'-CCA-3' motif | This work |
| pRt-A2S32 | Reference plasmid of pCC-A2S32; the PAM motif (5'-CCA-3') was replaced with the last 8 nt of the repeat sequence (5'-ATT GAA AG-3') | This work |
| pSe-Csa5 | pSeSD carrying the *S. islandicus csa5* gene | This work |
| pSe-Cas7 | pSeSD carrying the *S. islandicus cas7* gene | This work |
| pSe-Cas5 | pSeSD carrying the *S. islandicus cas5* gene | This work |
| pSe-Cas3′ | pSeSD carrying the *S. islandicus cas3′* gene | This work |
| pSe-Cas3″ | pSeSD carrying the *S. islandicus cas3″* gene | This work |
| pSe-Cas6 | pSeSD carrying the *S. islandicus cas6* gene | This work |

cloned to pSeSD at the compatible sites, yielding pSe-csa5, -cas5, -cas7, -cas3', -cas3" and -cas6.

**Plasmid interference assay of PAM-mediated DNA targeting.** Artificial interference plasmids were constructed with two spacers, spacer 6 of CRISPR array 1 (A1S6) and spacer 32 of array 2 (A2S32). Oligonucleotides containing the complete sequences of the two spacers that were preceded with 5'-CCA-3' or with the last five or eight nucleotides of the repeat were synthesized from TAG Copenhagen (**Table 3**). Each pair of primers was annealed with each other by heating to 95°C for 10 min and then cooling down gradually to room temperature, yielding double-stranded DNAs with one blunt end and one protruding end, the latter of which matched the DNA ends after Mlu I digestion. These DNA fragments were cloned to pSeSD at Stu I and Mlu I sites individually, yielding plasmids containing an artificial protospacer (interference plasmids) and those containing a repeat-spacer module (reference plasmids) (**Table 2**).

**RNA preparation and northern analysis.** Total RNAs were extracted using the Trizol reagent (Invitrogen) following the protocol provided by the manufacturer. Thirty μg RNA was loaded on a 6% polyacrylamide-SDS gel for each sample and fractionated. RNAs in the gel were then transferred onto a nylon membrane and immobilized. Procedures for pre-hybridization, end labeling of complementary nucleotides and hybridization in northern analysis were described earlier[20] and the probes used in this work included a probe of the repeat sequence (5'-CTT TCA ATT CTA TAG TAG ATT AGC NNN N-3') for detecting intermediates of pre-crRNA processing and a probe of spacer 32 of array 2 (5'-GCG GGA TTA GTA GGG ATT CCC ATA GGG CTC TAT GAA CT-3') hybridizing to mature crRNAs.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

