

CRISPR-mediated defense mechanisms in the hyperthermophilic archaeal genus *Sulfolobus*

Andrea Manica* and Christa Schleper

University of Vienna; Department of Genetics in Ecology; Vienna, Austria

Keywords: *Sulfolobales*, archaea, virus defense, CRISPR-Cas system, small RNAs

CRISPR (clustered regularly interspaced short palindromic repeats)-mediated virus defense based on small RNAs is a hallmark of archaea and also found in many bacteria. Archaeal genomes, and in particular, organisms of the extremely thermoacidophilic genus *Sulfolobus*, carry extensive CRISPR loci each with dozens of sequence signatures (spacers) able to mediate targeting and degradation of complementary invading nucleic acids. The diversity of CRISPR systems and their associated protein complexes indicates an extensive functional breadth and versatility of this adaptive immune system. *Sulfolobus solfataricus* and *S. islandicus* represent two of the best characterized genetic model organisms in the archaea not only with respect to the CRISPR system. Here we address and discuss in a broader context particularly recent progress made in understanding spacer recruitment from foreign DNA, production of small RNAs, in vitro activity of CRISPR-associated protein complexes and attack of viruses and plasmids in in vivo test systems.

Introduction

Small RNAs play a key role in cell regulation, stability and defense in all three domains of life.¹⁻⁶ Many different types of small RNAs were identified and their mode of action in gene regulation, epigenetics and defense were extensively studied in eukaryotes and prokaryotes over the past years.⁷⁻⁹

Recently, a new type of small RNAs called crRNAs (from CRISPR RNAs, clustered regularly interspaced short palindromic repeats) was discovered in Bacteria and Archaea, which has been shown to be active against invading genetic elements.¹⁰⁻¹⁵ This defense system is found in almost every genome of archaea and in about half of the bacterial genomes (www.crispi.genouest.org/).¹⁶ Although CRISPR-mediated defense is a fundamentally new process, its mode of action resembles formally the basic principles of siRNA activities in Eukaryotes,¹⁷⁻¹⁹ including (1) recognition of the invading genetic element, (2) amplification of the effector molecule, (3) processing of the precursor RNA and (4) targeting of the invader through base-pair recognition. However, there are also many remarkable differences between the two systems. For example, crRNAs being processed from a long precursor RNA (pre-crRNA) transcribed from a CRISPR locus, i.e., a

DNA region with dozens of unique DNA sequences of 30–45 nt (called spacers) that are separated by identical direct repeats.²⁰⁻²⁷ Upon viral or plasmid invasion, new short DNA sequences from the invader can be incorporated as spacers in the leader-proximal part of a CRISPR locus.^{10,28,29} The mechanism of recognition of the extra-chromosomal element as invader is still poorly understood. However, based on the enzymatic properties of the proteins involved in the recognition/integration of new spacers,²⁸⁻³³ this process unlike the eukaryotic RNAi system, should not involve the presence of dsRNAs.³⁴⁻³⁶

The CRISPR Cas system activity can be divided into three temporally and functionally distinct processes: adaptation, expression/processing and interference. All these steps involve different sets of Cas proteins (CRISPR-associated proteins), and in some systems, endogenous non-cas gene products.³⁷

During adaptation, two proteins defined as Cas1 and Cas2 (“core proteins”), ubiquitous to all CRISPR systems, recognize the invading genetic element and drive the incorporation of new DNA sequences into the CRISPR locus.³² In this way the invasion of an extra-chromosomal element is “recorded” in the host chromosome allowing the system to be prepared for future infections of the same invader. Once integrated, the new spacer is co-transcribed with all other spacers located in the same CRISPR locus. Transcription initiates at a promoter located within the leader sequence (upstream of the repeat cluster) and allows the CRISPR-Cas system to amplify the signal by producing a multitude of crRNAs from a single spacer. CRISPR and Cas expression is often under the control of cellular regulators³⁸⁻⁴⁰ and seems to be inducible by abiotic and/or biotic stress.^{38,41,42}

Upon transcription, the long CRISPR pre-RNA is processed by other Cas proteins (i.e., Cas6) or by endogenous proteins, to the short mature crRNAs.^{11,43,44} A mature crRNA is composed of a spacer derived from the extra-chromosomal genetic element, an 8 nt 5' tag derived from the preceding repeat and a 3' end handle of variable length that stems from the downstream repeat.^{11,13,43-45} Once mature, the crRNA is incorporated into the “interference” complex formed by several CRISPR-associated proteins and guides the enzymatic machinery to the targeted foreign nucleic acid via base-pairing interaction.

Several classifications of the highly divergent CRISPR-cas systems were proposed based on the type of repeat,²⁵ and/or Cas proteins encoded near the locus. Haft and colleagues initially identified several CRISPR type-specific proteins which allowed the classification into more than 45 families.²⁴ Recently, a new

*Correspondence to: Andrea Manica; Email: andrea.manica@univie.ac.at
Submitted: 01/25/13; Revised: 02/26/13; Accepted: 02/28/13
<http://dx.doi.org/10.4161/rna.24154>

©2012 Landes Bioscience. Do not distribute.

classification merged the different CRISPR families into three major types: type I, type II, type III, which can be further subdivided.⁴⁶ Several CRISPR types can be present simultaneously on the chromosome of the same organism, indicating that each type could have different activities against various genetic elements or nucleic acids. CRISPR systems were extensively studied in different organisms and many molecular details of their activities were elucidated. CRISPR type I (*Escherichia coli*, *S. solfataricus*, *Pseudomonas aeruginosa*), type II (*Streptococcus thermophilus*) and type IIIA (*Staphylococcus epidermidis*) seem to target principally DNA whereas type IIIB CRISPR (*Pyrococcus furiosus*, *S. solfataricus*) might be involved in the targeting of invading RNAs.^{10-15,45,47} Interestingly, type II CRISPR are only present in bacterial genomes³² while type I and III CRISPR are widespread within bacteria and archaea.³² Based on comparative studies, it has been inferred that intra- as well as interdomain transfer of CRISPR loci between archaea and bacteria must have occurred.^{32,48} The latter is particularly remarkable, considering the fundamental differences of transcriptional and translational mechanisms in the two domains, as well as of their cell wall structures and transferable genetic elements.^{49,50}