**Table 3.** Oligonucleotides used in this study

| Oligonucleotide | Sequence (5'-3'; PAM motif or repeat sequence underlined, restriction site in bold) |
|---|---|
| A1S6fwd-MluI | **CGC G**TC CAG TTG GCA TAA TCC TCA ACA TAC TTT CTT GTG TCG GGT ATG C |
| A1S6rev | GCA TAC CCG ACA CAA GAA AGT ATG TTG AGG ATT ATG CCA ACT GGA |
| ReA1S6fwd-MluI | **CGC G**AA AGG TTG GCA TAA TCC TCA ACA TAC TTT CTT GTG TCG GGT ATG C |
| ReA1S6rev | GCA TAC CCG ACA CAA GAA AGT ATG TTG AGG ATT ATG CCA ACC TTT (C) |
| A2S32fwd-MluI | **CGC G**TC CCA AGT TCA TAG AGC CCT ATG GGA ATC CCT ACT AAT CCC GC |
| A2S32rev | GCG GGA TTA GTA GGG ATT CCC ATA GGG CTC TAT GAA CTT GGG A |
| ReA2S32fwd-MluI | **CGC G**AT TGA AAG AGT TCA TAG AGC CCT ATG GGA ATC CCT ACT AAT CCC GC |
| ReA2S32rev | GCG GGA TTA GTA GGG ATT CCC ATA GGG CTC TAT GAA CTC TTT CAA T |
| Csa5fwd-MluI | CC**A CGC GT**T GGC ACA ACA AGT AAA AGA AG |
| Csa5rev-StuI | A**AG GCC T**TT TCT TCT CAC CAC CAC CTT G |
| Cas7fwd-NdeI | GTT GTT C**CA TAT G**AT AAG TGG TTC AGG TAG ATT T |
| Cas7rev-MluI | CG**A CGC GT**C TAC TTT TCC TTT AAT TTA CC |
| Cas5fwd-MluI | GG**A CGC GT**T GAT CTA CTC TAA GGT TTT TT |
| Cas5rev-StuI | G**AG GCC T**GC TAA AGA CAA CAT ATT CTC C |
| Cas3′fwd-NdeI | CCC AAA G**CA TAT G**TT GTC TTT AGC TGA CTT CTA T |
| Cas3′rev-MluI | CG**A CGC GT**T CAA TAC ACA CCA CCT ATT TC |
| Cas3′′fwd-NdeI | AAT AAA T**CA TAT G**AT CAA GCC TTG CGC TTA TGA G |
| Cas′′rev-MluI | CG**A CGC GT**T TAT AAA GTG GAA CCT CCA TT |
| Cas6fwd-NdeI | GGG ACC G**CA TAT G**CC ATT AAT TTT CAA GAT A |
| Cas6rev-MluI | GT**A CGC GT**A CCT TTA AAG TCT GAG GA |
| Cmr-α-Gfwd-SalI | AAT T**GT CGA C**GC AGG AAT GGT GGT AGA GTC T |
| Cmr-α-Grev-MluI | CC**A CGC GT**C CTC CCC ACT TGA ATA CTA CC |
| Cmr-α-Lfwd-NcoI | AAT T**CC ATG G**CA GGA AAG CAG TAG AGA AGG A |
| Cmr-α-Lrev-XhoI | ATC **CTC GAG** AAC TTG TAA CCC TAC GTC GTT |
| Cmr-α-Rfwd-XhoI | ATC **CTC GAG** ACT AAG ATG ACT AAG GAA GGC |
| Cmr-α-Rrev-SphI | AAT T**GC ATG C**AC AAA AAC ATG AAC GTA TGA G |
| Cmr-β-Gfwd-SalI | AAT T**GT CGA C**AT AGT CTT CAT CGG CAC ATA C |
| Cmr-β-Grev-MluI | CC**A CGC GT**T CAT CTG TAG AGG AGG AAA TA |
| Cmr-β-Lfwd-NcoI | AAT T**CC ATG G**TC AGC AGG CAG TAA AGG AAG A |
| Cmr-β-Lrev-XhoI | ATC **CTC GAG** CAA TGG GTA AGG TGG ATA AGA |
| Cmr-β-Rfwd-XhoI | ATC **CTC GAG** TTG CAT CGG GAC TAC TGT AAG |
| Cmr-β-Rrev-SphI | AAT T**GC ATG C**AG AGG GTG GGT AAG TGA TGA T |
| Cassis-Gfwd-SalI | AAT T**GT CGA C**AT GTT CTC CTT CAT GCC CGT T |
| Cascis-Grev-MluI | CC**A CGC GT**G TCC GTT CTA ATA CCG CTC CT |
| Cascis-Lfwd-NcoI | AAT T**CC ATG G**TT CAT CAA GAG GAA AAT CGT C |
| Cascis-Lrev-XhoI | ATC **CTC GAG** TTC GTT AGA AAC ACT CGC TAG |
| Cascis-Rfwd-XhoI | ATC **CTC GAG** TGT ATT CAT AAG CCT CAT TCC |
| Cascis-Rrev-SphI | AAT T**GC ATG C**CA ACG GAA ATA GTA GGG AAC A |
| Cas3′-Gfwd-SalI | AAT T**GT CGA C**AA GGT AAC GAA ACT AAA GAC T |
| Cas3′-Grev-MluI | CC**A CGC GT**C ACA CCA CCT ATT TCA CTA TT |
| Cas3′-Lfwd-NcoI | AAT T**CC ATG G**GC TTA CAC TGA GGA CAT TGT T |
| Cas3′-Lrev-XhoI | ATC **CTC GAG** ACC ATG ACG AGT TAT GAA ATC |
| Cas3′-Rfwd-XhoI | ATC **CTC GAG** TGA TCA AGC CTT GCG CTT ATG |
| Cas3′-Rrev-SphI | AAT T**GC ATG C**TA TCC CCC TAA TTG CGT TCA G |
| Cas3″-Gfwd-SalI | AAT T**GT CGA C**GG TGA TTG ATT CGA TTC TGG C |
| Cas3″-Grev-MluI | CC**A CGC GT**A GTG GAA CCT CCA TTT AAC TC |

**Table 3.** Oligonucleotides used in this study

| | |
|---|---|
| Cas3″-Lfwd-NcoI | AAT T**CC ATG G**GG TGA TTG ATT CGA TTC TGG C |
| Cas3″-Lrev-XhoI | ATC **CTC GAG** CTC ATA AGC GCA AGG CTT GAT |
| Cas3″-Rfwd-XhoI | ATC **CTC GAG** TTA TAA TAT ATT TGG AGA TAA |
| Cas3″-Rrev-SphI | AAT T**GC ATG C**TG AAT CAA ACT GTA GAG AAT A |
| Cas5-Gfwd-SalI | AAT T**GT CGA C**AG TGA TCC CTA CTG GTA AAG T |
| Cas5-Grev-MluI | CC**A CGC GT**A GAA GTC AGC TAA AGA CAA CA |
| Cas5-Lfwd-NcoI | AAT T**CC ATG G**GT TAG AAA GGC AAA AGG CTC G |
| Cas5-Lrev-XhoI | ATC **CTC GAG** TAA CTA CGG AAA AAC CCC AAT |
| Cas5-Rfwd-XhoI | ATC **CTC GAG** TTA GCT GAC TTC TAT AAC GAT |
| Cas5-Rrev-SphI | AAT T**GC ATG C**GT CTT GAG CAC CAA TAC TTT T |
| Cas6-Gfwd-SalI | AAT T**GT CGA C**TT GCC ATC AGT TGG CTT AAT T |
| Cas6-Grev-MluI | CC**A CGC GT**C CTT AAC CTT TAA AGT CTG AG |
| Cas6-Lfwd-NcoI | AAT T**CC ATG G**GT TAG AAA GGC AAA AGG CTC G |
| Cas6-Lrev-XhoI | ATC **CTC GAG** TGA AAC CGA GAT GGG ATA TAA |
| Cas6-Rfwd-XhoI | ATC **CTC GAG** CCT CAG ACT TTA AAG GTT AAG |
| Cas6-Rrev-SphI | AAT T**GC ATG C**GG AGG GGA GAT GTT CCG ATG C |
| Cas7-Gfwd-SalI | AAT T**GT CGA C**AA GCC TTA GCT CAT GCC TAT |
| Cas7-Grev-MluI | CC**A CGC GT**T CTT GCC CCT CTA CTT TTC CT |
| Cas7-Lfwd-NcoI | AAT T**CC ATG G**TC CAG GTT TAT GAA AAA TCT T |
| Cas7-Lrev-XhoI | ATC **CTC GAG** ATC ATT TCT TCT CAC CAC CAC |
| Cas7-Rfwd-XhoI | ATC **CTC GAG** GTA GAG GGG CAA GAG AGA TTT |
| Cas7-Rrev-SphI | AAT T**GC ATG C**CT CAA CTT CAA CCT CCT TGC T |
| Csa5-Gfwd-SalI | AAT T**GT CGA C**GA GGC TGT GAC TAA GGT TCT A |
| Csa5-Grev-MluI | CC**A CGC GT**T ATC ATT TCT TCT CAC CAC CA |
| Csa5-Lfwd-NcoI | AAT T**CC ATG G**GG AGG GAT GAG ATT AAT GAT T |
| Csa5-Lrev-XhoI | ATC **CTC GAG** CTT CTT TTA CTT GTT GTG CCA |
| Csa5-Rfwd-XhoI | ATC **CTC GAG** AGA AAT GAT AAG TGG TTC AGG |
| Csa5-Rrev-SphI | AAT T**GC ATG C**TA CAT GAA ACC TCC AAT ATC A |


### Acknowledgments

This work was supported by the Danish Council for Independent Research: Technology and Production Sciences (11-106683) and the Natural Science Foundation of China (grant no. 31128011, no. 31100050 and no. 31100096). We thank our colleagues at the laboratories in Copenhagen and in Wuhan for fruitful discussions.


### Author Contributions

Conceived and designed the experiments: Q.S., Y.X.L., W.P. and N.P. performed experiments: W.P., H.L. and S.H. analyzed data: Q.S., Y.X.L., W.P., N.P., H.L. and S.H. wrote the paper: Q.S. and W.P.