S. solfataricus and *S. islandicus* currently represent the most intensely studied archaea regarding the CRISPR-Cas system. Expression and genomic arrangements of the CRISPR loci have been analyzed,^{26,48,51-56} in vivo and in vitro activity of several complexes and proteins have been characterized and crystal structures of cas proteins have been solved.^{44,45,53,54,57,58} *S. solfataricus* possesses an extensive and complex CRISPR system, which includes six different CRISPR loci, two “adaptation” Cas cassettes and five “interference” Cas protein complexes, the latter belonging to CRISPR type I (three clusters), CRISPR type IIIA and type IIIB, respectively. The complexity of the CRISPR-cas systems might be correlated with the high diversity and abundance of genetic elements and viruses in archaea: Viruses from eight out of 12 structurally and genomically highly divergent virus families infect *Sulfolobales* or closely related thermoacidophilic organisms.⁵⁹ In the 3 Mbp genome of *S. solfataricus* P2, a total of 344 putatively mobilizable genetic elements can be found that make up about 10% of the whole genome.⁵¹ Besides the large numbers of mobile genetic elements, genetic tools have been developed for several *Sulfolobus* species,^{60,61} which allow to manipulate and functionally study the virus-host systems^{62,63} as well as the CRISPR-based defense mechanism.^{12,14,28}

CRISPR Locus Organization in *Sulfolobales*

Within the *Sulfolobales*, more than 80 repeat clusters were identified, with a total of ~4,500 unique spacer sequences.^{52,64,65} The length of individual CRISPR spacers is fairly conserved within the different CRISPR loci ranging between 34–44 nucleotides. The repeat-clusters can be several kilobases in length carrying more than 100 different spacers. CRISPR loci are usually conserved with respect to their repeat sequence and spacer length, although in some CRISPR loci, repeat sequences are not always uniform.⁴⁸ Interestingly, also two self-transmissible plasmids (pNOB8 and pKEF9), which can spread among *Sulfolobales*,

carry small repeat-clusters in their genomes. However, they lack the adjacent cas genes.⁵⁵

The majority of the spacers in a CRISPR locus are unique, although duplicate spacers can also be found in large repeat clusters.⁵⁶ Different strains of *S. solfataricus* share many identical spacers in the leader distal part of the CRISPR locus, whereas different strains of *S. islandicus* isolated from geographically distant hot-springs show no spacer conservations.⁴⁸ Moreover, only low CRISPR spacer conservation was found within a population of *S. islandicus* strains isolated from the same hot-spring in Kamchatka.⁶⁵

Many *Sulfolobus* species carry several CRISPR loci on their chromosome, which do not always share the same repeat type. A particular type of repeat is often associated with a particular cluster of core proteins (Cas1/Cas2). Also, the upstream region (leader) is often conserved in CRISPR loci that share the same repeat type.^{26,55}

CRISPR Transcription and Regulation in *Sulfolobales*

The leader sequence provides the DNA elements for CRISPR transcription. In *S. acidocaldarius*, transcriptional CRISPR start sites were mapped directly before or 17–21 nt upstream of the first repeat sequence and they were always preceded by a typical archaeal BRE/TATA box promoter motif.^{55,66} Along the leader sequence, additional motifs were found to be conserved between the different CRISPR families, although the role of these motifs remains unknown.⁵⁵

The long pre-crRNA transcript covering the full length of the locus is successively processed by Cas6 into small mature crRNAs.⁴⁴ Full-length transcripts were detected by northern blot assay for each CRISPR locus of *S. acidocaldarius* using oligonucleotide probes complementary to different CRISPR spacers and repeats.^{26,55} Small RNAs of 50–70 nt were also detected, compatible with the processing of the pre-crRNA within the repeat sequence.^{44,55} Furthermore, RNAs of shorter size, i.e., ~40 nt in length were detected using spacer-specific probes, but not with repeat-based probes, demonstrating that some crRNAs undergo a further processing of their repeat by an unknown exonuclease.⁵⁵ These results were supported by sequencing of crRNAs that co-purified with the Cmr complex of *S. solfataricus*.⁴⁵

Northern blot analysis of CRISPR pre-RNA revealed the presence of antisense transcripts for each of the *S. acidocaldarius* CRISPR loci.⁵⁵ Bi-directional transcription was also experimentally verified in *S. solfataricus* CRISPR loci^{26,55,67,68} and was confirmed by analyzing publicly available transcriptome data.⁶⁹ However, the loci were primarily transcribed from the leader strand.^{55,67,68} Additionally, internal transcription start sites have also been identified for some *Sulfolobus* CRISPR loci, probably caused by the incorporation of spacer sequences carrying promoter-like motifs.⁵² The presence of bidirectional transcripts has been found in few archaeal species: *Sulfolobus* spp., *P. furiosus* and *Methanococcus marisaludis*.^{13,70} In bacteria, with the exception of *Clostridium thermocellum*, no antisense transcripts were detected and CRISPR transcription seems to be mostly