### References


1. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 2005; 60:174-82; PMID:15791728; http://dx.doi.org/10.1007/s00239-004-0046-3.
2. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 2005; 151:653-63; PMID:15758212; http://dx.doi.org/10.1099/mic.0.27437-0.
3. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 2005; 151:2551-61; PMID:16079334; http://dx.doi.org/10.1099/mic.0.28048-0.
4. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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 2006; 1:7; PMID:16545108; http://dx.doi.org/10.1186/1745-6150-1-7.
5. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science 2010; 327:167-70; PMID:20056882; http://dx.doi.org/10.1126/science.1179555.
6. Terns MP, Terns RM. CRISPR-based adaptive immune systems. Curr Opin Microbiol 2011; 14:321-7; PMID:21531607; http://dx.doi.org/10.1016/j.mib.2011.03.005.
7. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 2011; 9:467-77; PMID:21552286; http://dx.doi.org/10.1038/nrmicro2577.
8. Garrett RA, Vestergaard G, Shah SA. Archaeal CRISPR-based immune systems: exchangeable functional modules. Trends Microbiol 2011; 19:549-56; PMID:21945420; http://dx.doi.org/10.1016/j.tim.2011.08.002.
9. Westra ER, Swarts DC, Staals RH, Jore MM, Brouns SJ, van der Oost J. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu Rev Genet 2012; 46:311-39; PMID:23145983; http://dx.doi.org/10.1146/annurev-genet-110711-155447.
10. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature 2012; 482:331-8; PMID:22337052; http://dx.doi.org/10.1038/nature10886.
11. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput Biol 2005; 1:e60; PMID:16292354; http://dx.doi.org/10.1371/journal.pcbi.0010060.