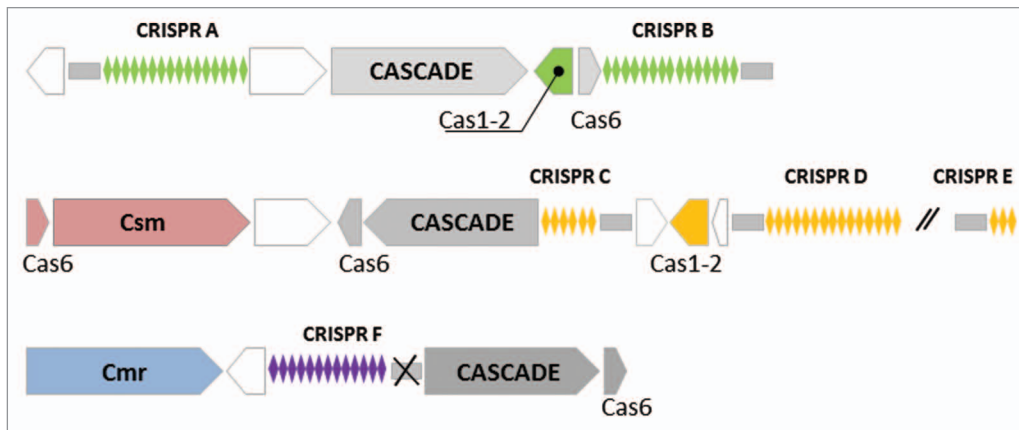


Figure 1. Schematic representation of the CRISPR-Cas systems in *Sulfolobus solfataricus*. The six CRISPR loci are named from A through F, the small gray rectangle in front of each locus represents the leader sequence. Loci with the same repeats share the same color. The putative “integration cassette” Cas1-Cas2 is highlighted with the same color as its adjacent locus. In dark gray, the three CASCADE cassettes located adjacent to a CRISPR locus, and the Cmr/Csm cassettes near the CRISPR F (light blue) and C (light red), respectively. In white, additional CRISPR-related proteins.

uni-directional.^{11,39,40,70} Although several hypotheses have been formulated, the biological importance of bidirectional transcription remains unclear and no experimental data have so far elucidated the role of antisense RNAs in CRISPR regulation or interference.^{55,71}

The presence of several putative regulatory motifs within the leader sequence, and the high energetic cost related to the possible constitutive expression of a long non-coding RNA (pre-crRNA) in the absence of an invading genetic element, speaks for a complex transcriptional regulation of the CRISPR locus. In *E. coli*, CRISPR locus and associated Cas gene transcription is under the control of the repressor H-NS, which binds particular DNA motifs within the leader sequence and in the promoter of some cas genes silencing their expression.³⁹ Transcription can be restored only in the presence of the transcriptional activator LeuO.⁴⁰ Within *Sulfolobales* (or other archaea), no such type of regulation has so far been demonstrated. However, Deng and co-workers identified a protein, Cbp1 (SSO0454), which directly interferes with the CRISPR locus transcription by binding to the CRISPR DNA repeats of *S. solfataricus* and of the conjugative plasmid pNOB8.^{71,72} Mutation and expression studies performed in *S. solfataricus* and *S. islandicus*, respectively, imply a role of this protein in the transcriptional regulation of the CRISPR locus.⁷¹ Cbp1-knockout mutants of *S. islandicus* showed strong reduction in CRISPR pre-crRNA levels (> 600 nt) compared with the wild-type strain and a Cbp1 complemented strain.⁷¹ Conversely, over-expression of the Cbp1 protein in *S. solfataricus* strain P2 revealed higher pre-crRNA (> 600 nt long) transcription levels compared with the wild-type strain.⁷¹

Cbp1 showed different binding affinity depending upon the repeat type, with higher affinity to the repeat sequence of locus C and D. In contrast to its enhancing activity on the leader transcript, no increase of CRISPR antisense transcripts and a decrease in transcription starting at internal sites was detected.⁷¹ Interestingly, microarray analysis of *Sulfolobus* P2 cells over-expressing Cbp1 showed significant reductions of the SSO1101 transcript, encoding one of the few proteins found to be involved

in biofilm formation in *Sulfolobales* sp.^{71,73} As postulated by the authors, the reduction of expression of proteins involved in biofilm formation could be a synergistic mechanism to block invader colonization of the neighboring cells.⁷¹ If correct, this would imply another difference to the eukaryotic RNAi system in which the interference signal can spread from cell to cell through the whole organism helping the recognition and the targeting of the invaders even in cells located very distantly from the infection site.⁷⁴

Taken together, the above results indicate a role of Cbp1 in the transcriptional regulation of *Sulfolobus* CRISPR loci, although its mode of action is still not fully understood.

Complexity and Adaptation of the *Sulfolobus* CRISPR-Cas System

Locus A and B of the six CRISPR loci (termed A through F) in *S. solfataricus* share the same repeat and similar leader sequences,^{26,55} which are distinct from those shared among loci C, D and E. In the proximity of locus A, B, C and D, two “adaptation” cassettes (Cas1-Cas2) are present (Fig. 1). The originally suggested classification by Lillestol and co-workers (2009), which comprised eight different CRISPR families for archaea, was based on leader sequences, repeat type and Cas1 gene phylogeny.^{48,55,56} Since it has been shown that Cas1 is involved in the integration of new spacers,^{29,30} this first classification could also reflect different spacer acquisition mechanisms of the different archaeal CRISPR families, whereas the classification of Makarowa et al.³⁰ is based primarily on the phylogeny of Cas proteins involved in the processing and interference process.

Although it is still unknown how the CRISPR system discriminates the extra-chromosomal element from the host chromosome during acquisition of new spacers, this process has been demonstrated under laboratory conditions for *E. coli*,^{29,33} *S. thermophiles*¹⁰ and, recently, *S. solfataricus*.²⁸ The process was shown to be dependent on a small recognition motif called PAM (protospacer adjacent motif)^{10,12,47,75} located in the vicinity of the

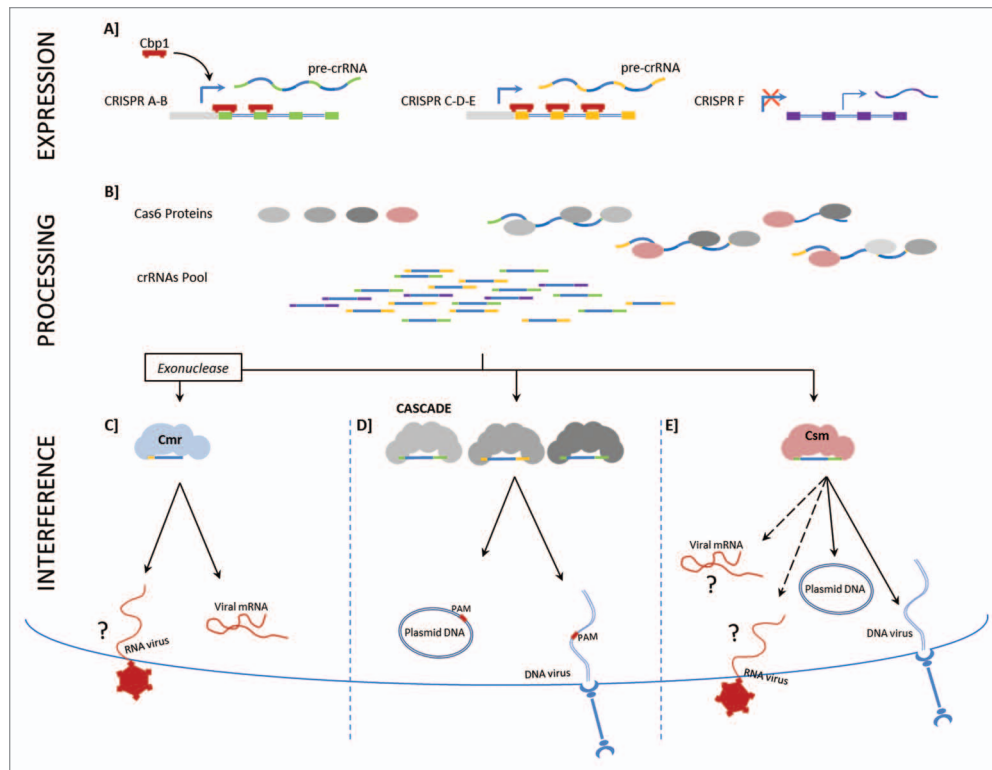


Figure 2. Overview of the CRISPR-Cas system from *Sulfolobus solfataricus* (SSO). CRISPR expression and processing in (A and B) and CRISPR interference in (C–E). CRISPR EXPRESSION: (A) Representation of the different CRISPR loci in *Sulfolobus solfataricus*. Leader transcriptional start sites are indicated. In dark red the CRISPR regulator Cbp1, which binds specifically CRISPR repeats influencing CRISPR transcription.⁷¹ Putative internal transcriptional start sites are reported for locus F. (B) The pre-crRNA is recognized by Cas6 and processed.⁴⁴ The four orthologous Cas6 proteins are colored based on their proximity to the respective Cas module. An unknown exonuclease could be involved in the processing of the crRNAs before these are integrated into the Cmr complex. CRISPR INTERFERENCE: different crRNAs are loaded into the different protein complexes. (C) CRISPR type IIIA: Csm module, due to its similarity with the *S. epidermidis* system its most probable target is dsDNA.¹⁵ Question marks denote possible, but still unknown functions of the complex. (D) CRISPR type IA: The three CASCADE complexes interact with crRNAs coming from different loci targeting extra-chromosomal DNA.⁴⁴ (E) CRISPR type IIIB: the Cmr module binds crRNAs lacking the 3' end of the repeat⁴⁵ cleaving RNAs. Question marks denote putative, but still not identified archaeal RNA viruses.

protospacer DNA in the foreign genetic elements, i.e., adjacent to the segment that is incorporated as a new spacer.

In *S. solfataricus*, active uptake of new spacers was demonstrated after challenging the host with environmental samples. In her study, Erdmann found that only locus C, D and E (CRISPR family I) incorporated new spacers. While the inactivity was expected for locus F (family I), which lacks the Cas1 and Cas2 genes and the leader sequence upstream of the CRISPR locus,^{48,55} it was more surprising that no spacers were acquired by CRISPR loci A and B (family II), which show a complete integration cassette formed by Cas4, Cas1 and Cas2, and which were already shown to be active during the interference process.^{12,14} The newly acquired spacers were incorporated after the first repeat in locus C and D similar to the adaptation mode found in *E. coli*.²⁹ Unexpectedly, new spacers were also integrated at different positions within the CRISPR locus E.²⁸ All these three CRISPR loci (C, D and E) carry the same direct repeat, underlining the possibility that different integration mechanisms could be active in the different CRISPR families.²⁸ Remarkably, almost all of the newly integrated spacers were found to match plasmid-like sequences and no viral spacers; although the presence of plasmids and

viruses in the same culture was demonstrated.²⁸ This result suggests a bias toward the incorporation of plasmid-like sequences in the CRISPR family I, although previous bioinformatic analyses of the spacers located within *S. solfataricus* and *S. islandicus* CRISPR loci do not indicate preferential integration of plasmid-derived sequences.^{26,76}

The incorporation of “exclusively” plasmid-derived spacers could be explained by the presence of a CRISPR-resistant virus and an aggressive conjugative plasmid in the enrichment culture.^{26,28,71,76} It is not difficult to think that the co-evolutionary arms-race between host and invaders could have led to the selection of mechanisms to avoid CRISPR interference or/and to block the incorporation of new spacers.⁷⁷ Integration of the viral genome into the host chromosome or modification of the nucleic acid via e.g., methylation, could represent simple but effective mechanisms to avoid CRISPR spacer integration and interference.