12. Plagens A, Tjaden B, Hagemann A, Randau L, Hensel R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax.* J Bacteriol 2012; 194:2491-500; PMID:22408157; http://dx.doi.org/10.1128/JB.00206-12.
13. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007; 315:1709-12; PMID:17379808; http://dx.doi.org/10.1126/science.1138140.
14. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli.* Nucleic Acids Res 2012; 40:5569-76; PMID:22402487; http://dx.doi.org/10.1093/nar/gks216.
15. Swarts DC, Mosterd C, van Passel MW, Brouns SJ. CRISPR interference directs strand specific spacer acquisition. PLoS One 2012; 7:e35888; PMID:22558257; http://dx.doi.org/10.1371/journal.pone.0035888.
16. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Commun 2012; 3:945; PMID:22781758; http://dx.doi.org/10.1038/ncomms1937.
17. Erdmann S, Garrett RA. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. Mol Microbiol 2012; 85:1044-56; PMID:22834906; http://dx.doi.org/10.1111/j.1365-2958.2012.08171.x.
18. Tang TH, Bachellerie JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus.* Proc Natl Acad Sci USA 2002; 99:7536-41; PMID:12032318; http://dx.doi.org/10.1073/pnas.112047299.
19. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 2008; 321:960-4; PMID:18703739; http://dx.doi.org/10.1126/science.1159689.
20. Lillestøl RK, Shah SA, Brügger K, Redder P, Phan H, Christiansen J, et al. CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. Mol Microbiol 2009; 72:259-72; PMID:19239620; http://dx.doi.org/10.1111/j.1365-2958.2009.06641.x.
21. Hale C, Kleppe K, Terns RM, Terns MP. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus.* RNA 2008; 14:2572-9; PMID:18971321; http://dx.doi.org/10.1261/rna.1246808.
22. Hatoum-Aslan A, Maniv I, Marraffini LA. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proc Natl Acad Sci USA 2011; 108:21218-22; PMID:22160698; http://dx.doi.org/10.1073/pnas.1112832108.
23. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. Binding and cleavage of CRISPR RNA by Cas6. RNA 2010; 16:2181-8; PMID:20884784; http://dx.doi.org/10.1261/rna.2230110.
24. Garside EL, Schellenberg MJ, Gesner EM, Bonanno JB, Sauder JM, Burley SK, et al. Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. RNA 2012; 18:2020-8; PMID:23006625; http://dx.doi.org/10.1261/rna.033100.112.
25. Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, Naismith JH, et al. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). J Biol Chem 2011; 286:21643-56; PMID:21507944; http://dx.doi.org/10.1074/jbc.M111.238485.
26. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science 2010; 329:1355-8; PMID:20829488; http://dx.doi.org/10.1126/science.1192272.
27. Richter H, Zoephel J, Schermuly J, Maticzka D, Backofen R, Randau L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis.* Nucleic Acids Res 2012; 40:9887-96; PMID:22879377; http://dx.doi.org/10.1093/nar/gks737.
28. Gesner EM, Schellenberg MJ, Garside EL, George MM, Macmillan AM. Recognition and maturation of effector RNAs in a CRISPR interference pathway. Nat Struct Mol Biol 2011; 18:688-92; PMID:21572444; http://dx.doi.org/10.1038/nsmb.2042.
29. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 2011; 471:602-7; PMID:21455174; http://dx.doi.org/10.1038/nature09886.
30. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat Struct Mol Biol 2011; 18:529-36; PMID:21460843; http://dx.doi.org/10.1038/nsmb.2019.
31. Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJ, van der Oost J, et al. Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature 2011; 477:486-9; PMID:21938068; http://dx.doi.org/10.1038/nature10402.
32. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 2009; 139:945-56; PMID:19945378; http://dx.doi.org/10.1016/j.cell.2009.07.040.
33. Westra ER, van Erp PB, Künne T, Wong SP, Staals RH, Seegers CL, et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol Cell 2012; 46:595-605; PMID:22521689; http://dx.doi.org/10.1016/j.molcel.2012.03.018.
34. Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 2010; 468:67-71; PMID:21048762; http://dx.doi.org/10.1038/nature09523.
35. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012; 337:816-21; PMID:22745249; http://dx.doi.org/10.1126/science.1225829.
36. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 2008; 322:1843-5; PMID:19095942; http://dx.doi.org/10.1126/science.1165771.
37. Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, et al. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Mol Cell 2012; 45:292-302; PMID:22227116; http://dx.doi.org/10.1016/j.molcel.2011.10.023.
38. Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, Graham S, et al. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. Mol Cell 2012; 45:303-13; PMID:22227115; http://dx.doi.org/10.1016/j.molcel.2011.12.013.
39. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 2009; 155:733-40; PMID:19246744; http://dx.doi.org/10.1099/mic.0.023960-0.
40. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli.* Nucleic Acids Res 2011; 39:9275-82; PMID:21813460; http://dx.doi.org/10.1093/nar/gkr606.
41. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci USA 2011; 108:10098-103; PMID:21646539; http://dx.doi.org/10.1073/pnas.1104144108.
42. Gudbergsdottir S, Deng L, Chen Z, Jensen JV, Jensen LR, She Q, et al. Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. Mol Microbiol 2011; 79:35-49; PMID:21166892; http://dx.doi.org/10.1111/j.1365-2958.2010.07452.x.
43. Fischer S, Maier LK, Stoll B, Brendel J, Fischer E, Pfeiffer F, et al. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. J Biol Chem 2012; 287:33351-63; PMID:22767603; http://dx.doi.org/10.1074/jbc.M112.377002.
44. Deng L, Kenchappa CS, Peng X, She Q, Garrett RA. Modulation of CRISPR locus transcription by the repeat-binding protein Cbp1 in *Sulfolobus.* Nucleic Acids Res 2012; 40:2470-80; PMID:22139923; http://dx.doi.org/10.1093/nar/gkr1111.
45. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 2010; 463:568-71; PMID:20072129; http://dx.doi.org/10.1038/nature08703.
46. Garrett RA, Shah SA, Vestergaard G, Deng L, Gudbergsdottir S, Kenchappa CS, et al. CRISPR-based immune systems of the Sulfolobales: complexity and diversity. Biochem Soc Trans 2011; 39:51-7; PMID:21265746; http://dx.doi.org/10.1042/BST0390051.
47. Guo L, Brügger K, Liu C, Shah SA, Zheng H, Zhu Y, et al. Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. J Bacteriol 2011; 193:1672-80; PMID:21278296; http://dx.doi.org/10.1128/JB.01487-10.
48. She Q, Zhang C, Deng L, Peng N, Chen Z, Liang YX. Genetic analyses in the hyperthermophilic archaeon *Sulfolobus islandicus.* Biochem Soc Trans 2009; 37:92-6; PMID:19143609; http://dx.doi.org/10.1042/BST0370092.
49. Zhang C, Tian B, Li S, Ao X, Dalgaard K, Gökce S, et al. Genetic manipulation in *Sulfolobus islandicus* and functional analysis of DNA repair genes. Biochem Soc Trans 2013; 41:405-10; PMID:23356319; http://dx.doi.org/10.1042/BST20120285.
50. Deng L, Garrett RA, Shah SA, Peng X, She Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus.* Mol Microbiol 2013; In press; PMID:23320564; http://dx.doi.org/10.1111/mmi.12152.
51. Lillestøl RK, Redder P, Garrett RA, Brügger K. A putative viral defence mechanism in archaeal cells. Archaea 2006; 2:59-72; PMID:16877322; http://dx.doi.org/10.1155/2006/542818.
52. Peng N, Deng L, Mei Y, Jiang D, Hu Y, Awayez M, et al. A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon *Sulfolobus islandicus.* Appl Environ Microbiol 2012; 78:5630-7; PMID:22660711; http://dx.doi.org/10.1128/AEM.00855-12.
53. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc Natl Acad Sci USA 2011; 108:10092-7; PMID:21536913; http://dx.doi.org/10.1073/pnas.1102716108.