In her experiment, Erdmann reported the integration of hundreds of new spacers into the CRISPR loci increasing our understanding of the protospacer adjacent motif (PAM) in *Sulfolobus*. The mapping of the newly acquired protospacer on the plasmid

genome has confirmed the short di-nucleotide sequence GG for loci C, D and E as PAM, the same di-nucleotide sequence previously identified by bioinformatic analyses.^{26,28,55,71,75,76}

CAS Protein Complexes of *Sulfolobus*

Clusters of genes encoding CRISPR-associated proteins that provide the enzymatic machinery of the system are present in the proximity of CRISPR loci. On the *S. solfataricus* chromosome, two “integration cassettes,” five “interference complexes” as well as other CRISPR-related proteins are encoded near the different loci. Loci A-B, C-D and F carry in their proximity a cas cassette with high similarity to the CASCADE complex found in *E. coli*,⁴⁸ additionally a Csm and Crm cassette (the latter formerly also referred to as RAMP proteins and according to Makarova et al.,³² now referred to as CRISPR type III A or B module) can be found in the vicinity of CRISPR C and F, respectively^{14,48,52} (Fig. 2).

The Type I CRISPR is the most diverse type with six different subtypes and its principle characteristic is the presence of the Cas3 protein, a helicase-nuclease involved in the targeting and degradation of foreign ssDNA.⁷⁸ CRISPR type I was intensively studied in *E. coli* and *P. auruginosa*.^{11,79-82} In those systems, a long precursor CRISPR RNA is processed to crRNA by Cas6, an endonuclease which recognizes and cleaves the pre-crRNA transcript at a particular site within the repeat. Once formed, the crRNA guides a complex of several Cas proteins called CASCADE (CRISPR associated complex for antiviral defense)¹¹ to the target DNA, blocking the invasion of genetic elements that carry DNA sequences (protospacer) complementary to the guide crRNA. In the CRISPR type I, the pairing between crRNA and its target and the presence of a PAM site⁷⁵ are essential conditions for an efficient DNA degradation.^{12,47} The presence of PAM originally identified as being essential for de-novo spacer acquisition was demonstrated to be an essential feature also for the interference process in CRISPR type I systems.⁴⁷ Furthermore, in vivo and in vitro studies have demonstrated the importance of the first seven to eight nucleotides at the 5' end of the crRNA (“seed” sequence) during the recognition process.^{47,81} Mutations in the seed sequence are not tolerated and lead to a loss of immunity. In *S. solfataricus*, three different CASCADE complexes are located in the proximity of CRISPR locus A-B, C-D and F.

Two of the Cas proteins located near the CRISPR locus C and D show high similarity to those of the *E. coli* CASCADE complex. This Cas cluster encodes proteins homologous to CasC (Csa2; SSO1442 also known as Cas7) and CasD (Cas5a, SSO1441), which co-purify in *S. solfataricus* with the processed crRNA and can bind specifically crRNAs and recognize single-stranded DNA when expressed and purified from *E. coli*.⁴⁴ The complex binds crRNAs of 60–70 nt in length and sequencing of the cloned small RNA fragments revealed the presence of full repeat-spacer units with 8 nt of the repeat at the 5'-end of the spacer sequence followed by the unique spacer with the remaining 16–17 nt of the repeat at its 3' end.⁴⁴ Interestingly, the crRNAs that co-purified with the archaeal aCASCADE originated from all the different *S. solfataricus* CRISPR loci and not only from

the adjacent loci C and D, indicating that this complex might be able to exert its activity with all or most crRNAs of the organism.

The Cas7(Csa2)-Cas5a complex (CASCADE) also co-purified with Cas6 (SSO1437), Csa5 (SSO1443) and possibly with Csa4.⁴⁴ Cas6 even seemed to be physically linked with the Cas7-Cas5a complex and was directly involved in the processing of the direct repeat, as previously described for Cas6 proteins from the archaeon *P. furiosus* and from *E. coli*.^{11,43,44}

S. solfataricus possesses four Cas6 homologs located in the proximity of the three putative CASCADE complexes and one in the vicinity of the Csm module.²⁴ One of those (heterologously expressed) Cas6 proteins (SSO2004), which is encoded adjacent to the CASCADE complex of locus F, showed the ability to cleave specifically within the direct repeats in a sequence-specific manner.⁴⁴ The presence of four different Cas6 genes in *S. solfataricus*, primarily located at the beginning or at the end part of the different Cas cassettes, reinforces the observation that each Cas6 protein is linked to one of the different CASCADE and/or Csm complexes. But the presence of crRNAs derived from all the different CRISPR loci in the CASCADE complex analyzed from Lintner and co-workers rises the possibility that a single Cas6 gene can recognize repeats from all different CRISPR loci.⁴⁴ Remarkably, a recent publication has shown that in *S. islandicus* REY15A, three different Cas interference cassette complexes (CRISPR type I and two type IIIB) co-exist with a single Cas6 protein. This observation suggests that in some CRISPR-Cas systems, a single Cas6 protein can interact with different Cas interference complexes or that the pre-crRNA processing by Cas6 takes place without a physical interaction between Cas6 and the interference complexes.⁸³

Additionally, during the isolation of the *Sulfolobus* CASCADE complex, paralogous proteins of Cas7 (Csa2; SSO1399) and Cas5a (SSO1400) genes were co-purified.⁴⁴ The ability of a single CASCADE complex to interact with crRNAs coming from all different CRISPR loci and the possibility of the formation of “mixed” CASCADE complexes, underline the complexity of the *Sulfolobus* CRISPR-Cas system. This complexity could be important to add further versatility and flexibility to the system. The possibility that the different interference complexes could use any spacers located on the CRISPR chromosome might increase the flexibility of the immune system. Furthermore, the presence of multiple putative CASCADE cassettes and Cmr/Csm modules, not only constitutes an extended protection for the cell, but is also a fertile source for the evolution of new CRISPR-Cas variants.

In vitro, the *Sulfolobus* CASCADE showed the ability to bind single-stranded DNA upon interaction with a crRNA, but did not convey DNA cleavage.⁴⁴ However, in vivo studies in *Sulfolobus* demonstrate that CRISPR crRNA-based DNA cleavage indeed occurs.^{12,14}

The presence of several CRISPR types suggests a specialization of the different complexes in targeting different extra-chromosomal elements and/or nucleic acids. The type III CRISPR system has been intensively studied from *Staphylococcus epidermidis* (CRISPR type IIIA),¹⁵ *P. furiosus*¹³ and, recently, *S. solfataricus*⁴⁵ (CRISPR Type IIIB). It is characterized by the presence of Cas10 proteins possibly involved in crRNA processing and nucleic

acid targeting. Furthermore, it usually contains a Cas6 protein shown to be important for crRNA maturation. Interestingly, the CRISPR type IIIA system targets DNA *in vivo* but not RNA,¹⁵ whereas the type IIIB system of the archaea *Pyrococcus* and *Sulfolobus* (Cmr module SSO1986-92) cleaves RNA *in vitro*.^{13,45} Both type IIIA and IIIB have shown no need for PAM sequences to cleave RNA or DNA, respectively.^{13,84} Nevertheless, a recent study on one of the two Cmr complexes of *S. islandicus* REY15A (Cmr- α), which is phylogenetically very distant from the previous Cmr complexes analyzed *in vitro*, demonstrates that type IIIB CRISPR-Cas systems can also target DNA *in vivo*, although its activity seems to be dependent on protospacer transcription.⁸³

The Cmr complex of *Sulfolobus*, like the CASCADE, showed the ability to interact with spacers from many CRISPR loci, but it interacted preferentially with crRNAs transcribed from CRISPR A and D.⁴⁵ Additionally, the crRNAs that co-precipitated with the Cmr complex showed the characteristic 8 nt 5'tag but a shorter, often completely absent 3' handle,⁴⁵ as previously observed in *Pyrococcus furiosus*, although the trimming of the crRNA handle seems to involve a different mechanism in that archaeon.¹³ This result fits with Cas6 cleavage of the repeat followed by exonucleolytic digestion of the spacer 3' handle from an unknown nuclease.⁴⁵

The presence of the crRNA 8 nt 5'tag and an unpaired sequence at the 3' end of the protospacer were pre-requisites for the Cmr-mediated RNA cleavage *in vitro*. Furthermore, no need of PAM sequences in the targeted RNA was necessary. Not only cleavage of the target RNA was observed in *Sulfolobus* but also a concomitant cleavage of crRNA albeit at lower levels. RNA cleavage sites were mapped on AU dinucleotides⁴⁵ and the absence of AU sites in the target RNA abolished cleavage. Interestingly, cleavage of the Cmr complex in the archaeon *Pyrococcus* follows a different mechanism, as the cutting site is located at a fixed position measured from the 3' end of the cognate crRNA.¹³

Targeting the Invader: In Vivo Studies

CRISPR-mediated immunity was recently demonstrated and characterized *in vivo* in *Sulfolobus* cells using recombinant viruses¹⁴ and plasmids,¹² respectively, which contained protospacer sequences compatible with crRNAs of the chromosome.

Spacers of CRISPR locus A and D conferred immunity against extra-chromosomal elements carrying homologous protospacers whereas spacers located in the CRISPR locus F did not.^{12,14} The absence of interference activity of *S. solfataricus* CRISPR F can be correlated with its lower pre-RNA transcription that might in turn be caused by the absence of a leader sequence.^{48,55,56,69} Nevertheless, processed crRNAs of this potentially "inactive" locus were found in the co-purified CASCADE crRNA pool.⁴⁴

The DNA targeting activity of the *Sulfolobus* CRISPR system was quantified and characterized in detail. Both viruses and plasmids carrying perfectly matching protospacers showed

transformation/transfection efficiencies that were orders of magnitudes lower than those of control vectors or were not even detectable.^{12,14} Furthermore, challenging the CRISPR system with perfectly matching protospacers under selective pressure, led to partial deletion of the respective CRISPR locus.¹² Similarly, an artificial miniCRISPR locus encoded on the incoming virus and targeting an endogenous non-essential gene was not stable in culture and underwent extensive recombination that led to loss of the targeting spacers.¹⁴ Taken together, these results indicate that DNA was targeted by the CRISPR-Cas system. Interestingly, few nucleotide mutations between crRNA and protospacers did not strongly abolish the DNA interference.^{12,14} The system tolerated up to four mutations at the 3' half of the crRNA and a more recent study in our laboratory demonstrates that many more mutations are tolerated by the system (Manica et al., submitted). Despite the high tolerance of the CRISPR-Cas system for mutations located at the crRNA 3' half, only six mutations in the crRNA 5' half ("seed" sequence region) are sufficient to abolish the interference (Manica et al., submitted). These findings support the observation made by Semenova and co-workers⁴⁷ in *E. coli*, where the first 8 nt of the crRNA 5' half (seed) were crucial for crRNA-target recognition.

Conclusion

Viruses, plasmids and other mobile genetic elements have been described in most archaea, but occur in extraordinary numbers and diversity in members of the genus *Sulfolobus*. It might therefore not be too surprising that species of this genus also harbor extensive and diverse CRISPR loci that help to defend against or cope with these elements. While it seems ad hoc unfavorable to study specific functions of the archaeal immune system in such a complex genomic context, the studies made in *Sulfolobus* clearly demonstrate the advantages of this complexity: the activity of clusters from different CRISPR families can be directly compared and their interactions studied, as demonstrated for spacer recruitment, recombination of loci, CASCADE complex formation and crRNA sorting. Specificity of proteins from different CRISPR families can be investigated in parallel and both DNA degradation (so far demonstrated only *in vivo*) and RNA degradation (so far only demonstrated *in vitro*) of invading nucleic acids can be studied in the same system. It will be important to obtain more mechanistic details on these functions in the future and to analyze the ecological relevance of the virus defense system for the host organisms and their invading genetic elements. Equally important will be the investigation of CRISPR systems in other genetically and/or biochemically tractable archaea in order to obtain a comprehensive picture of the function of diverse and abundant CRISPR RNAs in archaea.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

- Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* 2009; 136:215-33; PMID:19167326; <http://dx.doi.org/10.1016/j.cell.2009.01.002>
- Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 2009; 10:94-108; PMID:19148191; <http://dx.doi.org/10.1038/nrg2504>
- Hüttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope or hype? *Trends Genet* 2005; 21:289-97; PMID:15851066; <http://dx.doi.org/10.1016/j.tig.2005.03.007>
- Marchfelder A, Fischer S, Brendel J, Stoll B, Maier LK, Jäger D, et al. Small RNAs for defence and regulation in archaea. *Extremophiles* 2012; 16:685-96; PMID:22763819; <http://dx.doi.org/10.1007/s00792-012-0469-5>
- Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nuclear RNAs. *Nat Rev Mol Cell Biol* 2007; 8:209-20; PMID:17318225; <http://dx.doi.org/10.1038/nrm2124>
- Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell* 2009; 136:615-28; PMID:19239884; <http://dx.doi.org/10.1016/j.cell.2009.01.043>
- Alvarado V, Scholthof HB. Plant responses against invasive nucleic acids: RNA silencing and its suppression by plant viral pathogens. *Semin Cell Dev Biol* 2009; 20:1032-40; PMID:19524057; <http://dx.doi.org/10.1016/j.semcdb.2009.06.001>
- Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 2011; 43:880-91; PMID:21925377; <http://dx.doi.org/10.1016/j.molcel.2011.08.022>
- Zhou Y, Xie J. The roles of pathogen small RNAs. *J Cell Physiol* 2011; 226:968-73; PMID:20945366; <http://dx.doi.org/10.1002/jcp.22483>
- 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>
- Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuys 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>
- 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>
- 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>
- Manica A, Zebec Z, Teichmann D, Schleper C. In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol Microbiol* 2011; 80:481-91; PMID:21385233; <http://dx.doi.org/10.1111/j.1365-2958.2011.07586.x>
- 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>
- Rousseau C, Gonnet M, Le Romancer M, Nicolas J. CRISPI: a CRISPR interactive database. *Bioinformatics* 2009; 25:3317-8. PMID:19846435; <http://dx.doi.org/10.1093/bioinformatics/btp586>
- Brodersen P, Voinnet O. The diversity of RNA silencing pathways in plants. *Trends Genet* 2006; 22:268-80; PMID:16567016; <http://dx.doi.org/10.1016/j.tig.2006.03.003>
- Tolia NH, Joshua-Tor L. Slicer and the argonauts. *Nat Chem Biol* 2007; 3:36-43; PMID:17173028; <http://dx.doi.org/10.1038/nchembio848>
- Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 2010; 11:181-90; PMID:20125085; <http://dx.doi.org/10.1038/nrg2749>
- Jansen R, van Embden JD, Gastra W, Schouls LM. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* 2002; 6:23-33; PMID:11883425; <http://dx.doi.org/10.1089/15362310252780816>
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 2005; 151:2551-61; PMID:16079334; <http://dx.doi.org/10.1099/mic.0.28048-0>
- Godde JS, Bickerton A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J Mol Evol* 2006; 62:718-29; PMID:16612537; <http://dx.doi.org/10.1007/s00239-005-0223-z>
- Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 2007; 8:172; PMID:17521438; <http://dx.doi.org/10.1186/1471-2105-8-172>
- 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>
- Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* 2007; 8:R61; PMID:17442114; <http://dx.doi.org/10.1186/gb-2007-8-4-r61>
- 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>
- 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>
- 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>
- 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>
- Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, et al. A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol Microbiol* 2011; 79:484-502; PMID:21219465; <http://dx.doi.org/10.1111/j.1365-2958.2010.07465.x>
- 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>
- 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>
- Swarts DC, Mosterd C, van Passel MW, Brouns SJ. CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* 2012; 7:e55888; PMID:22558257; <http://dx.doi.org/10.1371/journal.pone.0035888>
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; 391:806-11; PMID:9486653; <http://dx.doi.org/10.1038/35888>
- Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 1999; 286:950-2; PMID:10542148; <http://dx.doi.org/10.1126/science.286.5441.950>
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; 409:363-6; PMID:11201747; <http://dx.doi.org/10.1038/35053110>
- 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>
- Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J Mol Biol* 2010; 395:270-81; PMID:19891975; <http://dx.doi.org/10.1016/j.jmb.2009.10.057>
- Pul U, Würm R, Arslan Z, Geissen R, Hofmann N, Wagner R. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol* 2010; 75:1495-512; PMID:20132443; <http://dx.doi.org/10.1111/j.1365-2958.2010.07073.x>
- Westra ER, Pul U, Heidrich N, Jore MM, Lundgren M, Strattmann T, et al. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol* 2010; 77:1380-93; PMID:20659289; <http://dx.doi.org/10.1111/j.1365-2958.2010.07315.x>
- 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>
- Perez-Rodríguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, Ke A, et al. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol* 2011; 79:584-99; PMID:21255106; <http://dx.doi.org/10.1111/j.1365-2958.2010.07482.x>
- 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>
- Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, Naismith JH, et al. Structural and functional characterization of an archaeal CASCADE complex for CRISPR-mediated viral defense. *J Biol Chem* 2011; 286:21643-56; PMID:21507944; <http://dx.doi.org/10.1074/jbc.M111.238485>
- 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>
- Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol Direct* 2011; 6:38; PMID:21756346; <http://dx.doi.org/10.1186/1745-6150-6-38>