54. Cady KC, O'Toole GA. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. J Bacteriol 2011; 193:3433-45; PMID:21398535; http://dx.doi.org/10.1128/JB.01411-10.
55. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, et al. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 2001; 98:7835-40; PMID:11427726; http://dx.doi.org/10.1073/pnas.141222098.
56. Kawarabayasi Y, Hino Y, Horikawa H, Jin-no K, Takahashi M, Sekine M, et al. Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. DNA Res 2001; 8:123-40; PMID:11572479; http://dx.doi.org/10.1093/dnares/8.4.123.
57. Chen L, Brügger K, Skovgaard M, Redder P, She Q, Torarinsson E, et al. The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. J Bacteriol 2005; 187:4992-9; PMID:15995215; http://dx.doi.org/10.1128/JB.187.14.4992-4999.2005.
58. Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ. Biogeography of the *Sulfolobus islandicus* pan-genome. Proc Natl Acad Sci USA 2009; 106:8605-10; PMID:19435847; http://dx.doi.org/10.1073/pnas.0808945106.
59. Beloglazova N, Petit P, Flick R, Brown G, Savchenko A, Yakunin AF. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. EMBO J 2011; 30:4616-27; PMID:22009198; http://dx.doi.org/10.1038/emboj.2011.377.
60. Bernick DL, Cox CL, Dennis PP, Lowe TM. Comparative genomic and transcriptional analyses of CRISPR systems across the genus *Pyrobaculum*. Front Microbiol 2012; 3:251; PMID:22811677; http://dx.doi.org/10.3389/fmicb.2012.00251.
61. Sashital DG, Wiedenheft B, Doudna JA. Mechanism of foreign DNA selection in a bacterial adaptive immune system. Mol Cell 2012; 46:606-15; PMID:22521690; http://dx.doi.org/10.1016/j.molcel.2012.03.020.
62. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev 2008; 22:3489-96; PMID:19141480; http://dx.doi.org/10.1101/gad.1742908.
63. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, Wanner BL, et al. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. Mol Microbiol 2010; 77:1367-79; PMID:20624226; http://dx.doi.org/10.1111/j.1365-2958.2010.07265.x.
64. Contursi P, Jensen S, Aucelli T, Rossi M, Bartolucci S, She Q. Characterization of the *Sulfolobus* host-SSV2 virus interaction. Extremophiles 2006; 10:615-27; PMID:16896526; http://dx.doi.org/10.1007/s00792-006-0017-2.
65. Deng L, Zhu H, Chen Z, Liang YX, She Q. Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. Extremophiles 2009; 13:735-46; PMID:19513584; http://dx.doi.org/10.1007/s00792-009-0254-2.
66. Zhang C, Guo L, Deng L, Wu Y, Liang Y, Huang L, et al. Revealing the essentiality of multiple archaeal *pcna* genes using a mutant propagation assay based on an improved knockout method. Microbiology 2010; 156:3386-97; PMID:20705666; http://dx.doi.org/10.1099/mic.0.042523-0.