47. 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>
48. Shah SA, Garrett RA. CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems. *Res Microbiol* 2011; 162:27-38; PMID:20863886; <http://dx.doi.org/10.1016/j.resmic.2010.09.001>
49. Cavicchioli R. *Archaea: molecular and cellular biology*. ASM Press, 2007
50. Garrett RA, Klenk HP. *Archaea: evolution, physiology, and molecular biology*. Blackwell, 2007
51. Brügger K, Torarinsson E, Redder P, Chen L, Garrett RA. Shuffling of *Sulfolobus* genomes by autonomous and non-autonomous mobile elements. *Biochem Soc Trans* 2004; 32:179-83; PMID:15046567; <http://dx.doi.org/10.1042/BST0320179>
52. Garrett RA, Shah SA, Vestergaard G, Deng L, Gudbergstottir 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>
53. Han D, Krauss G. Characterization of the endonuclease SSO2001 from *Sulfolobus solfataricus* P2. *FEBS Lett* 2009; 583:771-6; PMID:19174159; <http://dx.doi.org/10.1016/j.febslet.2009.01.024>
54. Han D, Lehmann K, Krauss G. SSO1450—a CAS1 protein from *Sulfolobus solfataricus* P2 with high affinity for RNA and DNA. *FEBS Lett* 2009; 583:1928-32; PMID:19427858; <http://dx.doi.org/10.1016/j.febslet.2009.04.047>
55. 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>
56. Shah SA, Hansen NR, Garrett RA. Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem Soc Trans* 2009; 37:23-8; PMID:19143596; <http://dx.doi.org/10.1042/BST0370023>
57. Lintner NG, Frankel KA, Tsutakawa SE, Alsbury DL, Copié V, Young MJ, et al. The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. *J Mol Biol* 2011; 405:939-55; PMID:21093452; <http://dx.doi.org/10.1016/j.jmb.2010.11.019>
58. Zhang J, Kasciukovic T, White MF. The CRISPR associated protein Cas4 Is a 5' to 3' DNA exonuclease with an iron-sulfur cluster. *PLoS ONE* 2012; 7:e47232; PMID:23056615; <http://dx.doi.org/10.1371/journal.pone.0047232>
59. Pina M, Bize A, Forterre P, Prangshvili D. The archaeo-viruses. *FEMS Microbiol Rev* 2011; 35:1035-54; PMID:21569059; <http://dx.doi.org/10.1111/j.1574-6976.2011.00280.x>
60. 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>
61. Wagner M, Berkner S, Ajon M, Driessen AJ, Lipps G, Albers SV. Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. *Biochem Soc Trans* 2009; 37:97-101; PMID:19143610; <http://dx.doi.org/10.1042/BST0370097>
62. Iverson E, Stedman K. A genetic study of SSV1, the prototypical fusellovirus. *Front Microbiol* 2012; 3:200; PMID:22679442; <http://dx.doi.org/10.3389/fmicb.2012.00200>
63. Maaty WS, Steffens JD, Heinemann J, Ortmann AC, Reeves BD, Biswas SK, et al. Global analysis of viral infection in an archaeal model system. *Front Microbiol* 2012; 3:411; PMID:23233852; <http://dx.doi.org/10.3389/fmicb.2012.00411>
64. 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.2011.08.020>
65. Held NL, Herrera A, Cadillo-Quiroz H, Whitaker RJ. CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLoS ONE* 2010; 5; PMID:20927396; <http://dx.doi.org/10.1371/journal.pone.0012988>
66. Torarinsson E, Klenk HP, Garrett RA. Divergent transcriptional and translational signals in Archaea. *Environ Microbiol* 2005; 7:47-54; PMID:15643935; <http://dx.doi.org/10.1111/j.1462-2920.2004.00674.x>
67. Tang TH, Polacek N, Zywicki M, Huber H, Brügger K, Garrett R, et al. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Mol Microbiol* 2005; 55:469-81; PMID:15659164; <http://dx.doi.org/10.1111/j.1365-2958.2004.04428.x>
68. Tang TH, Rozhdetsvensky TS, d'Orval BC, Bortolin ML, Huber H, Charpentier B, et al. RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Res* 2002; 30:921-30; PMID:11842103; <http://dx.doi.org/10.1093/nar/30.4.921>
69. Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R. A single-base resolution map of an archaeal transcriptome. *Genome Res* 2010; 20:133-41; PMID:19884261; <http://dx.doi.org/10.1101/gr.100396.109>
70. Richter H, Zoepfel 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>
71. 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>
72. Peng X, Brügger K, Shen B, Chen L, She Q, Garrett RA. Genus-specific protein binding to the large clusters of DNA repeats (short regularly spaced repeats) present in *Sulfolobus* genomes. *J Bacteriol* 2003; 185:2410-7; PMID:12670964; <http://dx.doi.org/10.1128/JB.185.8.2410-2417.2003>
73. Koerdt A, Orell A, Pham TK, Mukherjee J, Wlodkowski A, Karunakaran E, et al. Macromolecular fingerprinting of *Sulfolobus* species in biofilm: a transcriptomic and proteomic approach combined with spectroscopic analysis. *J Proteome Res* 2011; 10:4105-19; PMID:21761944; <http://dx.doi.org/10.1021/pr2003006>
74. Hyun TK, Uddin MN, Rim Y, Kim JY. Cell-to-cell trafficking of RNA and RNA silencing through plasmodesmata. *Protoplasma* 2011; 248:101-16; PMID:21042816; <http://dx.doi.org/10.1007/s00709-010-0225-6>
75. 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>
76. Held NL, Whitaker RJ. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ Microbiol* 2009; 11:457-66; PMID:19196276; <http://dx.doi.org/10.1111/j.1462-2920.2008.01784.x>
77. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 2013; 493:429-32; PMID:23242138; <http://dx.doi.org/10.1038/nature11723>
78. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* 2011; 30:1335-42; PMID:21343909; <http://dx.doi.org/10.1038/emboj.2011.41>
79. Haurwitz RE, Sternberg SH, Doudna JA. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J* 2012; 31:2824-32; PMID:22522703; <http://dx.doi.org/10.1038/emboj.2012.107>
80. 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>
81. 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>
82. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J Bacteriol* 2012; 194:5728-38; PMID:22885297; <http://dx.doi.org/10.1128/JB.01184-12>
83. 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; 87:1088-99; PMID:23320564; <http://dx.doi.org/10.1111/mmi.12152>
84. 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>